KU LEUVEN

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

Citation/Reference	Verdyck K., Lefevre Y., Tsiaflakis P., Moonen M. (2021), **Signal Impropriety in Discrete Multi-Tone Systems and Widely Linear Per-Tone Equalization** IEEE Open Journal of the Comms. Society, vol. 2, Feb. 2021, pp. 367-383.
Archived version	Author manuscript: the content is identical to the content of the published paper, but without the final typesetting by the publisher
Published version	https://doi.org/10.1109/OJCOMS.2021.3058177
Journal homepage	https://www.comsoc.org/publications/journals/ieee-ojcoms
Author contact	jeroen.verdyck@esat.kuleuven.be *+ 32 (0)16 32 47 23*
Abstract	In the context of digital subscriber line (DSL) systems, where the long-reach (LR) extension of G.fast has recently been proposed, interest in techniques dealing with long channel impulse responses (CIRs) without increasing the cyclic prefix (CP) length of the discrete multi-tone (DMT) modulation has recently resurfaced. The technique under consideration in this paper is referred to as *channel shortening*, and applies FIR filters to the received signal to shorten the apparent CIR. Time-domain equalization (TEQ) as well optimal per-tone equalization (PTEQ) FIR filters will be considered. When channel shortening techniques are analyzed, it is often overlooked that the received signals after DMT demodulation are generally *improper* when the CP is too short. Hence, the state-of-the-art signal-to-interference-plus-noise ratio (SINR) and bit loading expressions employed to assess system performance, which implicitly assume the

	received signals to be proper, misrepresent the true achievable performance. New expressions for the SINR and bit loading are therefore presented that explicitly take signal impropriety into account. Based on these expressions, it is then observed that – if the received signals are improper – the PTEQ FIR filters are no longer optimal, and that the achievable bit loading depends on a particular phase shift experienced by the transmitted signals. This paper therefore introduces a novel widely linear PTEQ – which is again optimal when the received signals are improper – and additionally proposes to optimally rotate signals in the complex plane prior to transmission. Finally, this paper assesses the performance increase obtainable by explicitly accounting for signal impropriety.
IR	n/a

(article begins on next page)

Signal Impropriety in Discrete Multi-Tone Systems and Widely Linear Per-Tone Equalization

Jeroen Verdyck, Yannick Lefevre, Paschalis Tsiaflakis, Marc Moonen *Fellow, IEEE*

In the context of digital subscriber line (DSL) systems, where the long-reach (LR) extension of G.fast has recently been proposed, interest in techniques dealing with long channel impulse responses (CIRs) without increasing the cyclic prefix (CP) length of the discrete multi-tone (DMT) modulation has recently resurfaced. The technique under consideration in this paper is referred to as *channel shortening*, and applies FIR filters to the received signal to shorten the apparent CIR. Time-domain equalization (TEQ) as well optimal per-tone equalization (PTEQ) FIR filters will be considered.

When channel shortening techniques are analyzed, it is often overlooked that the received signals after DMT demodulation are generally *improper* when the CP is too short. Hence, the state-of-the-art signal-to-interference-plus-noise ratio (SINR) and bit loading expressions employed to assess system performance, which implicitly assume the received signals to be proper, misrepresent the true achievable performance. New expressions for the SINR and bit loading are therefore presented that explicitly take signal impropriety into account. Based on these expressions, it is then observed that — if the received signals are improper — the PTEQ FIR filters are no longer optimal, and that the achievable bit loading depends on a particular phase shift experienced by the transmitted signals. This paper therefore introduces a novel widely linear PTEQ — which is again optimal when the received signals are improper — and additionally proposes to optimally rotate signals in the complex plane prior to transmission. Finally, this paper assesses the performance increase obtainable by explicitly accounting for signal impropriety.

Index Terms—**Augmented complex, channel shortening, composite real, discrete multi-tone (DMT), digital subscriber lines (DSL), improper signals, proper signals, PTEQ, TEQ, widely linear filters**

I. Introduction

G.fast is a digital subscriber line (DSL) standard [1], [2] for discrete multi-tone (DMT) based transmission over short local loops. An extension of the G.fast standard has recently been proposed, enabling its use on DSL loops with a length in excess of the originally intended $300\,\mathrm{m}$ limit [3]–[7]. These longer lines are currently operated with legacy DSL technologies (such as VDSL). The ability to let them be operated with G.fast would enable convergence to a single DSL standard, thereby eliminating co-existence problems between different DSL generations. Key enabling factors for this so-called long-reach (LR) extension of G.fast will be larger constellation sizes [6], a higher maximum aggregate transmit power [3], [6], the application of a looser spectral mask at low frequencies [3], and a longer cyclic prefix (CP) to support longer lines [3].

The CP — which is added to each DMT symbol prior to transmission — serves a dual purpose: 1) it eliminates inter-symbol-interference (ISI) between subsequent DMT symbols, and 2) it eliminates inter-carrier-interference (ICI) between subcarriers of the same DMT symbol. This elimination of the ISI/ICI is however contingent on the CP length plus one being at least equal to the channel impulse response (CIR) length. Supporting longer CPs as part of the LR extension of G.fast thus seems natural, as longer lines are indeed more dispersive.

This research work was carried out at the ESAT Laboratory of KU Leuven, in the frame of 1) Fonds de la Recherche Scientifique (FNRS) and Fonds Wetenschappelijk Onderzoek - Vlaanderen (FWO) EOS Project no 30452698 "(MUSE-WINET) MUlti-SErvice WIreless NETwork", 2) Research Project FWO nr. G.0B1818N "Real-time adaptive cross-layer dynamic spectrum management for fifth generation broadband copper access networks", and 3) VLAIO O&O Project nr. HBC.2017.1007 "Multi-gigabit Innovations in Access". The scientific responsibility is assumed by its authors.

J. Verdyck and M. Moonen are with the STADIUS Center for Dynamical Systems, Signal Processing, and Data Analytics of the Department of Electrical Engineering (ESAT), KU Leuven, Leuven, Belgium.

Y. Lefevre and P. Tsiaflakis are with Nokia Bell Labs, Antwerp, Belgium.

However, there is an argument to be made against the use of a long CP as well. This argument relates to the fact that a G.fast system requires the same CP length to be applied across all users in the system both in upstream and in downstream [2]. When a mix of long and short lines is connected to the same distribution point unit (DPU), using a longer CP to accommodate the more dispersive long lines will inevitably reduce the data rate on shorter lines. Based on this observation, interest in techniques that can deal with long CIRs without increasing the CP length has recently resurfaced.

The technique under consideration in this paper is referred to as channel shortening. Proposed first by Chow *et al.* [8], early channel shortening methods consisted of applying a time-domain equalization (TEQ) FIR filter to the received signal prior to DMT demodulation. The coefficients of this FIR filter were then chosen such that the channel and TEQ combined have an impulse response which is shorter than the original CIR. Optimal TEQ filter design has however proven to be challenging, leading to a plethora of design heuristics [8]–[15]. In order to simplify the filter design, Van Acker *et al.* [16] proposed moving the TEQ filter to the frequency domain — i.e. to after the DMT demodulation. The resulting channel shortening method is referred to as per-tone equalization (PTEQ), and has two important advantages over TEQ. Firstly, PTEQ generalizes TEQ — i.e. any TEQ filter can be implemented as a PTEQ filter, while the reverse is not true — without incurring a significant complexity increase. Secondly, the PTEQ filter coefficients that maximize the data rate can be calculated analytically, whereas TEQ design is notoriously challenging.

When channel shortening performance is analyzed in the context of DMT systems [8]–[18], there is a particular aspect that is often overlooked: if the CP is too short, then the received signal after DMT demodulation (prior to a soft/hard decision operation) is generally *improper* [19], [20]. A zero-

mean complex random variable Z is proper if its pseudo-variance vanishes [21] — i.e. if $\mathbb{E}\{Z^2\} = 0$ — or, equivalently, if its real and imaginary part are uncorrelated and have equal variance — i.e. if $\mathbb{E}\{\Re(Z)\,\Im(Z)\} = 0$ and $\mathbb{E}\{\Re(Z)^2\} = \mathbb{E}\{\Im(Z)^2\}$.[1] This definition can be extended to a random vector $\boldsymbol{Z}$, which is called proper if its pseudo-covariance matrix vanishes [21] — i.e. if $\mathbb{E}\{\boldsymbol{Z}\,\boldsymbol{Z}^{\mathrm{T}}\} = \mathbf{0}$. Examples of processes that produce proper random signals include the circularly symmetric complex Gaussian noise process, and randomly drawing symbols from a square 16-QAM constellation with uniform probability. An important property of propriety is that it is preserved under complex linear mappings — i.e. if $\mathbf{A} \in \mathbb{C}^{N\times M}$ and $\boldsymbol{Z}$ is a proper random vector, then $\mathbf{A}\,\boldsymbol{Z}$ is proper as well [22]. Consequently, a system's output will be proper if it can be written as a linear function of a proper input.

