Joint Precoding and Pre-equalization for Faster-Than-Nyquist Transmission over Multipath Fading Channels

Shan Wen, *Student Member, IEEE,* Guanghui Liu, *Senior Member, IEEE,* Chengxiang Liu, Huiyang Qu, Lei Zhang, *Senior Member, IEEE*, and Muhammad Ali Imran, *Senior Member, IEEE*

***Abstract*—Faster-than-Nyquist signaling (FTNS) has emerged as a promising technique to increase communication capacity in bandwidth-limited channels. However, the presence of FTN-induced inter-symbol interference (FTN-ISI) in the received observations, is detrimental to channel estimation (CE) and data detection in terms of computational complexity and performance. This paper copes with these problems by incorporating linear pre-equalization (LPE) and composite precoding formed by linear spectral precoding and Tomlinson-Harashima precoding (THP), into the FTNS. Specifically, LPE completely pre-equalizes the FTN-ISI, while spectral precoding resolves the LPE-caused signal spectral broadening by introducing proper artificial ISI, which is pre-equalized by THP. Channel-induced ISI, as the only ISI component in the observations, is estimated and equalized using the classical frequency-domain low-complexity schemes. We show that there are four advantages of the LPE-aided CE over the CE designed for the FTN transmissions without FTN-ISI pre-equalization, namely lower pilot overhead, simpler yet optimal pilot sequence design, lower mean-squared error of CE, and more robust against the FTN-ISI. Simulation results show that our scheme improves the performance of CE and detection, compared to existing FTN frequency-domain CE and equalization schemes.**

***Index Terms*—Faster-than-Nyquist signaling (FTNS), multi-path fading channel, channel estimation and data detection, pre-equalization, precoding, waveform design.**

I. Introduction

With the explosive growth of connected services in the fifth-generation networks, such as enhanced mobile broadband and massive Internet of Things, the scarcity of frequency band resources becomes even more prominent [1]. Under this circumstance, faster-than-Nyquist signaling (FTNS), as a non-orthogonal spectrally efficient signaling technique, has re-gained increased interest in both academic and industry in recent times after a period of silence [2]–[8]. In the traditional Nyquist transmissions, symbol rate is limited by the Nyquist rate to avoid inter-symbol interference (ISI). Faster-than-Nyquist (FTN) transmissions break through this limit in order to achieve higher capacity, but at the cost of introducing ISI, called FTN-ISI [2], [3], [7].

Numerous schemes have been developed to combat the FTN-ISI and then retrieve the transmitted FTN symbols. These can be roughly classified into three categories: trellis-search [9]–[12]; post-equalization [12]–[18]; hybrid pre- and post-equalization [8], [19]–[26]. In the first two categories, detection and/or FTN-ISI equalization are done at the receiver. The third category performs FTN-ISI pre-equalization at the transmitter. As a result, the FTN-ISI density in the received samples is reduced, and the post-equalization eliminates any residual FTN-ISI. The third category achieves better complexity-performance tradeoffs than the first two [20], [22], [24]. The papers cited above mainly assume additive white Gaussian noise (AWGN) channels, which can be impractical in real environments. When multipath fading channel is taken into consideration, an extra source of interference appears, resulting in a very complex FTN-ISI and channel-induced ISI (channel-ISI) coexisting. To date, there have been some works on the FTN transmission over the ISI channels.

A. Prior Art

Given that the receiver knows the channel impulse response (CIR) perfectly, several frequency-domain equalization (FDE) algorithms for block-wise FTN transmission with cyclic prefix (CP) have been proposed in [13], [27], [28] to combat the combined ISI due to the FTNS and the ISI channels. As the combined ISI affects more symbols than the channel-ISI, the FDE-assisted FTN transceiver requires a longer CP in building the circulant in comparison to its Nyquist counterpart, thereby increasing the transmission overhead. Besides, these FDE algorithms are inadequate for combating serious FTN-ISI and strong residual ISI results in an error floor. Jana *et al*. [20] considered the case in which the CIR is known to the transmitter and put forward two schemes running at the transmitter. One applies Tomlinson-Harashima precoding (THP) [29] to pre-equalize the combined ISI. The other employs linear pre-equalization (LPE) and THP to pre-equalize the FTN-ISI and the channel-ISI, respectively.

Generally, neither the receiver nor the transmitter knows the CIR perfectly. The CIR estimation is a prerequisite to equalization and/or detection. In [30], Sugiura *et al.* considered a pilot cyclic prefixed (PCP) FTN transmitting block structure and presented iterative frequency-domain channel

Shan Wen, Guanghui Liu (Corresponding author), Chengxiang Liu, and Huiyang Qu are with the School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, 611731, China (e-mail: shanwen@std.uestc.edu.cn; guanghuiliu@uestc.edu.cn; {cxliu, hyqu}@std.uestc.edu.cn).

Lei Zhang and Muhammad Ali Imran are with the James Watt School of Engineering, University of Glasgow, Glasgow, G12 8QQ, U.K. (e-mail: {Lei.Zhang, Muhammad.Imran}@glasgow.ac.uk).

estimation (CE) and ISI equalization schemes. However, a good detection performance requires many iterations, which increase the computational complexity and decoding latency dramatically. Also, the pilot overhead is at least twice that of the PCP Nyquist transmission [31]. In [32], Shi *et al.* adopted the superimposed pilot assignment and implemented the iterative frequency-domain joint CE and decoding based on the generalized approximated message passing algorithm. In comparison to [30], [32] requires a much higher complexity due to the lack of conversion from linear convolution to cyclic convolution. Recently, Li *et al.* [33] applied G-to-minus-half (GTMH) pre-equalization [23], [24] to the schemes of [30]. The channel estimator performs better since half of the FTN-ISI is pre-equalized at the transmitter. Nevertheless, GTMH incurs spectral broadening [18], causing the FTN signal to violate the spectrum regulations known as the spectral emission mask (SEM).

It is widely recognized that the choice of pilot sequence is crucial to the channel estimator's performance. In [30], [33], [34], the estimator operates in the presence of FTN-ISI. According to [30], the pilot sequences that are optimal for the Nyquist system, such as Chu sequence [35], would incur a significant performance loss. [34] derived the spectrum property of the optimal pilot sequence for the least-squares (LS) channel estimator. It was found that the optimal pilot sequence depends on the FTN parameters, i.e., shaping pulse and FTN time acceleration factor τ. For the estimator based on the minimum mean squared-error (MMSE) criterion, as far as we know, there is no literature reporting whether the optimal pilot sequences exist and what are their properties and closed-form expressions if they exist.[1] In [30], [33], the nonoptimal pilot sequences are adopted. Pilots are limited to binary phase-shift keying (BPSK) symbols. Then, all possible BPSK sequences are searched exhaustively in order to find a sequence with the lowest mean squared-error (MSE) of CE. Consider the twenty-symbol pilot sequence, there will be 2^{20} possible BPSK sequences. Thus, with a long pilot sequence, the search is rather time-consuming.

B. Contributions and Outline

This paper studies the CE and equalization for the FTN transmissions over the ISI channels. As have seen above, because of the presence of FTN-ISI in the received observations, the pilot overhead is increased and the processes of pilot sequence design, CE, and equalization are complicated. In this study, to keep the receiver structure simple but effective, the LPE proposed in [20] and precoding are incorporated into the FTNS. In contrast to GTMH, LPE pre-equalizes the FTN-ISI completely. Then, the receiver observes only the channel-ISI as if a Nyquist signal transmits through an ISI channel. We estimate and equalize the channel-ISI using the frequency-domain linear scheme. Note that LPE for unprecoded FTN system has been studied by Jana *et al*. [20]. However, they assumed that the CIR is perfectly known at the transmitter and pre-equalized the channel-ISI at the transmitter by THP. How the CE works is not elaborated. Moreover, the spectral broadening issue arising from the LPE is not addressed.

The main contributions of this paper are summarized below:

1) The LPE is introduced to aid the frequency-domain CE and equalization. Since the observations are free of the FTN-ISI, the pilot overhead used to build the circular convolution and then facilitate low-complexity frequency processing, can be reduced by at least half, compared to [30], [33], [34]. CE performance analysis proves that the pilot sequences optimal (from the MMSE perspective) for the Nyquist system now are optimal for our precoded LPE-FTN system. Thus, off-the-shelf pilot sequences, like the Chu sequence, can be directly used without redesign, thereby simplifying the design of pilot sequences. Furthermore, the LPE-aided channel estimator achieves a lower MSE than the existing one designed for the FTN transmission without FTN-ISI pre-equalization. Finally, both the channel estimator and equalizer are more robust against the FTN-ISI.
2) We propose linear spectral precoding to handle the spectral regrowth of the LPE-FTN signal, and the precoded LPE-FTN signal conforms to the given SEM. In this regard, the proposed FTN waveform is practical for use. The ISI due to the spectral precoding is pre-equalized by THP. Additionally, the composite precoding formed by the spectral precoding and THP is optimized for improved reception performance. The resulting optimization is shown to be convex.
3) A comprehensive investigation of FTN transmission over fading channels is carried out by extensive simulations, including measures of bit error rate (BER), power spectral density (PSD) and peak-to-average power ratio (PAPR) of the transmitted signal. We also analyze the computational complexity. In addition, comparisons are made among the schemes in [30], [33] and ours. The comparisons show that: the proposed scheme provides better CE and BER performances than the two benchmark schemes; the proposed scheme does not need complex iterations to obtain acceptable good performance, thus also having an advantage in complexity over [30]; in addition to meeting spectrum regulations, the proposed waveforms have dramatically smaller PAPRs when compared with the waveforms designed in [33].

The rest of this paper is organized as follows. Sec. II introduces the precoded LPE-FTN system model and lists the problems to be solved in this paper. In Sec. III, channel estimator and equalizer aided by the LPE are derived and performance analysis is performed. In Sec. IV, the benefits of the LPE to the CE are investigated. Sec. V first illustrates the spectral precoding based spectrum compression and then devotes to the THP-based pre-equalization and the precoding optimization. Simulation results are presented in Sec. VI. Conclusions are given in Sec. VII.

Notation: we use the upper-case and lower-case boldface letters to denote matrices and vectors, respectively. $\|\mathbf{x}\|_2$ is the ℓ_2-norm of $\mathbf{x}$ and $\mathrm{diag}\{\mathbf{x}\}$ is a diagonal matrix with $\mathbf{x}$ on the main diagonal. The superscripts $(\cdot)^*$, $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$ represent

[1]As will be seen in Sec. IV, it is hard to determine whether the MSE of the MMSE estimator is a convex function over the pilot sequence.

the complex conjugate, transpose and Hermitian transpose, respectively. $\mathbb{E}\{\cdot\}$, $\lfloor\cdot\rfloor$, and $\star$ indicate the expectation, floor and convolution operators, respectively. Let $\mathbf{F}_N$ be the $N\times N$ unitary discrete Fourier transform (DFT) matrix, whose (k,n)th element is given by $N^{-0.5}\exp\left(-j2\pi\left(k-1\right)\left(n-1\right)/N\right)$ for $k,i=0,1,\ldots,N-1$. We denote the kth entry of vector $\mathbf{x}$ as $x\left[k\right]$, and the notation $\left\{x\left[k\right]\right\}_{k=N_1}^{N_2}$ stands for the column-vector $\left[x\left[N_1\right],\ldots,x\left[N_2\right]\right]^{\mathrm{T}}$. Finally, $\mathcal{F}\left\{x\left(t\right)\right\}\triangleq\int x\left(t\right)e^{-j2\pi ft}dt$ means the Fourier transform of signal $x\left(t\right)$, while $\mathcal{Z}\left\{\mathbf{x}\right\}\triangleq\sum_{n=0}^{N-1}x\left[k\right]z^{-k}$ corresponds to the z-transform of a length-N vector $\mathbf{x}$.

II. System Model and Problem Statement

In this section, after establishing the signal model of the precoded LPE-FTN system in a multipath channel scenario, we derive the PSD and the average power of transmitted signal. Finally, the problems to be tackled in this paper are presented.

A. Precoded LPE-FTN System

Fig. 1 depicts the single-carrier precoded LPE-FTN baseband communication system. At the transmitter, the information bits are encoded and fed to the quadrature amplitude modulation (QAM) mapper[2], producing a symbol sequence $\mathbf{a}\triangleq\left\{a\left[n\right]\right\}$ with entry $a\left[n\right]$ taken independently from a unit variance M-ary constellation $\mathcal{A}$. This sequence is segmented into length-$\tilde{N}_s$ blocks $\mathbf{a}_i\triangleq\left\{a\left[i\tilde{N}_s+\kappa\right]\right\}_{\kappa=0}^{\tilde{N}_s-1}$. Each block is precoded, and padded with N_p pilot symbols. Time-division multiplexing (TDM) of pilots and data is considered here. Let $\mathbf{c}_i$ denote the entire precoder output of length N_s, where $N_s\geq\tilde{N}_s$. Accordingly, the padded block is of length $N=N_s+N_p$ and we use $\bar{\mathbf{c}}_i\triangleq\left\{\bar{c}_i\left[k\right]\right\}_{k=0}^{N-1}$ to denote such a block. Then, the blocks $\bar{\mathbf{c}}_i$ are fitted together in turn and the resulting sequence $\bar{\mathbf{c}}\triangleq\left\{\bar{c}\left[n\right]\right\}$ with the nth element $\bar{c}\left[n\right]=\bar{c}_i\left[k\right]$ if $n=iN+k$ enters the LPE unit for the FTN-ISI pre-cancellation. As shown in Fig. 1, the LPE is composed of an adder and a filter $1-F'\left(z\right)$ [20, Fig. 3]. Using length-$(\nu+1)$ vector $\left\{f\left[l\right]\right\}_{l=0}^{\nu}$ with elements $f\left[l\right]$ given by (7) to represent the FTN-ISI tap sequence in the Forney model [3], and denoting its z-transform as $F\left(z\right)$, we have

$$F'\left(z\right)\triangleq\frac{F\left(z\right)}{f\left[0\right]}=1+\frac{1}{f\left[0\right]}\sum\nolimits_{l=1}^{\nu}f\left[l\right]z^{-l}. \tag{1}$$

The nth output $x\left[n\right]$ of the LPE can be written as

$$x\left[n\right]=\bar{c}\left[n\right]-\frac{1}{f\left[0\right]}\sum\nolimits_{l=1}^{\nu}x\left[n-l\right]f\left[l\right], \tag{2}$$

where $x\left[n\right]=0$ for $n<0$. After the LPE, the signal $x\left[n\right]$ is pulse shaped by a unit-energy transmitter filter $\phi\left(t\right)$ to form the FTN signal:

$$s\left(t\right)=\rho_0\sum\nolimits_{n}x\left[n\right]\phi\left(t-n\tau T\right), \tag{3}$$

where T is the reciprocal of the Nyquist rate. To maintain the same average transmit power as in the Nyquist signaling, the FTN signal is normalized by factor ρ_0. It is associated with the specific precoding and the LPE-FTNS as well. Throughout this paper, $\phi\left(t\right)$ with Fourier transform $\psi\left(f\right)=\mathcal{F}\left\{\phi\left(t\right)\right\}$, is fixed as a T-orthogonal root raised cosine (RRC) pulse with roll-off factor β and time support interval $\left[0,40T\right]$, similar to prior FTN research [8]–[15], [19]–[22].