A first cause of impropriety in DMT systems with a too-short CP is that the received signal after DMT demodulation is not a linear function of the transmitted signal. Generally, it is a widely linear function of the transmitted signal, i.e. the relation between the input of the DMT system — represented by a vector $\boldsymbol{X}$ — and the output after DMT demodulation — represented by a vector $\boldsymbol{Y}$ — can be written as[2]

$$\boldsymbol{Y} = \mathbf{A}_1\,\boldsymbol{X} + \mathbf{A}_2\,\boldsymbol{X}^*. \tag{1}$$

Unlike linear mappings, widely linear mappings do not preserve propriety in general. As an example, consider the random variable $Y = X + X^* = 2\,\Re(X)$. This random variable is clearly improper, even if X is proper. In fact, Y in this example is said to be *maximally improper* because the support of its probability density function degenerates into a line in the complex plane [21]. Likewise in DMT systems, the transmitter's output signal is forced to be real-valued by letting some subcarriers transmit $\boldsymbol{X}$, and letting other subcarriers transmit $\boldsymbol{X}^*$. With a too-short CP, and hence ISI/ICI, this in particular will result in improper ISI/ICI and possibly a loss of orthogonality and difference in scaling between the real and imaginary parts of the transmitted signal.

A second cause of impropriety in DMT systems is the real-valued time-domain noise. If this noise is white, as is often assumed, then the noise after DMT demodulation will be proper [19]. If a TEQ filter is applied before the DMT demodulation however, then the resulting spectrally shaped time-domain noise may yield noise after the DMT demodulation that is improper [19]. A PTEQ can similarly render the noise improper, as the PTEQ output noise is a linear combination of the proper DFT output noise and multiple independent maximally improper noise contributions (see below). It can be concluded that, even if a channel shortening (P)TEQ filter is able to eliminate all ISI/ICI, improper noise can still be present in the received signal after DMT demodulation.

[1]The mathematical notation employed in these paragraphs is introduced at the end of Section I.

[2]Equation (1) is intentionally kept a bit vague at this point. The purpose of this equation is to introduce the sources of signal impropriety in DMT systems, not to give a detailed system model of the considered DMT systems. A more detailed DMT system model will be provided in Section II.

Even though signal impropriety seems to be inherent to DMT systems with a too-short CP using either TEQ or PTEQ, its effects are not taken into account in [8]–[18]. As a result, the state-of-the-art signal-to-interference-plus-noise ratio (SINR) and bit loading expressions employed to assess system performance in these works — which implicitly assume that the received signal after DMT demodulation is proper — mis-represent the true achievable performance. It has also been shown in [19] that explicitly accounting for the noise impropriety can increase performance in DMT systems, even in the absence of ISI/ICI. Even though it is assumed in [19] that the CP is sufficiently long and hence ISI/ICI and channel shortening are not considered, the results in [19] suggest that explicitly accounting for signal impropriety in channel shortening filter design may also increase performance.

The objective of this paper will therefore be twofold. The first objective is to present expressions for the SINR and bit loading that explicitly take all signal impropriety into account — i.e. the impropriety of the ISI/ICI and of the noise. These expressions will be instrumental in assessing the performance of channel shortening methods in LR G.fast systems. It is noted that an accurate performance prediction is crucial to determine the number of subscribers that can be served with a minimum data rate service, a key factor to assess the economic viability of LR G.fast. The second objective is to use these new SINR and bit loading expressions for PTEQ filter design, and to assess the performance increase that can be obtained in LR G.fast systems by explicitly accounting for signal impropriety.

Contributions & Outline

In Section II and Section III, the system model and problem statement will be introduced. In Section IV, new SINR and bit loading expressions will be presented for DMT systems with (P)TEQ that are valid when the received signal after DMT demodulation is improper. In Section V, system performance — which can now be evaluated using the expressions from Section IV — will be optimized.

The performance optimization in Section V will be based on two observations: 1) when the SINR and bit loading expressions from Section IV are employed, no analytical expression is available for the optimal PTEQ filter coefficients, and 2) if the ISI/ICI is improper, then the achievable bit loading of each subcarrier depends on a particular phase shift experienced by the transmitted signal. To address the first observation, it will be proposed in Section V-A to further generalize the PTEQ to a widely linear (WL) PTEQ. The main advantage of WL PTEQ over PTEQ — apart from generalizing PTEQ without incurring a significant complexity increase — will also be introduced: if the SINR and bit loading expressions from Section IV are employed, then the optimal WL PTEQ coefficients can indeed be calculated analytically. To tackle the second problem, it will be proposed in Section V-B to optimally rotate signals in the complex plane prior to transmission.

Finally, in Section VI, LR G.fast DSL simulation results will demonstrate that explicitly taking signal impropriety into account leads to an increased achievable bit loading when residual ISI/ICI is present in the system. When the employed channel shortening method is able to eliminate most of the ISI/ICI, the impact of impropriety is reduced and the gap

between the state-of-the-art SINR and bit loading expressions from [8]–[18] and the newly developed expressions disappears.

Notation

Throughout the paper, matrices will be (blackboard) bold-faced, capitalized, and non-italicized. Vectors will be bold-faced and italicized. Zero-based indexing will be used when addressing vector and matrix elements, i.e. an $N \times M$ matrix $\mathbf{A}$ is defined as follows:

$$\begin{aligned}\mathbf{A}^{\mathrm{T}} &= [(\mathrm{row}_0(\mathbf{A}))^{\mathrm{T}}, \ldots, (\mathrm{row}_{N-1}(\mathbf{A}))^{\mathrm{T}}]\,;\\ \mathbf{A} &= [\mathrm{col}_0(\mathbf{A}), \ldots, \mathrm{col}_{M-1}(\mathbf{A})]\,;\\ [\mathbf{A}]_{n,m} &= \mathrm{row}_n\big(\mathrm{col}_m(\mathbf{A})\big).\end{aligned}$$

The matrices $\mathbf{I}_N$ and $\mathbf{0}_{N\times M}$ will respectively denote the $N \times N$ identity matrix and the $N \times M$ all-zeros matrix. The dimension subscripts in $\mathbf{I}_N$ and $\mathbf{0}_{N\times M}$ will be omitted when they can be derived from the context. Moreover, $\mathbf{A}^{\mathrm{T}}$, $\mathbf{A}^*$, and $\mathbf{A}^{\mathcal{H}}$ will respectively denote the transpose, the complex conjugate, and the Hermitian transpose of $\mathbf{A}$. The real and complex parts of a matrix $\mathbf{A}$ will be denoted as $\Re(\mathbf{A})$ and $\Im(\mathbf{A})$, such that $\mathbf{A} = \Re(\mathbf{A}) + j\,\Im(\mathbf{A})$. The Frobenius norm of a matrix will be denoted as $|\bullet|$. The expected value operator will be denoted as $\mathbb{E}\{\bullet\}$. Finally, the sets of complex, real, positive real, and strictly positive real numbers will respectively be denoted as $\mathbb{C}$, $\mathbb{R}$, $\mathbb{R}_+$, and $\mathbb{R}_{++}$.

II. DMT System Model & Per-Tone Equalization

This section introduces a general model for DMT systems with per-tone equalization (PTEQ), and presents expressions for the achieved SINR and achievable bit loading as they are commonly found in literature [14]–[18].[3]

A. DMT Transceiver

Fig. 1 depicts the signal flow graph of a general DMT system consisting of an encoder $\mathbf{T}^{\mathcal{H}}$, a decoder $\mathbf{R}$, a serial-to-parallel (S/P) and a parallel-to-serial (P/S) converter, and a frequency-selective channel that is modeled as an FIR filter with real coefficients $H(z) = \sum_{k'=0}^{L} z^{-k'}\, h[k']$. It is noted that L does not define the CIR length but rather the channel order. The CIR length is thus $L+1$. The complex vectors $\boldsymbol{X}[k] \triangleq [X_0[k], X_1[k], , \ldots, X_{N-1}[k]]^{\mathrm{T}}$ and $\hat{\boldsymbol{X}}[k] \triangleq [\hat{X}_0[k], \hat{X}_1[k], \ldots, \hat{X}_{N-1}[k]]^{\mathrm{T}}$ respectively denote the frequency-domain input — with each $X_n[k]$ being a quadrature amplitude modulation (QAM) symbol — and the received signal after DMT demodulation (prior to a soft/hard decision operation), where N is the employed DFT size (see below) and k is the DMT symbol index. A different sample index k' is used between the P/S converter and the S/P converter, as these respectively execute an upsampling and a downsampling operation. It will be assumed that $X_n[k]$ is proper — i.e. that $\mathbb{E}\{(X_n[k])^2\} = 0$ — and that the time-domain noise $z_{\mathrm{t}}[k']$ is white. The subscript 't' is used for time-domain signals. The subcarrier power and time-domain noise power will respectively be denoted as $S_n \triangleq \mathbb{E}\{X_n[k]\, X_n^*[k]\}$ and $\sigma_{\mathrm{t}}^2 = \mathbb{E}\{|z_{\mathrm{t}}[k']|^2\}$.

CP-based DMT will be considered, in which the transmitter modulates the complex vector $\boldsymbol{X}[k]$ by application of an inverse discrete Fourier transform (IDFT), by adding a CP and a cyclic suffix (CS) to the IDFT output, and by windowing the result. The k-th transmitted DMT symbol is thus

$$\boldsymbol{x}[k] = \mathbf{T}^{\mathcal{H}}\,\boldsymbol{X}[k] = \mathbf{C}\left[\begin{array}{c|c} \mathbf{0}_{\nu_{\mathrm{p}}\times(N-\nu_{\mathrm{p}})} & \mathbf{I}_{\nu_{\mathrm{p}}} \\ \hline \multicolumn{2}{c}{\mathbf{I}_N} \\ \hline \mathbf{I}_{\nu_{\mathrm{s}}} & \mathbf{0}_{\nu_{\mathrm{s}}\times(N-\nu_{\mathrm{s}})} \end{array}\right] \mathbf{F}_N^{-1}\,\boldsymbol{X}[k], \tag{2}$$

with ν_{p} the CP length, ν_{s} the CS length, $\mathbf{C}$ the diagonal windowing matrix, and $\mathbf{F}_N^{-1}$ the N-point IDFT matrix. The following definition is used for the N-point discrete Fourier transform (DFT) matrix $\mathbf{F}_N$.

$$[\mathbf{F}_N]_{n,m} \triangleq \alpha^{-n\,m} \qquad \alpha \triangleq e^{2\pi j/N} \tag{3}$$

It is assumed that the windowing matrix $\mathbf{C}$ satisfies

$$[\mathbf{C}]_{i,i} = 1 \qquad \text{for } \nu_{\mathrm{s}} \le i \le N + \nu_{\mathrm{p}} - 1, \tag{4a}$$

$$[\mathbf{C}]_{i,i} = 1 - [\mathbf{C}]_{j,j} \quad \text{for } 0 \le i \le \nu_{\mathrm{s}} - 1 \text{ and } j = N + \nu_{\mathrm{p}} + i. \tag{4b}$$

The employed window is thus symmetric, and its flat part is $N + \nu_{\mathrm{p}} - \nu_{\mathrm{s}}$ samples long.[4] Prior to transmission, subsequent windowed DMT symbols $\boldsymbol{x}[k]$ are overlapped by ν_{s} samples and summed together to form the time-domain symbol stream $x_{\mathrm{t}}[k']$.[5] Taking this overlap of DMT symbols in the time domain into account, it is seen that the DMT symbol period can be denoted as $s = N + \nu_{\mathrm{p}}$.

At the receiver, the cyclic prefix and suffix are discarded, yielding the received signal vector $\boldsymbol{y}[k] \triangleq \big[y_{\mathrm{t}}[ks], y_{\mathrm{t}}[ks+1], \ldots, y_{\mathrm{t}}[ks+N-1]\big]^{\mathrm{T}}$. DMT demodulation further consists of the application of a DFT and a frequency-domain equalization (FEQ), such that the output $\hat{\boldsymbol{X}}[k]$ can be expressed as

$$\hat{\boldsymbol{X}}[k] = \mathbf{E}\,\mathbf{R}\,\boldsymbol{y}[k] = \mathbf{E}\,\mathbf{F}_N\,\boldsymbol{y}[k], \tag{5}$$

with $\mathbf{E}$ the diagonal FEQ matrix. Assuming perfect synchronization between the transmitter-side P/S converter and receiver-side S/P converter, each $\hat{X}_n[k]$ will be free of ISI/ICI if $\nu_{\mathrm{p}} \ge L + \nu_{\mathrm{s}}$.

DMT systems operate at baseband, resulting in the requirement that the transmitted signal $\boldsymbol{x}[k]$ be real-valued. In (2), this requirement is satisfied if and only if $\boldsymbol{X}[k]$ admits the following Hermitian symmetric structure:

$$X_{n^*}[k] = X_n^*[k],\ \forall n, \tag{6}$$

where — with a slight abuse of notation — n^* is the index of the subcarrier that is the Hermitian symmetric of subcarrier

[3] References [8]–[13] have not been included here, as the SINR expressions in these works do not take part of the desired signal or ISI/ICI into account.

[4] At this point, one could also explicitly give the elements of the windowing matrix rather than defining the constraint as in (4). However, the G.fast standard [2] mentions that the "values of the window samples are vendor discretionary." By defining the exact values of the window matrix elements only in Section VI, the developed theoretical derivations in the ensuing sections are — like the G.fast standard — valid regardless of the employed transmit window.

[5] A more detailed mathematical representation of this overlap-and-sum operation is given in (12), see below.

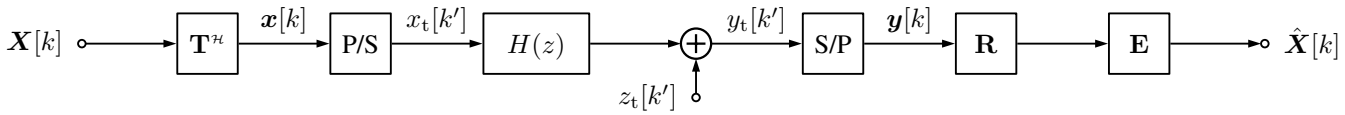


Fig. 1: Signal flow graph of a general DMT system.

n, i.e. where $n^* \triangleq (N-n) \mod N$. The DFT size N is assumed to be even.

B. Channel Shortening - TEQ & PTEQ

PTEQ [16] is introduced here by first adding a time-domain equalization (TEQ) filter to the receiver, and then deriving PTEQ from TEQ. The idea of TEQ is to apply a T-order FIR filter $W(z) = \sum_{k'=0}^{T} z^{-k'}\, w[k']$ with real coefficients to $y_{\mathrm{t}}[k']$ in order to shorten the channel impulse response (CIR), thereby reducing the required ν_{p} for ISI/ICI-free transmission and hence decreasing the corresponding overhead. The DMT receiver with TEQ thus convolves the received signal $y_{\mathrm{t}}[k']$ with the filter coefficients of $W(z)$ before applying the DMT decoding and equalization matrices $\mathbf{R}$ and $\mathbf{E}$. The operation executed by the DMT receiver with TEQ can thus be expressed as

$$\hat{\boldsymbol{X}}[k] = \mathbf{E}\,\mathbf{R} \begin{bmatrix} \boxed{\boldsymbol{w}^{\mathrm{T}}} & 0 & \cdots & 0 \\ 0 & \boxed{\boldsymbol{w}^{\mathrm{T}}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \boxed{\boldsymbol{w}^{\mathrm{T}}} \end{bmatrix} \bar{\boldsymbol{y}}[k], \qquad (7)$$

where

$$\bar{\boldsymbol{y}}[k] \triangleq \big[y_{\mathrm{t}}[ks-T], y_{\mathrm{t}}[ks-T+1], \dots, y_{\mathrm{t}}[ks+N-1]\big]^{\mathrm{T}}$$

is the extended received signal vector, and the $N \times (N+T)$ Toeplitz matrix contains N shifted copies of the TEQ filter coefficient vector $\boldsymbol{w}^{\mathrm{T}} \triangleq \big[w[T], w[T-1], \dots, w[0]\big]$.

PTEQ is obtained from TEQ by transferring the filter $W(z)$ to the frequency domain — i.e. to after the DFT operation [16] — which can be done based on the following reformulation of equation (7), where $E_n \triangleq \big[\mathbf{E}\big]_{n,n}$.

$$\hat{X}_n[k] = \underbrace{(E_n\, \boldsymbol{w}^{\mathrm{T}})}_{\triangleq \boldsymbol{w}_n^{\mathcal{H}}} \begin{bmatrix} \boxed{\mathrm{row}_n(\mathbf{R})} & 0 & \cdots & 0 \\ 0 & \boxed{\mathrm{row}_n(\mathbf{R})} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \boxed{\mathrm{row}_n(\mathbf{R})} \end{bmatrix} \bar{\boldsymbol{y}}[k] \qquad (8)$$

The $(T+1) \times (N+T)$ Toeplitz matrix in (8) contains $T+1$ shifted copies of $\mathrm{row}_n(\mathbf{R})$. Intuitively, the equivalence between (7) and (8) can be understood by interpreting the multiplication with $\mathbf{R}$ in (7) as applying a filter bank to the output of the TEQ — with FIR filters given as $R_n(z) = \sum_{k'=0}^{N-1} z^{-k'}\, \mathrm{col}_{N-1-k'}\,(\mathrm{row}_n(\mathbf{R}))$ — followed by a downsampling operation, and considering the fact that the filters $R_n(z)$ and $W(z)$ commute. Equation (8) reveals that the receiver with PTEQ allows each subcarrier to have its own TEQ, which is then implemented as a T-th order FEQ filter with filter coefficients $\boldsymbol{w}_n^{\mathcal{H}} \triangleq \big[w_n[T], w_n[T-1], \dots, w_n[0]\big]$.