The signal $s\left(t\right)$ traverses an L_c-tap frequency-selective Rayleigh fading channel, which is assumed to be invariant within one block but varies from one block to another. As in [30], [32], the channel is modeled as

$$h_i\left(t\right)=\sum\nolimits_{\ell=0}^{L_c-1}h_i\left[\ell\right]\delta\left(t-\ell\tau T\right), \tag{4}$$

where $\mathbf{h}_i\triangleq\left\{h_i\left[\ell\right]\right\}_{\ell=0}^{L_c-1}$ denotes the CIR in the ith block and $\delta\left(t\right)$ is the Kronecker delta function. Moreover, the overall channel gain is unity.

Let $\mathbf{y}\triangleq\left\{y\left[n\right]\right\}$ denote the sample sequence for the estimation of $\mathbf{a}$. It is generated by passing the noisy channel output signal $r\left(t\right)$ through the whitened matched filter, which consists of a filter matched to $\phi\left(t\right)$ followed by an FTN-rate sampler with perfect timing, and a digital whitening filter with transfer function $1/F^*\left(1/z^*\right)$ [2], [8], [20], [26]. Here we consider the block-wise estimation. To this end, $\mathbf{y}$ is parsed into length-N blocks $\mathbf{y}_i\triangleq\left\{y_i\left[k\right]\right\}_{k=0}^{N-1}$ with elements $y_i\left[k\right]=y\left[iN+k\right]$. Then, the block $\mathbf{y}_i$, which is associated with the channel $\mathbf{h}_i$ and the symbol block $\mathbf{a}_i$, is processed by the channel estimator and detector to extract an channel estimate $\hat{\mathbf{h}}_i$ and a block estimate $\hat{\mathbf{a}}_i$. Finally, $\hat{\mathbf{a}}_i$ and other symbol block estimates are fitted together and delivered to the QAM demapper, which outputs log-likelihood ratios (LLRs) feeding the channel decoder.

Defining the FTN-rate sampled autocorrelation of $\phi\left(t\right)$ as

$$g\left[l\right]=\int\phi\left(t\right)\phi^*\left(t-l\tau T\right)dt, \tag{5}$$

then the sample $y_i\left[k\right]$ can be represented as

$$y_i\left[k\right]=\rho_0\sum_{\ell=0}^{L_c-1}h_i\left[\ell\right]\sum_{l=0}^{\nu}x\left[iN+k-l-\ell\right]f\left[l\right]+\omega_i\left[k\right], \tag{6}$$

where $\omega_i\left[k\right]$ is a zero-mean white Gaussian noise with variance σ_ω^2 due to the noise whitening [18], [20], [26], and

$$g\left[l\right]=f\left[l\right]\star f^*\left[-l\right]. \tag{7}$$

Note that $\left\{f\left[l\right]\right\}_{l=0}^{\nu}$ can be obtained by the inverse z-transform of $F\left(z\right)$, where $F\left(z\right)$ is the solution to the spectral factorization[3] $F\left(z\right)F^*\left(1/z^*\right)=G\left(z\right)$ and $G\left(z\right)\triangleq\mathcal{Z}\left\{\left\{g\left[l\right]\right\}_{l=-\nu}^{\nu}\right\}$. As in [20], [21], [26], $F\left(z\right)$ is induced to be casual and minimum-phase[4].

From (2), it follows that

$$\sum\nolimits_{l=0}^{\nu}x\left[iN+k-l-\ell\right]f\left[l\right]=f\left[0\right]\bar{c}\left[iN+k-\ell\right]. \tag{8}$$

[2]The extension to more generic signal constellations is direct.

[3]A necessary and sufficient condition for the realization of the spectral factorization is $1/(1+\beta)\leq\tau\leq1$ [20].

[4]It can be shown [2], [29] that the 2ν roots of $G\left(z\right)$ have the symmetry that if ζ is a root, $1/\zeta^*$ is also a root. Moreover, $F\left(z\right)$ has the roots $\zeta_1,\ldots,\zeta_\nu$. There are 2^ν possible designs for $F\left(z\right)$. The minimum-phase version of $F\left(z\right)$ can be constructed from the roots with $\left|\zeta_i\right|\leq1,\forall i$.

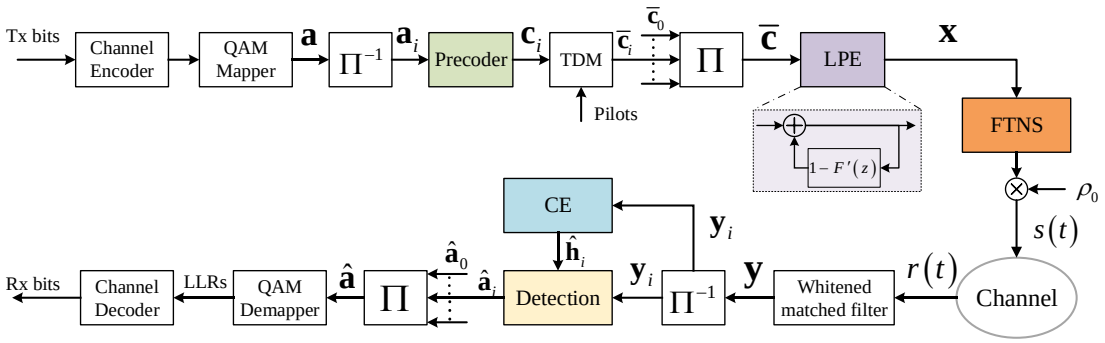


Fig. 1. Block diagram of the precoded LPE-FTN system model. We denote the sequence segmentation by Π^{-1} and the fitting of blocks together by Π.

By substituting (8) into (6) and defining $\rho_1 = \rho_0 f[0]$, we obtain

$$y_i[k] = \rho_1 \sum_{\ell=0}^{L_c-1} h_i[\ell]\, \bar{c}[iN + k - \ell] + \omega_i[k]. \quad (9)$$

Remark 1. From (9), we observe the absence of FTN-ISI in the samples $y_i[k]$. Such an observation model is similar to that for the Nyquist signal transmitting over a multipath channel. However, there are two subtle differences. First, for the LPE-FTN system without precoding, i.e. $\mathbf{a}_i = \mathbf{c}_i$, we have $\rho_1 = \sqrt{\tau}$, where $\rho_0 = \sqrt{\tau}/f[0]$ [20]. By contrast, $\rho_1 = 1$ in the Nyquist system. Second, given a physical channel $h_i(t)$, the CIR is different since $h_i[l] = h_i(l\tau T)$.

B. PSD and Average Power Analysis

Since in practice the pilots are independent of the data and $N_s \gg N_p$, we assume $\mathbf{c}_i = \bar{\mathbf{c}}_i$ (i.e., neglecting the pilots' effects on the PSD) to facilitate the analysis. In [20], the PSD of the LPE-FTN signal with $\mathbf{a}_i = \mathbf{c}_i$, is derived and given by [20, Proposition 4]

$$\Phi_s^{\mathrm{LPE}}(f) = \frac{\sigma_a^2 \rho_1^2}{\tau T} \Xi(f), \quad (10)$$

where σ_a^2 is the variance of QAM symbol $a[k]$ and equals 1 as we consider the unit variance alphabet; $\Xi(f) = \tau T \Psi(f) / \sum_k \Psi\left(f - \frac{k}{\tau T}\right)$ with $\Psi(f)$ equal to $|\psi(f)|^2$. It is worth mentioning that if $a[n] = x[n]$, which implies that the FTN signal generation neither involves the precoding nor LPE, the PSD can be expressed as [2, eq. (3.2)]

$$\Phi_s^{\mathrm{CFTN}}(f) = \frac{\rho_0^2}{\tau T} \Psi(f). \quad (11)$$

In the sequel, we refer to the FTNS with $a[n] = x[n]$ as the classical FTNS (CFTNS).

Now we discuss the case with precoding. The following proposition gives the PSD and average power expressions.

Proposition 1. *The precoded LPE-FTN signal $s(t)$ given in (3) has PSD*

$$\Phi_s(f) = \frac{\rho_1^2}{\tau T} \Phi_c(2\pi f \tau T)\, \Xi(f), \quad (12)$$

where

$$\Phi_c(\lambda) = \sum_q R_c[q] e^{-jq\lambda} \quad (13)$$

with the autocorrelation

$$R_c[q] = \mathbb{E}\{c_i[k]\, c_i^*[k-q]\}. \quad (14)$$

The average power is

$$P_{avg} = \frac{\rho_1^2}{\tau T} \sigma_c^2, \quad (15)$$

where $\sigma_c^2 = R_c[0]$.

Proof. Please see Appendix A. □

Comparisons among $\Phi_s^{\mathrm{LPE}}(f)$, $\Phi_s^{\mathrm{CFTN}}(f)$ and $\Phi_s(f)$ indicate that 1) the PSD shape of the signal generated by the CFTNS is only determined by $\psi(f)$, and independent of τ; 2) precoding and LPE both would modify the PSD shape and render it dependent on τ.

In Fig. 2, the normalized spectra of the 16-QAM FTN signals with different pairs of (β, τ) are plotted to illustrate the modifications introduced by the LPE. We assume $\mathbf{a}_i = \mathbf{c}_i$. As a reference, the PSD shapes of the signals obtained by means of the CFTNS are provided. Note that the legend "CFTNS, (β, τ_1) or (β, τ_2)" in Fig. 2 means that the CFTNS with (β, τ_1) leads to the same PSD shape as that with (β, τ_2). This supports the first indication above. Also shown in this figure is the SEM specified for the fixed microwave backhaul links with 56 MHz bandwidth [36]. From Fig. 2, we observe that the CFTNS-generated signals comply with the SEM. For the LPE-FTN signals, their spectra depend on the pair (β, τ). Moreover, LPE results in a broadening in the transition band and part of the spectra breaches the SEM. It is well understood that FTNS is capable of improving the system capacity without leasing additional frequency spectrum [3], [5]. Therefore, the spectral broadening contradicts the philosophy of the FTNS and neutralizes the FTN gain by a certain amount.

C. Problem Statement

As noted in *Remark 1*, we are dealing with an observation model (9) that is frequently encountered in the Nyquist transmission. Conceptually, the methods of CE and channel-ISI equalization initially developed for the Nyquist transmission can be appropriately adapted for the LPE-FTN system. In this paper, we consider linear frequency-domain methods, and perform CE and equalization separately since such "modular" frequency-domain processing is simple to implement

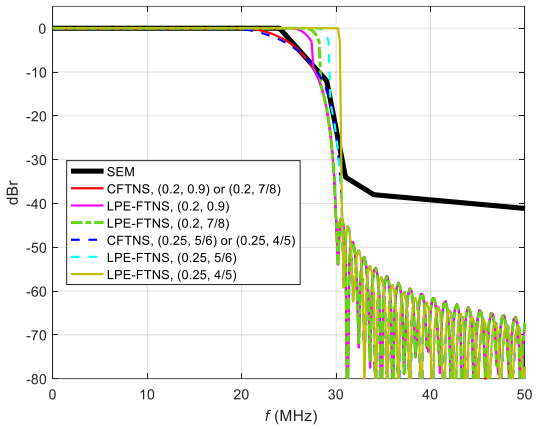


Fig. 2. Spectra of the 16-QAM LPE-FTN signals with different (β, τ) pairs. dBr is the abbreviation of dB relative.

and prevalent in wideband wireless communications with Nyquist signaling, although there also exist joint estimation-equalization receivers in the literature (see, e.g., [32], [37] and the references therein). In these backgrounds, it is unknown

(i) whether there is a CE performance loss for the LPE-FTN system, compared to the Nyquist system? if the loss exists, how much it is and how it is related to the FTN parameters?

(ii) what benefits will the LPE bring to the frequency-domain CE?

In addition to these two problems, this work also intends to address the spectral broadening induced by the LPE. To be specific, we utilize linear spectral precoding (called the time-domain spectral shaping (TDSS) hereafter) to manage the correlation between the symbols $c_i[k]$ in such a way that the precoded LPE-FTN waveforms conform to a given SEM. It is a viable solution since the term $\Phi_c(2\pi f\tau T)$ contributes to the PSD shape, as shown in (12). For the TDSS, we need to solve two problems:

(iii) how to design it to obtain SEM-compliant LPE-FTN waveforms,

(iv) how to mitigate the correlation or equivalently the TDSS-induced ISI (TDSS-ISI)?

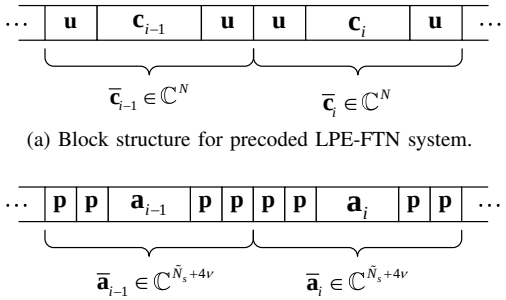
III. LPE-Aided Frequency-Domain Channel Estimation and Equalization

In this section, the LPE-aided frequency-domain channel-ISI estimator and equalizer are derived. To address the *Problem* (i), the CE performance in terms of MSE is analyzed, and the optimal pilot sequences that achieve the lower bound of the MSE are deduced.

A. LPE-Aided Channel Estimation and Equalization

We consider the PCP FTN block structure, as shown in Fig. 3 (a). Specifically, the block $\mathbf{u} = \{u[k]\}_{k=0}^{\tilde{N}_p-1}$ consists of $\tilde{N}_p$ $(= N_p/2)$ pilots. Moreover, $\tilde{N}_p \geq L_c$ is assumed. The ith transmitting block is $\bar{\mathbf{c}}_i = \left[\begin{array}{ccc}\mathbf{u}^{\mathrm{T}} & \mathbf{c}_i^{\mathrm{T}} & \mathbf{u}^{\mathrm{T}}\end{array}\right]^{\mathrm{T}}$. Note that the ith received block $\mathbf{y}_i$ has a circulant-matrix-based presentation:

$$\mathbf{y}_i = \rho_1 \mathbf{H}_i \left[\begin{array}{ccc}\mathbf{u}^{\mathrm{T}} & \mathbf{c}_i^{\mathrm{T}} & \mathbf{u}^{\mathrm{T}}\end{array}\right]^{\mathrm{T}} + \mathbf{w}_i, \tag{16}$$



(a) Block structure for precoded LPE-FTN system.

(b) Block structure studied in [30].

Fig. 3. Transmitting block structures.

where $\mathbf{H}_i$ is an $N \times N$ circulant matrix defined by its first column $\mathbf{h}'_i \triangleq \left[\mathbf{h}_i^{\mathrm{T}}, \mathbf{0}_{1\times(N-L_c)}\right]^{\mathrm{T}}$, and $\mathbf{w}_i \triangleq \{\omega_i[k]\}_{k=0}^{N-1}$. To prove (16), we use the relation $\bar{c}_i[k] = \bar{c}[iN+k]$, where $k = 0, 1, \ldots, N-1$, and replace the term $\bar{c}[iN+k-\ell]$ in (9) by

$$\bar{c}[iN+k-\ell] = \begin{cases} \bar{c}_i[k-\ell], & \text{if } k-\ell \geq 0 \\ \bar{c}_{i-1}[N+k-\ell], & \text{if } k-\ell < 0. \end{cases} \tag{17}$$

From Fig. 3 (a), we observe that $\bar{c}_{i-1}[N+k-\ell] = \bar{c}_i[N+k-\ell]$ for $-\tilde{N}_p \leq k-\ell < 0$ since $\tilde{N}_p \geq L_c$. Thus $\bar{c}[iN+k-\ell] = \bar{c}_i[\langle k-\ell\rangle_N]$, where $\langle k-\ell\rangle_N$ denotes $k-\ell$ modulo N, and (9) becomes

$$y_i[k] = \rho_1 \sum_{\ell=0}^{L_c-1} h_i[\ell]\, \bar{c}_i[\langle k-\ell\rangle_N] + \omega_i[k]. \tag{18}$$

It follows that $\mathbf{y}_i$ can be written as (16).

The first $\tilde{N}_p$ samples of $\mathbf{y}_i$, represented by the vector $\mathbf{y}_{i,p} \triangleq \{y_i[k]\}_{k=0}^{\tilde{N}_p-1}$, are used for the CE. For $0 \leq k \leq \tilde{N}_p - 1$, substituting for $\bar{c}_i[k-\ell]$ in (17) by $u[k-\ell]$ and $\bar{c}_{i-1}[N+k-\ell]$ by $u\left[\tilde{N}_p+k-\ell\right]$, we have

$$\bar{c}[iN+k-\ell] = \begin{cases} u[k-\ell], & \text{if } k-\ell \geq 0 \\ u\left[\tilde{N}_p+k-\ell\right], & \text{if } k-\ell < 0. \end{cases} \tag{19}$$

Then, $y_i[k]$ $(k = 0, 1, \ldots, \tilde{N}_p - 1)$ is given by

$$y_i[k] = \rho_1 \sum_{\ell=0}^{L_c-1} h_i[\ell]\, u\left[\langle k-\ell\rangle_{\tilde{N}_p}\right] + \omega_i[k] \tag{20}$$

or in a matrix form

$$\mathbf{y}_{i,p} = \rho_1 \mathbf{H}_{i,p}\mathbf{u} + \mathbf{w}_{i,p}, \tag{21}$$

where $\mathbf{H}_{i,p}$ is an $\tilde{N}_p \times \tilde{N}_p$ circulant matrix with $\mathbf{h}_{i,p} \triangleq \left[\mathbf{h}_i^{\mathrm{T}}, \mathbf{0}_{1\times(\tilde{N}_p-L_c)}\right]^{\mathrm{T}}$ as its first column, and $\mathbf{w}_{i,p} \triangleq \{\omega_i[k]\}_{k=0}^{\tilde{N}_p-1}$.

To determine the frequency-domain channel estimator, we first apply $\tilde{N}_p$-point DFT to both sides of (21) and obtain

$$\mathbf{y}_{i,p,f} \triangleq \mathbf{F}_{\tilde{N}_p}\mathbf{y}_{i,p} = \rho_1 \mathbf{\Lambda}_{i,p}\mathbf{F}_{\tilde{N}_p}\mathbf{u} + \mathbf{w}_{i,p,f}. \tag{22}$$

In (22), $\mathbf{w}_{i,p,f} = \mathbf{F}_{\tilde{N}_p}\mathbf{w}_{i,p}$ and $\mathbf{\Lambda}_{i,p}$ is a diagonal matrix given by $\mathbf{\Lambda}_{i,p} = \mathrm{diag}\{\mathbf{h}_{i,p,f}\}$, where

$$\mathbf{h}_{i,p,f} = \sqrt{\tilde{N}_p}\mathbf{F}_{\tilde{N}_p}\mathbf{h}_{i,p} \tag{23}$$

is the channel frequency response. Defining the diagonal matrix $\mathbf{\Lambda_u}$ with $\mathbf{u}_f \triangleq \mathbf{F}_{\tilde{N}_p}\mathbf{u}$ on the main diagonal, we can write (22) as

$$\mathbf{y}_{i,p,f} = \rho_1 \mathbf{\Lambda_u} \mathbf{h}_{i,p,f} + \mathbf{w}_{i,p,f}. \tag{24}$$

There exist two commonly used methods, namely LS and MMSE, to obtain an estimate of $\mathbf{h'}_{i,p}$ based on (24). For the LS method, we have

$$\hat{\mathbf{h}}^{\mathrm{LPE}}_{i,p,\mathrm{LS}} = \frac{1}{\rho_1\sqrt{\tilde{N}_p}}\mathbf{F}^{\mathrm{H}}_{\tilde{N}_p}\mathbf{\Lambda}^{-1}_{\mathbf{u}}\mathbf{y}_{i,p,f}. \tag{25}$$

When employing the MMSE estimator,

$$\hat{\mathbf{h}}^{\mathrm{LPE}}_{i,p,\mathrm{MMSE}} = \frac{1}{\sqrt{\tilde{N}_p}}\mathbf{F}^{\mathrm{H}}_{\tilde{N}_p}\mathbf{W}^{\mathrm{LPE}}_{p,\mathrm{MMSE}}\mathbf{y}_{i,p,f}, \tag{26}$$

where

$$\mathbf{W}^{\mathrm{LPE}}_{p,\mathrm{MMSE}} = \rho_1^{-1}\mathbf{Q}_{\mathrm{LPE}}\mathbf{\Lambda}^{-1}_{\mathbf{u}}. \tag{27}$$

In (27), $\mathbf{Q}_{\mathrm{LPE}} = \mathbf{R}_{\mathbf{h},f}\left[\mathbf{R}_{\mathbf{h},f} + \frac{\sigma^2_\omega}{\rho_1^2}\left(\mathbf{\Lambda}^{\mathrm{H}}_{\mathbf{u}}\mathbf{\Lambda}_{\mathbf{u}}\right)^{-1}\right]^{-1}$ with $\mathbf{R}_{\mathbf{h},f}$ being the autocorrelation matrix of $\mathbf{h}_{i,p,f}$, defined as

$$\mathbf{R}_{\mathbf{h},f} = \mathbb{E}\left\{\mathbf{h}_{i,p,f}\mathbf{h}^{\mathrm{H}}_{i,p,f}\right\} = \tilde{N}_p\mathbf{F'}_{\tilde{N}_p}\mathbf{R}_{\mathbf{h}}\mathbf{F'}^{\mathrm{H}}_{\tilde{N}_p}, \tag{28}$$

where $\mathbf{R}_{\mathbf{h}} \triangleq \mathbb{E}\left\{\mathbf{h}_i\mathbf{h}_i^{\mathrm{H}}\right\}$ and $\mathbf{F'}_{\tilde{N}_p}$ is the first L_c columns of the matrix $\mathbf{F}_{\tilde{N}_p}$.

The estimate $\hat{\mathbf{h}}_i$ is obtained by taking the first L_c elements of $\hat{\mathbf{h}}^{\mathrm{LPE}}_{i,p,\mathrm{LS}}$ or $\hat{\mathbf{h}}^{\mathrm{LPE}}_{i,p,\mathrm{MMSE}}$ and then used for the channel-ISI removal. Owing to the circulant-matrix-based structure in (16), the MMSE-FDE is apdoted. The estimate of $\bar{\mathbf{c}}_i$ is

$$\hat{\bar{\mathbf{c}}}_i = \mathbf{F}^{\mathrm{H}}_N\mathbf{W}^{\mathrm{LPE}}_{d,\mathrm{MMSE}}\mathbf{F}_N\mathbf{y}_i, \tag{29}$$

where

$$\mathbf{W}^{\mathrm{LPE}}_{d,\mathrm{MMSE}} = \frac{1}{\rho_1}\left(\hat{\mathbf{\Lambda}}^{\mathrm{H}}\hat{\mathbf{\Lambda}} + \frac{\sigma^2_\omega}{\rho_1^2\sigma_c^2}\mathbf{I}_N\right)^{-1}\hat{\mathbf{\Lambda}}^{\mathrm{H}} \tag{30}$$

denotes the frequency-domain MMSE equalizer, and $\hat{\mathbf{\Lambda}} = \sqrt{N}\mathrm{diag}\left\{\mathbf{F}_N\hat{\mathbf{h}}_i\right\}$ with $\hat{\mathbf{h}}'_i = \left[\hat{\mathbf{h}}_i^{\mathrm{T}}, \mathbf{0}_{1\times(N-L_c)}\right]^{\mathrm{T}}$. The estimate $\hat{\mathbf{c}}_i$ of $\mathbf{c}_i$ corresponds to the center N_s symbols of $\hat{\bar{\mathbf{c}}}_i$.

B. CE MSE Analysis and Pilot Sequence Design

We design the pilot sequence from the MMSE perspective. The MSE for the LS estimator is given by

$$\begin{aligned}\varepsilon^{\mathrm{LPE}}_{\mathrm{LS}} &= \frac{1}{\tilde{N}_p}\mathbb{E}\left\{\left\|\mathbf{h}_{i,p,f} - \hat{\mathbf{h}}^{\mathrm{LPE}}_{i,p,f,\mathrm{LS}}\right\|_2^2\right\} \\ &= \frac{\mathbb{E}\left\{\left\|\mathbf{\Lambda}^{-1}_{\mathbf{u}}\mathbf{w}_{i,p,f}\right\|_2^2\right\}}{\rho_1^2\tilde{N}_p} = \frac{\sigma^2_\omega}{\rho_1^2\tilde{N}_p}\mathrm{tr}\left[\left(\mathbf{\Lambda}^{\mathrm{H}}_{\mathbf{u}}\mathbf{\Lambda}_{\mathbf{u}}\right)^{-1}\right]\end{aligned} \tag{31}$$

where $\hat{\mathbf{h}}^{\mathrm{LPE}}_{i,p,f,\mathrm{LS}} = \sqrt{\tilde{N}_p}\mathbf{F}_{\tilde{N}_p}\hat{\mathbf{h}}^{\mathrm{LPE}}_{i,p,\mathrm{LS}}$. Under the power constraint of $\|\mathbf{u}\|_2^2 = \|\mathbf{u}_f\|_2^2 = \tilde{N}_p c_0$, where c_0 is positive constant, the vector $\mathbf{u}_f$ having the property [31, eq. (5)]

$$|u_f[0]| = \cdots = \left|u_f\left[\tilde{N}_p - 1\right]\right| = \sqrt{c_0}, \tag{32}$$

minimizes $\varepsilon_{p,\mathrm{LS}}$. Substituting (32) into (31), the lower bound of $\varepsilon^{\mathrm{LPE}}_{\mathrm{LS}}$, denoted as $\varepsilon^{\mathrm{LPE}}_{\mathrm{LS,min}}$, equals

$$\varepsilon^{\mathrm{LPE}}_{\mathrm{LS,min}} = \sigma^2_\omega/\left(\rho_1^2 c_0\right). \tag{33}$$

There are many sequences that can achieve the property given in (32), from the PAPR point of view, we adopt the Chu sequences [35] as they also have constant magnitude in the time domain. Particularly, the length-$\tilde{N}_p$ Chu sequence with entries $u_{\mathrm{Chu}}[n]$ for $n = 1, \ldots, \tilde{N}_p$ is given by

$$u_{\mathrm{Chu}}[n] = \begin{cases} e^{j\pi k_0 n^2/\tilde{N}_p}, & \text{for even } \tilde{N}_p \\ e^{j\pi k_0 n(n+1)/\tilde{N}_p}, & \text{for odd } \tilde{N}_p \end{cases} \tag{34}$$

where k_0 and $\tilde{N}_p$ are relatively prime.

Defining $\hat{\mathbf{h}}^{\mathrm{LPE}}_{i,p,f,\mathrm{MMSE}} = \sqrt{\tilde{N}_p}\mathbf{F}_{\tilde{N}_p}\hat{\mathbf{h}}^{\mathrm{LPE}}_{i,p,\mathrm{MMSE}}$, the MSE for the MMSE estimator, is computed as (35), shown at the top of next page. The first equality in (35a) is due to the orthogonality principle $\mathbb{E}\left\{\hat{\mathbf{h}}^{\mathrm{LPE}}_{i,p,f,\mathrm{MMSE}}\left(\mathbf{h}_{i,p,f} - \hat{\mathbf{h}}^{\mathrm{LPE}}_{i,p,f,\mathrm{MMSE}}\right)^{\mathrm{H}}\right\} = \mathbf{0}_{\tilde{N}_p}$. In (35b), we assume that $\mathbf{R}_{\mathbf{h},f}$ is positive definite.[5] (35c) follows from the singular value decomposition (SVD) of matrix $\mathbf{R}_{\mathbf{h},f}$, i.e., $\mathbf{R}_{\mathbf{h},f} = \mathbf{U}\mathbf{\Lambda}_{\mathbf{h}}\mathbf{U}^{\mathrm{H}}$, where $\mathbf{U}$ is a unitary matrix and $\mathbf{\Lambda_h}$ is a diagonal matrix having the eigenvalues $\lambda_n\left(n = 0, 1, \ldots, \tilde{N}_p - 1\right)$ of $\mathbf{R}_{\mathbf{h},f}$ as its diagonal elements. The fact that $\mathrm{tr}\left(\mathbf{U}\mathbf{X}\mathbf{U}^{\mathrm{H}}\right) = \mathrm{tr}\left(\mathbf{X}\right)$ when $\mathbf{U}$ is a unitary matrix, is used in (35d). In (35e), the inequality [38, eq. (20)]

$$\begin{aligned}&\mathrm{tr}\left[\left(\mathbf{\Lambda}^{-1}_{\mathbf{h}} + \mathbf{\Lambda}^{-1}_{\mathbf{h}}\mathbf{U}^{\mathrm{H}}\tfrac{\sigma^2_\omega}{\rho_1^2}\left(\mathbf{\Lambda}^{\mathrm{H}}_{\mathbf{u}}\mathbf{\Lambda}_{\mathbf{u}}\right)^{-1}\mathbf{U}\mathbf{\Lambda}^{-1}_{\mathbf{h}}\right)^{-1}\right] \\ &\le \textstyle\sum_{n=0}^{\tilde{N}_p-1}\frac{\lambda_n^2}{\lambda_n + \sigma^2_\omega/\left(\rho_1^2 c_0\right)}\end{aligned} \tag{36}$$

with equality if the pilot sequence $\mathbf{u}$ satisfies (32), is invoked.