It may appear that PTEQ incurs a significant run-time complexity increase w.r.t. TEQ, as the direct implementation of (8) would effectively require $T+1$ DFT operations (cf. number of rows in the Toeplitz matrix) to be executed [16]. Moreover, even though the NT time-domain multiplications per received DMT symbol required by TEQ have been replaced by $\frac{N}{2}+1$ times T multiplications in the frequency domain — apparently halving the required number of multiplications — these frequency-domain multiplications are complex-valued, e.g. requiring 4 real multiplications each and further adding to the run-time complexity of PTEQ. Fortunately, it has been shown in [16] that both effects can be alleviated by decomposing the Toeplitz matrix in (8) into a lower triangular matrix $\mathbf{L}_n$ and a sparse matrix $\mathbf{R}_n$, i.e.

$$\begin{bmatrix} \boxed{\mathrm{row}_n(\mathbf{R})} & 0 & \cdots & 0 \\ 0 & \boxed{\mathrm{row}_n(\mathbf{R})} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \boxed{\mathrm{row}_n(\mathbf{R})} \end{bmatrix} = \mathbf{L}_n\, \mathbf{R}_n \qquad (9)$$

where

$$\mathbf{L}_n \triangleq \begin{bmatrix} \alpha^{0n} & 0 & \cdots & 0 \\ \alpha^{1n} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \alpha^{Tn} & \cdots & \alpha^{1n} & \alpha^{0n} \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0}_{1\times T} \\ \mathbf{0}_{T\times 1} & \alpha^n\, \mathbf{I}_T \end{bmatrix}, \qquad (10a)$$

$$\mathbf{R}_n \triangleq \left[\begin{array}{cc|c} & \boxed{\mathrm{row}_n(\mathbf{R})} & \mathbf{0}_{1\times T} \\ -\mathbf{I}_T & \mathbf{0}_{T\times(N-T)} & \mathbf{I}_T \end{array}\right]. \qquad (10b)$$

With this decomposition, equation (8) becomes

$$\hat{X}_n[k] = \boldsymbol{v}_n^{\mathcal{H}}\, \mathbf{R}_n\, \bar{\boldsymbol{y}}[k], \qquad (11)$$

where $\boldsymbol{v}_n \triangleq \mathbf{L}_n^{\mathcal{H}}\, \boldsymbol{w}_n$ is merely an alternative PTEQ filter representation. In addition, $\boldsymbol{y}_n[k] \triangleq \mathbf{R}_n\, \bar{\boldsymbol{y}}[k]$ will denote the vector of PTEQ filter inputs on subcarrier n. The receiver with PTEQ can thus obtain $\hat{X}_n[k]$ as a linear combination of the output of a single DFT (cf. first row of $\mathbf{R}_n$) and T real-valued difference terms $y_{\mathrm{t}}[ks+N-t] - y_{\mathrm{t}}[ks-t]$ with $t \in \{1, 2, \dots, T\}$ (cf. last T rows of $\mathbf{R}_n$).

A more detailed overview of the run-time complexity of both TEQ and PTEQ is given in Table I. It has been assumed that the subcarriers corresponding to the DC and Nyquist frequency are inactive, and that w.l.o.g. the first coefficient of the TEQ filter equals 1. Moreover, $n_\times$ and n_+ represent the number of real multiplications and additions that are required to execute a complex multiplication. Note that the values of $n_\times$ and n_+ are implementation-dependent. The most common

TABLE I: Number of floating point operations required by channel shortening and equalization a DMT receiver.

	Nb. real mult.	Nb. real add.
TEQ	TN	TN
FEQ	$n_\times \left(\frac{N}{2}-1\right)$	$n_+ \left(\frac{N}{2}-1\right)$
T+F	$\left(\frac{N}{2}-1\right)(2T+n_\times)+2T$	$\left(\frac{N}{2}-1\right)(2T+n_+)+2T$
PTEQ	$\left(\frac{N}{2}-1\right)(2T+n_\times)$	$\left(\frac{N}{2}-1\right)(2T+n_+)+T$

implementations result in $n_\times = 4$ and $n_+ = 2$ [22], or in $n_\times = 3$ and $n_+ = 5$ [23, pp. 591 & 706]. Regardless of which implementation is considered, it can be concluded from Table I that PTEQ has a complexity that is comparable to that of the combination of a TEQ and a FEQ (T+F).

C. Signal-to-Interference-plus-Noise Ratio

An expression for the SINR in DMT systems using PTEQ can be derived as follows. The channel equation for a DMT system using PTEQ is given in (12a) at the top of the next page. In (12a), the $(N+T)\times(3s+\nu_\mathrm{s})$ Toeplitz matrix $\mathbf{H}$ contains $N+T$ shifted copies of the channel vector $\boldsymbol{h}^\mathrm{T} \triangleq \big[h[L], h[L-1], \ldots, h[0]\big]$. Moreover, the extended noise vector is denoted as $\bar{\boldsymbol{z}}[k] \triangleq \big[z_\mathrm{t}[ks-T], z_\mathrm{t}[ks-T+1], \ldots, z_\mathrm{t}[ks+N-1]\big]^\mathrm{T}$. Furthermore, the dimensions of the all-zero matrices $\mathbf{O}_1$ and $\mathbf{O}_2$ are respectively given as $(N+T)\times(s+\nu_\mathrm{p}-L-T+\delta)$ and $(N+T)\times(s+\nu_\mathrm{s}-\delta)$, where δ denotes the so-called synchronization delay [16].[6] By partitioning the Toeplitz matrix $\mathbf{H}$ into three overlapping $(N+T)\times(s+\nu_\mathrm{s})$ matrices $\{\mathbf{H}[l]\}_{l=\{-1,0,1\}}$, equation (12a) can be rewritten as

$$\bar{\boldsymbol{y}}[k] = \sum_{l=-1}^{1} \mathbf{H}[l] \sum_{m=0}^{N-1} \mathrm{row}_m(\mathbf{T})^\mathcal{H}\, X_m[k+l] + \bar{\boldsymbol{z}}[k]. \quad (12b)$$

Finally, by defining the residual channel vectors $\boldsymbol{h}_{nm}[l]$ and residual channel matrices $\mathbb{H}_{nm}$ as

$$\boldsymbol{h}_{nm}[l] \triangleq \mathbf{R}_n\, \mathbf{H}[l]\ \mathrm{row}_m(\mathbf{T})^\mathcal{H}, \quad (13a)$$

$$\mathbb{H}_{nm} \triangleq \begin{cases} \begin{bmatrix} \boldsymbol{h}_{nm}[-1] & \boldsymbol{h}_{nm}[0] & \boldsymbol{h}_{nm}[1] \end{bmatrix} \\ \qquad\qquad \text{if } m \notin \{n, n^*\} \\ \begin{bmatrix} \boldsymbol{h}_{nm}[-1] & \mathbf{0}_{(T+1)\times 1} & \boldsymbol{h}_{nm}[1] \end{bmatrix} \\ \qquad\qquad \text{if } m \in \{n, n^*\} \end{cases} \quad (13b)$$

and by introducing $\boldsymbol{h}_n \triangleq \boldsymbol{h}_{nn}[0]$, $\tilde{\boldsymbol{h}}_n \triangleq \boldsymbol{h}_{nn^*}[0]$, and $\boldsymbol{X}_m[k] \triangleq [X_m[k-1], X_m[k], X_m[k+1]]^\mathrm{T}$, equations (11) and (12b) can be joined into the single input-output equation (14) at the top of the next page.

From (14), the following expression for the SINR on subcarrier n can be obtained:

$$\gamma_n = \frac{\boldsymbol{v}_n^\mathcal{H}\ (S_n\, \boldsymbol{h}_n\, \boldsymbol{h}_n^\mathcal{H})\ \boldsymbol{v}_n}{\boldsymbol{v}_n^\mathcal{H}\ \left(\boldsymbol{\Psi}_{\boldsymbol{z}_n\boldsymbol{z}_n} + S_n\, \tilde{\boldsymbol{h}}_n\, \tilde{\boldsymbol{h}}_n^\mathcal{H}\right)\ \boldsymbol{v}_n}, \quad (15)$$

[6]It has implicitly been assumed that $s+\nu_\mathrm{p}+\delta \geq L+T$ and $s+\nu_\mathrm{s} \geq \delta$. The case where either condition is not satisfied can be accommodated for by extending the rows of $\mathbf{H}$ and including more interfering symbols $\boldsymbol{X}[k-a]$.

where $\boldsymbol{\Psi}_{\boldsymbol{z}_n\boldsymbol{z}_n}$ is the interference-plus-noise covariance matrix of subcarrier n, which is defined as

$$\boldsymbol{\Psi}_{\boldsymbol{z}_n\boldsymbol{z}_n} \triangleq \mathbb{E}\{\boldsymbol{z}_n\, \boldsymbol{z}_n^\mathcal{H}\} = \sum_{m=0}^{N-1} S_m\, \mathbb{H}_{nm}\, \mathbb{H}_{nm}^\mathcal{H} + \sigma_\mathrm{t}^2\, \mathbf{R}_n\, \mathbf{R}_n^\mathcal{H}. \quad (16)$$

Equivalent SINR expressions have been used for the design and/or performance evaluation of DMT systems in, among others, [14]–[18]. Despite this widespread adoption of the SINR expression as in (15), however, it seems ill-advised to model the ICI coming from subcarrier n^* as uncorrelated *noise* on subcarrier n. While the considered interfering signal — which is given as $X_n^*[k]$ due to (6) — is indeed uncorrelated with $X_n[k]$, it is not *independent* of $X_n[k]$ and will cause a loss of orthogonality and difference in scaling between the real and imaginary parts of $X_n[k]$ in $\hat{X}_n[k]$ rather than contributing to the total interference-plus-noise. This loss of orthogonality and difference in scaling is a manifestation of the received signals' impropriety. In the same vein, interference from both $X_m[k]$ and $X_m^*[k]$ may result in a loss of propriety for the ISI/ICI, which is not accounted for in (15) in general.