In what follows, we investigate the effects of τ on the MSE when $\mathbf{u}$ satisfies (32) and subsequently give two remarks.

Proposition 2. *When the optimal pilot sequence is used, the MSE $\varepsilon^{\mathrm{LPE}}_{\mathrm{LS,min}}$ or $\varepsilon^{\mathrm{LPE}}_{\mathrm{MMSE,min}}$ for the LPE-FTN system without precoding (i.e., $\mathbf{a}_i = \mathbf{c}_i$), increases monotonically as τ decreases.*

Proof. Given the pilot block $\mathbf{u}$ satisfying (32) and noise level σ^2_ω, $\varepsilon^{\mathrm{LPE}}_{\mathrm{LS,min}}$ and $\varepsilon^{\mathrm{LPE}}_{\mathrm{MMSE,min}}$ depend on ρ_1. Recall from *Remark 1* that $\rho_1 = \sqrt{\tau}$ in the LPE-FTN system without precoding, then it can be readily proved that $\varepsilon^{\mathrm{LPE}}_{\mathrm{LS,min}}$ and $\varepsilon^{\mathrm{LPE}}_{\mathrm{MMSE,min}}$ increase monotonically as τ decreases. □

Remark 2. As illustrated in Sec. II-A, ρ_1 is associated with the specific precoding and the LPE-FTNS as well. If $\rho_1 < 1$ and ρ_1 increases monotonically upon increasing τ, then Prop. 3 holds for the precoded LPE-FTN system.

Remark 3. Since the two frequency-domain channel estimators represented by (25) and (26) both are suitable for the Nyquist transmission just by setting $\rho_1 = 1$, with Prop. 3, we conclude that the LPE-FTN transmission without precoding has inferior CE performance to the Nyquist transmission.

With *Remark 3*, in the following proposition, the MSE performance gap between the precoded LPE-FTN and the

[5] In [38, Appendix A], the rationality of the assumption has been proved.

$$\varepsilon_{\mathrm{MMSE}}^{\mathrm{LPE}} = \frac{1}{\tilde{N}_p}\mathbb{E}\left\{\left\|\mathbf{h}_{i,p,f} - \hat{\mathbf{h}}_{i,p,f,\mathrm{MMSE}}^{\mathrm{LPE}}\right\|^2\right\} = \frac{1}{\tilde{N}_p}\mathrm{tr}\left[\mathbb{E}\left\{\left(\mathbf{h}_{i,p,f} - \hat{\mathbf{h}}_{i,p,f,\mathrm{MMSE}}^{\mathrm{LPE}}\right)\left(\mathbf{h}_{i,p,f} - \hat{\mathbf{h}}_{i,p,f,\mathrm{MMSE}}^{\mathrm{LPE}}\right)^{\mathrm{H}}\right\}\right]$$

$$= \frac{1}{\tilde{N}_p}\mathrm{tr}\left[\mathbb{E}\left\{\left(\mathbf{h}_{i,p,f} - \hat{\mathbf{h}}_{i,p,f,\mathrm{MMSE}}^{\mathrm{LPE}}\right)\mathbf{h}_{i,p,f}^{\mathrm{H}}\right\}\right] = \frac{1}{\tilde{N}_p}\mathrm{tr}\left[\mathbf{R}_{\mathbf{h},f} - \mathbf{Q}_{\mathrm{LPE}}\mathbf{R}_{\mathbf{h},f}\right] \tag{35a}$$

$$= \frac{1}{\tilde{N}_p}\mathrm{tr}\left[\mathbf{R}_{\mathbf{h},f} - \left(\mathbf{R}_{\mathbf{h},f}^{-1} + \mathbf{R}_{\mathbf{h},f}^{-1}\frac{\sigma_\omega^2}{\rho_1^2}\left(\mathbf{\Lambda}_{\mathbf{u}}^{\mathrm{H}}\mathbf{\Lambda}_{\mathbf{u}}\right)^{-1}\mathbf{R}_{\mathbf{h},f}^{-1}\right)^{-1}\right] \tag{35b}$$

$$= \frac{1}{\tilde{N}_p}\mathrm{tr}\left[\mathbf{U}\mathbf{\Lambda}_{\mathbf{h}}\mathbf{U}^{\mathrm{H}} - \left(\mathbf{U}\mathbf{\Lambda}_{\mathbf{h}}^{-1}\mathbf{U}^{\mathrm{H}} + \mathbf{U}\mathbf{\Lambda}_{\mathbf{h}}^{-1}\mathbf{U}^{\mathrm{H}}\frac{\sigma_\omega^2}{\rho_1^2}\left(\mathbf{\Lambda}_{\mathbf{u}}^{\mathrm{H}}\mathbf{\Lambda}_{\mathbf{u}}\right)^{-1}\mathbf{U}\mathbf{\Lambda}_{\mathbf{h}}^{-1}\mathbf{U}^{\mathrm{H}}\right)^{-1}\right] \tag{35c}$$

$$= \frac{1}{\tilde{N}_p}\mathrm{tr}\left[\mathbf{\Lambda}_{\mathbf{h}} - \left(\mathbf{\Lambda}_{\mathbf{h}}^{-1} + \mathbf{\Lambda}_{\mathbf{h}}^{-1}\mathbf{U}^{\mathrm{H}}\frac{\sigma_\omega^2}{\rho_1^2}\left(\mathbf{\Lambda}_{\mathbf{u}}^{\mathrm{H}}\mathbf{\Lambda}_{\mathbf{u}}\right)^{-1}\mathbf{U}\mathbf{\Lambda}_{\mathbf{h}}^{-1}\right)^{-1}\right] \tag{35d}$$

$$\geq \frac{1}{\tilde{N}_p}\sum\nolimits_{n=0}^{\tilde{N}_p-1}\left(\lambda_n - \frac{\lambda_n^2}{\lambda_n + (\sigma_\omega^2/\rho_1^2 c_0)}\right) = \frac{1}{\tilde{N}_p}\sum\nolimits_{n=0}^{\tilde{N}_p-1}\frac{\lambda_n}{(\lambda_n c_0\rho_1^2/\sigma_\omega^2)+1} = \varepsilon_{\mathrm{MMSE,min}}^{\mathrm{LPE}} \tag{35e}$$

Nyquist transmissions is quantified in terms of signal-to-noise ratio (SNR), defined as $1/\sigma_\omega^2$.[6]

Proposition 3. *To achieve a given value of CE MSE, the difference in the required SNR is about* $10\log_{10}\rho_1$ *dB.*

Proof. Note that $\varepsilon_{\mathrm{LS}}^{\mathrm{LPE}}$ or $\varepsilon_{\mathrm{MMSE}}^{\mathrm{LPE}}$ can be regarded as the CE MSE of the Nyquist transmission under the noise level σ_ω^2/ρ_1. Then, the SNR difference can be easily identified. □

IV. The Benefits for Frequency-Domain CE from the LPE

In this section, we address the *Problem* (ii). The frequency-domain CE in [30] is thoroughly investigated and compared with the LPE-aided CE. We illustrate the benefits from three perspectives, namely lower pilot overhead, simpler yet optimal pilot sequence design, and lower CE MSE.

The transmitting block structure studied in [30, Fig. 2] is shown in Fig. 3 (b). $\mathbf{p} \triangleq [p[0], p[1], \ldots, p[\nu-1]]^{\mathrm{T}}$ denotes the FTN pilot sequence and $\bar{\mathbf{a}}_i = \left[\mathbf{p}^{\mathrm{T}}, \mathbf{p}^{\mathrm{T}}, \mathbf{a}_i^{\mathrm{T}}, \mathbf{p}^{\mathrm{T}}, \mathbf{p}^{\mathrm{T}}\right]^{\mathrm{T}}$ constructs a transmitting block. Thus, the number of pilots in a transmitting block is $N_p = 4\nu$. Recall that $N_p = 2\tilde{N}_p$ in the LPE-aided CE scheme. As suggested in [30], the value of ν should be sufficiently larger than the delay spread L_c. $\tilde{N}_p$, however, has a minimum value of L_c. As a result, the LPE-aided CE scheme has an advantage of reducing the pilot overhead by at least half.

At the receiver, the first ν samples of the ith received block that corresponds to $\bar{\mathbf{a}}_i$ are used for the CE. According to [30, eq. (31)], the frequency-domain presentation of the first ν samples can be written in matrix form as

$$\mathbf{y}_{i,p,f}^{\mathrm{w/o\,pre-eq}} = \sqrt{\tau}\mathbf{\Gamma}\mathbf{h}'_{i,p,f} + \boldsymbol{\eta}_{i,p,f}. \tag{37}$$

Using $\mathbf{F}_\nu$ to denote the ν-point DFT matrix and defining the vector $\mathbf{g}_p \triangleq [g[-\nu] + g[0] + g[\nu], g[1] + g[-\nu+1], \ldots, g[\nu-1] + g[-1]]^{\mathrm{T}}$ of length ν, then the diagonal matrix $\mathbf{\Gamma} = \mathrm{diag}\{\mathbf{p}_f\mathbf{g}_{p,f}\}$, where $\mathbf{p}_f = \mathbf{F}_\nu\mathbf{p}$ and $\mathbf{g}_{p,f} = \sqrt{\nu}\mathbf{F}_\nu\mathbf{g}_p$, and $\mathbf{h}'_{i,p,f} = \sqrt{\nu}\mathbf{F}_\nu\left[\mathbf{h}_i^{\mathrm{T}}, \mathbf{0}_{1\times(\nu-L_c)}\right]^{\mathrm{T}}$. Moreover, $\boldsymbol{\eta}_{i,p,f} = \mathbf{F}_\nu\boldsymbol{\eta}_{i,p}$, where $\boldsymbol{\eta}_{i,p}$ is a length-ν colored noise sequence with zero-mean and autocorrelation function $\mathbb{E}\left\{\eta_{i,p}[n]\,\eta_{i,p}^*[n-l]\right\} = \sigma_\omega^2 g[l]$.

In the following, we examine the CE schemes using the LS and MMSE criteria. Note that in [30], Sugiura *et al*. only presented the channel estimator based on the MMSE criterion, while the authors in [34] derived the MSE and the optimal pilot sequences for the LS estimator under the assumption of white noise, which is not the case in the Ungerboeck model [2]. We begin with the LS estimator. Similar to (31), the CE MSE of the LS estimator is computed as

$$\begin{aligned}\varepsilon_{\mathrm{LS}} &= \frac{1}{\tau\nu}\mathbb{E}\left\{\left\|\mathbf{\Gamma}^{-1}\boldsymbol{\eta}_{i,p,f}\right\|^2\right\} \\ &= \frac{\sigma_\omega^2}{\tau\nu}\mathrm{tr}\left[\mathbf{\Gamma}^{-1}\mathbf{F}_\nu\mathbf{R}_{\boldsymbol{\eta}}\mathbf{F}_\nu^{\mathrm{H}}\left(\mathbf{\Gamma}^{\mathrm{H}}\right)^{-1}\right] \\ &= \frac{\sigma_\omega^2}{\tau\nu}\sum\nolimits_{n=0}^{\nu-1}\Phi[n]\,/\left(|p_f[n]|^2|g_{p,f}[n]|^2\right).\end{aligned} \tag{38}$$

In (38), the second equality is due to $\mathbb{E}\left\{\boldsymbol{\eta}_{i,p}\boldsymbol{\eta}_{i,p}^{\mathrm{H}}\right\} = \sigma_\omega^2\mathbf{R}_{\boldsymbol{\eta}}$, where $\mathbf{R}_{\boldsymbol{\eta}}$ is a symmetric Toeplitz matrix with the first row being $[g[0], g[1], \ldots, g[\nu-1]]$. $\Phi[n]$ is the nth diagonal element of the matrix $\mathbf{F}_\nu\mathbf{R}_{\boldsymbol{\eta}}\mathbf{F}_\nu^{\mathrm{H}}$ and equal to [30, eq. (12)]

$$\Phi[n] = \frac{1}{\nu}\sum_{m_1=0}^{\nu-1}\sum_{m_2=0}^{\nu-1} g[m_1 - m_2]\, e^{j\frac{2\pi(m_1-m_2)n}{\nu}}, \tag{39}$$

while $p_f[n]$ and $g_{p,f}[n]$ are the nth element of $\mathbf{p}_f$ and $\mathbf{g}_{p,f}$, respectively.

Proposition 4. *Under the pilot power constraint* $\|\mathbf{p}\|_2^2 = \|\mathbf{p}_f\|_2^2 = \nu c_0$*, the pilot sequence for minimizing* $\varepsilon_{\mathrm{LS}}$ *satisfies the property as follows:*

$$|p_f[n]|^2 = \frac{\nu c_0\sqrt{\Phi_g[n]}}{\sum_{n=0}^{\nu-1}\sqrt{\Phi_g[n]}}, \tag{40}$$

where $\Phi_g[n] = \Phi[n]/|g_{p,f}[n]|^2$.

[6] Since we consider the precoded LPE-FTN and the Nyquist transmissions with the same average transmit power, the SNR definition is not affected by the signaling. The definition given here corresponds to the transmission of the Nyquist signal on the unit-power ISI channel with AWGN and is widely used in the literature on the Nyquist transmissions (e.g., [31], [39]).

Proof. (40) can be obtained by using the Lagrange multiplier method [40]. The details are omitted for brevity. □

By the definitions of $\Phi[n]$ and $\mathbf{g}_p$, we note that the optimal sequence does not have a flat spectrum and depends on a triple (β, τ, ν). By contrast, we see from (32) that the optimal sequence for the LPE-aided LS estimator has a flat spectrum and is independent of these three parameters.