D. Bit Loading

A relevant performance metric for DMT systems is the total achievable bit loading, where the achievable bit loading on subcarrier n is often expressed as

$$b_n = \log_2\left(1 + \Gamma^{-1}\, \gamma_n\right) \quad (17)$$

with Γ the SNR gap-to-capacity. The value of Γ depends on the target bit error rate (BER). Adjustments are often made to the SNR gap to account for a noise margin and — in case coded transmission is considered — a coding gain. Similar bit loading expressions have been used for the design and/or performance evaluation of DMT systems in, among others, [8]–[18].

If the PTEQ filter coefficients are chosen to maximize the bit loading, then it is readily seen that (15) and (17) yield the following expression for the optimal PTEQ filter coefficients.

$$\boldsymbol{v}_n = \left(\boldsymbol{\Psi}_{\boldsymbol{z}_n\boldsymbol{z}_n} + S_n\, \tilde{\boldsymbol{h}}_n\, \tilde{\boldsymbol{h}}_n^\mathcal{H}\right)^{-1} \boldsymbol{h}_n \quad (18)$$

Equation (18) demonstrates the main advantage of PTEQ over TEQ — apart from generalizing TEQ without incurring a significant run-time complexity increase: the optimal PTEQ filter coefficients can be calculated *analytically* [16], whereas TEQ design is notoriously challenging.

III. Problem Statement

When the bit loading is calculated using (17), it is implicitly assumed that the received signals are proper. The fact that this assumption is not always satisfied, is illustrated in Fig. 2 which displays a received signal after DMT demodulation in a DMT system with a too-short cyclic prefix and with PTEQ filter vectors $\boldsymbol{v} = [1, 0, \ldots, 0]^\mathrm{T}$.[7] The impropriety of

[7]A detailed description of the Monte-Carlo simulations that yielded these figures can be found in Section VI-A.

$$\bar{\boldsymbol{y}}[k] = \underbrace{\left[\begin{array}{c|ccc|c} \boldsymbol{O}_1 & \boxed{\boldsymbol{h}^{\mathrm{T}}} & 0 & \cdots & 0 & \boldsymbol{O}_2 \\ & 0 & \boxed{\boldsymbol{h}^{\mathrm{T}}} & \ddots & \vdots & \\ & \vdots & \ddots & \ddots & 0 & \\ & 0 & \ldots & 0 & \boxed{\boldsymbol{h}^{\mathrm{T}}} & \end{array}\right]}_{\triangleq \mathbf{H}} \left[\begin{array}{ccc} \boxed{\mathbf{T}^{\mathcal{H}}} & \mathbf{0}_{s\times N} & \mathbf{0}_{s\times N} \\ \mathbf{0}_{s\times N} & \boxed{\mathbf{T}^{\mathcal{H}}} & \mathbf{0}_{s\times N} \\ \mathbf{0}_{s\times N} & \mathbf{0}_{s\times N} & \boxed{\mathbf{T}^{\mathcal{H}}} \end{array}\right] \left[\begin{array}{c} \boldsymbol{X}[k-1] \\ \boldsymbol{X}[k] \\ \boldsymbol{X}[k+1] \end{array}\right] + \bar{\boldsymbol{z}}[k] \tag{12a}$$

$$\hat{X}_n[k] = \boldsymbol{v}_n^{\mathcal{H}} \left(\underbrace{\left[\begin{array}{cc} \boldsymbol{h}_n & \tilde{\boldsymbol{h}}_n \end{array}\right]}_{\triangleq \mathbf{H}_n} \left[\begin{array}{c} X_n[k] \\ X_n^*[k] \end{array}\right] + \underbrace{\sum_{m=0}^{N-1} \mathbb{H}_{nm}\, \boldsymbol{X}_m[k] + \mathbf{R}_n\, \bar{\boldsymbol{z}}[k]}_{\triangleq \boldsymbol{z}_n[k]} \right). \tag{14}$$

TABLE II: Achieved SINR and maximum bit loading for the received signals in Fig. 2.

	Fig. 2a: Original			Fig. 2b: Rotated			Fig. 2c: Proper		
	$\gamma_n^{\Re}$ (dB)	$\gamma_n^{\Im}$ (dB)	b_n (bit)	$\gamma_n^{\Re}$ (dB)	$\gamma_n^{\Im}$ (dB)	b_n (bit)	$\gamma_n^{\Re}$ (dB)	$\gamma_n^{\Im}$ (dB)	b_n (bit)
(15) and (17)	35.5		8.61	35.5		8.61	35.5		8.61
(26) and (29)	35.1	35.8	8.61	32.5	50.9	10.68	35.5	35.5	8.61

the interference-plus-noise in these examples is caused by the DMT symbols' Hermitian symmetric structure, which is described in (6). Due to this Hermitian symmetric structure, the ISI/ICI is a widely linear function of the transmitted proper QAM symbols. As widely linear mappings do not preserve propriety, the resulting ISI/ICI can be improper.

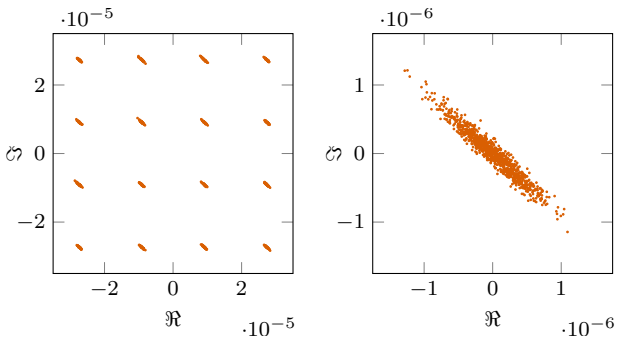
Equation (17) establishes a relation between the SINR and achievable bit loading for a given target BER, and is derived under the assumption that interference-plus-noise is proper and has a Gaussian distribution. The problem is not just that (17) is invalid when the received signals are improper, but also that (15) contains no information regarding the signal impropriety. This fact is reflected in the SINR values yielded by (15) for the three received signals in Fig. 2 — given in the first row of Table II. The SINR values are all equal, even though the interference-plus-noise of Fig. 2b will result in a higher (i.e. worse) BER than that in Fig. 2a and Fig. 2c.[8]

For a given constellation size and SNR, determining the achieved BER thus requires knowledge of both the interference-plus-noise variance and pseudo-variance (i.e. requires a full second order statistical characterization of the interference-plus-noise). Equivalently, if the interference-plus-noise is improper and a given target BER is to be achieved, then the maximum achievable bit loading cannot be determined given only the interference-plus-noise variance (which is expressed relatively to the direct signal power in the SINR). The motivation for why (15) and (17) cannot be considered valid, can thus also be formulated as follows: if some aspect of the interference-plus-noise distribution influences the relation between the achievable bit loading (i.e. the QAM size), the SINR, and target BER but is left un-modeled, then the highest bit loading that is guaranteed to achieve the target BER cannot be determined reliably.

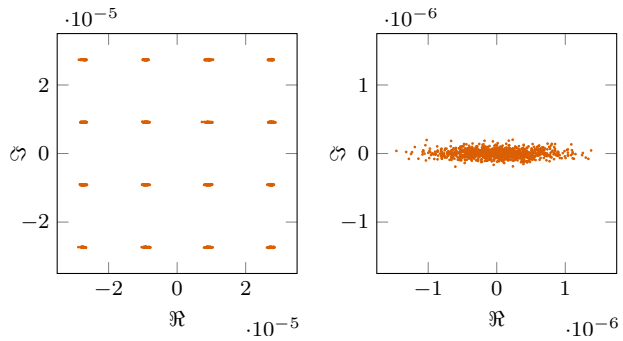
Signal impropriety however need not be a source of performance degradation [19]. In order to take advantage of signal impropriety, this paper follows [19] by decomposing each transmitted QAM symbol $X_n[k]$ into two independent pulse-amplitude modulation (PAM) symbols and by applying the PAM decoding decision independently to the real signals $\Re(X_n[k])$ and $\Im(X_n[k])$. Similar equations to (15) and (17) will be developed for PAM decoding in Section IV, which again establish a relation between the SINR and achievable bit loading given a target BER. These equations are derived under the assumption that the interference-plus-noise in $\Re(X_n[k])$ (or $\Im(X_n[k])$) has a 1-D Gaussian distribution. Note that these assumptions include no statement about the interference-plus-noise being proper. Contrary to what was the case for QAM, no such assumption is required here as the distribution of a zero-mean real Gaussian signal is fully determined by its variance — which, for the considered interference-plus-noise signal, is expressed relatively to the direct signal power in the SINR. Contrary to what was the case for QAM decoding, it is thus no longer possible that some un-modeled aspect of the interference-plus-noise distribution in $\Re(\hat{X}_n[k])$ (or $\Im(\hat{X}_n[k])$) influences the relation between the achievable bit loading, the SINR, and the target BER. As a result, the highest bit loading that is guaranteed to achieve the target BER can be determined reliably based solely on the SINR. In the example of Fig. 2b, these new expressions replacing (15) and (17) will not only guarantee that the target BER is achieved: employing a different PAM constellation for the real and imaginary part of $X_n[k]$ will additionally enable a larger number of bits to be loaded in the $\Im$-direction of $X_n[k]$ than in the $\Re$-direction, thereby taking advantage of the signal impropriety and yielding a higher overall bit loading.