Inserting (40) into (38), the minimum value of $\varepsilon_{\mathrm{LS}}$, denoted as $\varepsilon_{\mathrm{LS,min}}$, is

$$\varepsilon_{\mathrm{LS,min}} = \frac{\sigma_\omega^2}{\tau c_0}\rho_{\mathrm{LS}}^2, \tag{41}$$

in which $\rho_{\mathrm{LS}} = \left(\sum_{n=0}^{\nu-1}\sqrt{\Phi_g[n]}\right)/\nu$. Again, ρ_{LS} depends on the triple (β, τ, ν). Moreover, in the case of $\tau = 1$, since $g[l] = \delta[l]$ [2], [7], $\Phi_g[n] = 1$ and $\rho_{\mathrm{LS}} = 1$. Setting $\nu = 15$, plots of ρ_{LS} as a function of parameter pair (β, τ) are given in Fig. 4.[7] Note that for other choices of ν, similar plots can be obtained. From Fig. 4, we observe $\rho_{\mathrm{LS}} \geq 1$. Thus, $\varepsilon_{\mathrm{LS,min}} \geq \sigma_\omega^2/(\tau c_0)$. Recall that $\varepsilon_{\mathrm{LS,min}}^{\mathrm{LPE}} = \sigma_\omega^2/(\tau c_0)$ for the LPE-FTN transmission without precoding. Thus, the LPE-aided LS estimator can achieve better performance. Furthermore, the figure shows that for a given β, there exists a τ_β^0 when τ is smaller than it, ρ_{LS} is remarkably improved. For example, τ_β^0 can be 0.9 for $\beta = 0.1$. Then we come to the conclusion that the LS estimator operating with the model (37) is inadequate for the cases of small τ's or equivalently the cases with serious FTN-ISI.

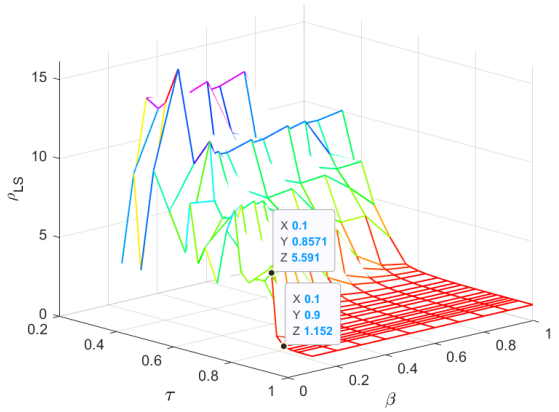


Fig. 4. Plots of ρ_{LS} in the case of $\nu = 15$.

We proceed to the MMSE case after observing the benefits of the LPE in the LS case. Based on (37), it can be readily shown that the MMSE estimator is

$$\mathbf{W}_{p,\mathrm{MMSE}} = \frac{1}{\sqrt{\tau}}\mathbf{Q}\mathbf{\Gamma}^{-1}, \tag{42}$$

where $\mathbf{Q} = \mathbf{R'}_{h,f}\left(\mathbf{R'}_{h,f} + \frac{\sigma_\omega^2}{\tau}\mathbf{\Gamma}^{-1}\mathbf{F}_\nu\mathbf{R}_{\boldsymbol{\eta}}\mathbf{F}_\nu^{\mathrm{H}}\left(\mathbf{\Gamma}^{\mathrm{H}}\right)^{-1}\right)^{-1}$ and $\mathbf{R'}_{\mathbf{h},f} = \mathbb{E}\left\{\mathbf{h'}_{i,p,f}\mathbf{h'}_{i,p,f}^{\mathrm{H}}\right\} = \nu\mathbf{F'}_\nu\mathbf{R}_{\mathbf{h}}\mathbf{F'}_\nu^{\mathrm{H}}$ with $\mathbf{F'}_\nu$ being the first L_c columns of the matrix $\mathbf{F}_\nu$. Similar to (35c), by using the SVD of $\mathbf{R'}_{h,f}$, i.e., $\mathbf{R'}_{h,f} = \mathbf{U}_1\tilde{\Lambda}_{\mathbf{h}}\mathbf{U}_1^{\mathrm{H}}$, the MSE can be expressed as

$$\begin{aligned}\varepsilon_{\mathrm{MMSE}} &= \frac{1}{\nu}\mathrm{tr}\left[\mathbf{R'}_{h,f} - \mathbf{Q}\mathbf{R'}_{h,f}\right] \\ &= \frac{1}{\nu}\mathrm{tr}\left[\tilde{\Lambda}_{\mathbf{h}} - \left(\tilde{\Lambda}_{\mathbf{h}}^{-1} + \mathbf{B}\right)^{-1}\right]\end{aligned} \tag{43}$$

where $\mathbf{B} = \frac{\sigma_\omega^2}{\tau}\tilde{\Lambda}_{\mathbf{h}}^{-1}\mathbf{U}_1^{\mathrm{H}}\mathbf{\Gamma}^{-1}\mathbf{F}_\nu\mathbf{R}_{\boldsymbol{\eta}}\mathbf{F}_\nu^{\mathrm{H}}\left(\mathbf{\Gamma}^{\mathrm{H}}\right)^{-1}\mathbf{U}_1\tilde{\Lambda}_{\mathbf{h}}^{-1}$. Since $\mathbf{B}$ can not be a diagonal matrix[8], $\varepsilon_{\mathrm{MMSE}}$ can not be further simplified. As a result, it is difficult to obtain the closed-form of the optimal pilot sequence (it may not exist) and compare the lower bound of $\varepsilon_{\mathrm{MMSE}}$ with $\varepsilon_{\mathrm{MMSE,min}}^{\mathrm{LPE}}$.

In practice, a search procedure is employed to find a nonoptimal pilot sequence that yields a $\varepsilon_{\mathrm{MMSE}}$ as low as possible. In [30], [33], the search is limited the BPSK sequence. Efforts on the search exert additional computational complexity in the CE. Besides, since the number of pilots, 4ν, and the pilot symbols, both depend on the FTN parameters, the search procedure has to be executed again when the FTN parameters change. By contrast, the optimal pilot sequences for the LPE-aided MMSE estimator are independent of the FTN parameters and the off-the-shelf sequences (e.g., the Chu sequence) remain optimal. Furthermore, the simulation results in Figs. 6 and 7 of Sec. VI-A show that $\varepsilon_{\mathrm{MMSE}}$ obtained with the nonoptimal BPSK pilot sequence is larger than $\varepsilon_{\mathrm{MMSE,min}}^{\mathrm{LPE}}$.

Remark 4. The actual MMSE estimator used in [30] is

$$\mathbf{W}_{p,\mathrm{Sugiura}} = \frac{1}{\sqrt{\tau}}\mathbf{\Gamma}^{\mathrm{H}}\left(\mathbf{\Gamma}\mathbf{\Gamma}^{\mathrm{H}} + \frac{\sigma_\omega^2}{\tau}\mathbf{\Phi}_p\right)^{-1}, \tag{44}$$

instead of $\mathbf{W}_{p,\mathrm{MMSE}}$, to reduce the complexity of CE. To obtain $\mathbf{W}_{p,\mathrm{Sugiura}}$, the matrix $\mathbf{F}_\nu\mathbf{R}_{\boldsymbol{\eta}}\mathbf{F}_\nu^{\mathrm{H}}$ in $\mathbf{Q}$ is approximated by a diagonal matrix $\mathbf{\Phi}_p = \mathrm{diag}\{\Phi[0], \ldots, \Phi[\nu-1]\}$ and $\mathbf{R'}_{\mathbf{h},f} = \mathbf{I}_{L_c}$ (or equivalently $\mathbf{R}_{\mathbf{h}} = \mathbf{I}_{L_c}/L_c$ and $L_c = \nu$) is assumed. Note that this assumption only suits for the channels having equal-gain power delay profile (PDP). If this assumption is not correct, there will be a performance loss. Using $\varepsilon_{\mathrm{MMSE}}^{\mathrm{Sugiura}}$ to denote the CE MSE corresponding to the estimator $\mathbf{W}_{p,\mathrm{Sugiura}}$, one can infer that $\varepsilon_{\mathrm{MMSE}}^{\mathrm{Sugiura}} \geq \varepsilon_{\mathrm{MMSE}}$ due to the approximation and/or the assumption not tenable.

For the LPE-aided estimator $\mathbf{W}_{p,\mathrm{MMSE}}^{\mathrm{LPE}}$ given in (27), a significant simplification occurs if we consider the equal-gain PDP (EG-PDP) and $L_c = \tilde{N}_p$. The simplified expression is

$$\mathbf{W}_{p,\mathrm{MMSE}}^{\mathrm{LPE,EG-PDP}} = \rho_1^{-1}\mathbf{\Lambda}_{\mathbf{u}}^{\mathrm{H}}\left[\mathbf{\Lambda}_{\mathbf{u}}\mathbf{\Lambda}_{\mathbf{u}}^{\mathrm{H}} + \frac{\sigma_\omega^2}{\rho_1^2}\mathbf{I}_{\tilde{N}_p}\right]^{-1}. \tag{45}$$

Despite the similar expressions, we remark that $\mathbf{W}_{p,\mathrm{MMSE}}^{\mathrm{LPE,EG-PDP}}$ has an advantage over $\mathbf{W}_{p,\mathrm{Sugiura}}$ in computational complexity. Specifically, the computation of $\mathbf{g}_{p,f}$ makes $\mathbf{\Gamma}$ require ν^3 more operations than $\mathbf{\Lambda}_{\mathbf{u}}$, while $\mathbf{\Phi}_p$ also has a computational complexity at the order of ν^3.

Remark 5. LPE also benefits the channel equalization. Our LPE-aided MMSE-FDE only suppresses the channel-ISI, whereas the MMSE-FDE in [30], [33] combats the FTN-ISI and channel-ISI concurrently. When the FTN-ISI is serious, the latter behaves poorly. As will be seen in Figs. 9 and 10,

[7] In certain cases, $|g_{p,f}[n]|$ may be very small, which results in a ρ_{LS} larger than 1000. To avoid this, the value of $|g_{p,f}[n]|$ is replaced by 0.01 if $|g_{p,f}[n]| < 0.01$ in the simulations.

[8] It can be proved by using the reduction to absurdity.

the existing MMSE-FDE results in BER floors even though the CIR is perfectly known at the receiver.

V. TDSS and THP-Based TDSS-ISI Pre-equalization

As previously mentioned in the Problem Statement, we resort to the TDSS to address the LPE-FTN signal spectral regrowth problem. Since the TDSS-ISI is known at the transmitter, we utilize THP to pre-equalize it. To sum up, the precoding module in Fig. 1 is realized as the cascade of THP and TDSS, as shown in Fig. 5. In this section, on the basis of (12), the PSD of the LPE-FTN signal with the proposed precoding is given and the *Problem* (iii) of designing the SEM-compliant LPE-FTN waveform is formulated as a constraint. Then the THP-based TDSS-ISI pre-equalization as a solution to *Problem* (iv) is described. Finally, the optimization for seeking the optimal precoding is constructed and solved.

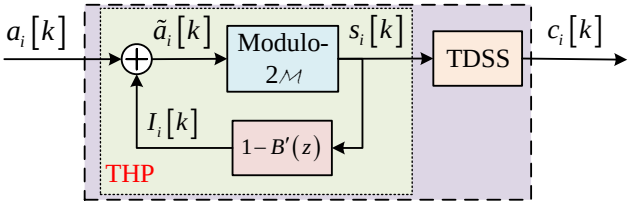


Fig. 5. Blcok diagram of the proposed precoding.

A. Further Analysis for PSD

Let $\mathbf{s}_i \triangleq \{s_i[k]\}_{k=0}^{N_s-1}$ denote the input (output) sequence of the TDSS (THP) block and consider that the TDSS is implemented as an length-(L_b+1) discrete-time filter with the input-output relation

$$c_i[k] = \sum_{l=0}^{L_b} s_i[k-l]b[l], \tag{46}$$

where $\mathbf{b} = \{b[k]\}_{k=0}^{L_b}$ is a set of tap coefficients and

$$\|\mathbf{b}\|_2^2 = 1 \tag{47}$$

to maintain the transmit power. By substituting for $c_i[k]$ in (14), we find that

$$\begin{aligned} R_c[q] &= \sum_{l_1=0}^{L_b}\sum_{l_2=0}^{L_b} \mathbb{E}\{s_i[k-l_1]s_i^*[k-q-l_2]\} b[l_1]b^*[l_2] \\ &= \sigma_s^2 \sum_{l=0}^{L_b} b^*[l]b[k+l] \end{aligned} \tag{48}$$

where in the second equality we have assumed that the THP output symbols $s_i[k]$ are zero-mean and statistically independent. It amounts to saying that $\mathbb{E}\{s_i[k]s_i^*[k-l]\} = \sigma_s^2\delta[l]$, where σ_s^2 denotes the variance of $s_i[k]$. Although not rigorously true, this assumption is reasonable for large constellations with size $M \geq 16$ [29]. Substituting (48) into (12) results in

$$\begin{aligned} \Phi_s(f) &= \frac{\rho_1^2}{\tau T}\left|\sum_{l=0}^{L_b} b[l]e^{j2\pi f l\tau T}\right|^2 \Xi(f) \\ &= \frac{\rho_1^2}{\tau T}\left|\mathbf{v}^{\mathrm{H}}(f;\tau;L_b)\mathbf{b}\right|^2 \Xi(f) \end{aligned} \tag{49}$$

with $\mathbf{v}(f;\alpha;L) = \left[1, e^{j2\pi f\alpha T}, \ldots, e^{j2\pi f L\alpha T}\right]^{\mathrm{T}}$. Since the TDSS is used to control the spectrum below the mask, denoted by $S_{\mathrm{mask}}(f)$, this imposes another constraint on $\mathbf{b}$ [18]:

$$\left|\mathbf{v}^{\mathrm{H}}(f;\tau;L_b)\mathbf{b}\right|^2 \Xi(f) \leq S_{\mathrm{mask}}(f), \forall f < W, \tag{50}$$

where W is the nominal bandwidth allocated by the regulatory bodies. Generally, $S_{\mathrm{mask}}(f)$ is defined in a consecutive range of frequencies. As a result, (50) constitutes infinitely many inequalities. To handle this, we sample the transmitted signal spectrum at P frequency locations $f_p \in \{f_0, f_1, \ldots, f_{P-1}\}$, as in [18], [26]. For each frequency bin f_p, we have

$$\left|\mathbf{v}^{\mathrm{H}}(f_p;\tau;L_b)\mathbf{b}\right|^2 \leq \tilde{S}_{\mathrm{mask}}(f_p), \tag{51}$$

where $\tilde{S}_{\mathrm{mask}}(f_p) = S_{\mathrm{mask}}(f_p)/\Xi(f_p)$.

B. THP-Based TDSS-ISI Pre-equalization

The motivation of using THP is explained. First, applying THP-based pre-equalization can mitigate the potential problems that might arise when TDSS-ISI equalization is implemented at the receiver, such as an increase in receiver complexity and a degraded performance due to noise enhancement. Second, we can observe from (49) that PSD is independent of THP. Lastly, THP uses a modulo-$2\mathcal{M}$ operator to limit the transmission power increase caused by pre-equalization.