The increased bit loading yielded by the newly developed expressions (cf. Section IV) however comes at a price: maxi-

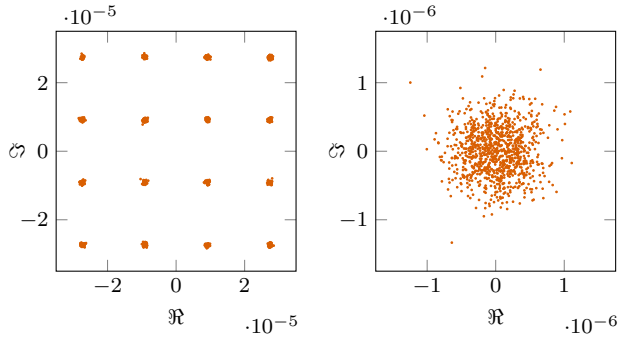
[8]Provided that the QAM decision regions are constructed under the assumption that the interference-plus-noise is proper. Note that the interference-plus-noise in Fig. 2a will lead to unexpected dependencies between error events, which can be detrimental to the BER when coded modulation is employed without taking the interference-plus-noise impropriety into account.



(a) Original signal with improper interference-plus-noise



(b) Signal with rotated improper interference-plus-noise



(c) Signal with proper interference-plus-noise

Fig. 2: Example of a received signal after DMT demodulation in a G.fast system with a 600 m loop length and $\nu_{\mathrm{p}} = 128$. In each pair of figures, the left figure depicts $[\boldsymbol{y}_n[k]]_0$ for a single subcarrier ($n = 100$), and the right figure depicts the interference-plus-noise component of the same signal. Fig. 2a displays the received signal as obtained through the Monte-Carlo simulation of the described DMT system. It can be observed that the interference-plus-noise is improper. Fig. 2b depicts the same received signal, but where the interference-plus-noise component has been rotated relative to the received QAM symbol (cf. Section V-B). Fig. 2c illustrates what the received signal looks like if the interference-plus-noise is proper with the same power as in Fig. 2a and Fig. 2b. Further details regarding the simulation setup can be found in Section VI-A.

mizing the data rate of the DMT system will no longer result in analytical expressions for the optimal PTEQ filter coefficients. To tackle this issue, it will be proposed to further generalize the PTEQ to a novel *widely linear* (WL) PTEQ. Section V will demonstrate the main advantage of WL PTEQ over PTEQ — apart from generalizing PTEQ without incurring a significant complexity increase: if the new SINR and rate expressions from Section IV are employed, then the optimal WL PTEQ coefficients can be calculated *analytically*.

Even when the newly developed bit loading expressions (cf. Section IV) and novel WL PTEQ filter (cf. Section V) are employed, the interference-plus-noise affecting the received PAM symbols on tone n may still be correlated (as e.g. in Fig. 2a). As a result, errors will not occur independently in these two PAM symbols. This dependence between the error events in the two PAM symbols in $X_n[k]$ will then complicate the design of the trellis coded modulation (TCM), the use of which is mandatory in G.fast [2]. Additionally, it will also be observed that the relative rotation of the interference-plus-noise distribution in $\hat{X}_n[k]$ strongly impacts the achievable bit loading. To further improve the achievable bit loading, it is therefore proposed to rotate each QAM symbol $X_n[k]$ prior to transmission.[9] It will be proven that the proposed QAM symbol rotation maximizes the achievable bit loading, and that such optimal rotation effectively decorrelates the noise in the real and imaginary parts of the received signal, thereby simplifying TCM design and analysis.

IV. Composite Real Channel Model

This section provides expressions for the SINR of the real and imaginary part of $\hat{X}_n[k]$ and an SNR gap approximation for the achievable bit loading. The expressions for the SINR will be based on a new composite real channel model, which expresses the real signals $\Re(\hat{X}_n[k])$ and $\Im(\hat{X}_n[k])$ as linear functions of the real signals $\Re(X_m[l])$, $\Im(X_m[l])$, and $\bar{\boldsymbol{z}}[k]$. It will be assumed that $X_0[k] \equiv 0$ and $X_{\frac{N}{2}}[k] \equiv 0$ in order to simplify notation.

A. Signal-to-Interference-plus-Noise Ratio

The starting point for the derivation of the SINR expressions for the real and imaginary part of $\hat{X}_n[k]$ will be the following

[9]This type of transmitter-side symbol rotation is compatible with the G.fast standard [2], as the standard specifies that a complex gain scaling should be applied to the transmitted symbols upon updating the vectoring matrix in order to absorb "FEQ shock".

augmented complex channel model (equivalent to (12b))

$$\bar{\boldsymbol{y}}[k] = \sum_{l=-1}^{1} \mathbf{H}[l] \sum_{m=1}^{N/2-1} \begin{bmatrix} \mathrm{row}_m(\mathbf{T}) \\ \mathrm{row}_m^*(\mathbf{T}) \end{bmatrix}^{\mathcal{H}} \begin{bmatrix} X_m[k+l] \\ X_m^*[k+l] \end{bmatrix} + \bar{\boldsymbol{z}}[k] \tag{19}$$

in conjunction with the augmented complex PTEQ filter equation (equivalent to (11))

$$\begin{bmatrix} \hat{X}_n[k] \\ \hat{X}_n^*[k] \end{bmatrix} = \begin{bmatrix} \boldsymbol{v}_n & \mathbf{0}_{(T+1)\times 1} \\ \mathbf{0}_{(T+1)\times 1} & \boldsymbol{v}_n^* \end{bmatrix}^{\mathcal{H}} \begin{bmatrix} \mathbf{R}_n \\ \mathbf{R}_n^* \end{bmatrix} \bar{\boldsymbol{y}}[k]. \tag{20}$$

In (19), $\mathrm{row}_n^*(\mathbf{T})$ denotes the complex conjugate of $\mathrm{row}_n(\mathbf{T})$. Equation (19) was obtained from (12b) by considering (6) and the relation $\mathrm{row}_n^*(\mathbf{T}) = \mathrm{row}_{n^*}(\mathbf{T})$. Similarly, equation (20) can be obtained from (11) by employing the fact that $\bar{\boldsymbol{y}}[k]$ is real-valued.

From the augmented complex channel model as in (19), the following *composite real* channel model can then be obtained.

$$\bar{\boldsymbol{y}}[k] = 2 \sum_{l=-1}^{1} \mathbf{H}[l] \sum_{m=1}^{N/2-1} \left(\mathrm{row}_m^{\mathbb{R}}(\mathbf{T})\right)^{\mathsf{T}} X_m^{\mathbb{R}}[k+l] + \bar{\boldsymbol{z}}[k] \tag{21}$$

In this equation, $\mathbf{A}^{\mathbb{R}}$ denotes the composite real representation of a complex matrix $\mathbf{A}$, which is defined as

$$\mathbf{A}^{\mathbb{R}} \triangleq \begin{bmatrix} \Re(\mathbf{A}) \\ \Im(\mathbf{A}) \end{bmatrix}$$

where $\Re(\mathbf{A})$ and $\Im(\mathbf{A})$ respectively denote the real and imaginary part of a matrix $\mathbf{A}$. In (21), $\mathrm{row}_m^{\mathbb{R}}(\mathbf{T})$ is thus given as

$$\mathrm{row}_m^{\mathbb{R}}(\mathbf{T}) = \begin{bmatrix} \Re(\mathrm{row}_m(\mathbf{T})) \\ \Im(\mathrm{row}_m(\mathbf{T})) \end{bmatrix}.$$

To obtain (21) from (19), the following relations are used:

$$\begin{bmatrix} \mathbf{A} \\ \mathbf{A}^* \end{bmatrix} = \mathbf{\Pi}\, \mathbf{A}^{\mathbb{R}}, \quad \mathbf{A}^{\mathbb{R}} = \tfrac{1}{2}\, \mathbf{\Pi}^{\mathcal{H}} \begin{bmatrix} \mathbf{A} \\ \mathbf{A}^* \end{bmatrix} \tag{22}$$

with

$$\mathbf{\Pi} \triangleq \begin{bmatrix} \mathbf{I} & j\,\mathbf{I} \\ \mathbf{I} & -j\,\mathbf{I} \end{bmatrix}.$$

Similarly, the following composite real PTEQ filter equation can be obtained from (20)

$$\hat{X}_n^{\mathbb{R}}[k] = \underbrace{\begin{bmatrix} \Re(\boldsymbol{v}_n) & -\Im(\boldsymbol{v}_n) \\ \Im(\boldsymbol{v}_n) & \Re(\boldsymbol{v}_n) \end{bmatrix}^{\mathsf{T}}}_{\triangleq \mathbf{U}_n^{\mathsf{T}}} \mathbf{R}_n^{\mathbb{R}}\, \bar{\boldsymbol{y}}[k]. \tag{23}$$

In (23), $\mathbf{R}_n^{\mathbb{R}}$ thus has dimensions $2(T+1)\times(N+T)$. Moreover, $\boldsymbol{y}_n^{\mathbb{R}}[k] = \mathbf{R}_n^{\mathbb{R}}\, \bar{\boldsymbol{y}}[k]$ will denote the composite real representation of PTEQ filter inputs $\boldsymbol{y}_n[k]$ on subcarrier n.