We now detail the THP. The feedback filter is designed using the zero-forcing criterion (i.e., complete TDSS-ISI pre-cancellation) and has a transfer function $1-B'(z)$, where $B'(z) = \mathcal{Z}\{\mathbf{b}\}/b[0]$. $I_i[k]$ and $\tilde{a}_i[k]$ denote the output of the feedback filter and the input of the modulo, respectively. Moreover, $\tilde{a}_i[k] = a_i[k] + I_i[k]$ (cf. Fig. 5). Using the definitions made above, the THP block works as follows

$$s_i[k] = \tilde{a}_i[k] + 2\mathcal{M}d_i[k] = a_i[k] + I_i[k] + 2\mathcal{M}d_i[k], \tag{52}$$

where $d_i[k] = -\lfloor(\tilde{a}_i[k]+\mathcal{M})/2\mathcal{M}\rfloor$. As in [20], [22], we set $\mathcal{M} = \sqrt{1.5M/(M-1)}$ and then $\sigma_s^2 = 2\mathcal{M}^2/3 = M/(M-1)$. Since $\sigma_c^2 = R_c[0] = \sigma_s^2\|\mathbf{b}\|^2 = \sigma_s^2$ under the constraint (47), P_{avg} given in (15) can be expressed as

$$P_{avg} = \frac{\rho_1^2}{\tau T}\sigma_s^2 = \frac{\rho_1^2}{\tau T}\frac{M}{M-1}. \tag{53}$$

Since $\sigma_c^2 = \sigma_a^2 = 1$ in the absence of precoding, the proposed precoding increases the transmit power. However, the power loss $M/(M-1)$ is negligible for a large value of constellation size M.

Plugging (52) into (46) and taking $I_i[k] = -\left(\sum_{l=1}^{L_b} b[l]s_i[k-l]\right)/b[0]$ into account, we obtain

$$c_i[k] = b[0](a_i[k] + 2\mathcal{M}d_i[k]). \tag{54}$$

At the receiver, the MMSE-FDE represented by (29) generates $\hat{c}_i[k]$, from which $a_i[k]$ and $d_i[k]$ can be estimated. Specifically, $\hat{d}_i[k]$, which is the estimate of $d_i[k]$, can be obtained by applying the same modulo function used in (52) to $\hat{c}_i[k]$, i.e. $\hat{d}_i[k] = \lfloor\hat{c}_i[k]/(2b[0]\mathcal{M}) + 1/2\rfloor$. Then, $2\mathcal{M}\hat{d}_i[k]b[0]$ is subtracted from $\hat{c}_i[k]$ before estimating $a_i[k]$. This estimation scheme, however, may suffer from the so-called "modulo-loss" for the coded system [20]. Precisely,

incorrect estimate $\hat{d}_i[k]$ seriously degrades the quality of LLR values, which then increases the BER. To alleviate the loss, as in [26], we employ the expanded a priori demapper (EAD) proposed in [20], circumventing the estimation of $d_i[k]$. We refer to [20, eq. (13)] for the computation of LLRs using the EAD.

C. Precoding Optimization

Since the THP block (or the feedback filter) is determined by the TDSS, the optimization of precoding actually turns to the TDSS design. Recall that the primary goal of TDSS is to render the LPE-FTN signal spectrum conform with the mask. However, the selection of $\mathbf{b}$ has an impact on the BER performance. If the EAD is used, Wen *et al.* [26] founded that larger value of $b[0]$ helps the EAD generate LLRs with higher quality, thus improving the performance. Inspired by this, we propose to maximize $b[0]$ subject to the two constraints given by (47) and (51), resulting in the optimization problem

$$\mathcal{P}1 \begin{cases} \max\limits_{\mathbf{b}} \quad b[0] \\ \text{subject to } (51)\, and\ \|\mathbf{b}\|_2^2 \le 1. \end{cases} \tag{55}$$

Note that the constraint (47) has been relaxed.

Proposition 5. *$\mathcal{P}1$ is a convex optimization problem.*

Proof. First, the objective can be regarded as the result of an affine function, i.e., $b[0] = [1, \mathbf{0}_{1\times L_b}]\,\mathbf{b}$. For the constraint (51), it is the composition of the absolute value function with an affine function. Finally, the constraint (47) corresponds to an ℓ_2-norm constraint. Thus, $\mathcal{P}1$ is convex [40]. This completes the proof. □

In this work, the optimal solution to $\mathcal{P}1$ is obtained using the open source software CVX [41] and then normalized to unit-power. A summary of the TDSS design procedure is given in Alg. 1.

Algorithm 1 Procedure of the TDSS Design

Input: L_b, τ, β, mask profile $S_{\text{mask}}(f_p)$, for $p = 0, \ldots, P-1$;

Output: An optimized solution $\mathbf{b}_{opt}$ to $\mathcal{P}1$ and $b_{opt}[0]$;

1: Set parameter T such that the T-orthogonal RRC pulse $\phi(t)$ with β satisfies the SEM;
2: **for** $p = 0, \ldots, P-1$ **do**
3: Compute $\mathbf{v}(f_p; \tau; L_b)$, $\Xi(f_p)$ and $\tilde{S}_{\text{mask}}(f_p)$;
4: **end for**
5: Construct the Problem $\mathcal{P}1$ and then find the optimal solution $\tilde{\mathbf{b}}_{opt}$;
6: Perform the normalization and obtain $\mathbf{b}_{opt} = \tilde{\mathbf{b}}_{opt} / \left\|\tilde{\mathbf{b}}_{opt}\right\|_2$ and $b_{opt}[0]$;

VI. Performance and Complexity Evaluation

In this section, simulations with different system parameters and fading channels are conducted in order to assess the performance of the precoded LPE-FTN system with frequency-domain linear MMSE CE[9] and equalization. The simulation parameters are summarized in Tab. I. Pilot and data symbols are assumed to have the same power. With this assumption, the term c_0 in (32) is set to 1 and we have the pilot power constraint $\|\mathbf{u}\|_2^2 = \tilde{N}_p$. As the Nyquist signal has an average transmit power of $1/T$ [2], comparing this result with P_{avg} in (53), we have $\rho_1 = \sqrt{\frac{\tau(M-1)}{M}}$ (or equivalently, $\rho_0 = \sqrt{\frac{\tau(M-1)}{M}}/f[0]$) to keep the average transmit power identical. Moreover, we consider the 11-ray block Rayleigh fading channels with equal-gain PDP or exponential PDP with decay factor 3.4744 dB [42].

TABLE I
Simulation Parameters

Transmitter	Channel code	Low-density parity-check code with rate 1/2 and length 3840 [43]
	Constellation size	$M = 64, 256$
	Number of symbols per block	$\tilde{N}_s = 512$
	Pilot sequence	Chu sequence with power constraint $\|\mathbf{u}\|_2^2 = \tilde{N}_p$
	Transmit Power normalization factor	$\rho_0 = \sqrt{\frac{\tau(M-1)}{M}}/f[0]$
SEM	Specified for fixed microwave backhaul links with 56 MHz bandwidth [36] and depicted in Fig. 2	
Channel	Fading type	Frequency-selective block Rayleigh
	PDP	equal-gain PDP and exponential PDP

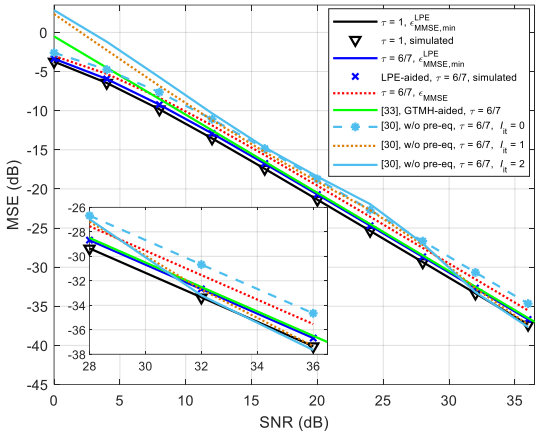
To demonstrate the superiority, we also compare our scheme with existing FTN frequency-domain CE and equalization schemes in [30] and [33]. The transmitter in [30] does not perform the FTN-ISI pre-equalization while the FTN-ISI is partly pre-equalized by the GTMH in [33]. Hereinafter, we refer to the benchmark CE schemes as the w/o pre-eq CE and the GTMH-aided CE, respectively. Additionally, in [30], the CE and channel equalization are performed iteratively. Specifically, tentative decisions of detected symbols are utilized to improve the CE performance. Then the combined ISI is equalized again using the updated channel estimates. We use I_{it} to denote the number of iterations. For the transmitting block adopted in [33], it contains $N_p = 5\nu$ pilots, thereby having the highest pilot overhead among the three schemes to be compared. Finally, the BPSK-based pilot sequences obtained through 2×10^5 trials are used for the two benchmark schemes as well as the estimator $\mathbf{W}_{p,\text{MMSE}}$ given by (42).

A. CE MSE performance

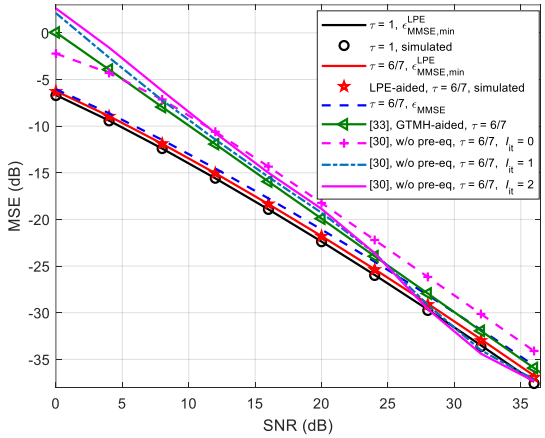
Fig. 6 shows the MSE versus SNR for the 64-QAM FTN transmission with $(\beta, \tau) = (0.25, 6/7)$. The FTN-ISI is moderate. We choose $\tilde{N}_p = \nu = 15$. The MSE is averaged over 50000 independent channel realizations. From Fig. 6, we see that the GTMH-aided and LPE-aided CE schemes deliver better performance than the w/o pre-eq CE with $I_{\text{it}} = 0$ under both two types of channels. Since the three FTN channel estimators have comparable computational complexity if $I_{\text{it}} = 0$, the FTN-ISI pre-equalization contributes to enhancing the accuracy of channel estimates without increasing the complexity of the receiver. Fig. 6 also shows that the LPE-aided performs better than the GTMH-aided. The advantage is

[9] In Sec. IV, the LPE has been proved to benefit for reducing the MSE of the LS estimator, thus we only consider the MMSE estimator in the simulations.

more obvious in the case of exponential PDP. For the w/o pre-eq scheme with $I_{\rm it} > 0$, improved performance is expected to be obtained at the cost of increased complexity. However, it is observed that the MSE of the w/o pre-eq CE scheme increases with the increase of iterations at low SNRs. This is because the equalizer behaves poorly at low SNRs and, hence, the tentative decisions would be highly unreliable. For moderate-to-high SNRs, the first iteration noticeably improves the performance. Furthermore, the w/o pre-eq CE scheme with two iterations has the best performance among all the comparison schemes when the SNR is high enough.



(a) channel with equal-gain PDP



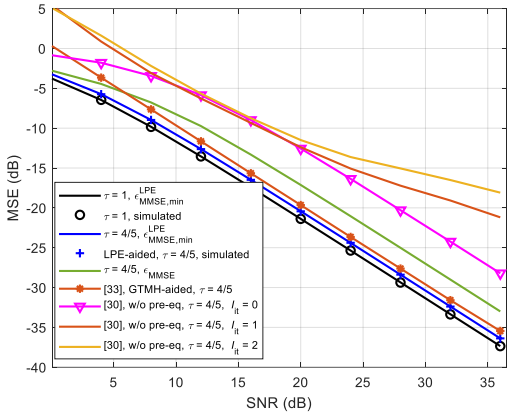
(b) channel with exponential PDP

Fig. 6. CE MSE performance comparisons among different schemes. $\mathbf{b}_{opt} = [0.9593, 0.1951, -0.1522, 0.1093, -0.0703, 0.0380, -0.0139, -0.0011, 0.0072, -0.0059]^{\rm T}$ for the proposed scheme.

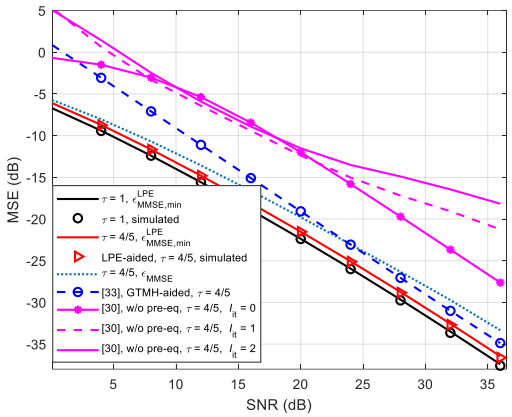
In Fig. 6, plots of the theoretical statistics $\varepsilon^{\rm LPE}_{\rm MMSE,min}$ and $\varepsilon_{\rm MMSE}$ as defined in (35e) and (43) are given. As can be seen, the simulated MSE results of the LPE-aided CE match the statistics $\varepsilon^{\rm LPE}_{\rm MMSE,min}$ very well. Moreover, $\varepsilon^{\rm LPE}_{\rm MMSE,min} < \varepsilon_{\rm MMSE}$, which indicates that the LPE helps the MMSE estimator reduce the MSE. On the other hand, the MSE of the w/o pre-eq CE scheme with $I_{\rm it} = 0$ is larger than $\varepsilon_{\rm MMSE}$, which implies that the reduced-complexity estimator $\mathbf{W}_{p,\rm Sugiura}$ given in (44) results in performance loss compared to the estimator $\mathbf{W}_{p,\rm MMSE}$. The performance loss is larger for the case of exponential PDP. As discussed in *Remark 4*, the reason is that equal-gain PDP is assumed in the derivation of $\mathbf{W}_{p,\rm Sugiura}$. Besides, we will see in Fig. 7 that the more serious FTN-ISI prompts a larger loss up to 5 dB.

Having compared the performance of several FTN channel estimators, we now compare the MSE of the precoded LPE-FTN transmission with that of the Nyquist transmission. Note that ρ_1 satisfies the condition given in *Remark 2*, then with *Remark 3* we know that the precoded LPE-FTN transmission has inferior CE performance to the Nyquist transmission. The results in Fig. 6 confirm this. In Prop. 3, their theoretical gap is provided. From Fig. 6, we observe that there is good agreement between the theoretical and simulated gap results. Particularly, the gap is about 0.7 dB at an MSE of 33 dB while the theoretical result is $0.6695 = -10\log\rho_1$ dB.

In Fig. 7, we examine the case of 64-QAM FTN transmission with $(\beta, \tau) = (0.25, 4/5)$. The FTN-ISI now is serious. In such a case, Fig. 7 provides more information than Fig. 6. First, the iterations result in performance degradation for the w/o pre-eq CE scheme at most SNRs. Actually, the MMSE-FDE in [30] is inadequate for combating the serious FTN-ISI, producing a BER floor at the level of 0.1. Thus, the loss in CE performance can be attributed to the error propagation in the iterative loop. Second, the gap between the LPE-aided and the GTMH-aided, the gap between $\varepsilon^{\rm LPE}_{\rm MMSE,min}$ and $\varepsilon_{\rm MMSE}$, both become larger, which implies that the LPE-aided CE is more suitable for the FTN transmissions with small τ's.



(a) channel with equal-gain PDP



(b) channel with exponential PDP

Fig. 7. CE MSE performance comparisons among different schemes. $\mathbf{b}_{opt} = [0.8569, 0.4184, -0.2640, 0.1345, -0.0444, -0.0045, 0.0205, -0.0174, 0.0083, -0.0019]^{\rm T}$ for the proposed scheme.

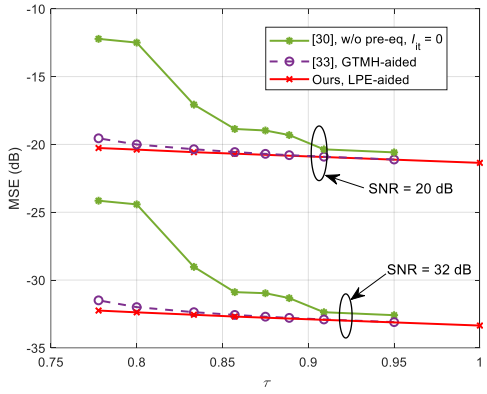


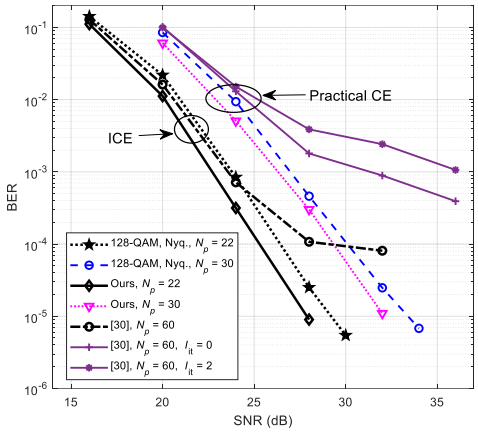
Fig. 8. The CE MSE versus τ for different CE schemes.

Fig. 8 shows the comparison of MSE performance as a function of the FTN acceleration factor τ where $\beta = 0.3$, $\tilde{N}_p = \nu = 15$, SNR $= 20$ and 32 dB. The channels with equal-gain PDP are used. τ ranging from 1 to 7/9 (slightly larger than $1/(1+\beta)$) is examined. To better illustrate the benefit from the FTN-ISI pre-equalization, $I_{\rm it} = 0$ is considered for the w/o pre-eq CE scheme. As we can see, the FTN-ISI degrades the performance of the w/o pre-eq CE scheme, and the degradation is significant when τ is small. For example, when SNR is 20 dB, the distance between the MSE values corresponding to $\tau = 7/9$ and $6/7$ is about 6.6 dB. By contrast, the MSE performance of the GTMH-aided and LPE-aided CE schemes is nearly independent of τ. Nevertheless, since the FTN-ISI is not completely pre-removed by the GTMH, a non-ignorable loss is observed when τ equals 7/9 because of the residual FTN-ISI. It can be inferred that the GTMH-aided CE scheme may suffer from a considerable CE accuracy degradation when the residual FTN-ISI is strong. Thus, we can expect that the performance improvement of the LPE-aided scheme over the benchmarks enlarges with the increase of FTN-ISI level.

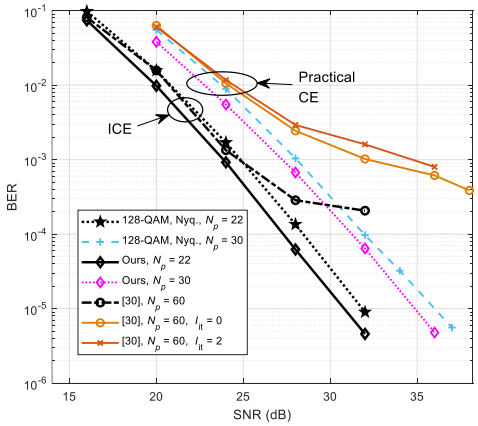
B. BER Performance

We first evaluate the 64-QAM FTNS with $(\beta, \tau) = (0.25, 6/7)$. The TDSS design agrees with the one in Fig. 6. To show the benefits of the LPE to the channel equalization, the BER curves obtained with ideal CE (ICE) are also given. Note that the FTN transmission reaches a capacity of 3.5 bits/channel use. Thus, the BER performance of the same-capacity 128-QAM Nyquist signaling is exhibited as a benchmark, in the figure.

From Fig. 9 (a) and Fig. 9 (b), we observe that in either the ICE or practice CE case, the proposed scheme outperforms the benchmark schemes.[10] Specifically, the proposed schemes gain about $1 \sim 1.35$ dB over the 128-QAM Nyquist benchmarks at BER = 10^{-5}. We attribute this to the higher robustness of 64-QAM than 128-QAM against the channel-ISI. Besides, the performance of the FTN benchmark [30] is comparable to that of the proposed scheme at low SNRs. However, the BER curves diverge at high SNRs. The FTN-ISI causes irreducible error floors for the FTN benchmark even considering the ICE. Moreover, the equalization performance is deteriorated after two iterations, although we have seen in Fig. 6 that the w/o pre-eq CE scheme with $I_{\rm it} = 2$ works better than that with $I_{\rm it} = 0$ at the high SNRs. A possible reason is that for some channels, the CE performance is seriously degraded after two iterations, thereby further deteriorating the detection performance and resulting in lots of detection errors, which increases the overall BER.



(a) channel with equal-gain PDP



(b) channel with exponential PDP

Fig. 9. BER performance comparisons among different schemes. $(\beta, \tau) = (0.25, 6/7)$ for FTNS while $(\beta, \tau) = (0.25, 1)$ for Nyquist signaling. Since the value of N_p has negligible effect on the BER performance under the ICE case, BER performance is only evaluated with a typical N_p for simplicity.

We also notice from Fig. 9 that the CE errors result in substantial performance degradation for all schemes assessed here. Moreover, the performance loss is larger in the higher SNR regime, although more accurate channel estimates are achieved and the CE MSE is smaller than -30 dB (see Fig. 6). There are two possible reasons for the performance degradation. First, the residual channel-ISI remains after the MMSE-FDE and may be very strong for the channels with deep fades. Second, the CE errors magnify the residual channel-ISI. According to these factors, in order to alleviate the performance loss, we can replace the linear MMSE-FDE by a nonlinear equalizer, such as (frequency-domain) decision feedback equalization [44], for further suppress the residual channel-ISI. A joint channel-ISI estimation and equalization is also worth investigating, which

[10] As will be seen in Fig. 11, GTMH results in a spectral regrowth, which is not addressed in [33]. Considering this fact, the BER comparisons between the scheme [33] and the proposed one are not conducted.

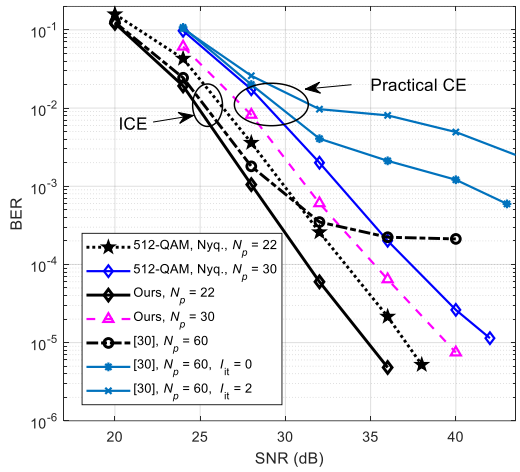


Fig. 10. BER performance comparisons among different schemes. $(\beta, \tau) = (0.25, 8/9)$ for FTNS while $(\beta, \tau) = (0.25, 1)$ for Nyquist signaling.

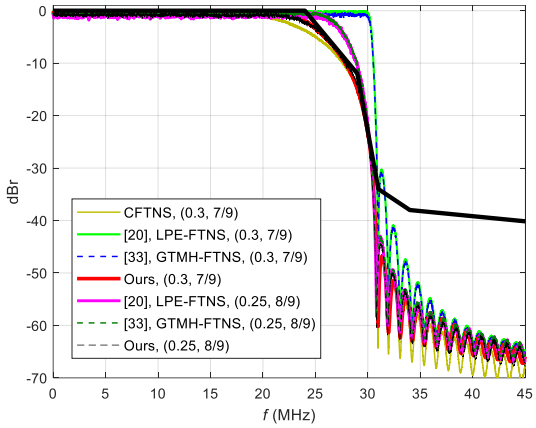


Fig. 11. The PSD plots of different 64-QAM FTN signals.

can minimize both CE errors and residual channel-ISI at the same time. We will, however, study these two performance enhancement methods in the future.

Fig. 10 depicts the BER curves of the 256-QAM FTNS with $(\beta, \tau) = (0.25, 8/9)$, which achieves the capacity of the 512-QAM Nyquist signaling scheme. The TDSS filter taps are $\mathbf{b}_{opt} = [0.9839, 0.1064, -0.0922, 0.0756, -0.0585, 0.0422, -0.0279, 0.0163, -0.0078, 0.0026]^{\mathrm{T}}$. Only the channels with the equal-gain PDP are considered. From Fig. 10, we come to same conclusions as that obtained from Fig. 9, except that the gain of the proposed FTN scheme over the Nyquist benchmarks increases to about 3 dB. Thus, FTN can enhance the system's robustness to channel-ISI by reducing the required constellation size while maintaining the capacity.

C. PSD of the Proposed FTN Signals

Fig. 11 plots the PSD of different 64-QAM FTN signals. $\mathbf{b}_{opt} = [0.8726, 0.3861, -0.2548, 0.142, -0.0577, 0.0054, 0.0184, -0.0209, 0.0118, -0.0004, -0.007, 0.0071]^{\mathrm{T}}$ for our scheme with $(\beta, \tau) = (0.3, 7/9)$. From Fig. 11, it is visible that the LPE-FTN [20] and GTMH-FTN [33] signals have similar PSD shapes, and both violate the SEM constraints. By contrast, the proposed FTN signal strictly complies to the SEM. Moreover, the PSD of proposed signal is closer to the SEM than that of the FTN signal generated by the CFTNS. This implies that the proposed FTNS attempts to fully utilize the spectrum available to transmit information.

D. PAPR Analysis

Fig. 12 compares the complementary cumulative distribution function (CCDF) of the PAPR of the proposed FTN waveforms with existing waveforms. It is observed that the use of FTNS results in a PAPR increase, compared with the conventional Nyquist signaling. Moreover, the proposed FTNS is inferior to the CFTNS, but performs better than other two FTNS schemes using pre-equalization, i.e. LPE-FTNS [20] and GTMH-FTNS [33]. Particularly, the proposed FTN signals feature a $0.2 \sim 0.6$ dB higher PAPR in comparison with the signals generated by the CFTNS when the value of CCDF is 10^{-4}. The PAPR results of GTMH-FTN signals seem to be independent to the QAM order and FTN parameters. And they are much poorer than ours. A possible reason is that the GTMH is a matrix-based pre-equalization and the matrix becomes ill-conditioned when the FTN-ISI is serious [23]. For the LPE-FTNS, the PAPR is quite high in the case of $(\beta, \tau) = (0.3, 7/9)$. This is because the LPE-FTNS with the T-orthogonal RRC is equivalent to the Nyquist signaling with the RRC that is τT-orthogonal and has a roll-factor $\hat{\beta} = (1+\beta)\tau - 1$ [20]. When $(\beta, \tau) = (0.3, 7/9)$, $\hat{\beta} = 0.0111$. It is well known that the Nyquist signal obtained by using the RRC pulse with small roll-off factor exhibits a high PAPR. To summarize, the superiority of precoded LPE-FTNS over the LPE-FTNS indicates that the proposed precoding also reduces the PAPR, although its primary goals are to compress the signal spectrum and pre-equalize the TDSS-ISI.

E. Complexity Analysis

Tab. II offers a detailed complexity of the precoded LPE-FTN transceiver, as well as for the benchmark scheme [30]. In particular, the implementation of precoding requires about $N_s(2L_b+1)$ complex-valued multiplications and $2N_sL_b$ additions, and N_s modulo operations for a length-$\tilde{N}_s$ signal block and $N_s = \tilde{N}_s$ in our precoding. The LPE needs νN complex-valued multiplications and additions. Under the condition when the optimal pilot sequences are used, we see that the channel estimator $\mathbf{W}_{p,\mathrm{MMSE}}^{\mathrm{LPE}}$ defined in (27) is independent of the pilot sequences and needs to be calculated once in an off-line manner if $\mathbf{R}_{\mathbf{h}}$ and σ_ω^2 are known beforehand. Thus, the main complexity of the LPE-aided CE comes from (22) and (25), which have a total complexity of $3\tilde{N}_p^2$ complex-valued multiplications and $3\tilde{N}_p\left(\tilde{N}_p - 1\right)$ additions. For the channel equalization, we assume that the conversions between the time-domain signals and their frequency-domain presentations are realized by efficient fast Fourier transform (FFT) and inverse FFT (IFFT). The computational complexity for an N-point FFT or IFFT is about $N \log N$ complex-valued multiplications and $(N/2)\log N$ additions. The complexity for the LPE-aided channel equalization involving (29) and (30) is $3N \log N + 3N$ complex-valued multiplications and $\frac{3}{2}N \log N + N$ additions.