By defining the residual channel vectors $\boldsymbol{g}_{nm}[l]$ and residual channel matrices $\mathbb{G}_{nm}$ as

$$\boldsymbol{g}_{nm}[l] \triangleq 2\, \mathbf{R}_n^{\mathbb{R}}\, \mathbf{H}[l]\, \mathrm{row}_m(\mathbf{T})^{\mathsf{T}}, \tag{24a}$$

$$\mathbb{G}_{nm} \triangleq \begin{cases} \begin{bmatrix} \boldsymbol{g}_{nm}[-1] & \boldsymbol{g}_{nm}[0] & \boldsymbol{g}_{nm}[1] \end{bmatrix} & \text{if } m \neq n \\ \begin{bmatrix} \boldsymbol{g}_{nm}[-1] & \mathbf{0}_{2(T+1)\times 1} & \boldsymbol{g}_{nm}[1] \end{bmatrix} & \text{if } m = n \end{cases}, \tag{24b}$$

and by introducing $\boldsymbol{g}_n = \boldsymbol{g}_{nn}[0]$, equations (23) and (21) can be joined into the single input-output equation (25) at the top of the next page (equivalent to (14)).

It was assumed that $X_n[k]$ is proper, such that the symbol power S_n is distributed equally over the real and imaginary component of $X_n[k]$.[10] From (25), the following expression for the SINR of $\Re(\hat{X}_n[k])$ can therefore be obtained:

$$\gamma_n^{\Re} = \frac{\mathrm{row}_0(\mathbf{U}_n^{\mathsf{T}}) \left(\frac{S_n}{2}\, \Re(\boldsymbol{g}_n)\, \Re(\boldsymbol{g}_n^{\mathsf{T}})\right) \mathrm{col}_0(\mathbf{U}_n)}{\mathrm{row}_0(\mathbf{U}_n^{\mathsf{T}}) \left(\mathbf{\Psi}_{\boldsymbol{z}_n^{\mathbb{R}}\boldsymbol{z}_n^{\mathbb{R}}} + \frac{S_n}{2}\, \Im(\boldsymbol{g}_n)\, \Im(\boldsymbol{g}_n^{\mathsf{T}})\right) \mathrm{col}_0(\mathbf{U}_n)}, \tag{26}$$

where $\mathbf{\Psi}_{\boldsymbol{z}_n^{\mathbb{R}}\boldsymbol{z}_n^{\mathbb{R}}} \triangleq \mathbb{E}\{\boldsymbol{z}_n^{\mathbb{R}}\, (\boldsymbol{z}_n^{\mathbb{R}})^{\mathsf{T}}\}$ is the composite real interference-plus-noise covariance matrix of subcarrier n, which can be calculated as

$$\mathbf{\Psi}_{\boldsymbol{z}_n^{\mathbb{R}}\boldsymbol{z}_n^{\mathbb{R}}} = \sum_{m=1}^{N/2-1} \frac{S_m}{2}\, \Re(\mathbb{G}_{nm}\, \mathbb{G}_{nm}^{\mathcal{H}}) + \sigma_{\mathrm{t}}^2\, \mathbf{R}_n^{\mathbb{R}}\, (\mathbf{R}_n^{\mathbb{R}})^{\mathsf{T}}. \tag{27}$$

A similar expression can be obtained for $\gamma_n^{\Im}$, the SINR of $\Im(\hat{X}_n[k])$.

B. Bit Loading

It is assumed that PAM symbols are transmitted on both the real and imaginary part of $X_n[k]$. The achievable bit loading for the real — or imaginary — part of $X_n[k]$ can then be expressed as

$$b = \frac{1}{2}\, \log_2\left(1 + \Gamma^{-1}\, \gamma\right). \tag{28}$$

If received at the same SINR, then uncoded M^2-QAM signals that are subject to proper interference-plus-noise yield the same BER performance as uncoded M-PAM signals. As such, it is intuitively clear that — if uncoded modulation and square QAM constellations are considered — the same value of Γ will result in the same BER, regardless of whether the case is considered in which the bit loading is determined using (17) and all signals are proper, or the case is considered in which the bit loading is determined using (28) and the received signals are allowed to be improper. In order to achieve the same BER performance, the value of Γ in (28) should therefore be the same as in (17).[11] With (26) and (28), the bit loading of subcarrier n is given as

$$b_n = \frac{1}{2}\, \log_2(1 + \Gamma^{-1}\, \gamma_n^{\Re}) + \frac{1}{2}\, \log_2(1 + \Gamma^{-1}\, \gamma_n^{\Im}). \tag{29}$$

[10] The ensuing derivations can be extended to improper QAM symbols by applying a 2×2 precoding matrix $\mathbf{P}_n$ to each $X_n^{\mathbb{R}}[k]$ prior to transmission, with $\mathbf{P}_n$ the Cholesky factor of the transmitted QAM symbol's desired covariance matrix. The extension to improper QAM symbols is however considered to be out-of-scope.

[11] While intuitively clear for uncoded PAM symbols, this statement is not true for coded modulation. As the interference-plus-noise experienced by the two PAM symbols in $\hat{X}_n[k]$ may be correlated, error events in these two PAM symbols will not be statistically independent. This statistical dependence should be taken into account when designing the coded modulation. If the interference-plus-noise experienced by the two PAM symbols in $\hat{X}_n[k]$ is uncorrelated however, then the TCM encoding employed in G.fast is straightforwardly extended to the proposed scheme of composing $\hat{X}_n[k]$ of two independent PAM symbols. Further details are given in Section V-B.

$$\hat{X}_n^{\mathbb{R}}[k] = \mathbf{U}_n^{\mathsf{T}} \Bigg(\underbrace{\begin{bmatrix} \Re(\boldsymbol{g}_n) & \Im(\boldsymbol{g}_n) \end{bmatrix}}_{\triangleq \mathbf{G}_n} X_n^{\mathbb{R}}[k] + \underbrace{\sum_{m=1}^{N/2-1} \begin{bmatrix} \Re(\mathbb{G}_{nm}) & \Im(\mathbb{G}_{nm}) \end{bmatrix} \boldsymbol{X}_m^{\mathbb{R}}[k+l] + \mathbf{R}_n^{\mathbb{R}} \bar{\boldsymbol{z}}[k]}_{\triangleq \boldsymbol{z}_n^{\mathbb{R}}[k]} \Bigg). \tag{25}$$

V. Optimizing Performance: Widely Linear PTEQ & Symbol Rotation

In Table II, the SINR and bit loading values corresponding to the received signals in Fig. 2 are given. From these examples, it is already seen that taking advantage of the signal impropriety yields a higher overall bit loading. Moreover, the examples of Fig. 2 and Table II indicate that the relative rotation between the axes of the received PAM symbols and the interference-plus-noise distribution in $\hat{X}_n[k]$ strongly impacts the achievable bit loading. In this section, it will therefore be shown that optimally rotating the symbols $X_n[k]$ in the complex plane prior to transmission can yield significant performance gains. First however, the computation of the PTEQ filter coefficients $\mathbf{U}_n$ that maximize the achievable bit loading as in (29) is considered.

A. Widely Linear Per-Tone Equalization

From (23), it is seen that — if $\hat{X}_n[k]$ is to be obtained by applying a linear filter to $\boldsymbol{y}_n[k]$ as in (11) — $\mathbf{U}_n$ should exhibit a block-skew-circulant structure [22], i.e. $\mathbf{U}_n$ should comply with

$$\mathbf{U}_n = \begin{bmatrix} \Re(\boldsymbol{v}_n) & -\Im(\boldsymbol{v}_n) \\ \Im(\boldsymbol{v}_n) & \Re(\boldsymbol{v}_n) \end{bmatrix} \tag{30}$$

for some vector $\boldsymbol{v}_n \in \mathbb{C}^{T+1}$. This constraint is however difficult to satisfy, as it couples the optimization of $\mathrm{col}_0(\mathbf{U}_n)$ and $\mathrm{col}_1(\mathbf{U}_n)$. Consequently, no analytical expression for the optimal $\mathbf{U}_n$ is readily available, thereby losing one of the main advantages of PTEQ over TEQ.

However, this problem can be circumvented by discarding the constraint specified in (30), and allowing $\mathbf{U}_n$ to be any matrix in $\mathbb{R}^{2(T+1)\times 2}$. The real and imaginary part of $\hat{X}_n[k]$ will then each be obtained by applying a separate linear filter to the composite real received signal $\boldsymbol{y}_n^{\mathbb{R}}[k]$, which is equivalent to applying a *widely linear* (WL) filter [24] to $\boldsymbol{y}_n[k]$. An analytical expression for the optimal $\mathbf{U}_n$, i.e. for the optimal WL PTEQ coefficients, is then readily derived.