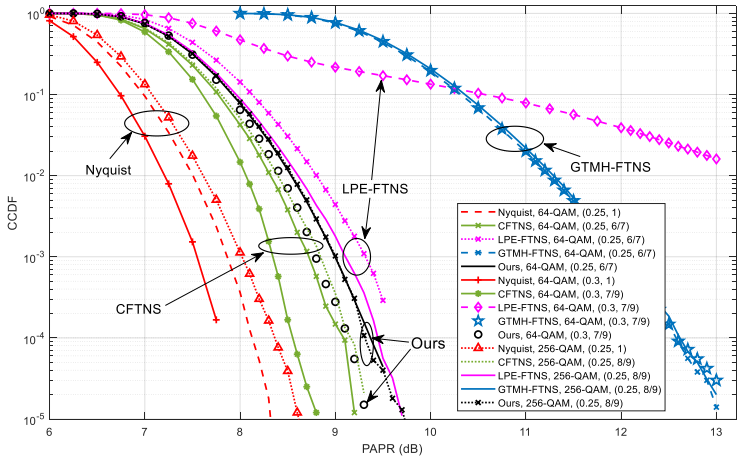


Fig. 12. PAPR comparisons among different signaling schemes.

Tab. II reveals that the overall complexity of the proposed scheme grows at the order of $N \log N$, while the benchmark FTN scheme has a complexity $\mathcal{O}\left((I_{\text{it}}+1)\left(\tilde{N}_s+2\nu\right)\log\left(\tilde{N}_s+2\nu\right)\right)$. Since N is approximately equal to $\tilde{N}_s+2\nu$, the proposed scheme also has lower complexity.

VII. Conclusion

In this paper, we address the CE and detection for the FTN transmissions over frequency-selective fading channels. LPE and precoding techniques are applied to the FTNS. Specifically, LPE completely pre-equalizes the FTN-ISI, while precoding resolves the LPE-caused signal spectral broadening. Owing to the transmitter pre-processing, channel-ISI is the only ISI component in the observations, and we estimate and equalize it using the classical frequency-domain low-complexity schemes. The lack of the FTN-ISI in the observations offers many benefits to the CE and equalization, e.g., having a lower CE MSE and more robust against the FTN-ISI. The simulation demonstrates that the transmitted signal meets the spectral mask. Additionally, our scheme delivers more accurate channel estimates and improved detection performance, compared with the existing frequency-domain CE and equalization schemes.

Appendix A
Proof of Proposition 1

When the input to the FTNS block is the sequence $\{x[n]\}$, the PSD of the resulting FTN signal is of the form [2], [20]

$$\Phi_s(f) = \frac{\rho_0^2}{\tau T}\Phi_x(2\pi f\tau T)\,\Psi(f), \tag{56}$$

where $\Phi_x(\lambda) = \sum_q R_x[q]e^{-jq\lambda}$ and $R_x[q]$ represents the autocorrelation function of $\{x[n]\}$.

According to [20, appendix C], $\Phi_x(2\pi f\tau T)$ equals

$$\Phi_x(2\pi f\tau T) = \tau T\frac{f^2[0]\,\Phi_c(2\pi f\tau T)}{\sum_l \Psi\left(f-\frac{l}{\tau T}\right)}. \tag{57}$$

Then, substituting (57) into (56) yields (12).

With the PSD expression (12) and defining $\varphi(t)$ as the inverse Fourier transform of $\Xi(f)$, the average power is computed as

$$P_{avg} = \int_{-\infty}^{\infty}\Phi_s(f)\,df = R_s(0) \tag{58}$$

with

$$R_s(t) = \frac{\rho_1^2}{\tau T}\sum\nolimits_q R_c[q]\,\varphi(t-q\tau T) \tag{59}$$

being the inverse Fourier transform of $\Phi_s(f)$. Note that $\varphi(t)$ is a τT-orthogonal Nyquist pulse [20]. That is, $\varphi(q\tau T) = \delta[q]$. This property leads to

$$R_s(0) = \frac{\rho_1^2}{\tau T}\sum\nolimits_q R_c[q]\,\varphi(-q\tau T) = \frac{\rho_1^2}{\tau T}R_c[0]. \tag{60}$$

Combining (58) and (60), we obtain (15).

References

[1] L. Zhang, A. Farhang, G. Feng, and O. Onireti, *Radio Access Network Slicing and Virtualization for 5G Vertical Industries*. Wiley, 2020.

[2] F. Rusek, "Partial response and faster-than-Nyquist signaling," Ph.D. dissertation, Dept. Elec. Inform. Tech., Lund Univ., Lund, Sweden, Sep. 2007.

[3] J. B. Anderson, F. Rusek *et al.*, "Faster-than-Nyquist signaling," *Proc. IEEE*, vol. 101, no. 8, pp. 1817–1830, Aug. 2013.

[4] A. Modenini, F. Rusek, and G. Colavolpe, "Faster-than-Nyquist signaling for next generation communication architectures," in *Proc. 22nd Eur. Signal Process. Conf. (EUSIPCO)*, Lisbon, Portugal, Sep. 2014, pp. 1856–1860.

[5] C. Le, M. Schellmann, M. Fuhrwerk *et al.*, "On the practical benefits of faster-than-Nyquist signaling," in *Proc. Int Conf. on Advanced Tech. Commun. (ATC)*, Hanoi, Vietnam, Oct. 2014, pp. 208–213.

[6] "From today to tomorrow: Huawei microwave & mm-wave whitepaper," Huawei, Shenzhen, China, Tech. Rep., 2016. [Online]. Available: https://www-file.huawei.com/-/media/CORPORATE/PDF/white%20paper/huawei-microwave-whitepaper-2016.pdf

[7] J. Fan, S. Guo *et al.*, "Faster-than-Nyquist signaling: an overview," *IEEE Access*, vol. 5, pp. 1925–1940, Feb. 2017.

[8] M. Maso and S. Tomasin, "Pre-equalized faster than Nyquist transmission for 5G cellular microwave backhaul," in *Proc. IEEE Int. Workshop Signal Process. Advances Wireless Commun. (SPAWC)*, Edinburgh, UK, July 2016, pp. 1–6.

[9] 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, Sep. 2012.

TABLE II
COMPARISONS OF COMPUTATIONAL COMPLEXITY FOR A SIGNAL BLOCK

Scheme	Complex Addition/Subtraction	Complex Multiplication/Division	Others
Ours	$2N_sL_b + 3\tilde{N}_p\left(\tilde{N}_p - 1\right)$ $+\frac{3}{2}N\log N + (\nu+1)N$	$N_s\left(2L_b+1\right) + 3\tilde{N}_p^2$ $+3N\log N + (\nu+3)N$	N_s modulo operations
[30]	$2\nu\left(\nu-1\right) + (2I_{\mathrm{it}}+1)\left(\tilde{N}_s+2\nu\right)$ $+\left(\frac{5}{2}I_{\mathrm{it}}+\frac{3}{2}\right)\left(\tilde{N}_s+2\nu\right)\log\left(\tilde{N}_s+2\nu\right)$	$\nu\left(2\nu+1\right) + (6I_{\mathrm{it}}+3)\left(\tilde{N}_s+2\nu\right)$ $+(5I_{\mathrm{it}}+3)\left(\tilde{N}_s+2\nu\right)\log\left(\tilde{N}_s+2\nu\right)$	$I_{\mathrm{it}}\left(\tilde{N}_s+2\nu\right)$ hard/soft decisions

[10] S. Li, B. Bai, J. Zhou *et al.*, "Reduced-complexity equalization for faster-than-Nyquist signaling: New methods based on Ungerboeck observation model," *IEEE Trans. Commun.*, vol. 66, no. 3, pp. 1190–1204, Mar. 2018.

[11] S. Li, J. Yuan, B. Bai *et al.*, "Code-based channel shortening for faster-than-Nyquist signaling: Reduced-complexity detection and code design," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 3996–4011, July 2020.

[12] S. Li, W. Yuan, J. Yuan *et al.*, "Time-domain vs. frequency-domain equalization for FTN signaling," *IEEE Trans. Veh. Tech.*, vol. 69, no. 8, pp. 9174–9179, June 2020.

[13] S. Sugiura, "Frequency-domain equalization of faster-than-Nyquist signaling," *IEEE Wireless Commun. Lett.*, vol. 2, no. 5, pp. 555–558, Oct. 2013.

[14] S. Sugiura and L. Hanzo, "Frequency-domain-equalization-aided iterative detection of faster-than-Nyquist signaling," *IEEE Trans. Veh. Tech.*, vol. 64, no. 5, pp. 2122–2128, May 2015.

[15] P. Sen, T. Aktas, and A. Ö. Yilmaz, "A low-complexity graph-based LMMSE receiver designed for colored noise induced by FTN-signaling," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Istanbul, Turkey, April 2014, pp. 642–647.

[16] S. Peng, A. Liu, X. Liu *et al.*, "MMSE turbo equalization and detection for multicarrier faster-than-Nyquist signaling," *IEEE Trans. Veh. Tech.*, vol. 67, no. 3, pp. 2267–2275, Mar. 2018.

[17] S. Wen, G. Liu, Q. Chen *et al.*, "Optimization of precoded FTN signaling with MMSE-based turbo equalization," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Shanghai, China, May 2019, pp. 1–6.

[18] ——, "Time-frequency compressed FTN signaling: A solution to spectrally efficient single-carrier system," *IEEE Trans. Commun.*, vol. 68, no. 5, pp. 3125–3139, May 2020.

[19] D. Chang, O. Omomukuyo, X. Lin *et al.*, "Robust faster-than-Nyquist PDM-mQAM systems with Tomlinson-Harashima precoding," *IEEE Photon. Tech. Lett.*, vol. 28, no. 19, pp. 2106–2109, Oct. 2016.

[20] M. Jana, A. Medra, L. Lampe *et al.*, "Pre-equalized faster-than-Nyquist transmission," *IEEE Trans. Commun.*, vol. 65, no. 10, pp. 4406–4418, Oct. 2017.

[21] M. Jana, L. Lampe, and J. Mitra, "Dual-polarized faster-than-Nyquist transmission using higher order modulation schemes," *IEEE Trans. Commun.*, vol. 66, no. 11, pp. 5332–5345, Nov. 2018.

[22] H. Wang, A. Liu *et al.*, "Iterative-detection-aided Tomlinson-Harashima precoding for faster-than-Nyquist signaling," *IEEE Access*, vol. 8, pp. 7748–7757, Jan. 2020.

[23] E. Ringh, "Low complexity algorithm for faster-than-Nyquist signaling: using coding to avoid an NP-hard problem," Master's thesis, School of Engineering Sciences, KTH Royal Institute of Technology, 2013.

[24] Q. Li, F. Gong, P. Song *et al.*, "Beyond DVB-S2X: Faster-than-Nyquist signaling with linear precoding," *IEEE Trans. Broadcast.*, vol. 66, no. 3, pp. 620–629, Sept. 2020.

[25] B. Qian, X. Wang, J. Wen *et al.*, "Novel intersymbol interference cancellation scheme to enable parallel computational and high-performance faster-than-Nyquist signaling," *IEEE Access*, vol. 5, pp. 24 758–24 765, Nov. 2017.

[26] S. Wen, G. Liu, C. Liu *et al.*, "Waveform design for high-order QAM faster-than-Nyquist transmission in the presence of phase noise," IEEE Trans. Wireless Commun., accepted.

[27] W. Yuan, N. Wu, H. Wang *et al.*, "Variational inference-based frequency-domain equalization for faster-than-Nyquist signaling in doubly selective channels," *IEEE Signal Process. Lett.*, vol. 23, no. 9, pp. 1270–1274, Sept. 2016.

[28] N. Wu, W. Yuan, H. Wang *et al.*, "Frequency-domain iterative message passing receiver for faster-than-Nyquist signaling in doubly selective channels," *IEEE Wireless Commun. Lett.*, vol. 5, no. 6, pp. 584–587, Dec. 2016.

[29] R. F. H. Fischer, *Precoding and Signal Shaping for Digital Transmission*. Wiley, 2002.

[30] T. Ishihara and S. Sugiura, "Iterative frequency-domain joint channel estimation and data detection of faster-than-Nyquist signaling," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6221–6231, Sep. 2017.

[31] Y. Zeng and T. S. Ng, "Pilot cyclic prefixed single carrier communication: channel estimation and equalization," *IEEE Signal Process. Lett.*, vol. 12, no. 1, pp. 56–59, Jan. 2005.

[32] Q. Shi, N. Wu, X. Ma *et al.*, "Frequency-domain joint channel estimation and decoding for faster-than-Nyquist signaling," *IEEE Trans. Commun.*, vol. 66, no. 2, pp. 781–795, Feb. 2018.

[33] Q. Li, F. Gong, P. Song *et al.*, "Joint channel estimation and precoding for faster-than-Nyquist signaling," *IEEE Trans. Veh. Tech.*, vol. 69, no. 11, pp. 13 139–13 147, Sep. 2020.

[34] S. Gao, L. Lin, Y. Lin *et al.*, "Pilot design for channel estimation in a long-tap faster-than-Nyquist signaling transmission," in *Proc. IEEE Int. Conf. Commun. Tech. (ICCT)*, Oct. 2017, pp. 1–5.

[35] D. Chu, "Polyphase codes with good periodic correlation properties," *IEEE Trans. Info. Theory*, vol. 18, no. 4, pp. 531–532, May 1972.

[36] ETSI, "Fixed Radio Systems; Characteristics and requirements for point-to-point equipment and antennas; Part 2-2: Digital systems operating in frequency bands where frequency co-ordination is applied," Tech. Spec. EN 302 217-2-2, Sep. 2012. [Online]. Available: https://www.etsi.org.

[37] C. Huang, L. Liu, C. Yuen, and S. Sun, "Iterative channel estimation using LSE and sparse message passing for MmWave MIMO systems," *IEEE Trans. Signal Processing*, vol. 67, no. 1, pp. 245–259, Jan. 2019.

[38] Y. Wang and X. Dong, "Frequency-domain channel estimation for SC-FDE in UWB communications," *IEEE Trans. Commun.*, vol. 54, no. 12, pp. 2155–2163, Dec. 2006.

[39] L. Deneire, B. Gyselinckx, and M. Engels, "Training sequence versus cyclic prefix-a new look on single carrier communication," *IEEE Commun. Lett.*, vol. 5, no. 7, pp. 292–294, July 2001.

[40] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[41] M. Grant and S. Boyd. (2014, Mar.) CVX: Matlab software for disciplined convex programming, version 2.1. [Online]. Available: http://cvxr.com/cvx

[42] Y. R. Zheng and C. Xiao, "Channel estimation for frequency-domain equalization of single-carrier broadband wireless communications," *IEEE Trans. Veh. Tech.*, vol. 58, no. 2, pp. 815–823, Feb. 2009.

[43] T. T. B. Nguyen, T. N. Tan, and H. Lee, "Effecient QC-LDPC encoder for 5G new radio," *Electronics*, vol. 8, no. 6, p. 668, June 2019.

[44] N. Benvenuto and S. Tomasin, "On the comparison between OFDM and single carrier modulation with a DFE using a frequency-domain feedforward filter," *IEEE Trans. Commun.*, vol. 50, no. 6, pp. 947–955, June 2002.