Before presenting this analytical expression, a remark is due regarding the structure of $\mathbf{R}_n^{\mathbb{R}}$. As the last T rows of $\mathbf{R}_n$ are real-valued, the last T rows of $\mathbf{R}_n^{\mathbb{R}}$ will contain only zeros. Consequently, the interference-plus-noise covariance matrix $\boldsymbol{\Psi}_{\boldsymbol{z}_n^{\mathbb{R}}\boldsymbol{z}_n^{\mathbb{R}}} = \mathbb{E}\{\boldsymbol{z}_n^{\mathbb{R}}[k]\,(\boldsymbol{z}_n^{\mathbb{R}}[k])^{\mathcal{H}}\}$ will be singular, causing problems when deriving the optimal WL PTEQ filter coefficients.[12] However, the all-zero rows of $\mathbf{R}_n^{\mathbb{R}}$ may readily be removed from the composite real receiver expression. To this end, $\underline{\mathbf{A}}$ — which has a bar accent below the symbol — will denote the matrix that is obtained by removing the last $m-1$ rows from a $2m \times n$ matrix $\mathbf{A}$. Likewise, $\mathbf{A}^{\mathbb{R}}$ will denote the matrix that is obtained by removing the last $m-1$ rows from the composite real representation of an $m \times n$ matrix $\mathbf{A}$. Using this new notation, the last T rows of $\mathbf{U}_n$ and $\mathbf{R}_n$ in (23) can effectively be removed by writing the equation as

$$\hat{X}_n^{\mathbb{R}}[k] = \underline{\mathbf{U}}_n^{\mathsf{T}}\, \mathbf{R}_n^{\mathbb{R}}\, \bar{\boldsymbol{y}}[k]. \tag{31}$$

Using the same notation to remove the last T rows from $\boldsymbol{g}_n$, $\boldsymbol{z}_n^{\mathbb{R}}$, and $\mathbb{G}_{nm}$, the expression for the SINR of $\Re(\hat{X}_n[k])$ can be rewritten as (equivalent to (26))

$$\gamma_n^{\Re} = \frac{\mathrm{row}_0(\underline{\mathbf{U}}_n^{\mathsf{T}}) \left(\frac{S_n}{2}\, \Re(\underline{\boldsymbol{g}}_n)\, \Re(\underline{\boldsymbol{g}}_n^{\mathsf{T}})\right) \mathrm{col}_0(\underline{\mathbf{U}}_n)}{\mathrm{row}_0(\underline{\mathbf{U}}_n^{\mathsf{T}}) \left(\boldsymbol{\Psi}_{\boldsymbol{z}_n^{\mathbb{R}}\boldsymbol{z}_n^{\mathbb{R}}} + \frac{S_n}{2}\, \Im(\underline{\boldsymbol{g}}_n)\, \Im(\underline{\boldsymbol{g}}_n^{\mathsf{T}})\right) \mathrm{col}_0(\underline{\mathbf{U}}_n)}, \tag{32}$$

where

$$\boldsymbol{\Psi}_{\boldsymbol{z}_n^{\mathbb{R}}\boldsymbol{z}_n^{\mathbb{R}}} = \sum_{m=1}^{N/2-1} \frac{S_m}{2}\, \Re(\mathbb{G}_{nm}\, \mathbb{G}_{nm}^{\mathcal{H}}) + \sigma_{\mathrm{t}}^2\, \mathbf{R}_n^{\mathbb{R}}\, (\mathbf{R}_n^{\mathbb{R}})^{\mathsf{T}}. \tag{33}$$

With this equivalent expression for the SINR, the optimal WL PTEQ filter coefficients are — up to scaling — given as

$$\mathrm{col}_0(\underline{\mathbf{U}}_n) = \left(\boldsymbol{\Psi}_{\boldsymbol{z}_n^{\mathbb{R}}\boldsymbol{z}_n^{\mathbb{R}}} + \frac{S_n}{2}\, \Im(\underline{\boldsymbol{g}}_n)\, \Im(\underline{\boldsymbol{g}}_n^{\mathsf{T}})\right)^{-1} \Re(\underline{\boldsymbol{g}}_n). \tag{34}$$

A similar expression can be obtained for $\gamma_n^{\Re}$, the SINR of $\Im(\hat{X}_n[k])$, and for the optimal WL PTEQ filter coefficients $\mathrm{col}_1(\underline{\mathbf{U}}_n)$.

It may appear that WL PTEQ amounts to a complexity increase w.r.t PTEQ, as the WL PTEQ — if implemented based on the composite real filter representation of (31) — has two $T+2$-tap filters for each subcarrier, whereas the PTEQ has only a single $T+1$-tap filter per subcarrier. However, all filter coefficients and signals in the WL PTEQ are real, such that WL PTEQ requires $2T+4$ real multiplications and $2T+2$ real additions per subcarrier. The computational complexity of WL PTQ is summarized in Table III. After comparing Table III to Table I, it can be concluded that the run-time complexity of WL PTEQ is exactly the same as that of PTEQ implemented with $n_\times = 4$ and $n_+ = 2$, and is comparable to that of PTEQ implemented with $n_\times = 3$ and $n_+ = 5$.

B. Transmitter-side Symbol Rotation

The examples of Fig. 2 and Table II indicate that the relative rotation between the axes of the received PAM symbols

[12] Compare to how, in (18), $\boldsymbol{\Psi}_{\boldsymbol{z}_n\boldsymbol{z}_n}$ being singular would cause problems.

TABLE III: Number of floating point operations required by channel shortening and equalization a DMT receiver (continued).

	Nb. real mult.	Nb. real add.
WL PTEQ	$\left(\frac{N}{2}-1\right)(2T+4)$	$\left(\frac{N}{2}-1\right)(2T+2)+T$

and the interference-plus-noise distribution in $\hat{X}_n[k]$ strongly impacts the achievable bit loading. Indeed, the results in Table II show that a higher bit loading can be achieved if the eigenvectors of the interference-plus-noise covariance matrix $\underline{\mathbf{U}}_n^{\mathrm{T}} \boldsymbol{\Psi}_{\boldsymbol{z}_n^{\mathbb{R}} \boldsymbol{z}_n^{\mathbb{R}}} \underline{\mathbf{U}}_n$ are well aligned with the axis of the received PAM symbols in $\hat{X}_n[k]$. To further improve the achievable bit loading, it is therefore proposed to rotate the symbols $X_n[k]$ prior to transmission — i.e. to multiply the complex symbols $X_n[k]$ on each subcarrier n with a complex scalar ρ_n satisfying $|\rho_n| = 1$.[13] Here, the appropriate symbol rotation will be derived.

In the composite real channel model of (25), the rotation of $X_n[k]$ can be included by multiplying $X_n^{\mathbb{R}}[k]$ with an orthogonal rotation matrix $\mathbf{P}_n$, yielding

$$\hat{X}_n^{\mathbb{R}}[k] = \underline{\mathbf{U}}_n^{\mathrm{T}} \left(\mathbf{G}_n \, \mathbf{P}_n \, X_n^{\mathbb{R}}[k] + \boldsymbol{z}_n^{\mathbb{R}}[k] \right). \tag{35}$$

It is noted that, as it was assumed that $X_n[k]$ is proper, this rotation operation has no influence on the transmitted symbol's covariance matrix.

$$\begin{aligned} \mathbb{E}\left\{ \mathbf{P}_n \, X_n^{\mathbb{R}}[k] \, (X_n^{\mathbb{R}}[k])^{\mathrm{T}} \, \mathbf{P}_n^{\mathrm{T}} \right\} &= \mathbf{P}_n \, \mathbb{E}\left\{ X_n^{\mathbb{R}}[k] \, (X_n^{\mathbb{R}}[k])^{\mathrm{T}} \right\} \mathbf{P}_n^{\mathrm{T}} \\ &= \mathbf{P}_n \, \frac{S_n}{2} \, \mathbf{I}_2 \, \mathbf{P}_n^{\mathrm{T}} = \frac{S_n}{2} \, \mathbf{I}_2 = \mathbb{E}\left\{ X_n^{\mathbb{R}}[k] \, (X_n^{\mathbb{R}}[k])^{\mathrm{T}} \right\} \end{aligned} \tag{36}$$

Consequently, rotating the transmitted symbols $X_n[k]$ does not impact the interference-plus-noise covariance matrix $\boldsymbol{\Psi}_{\boldsymbol{z}_m^{\mathbb{R}} \boldsymbol{z}_m^{\mathbb{R}}}$ for any m. This independence will be exploited when calculating the optimal rotation matrix $\mathbf{P}_n$ (see below).

Next, a WL FEQ is applied to align the axes of the two received PAM symbols with the real and imaginary axis. If this WL FEQ is absorbed into the WL PTEQ, then the obtained unbiased WL PTEQ filter is given by

$$\underline{\mathbf{U}}_{n,\mathrm{ub}}^{\mathrm{T}} = (\underline{\mathbf{U}}_n^{\mathrm{T}} \, \underline{\mathbf{G}}_n \, \mathbf{P}_n)^{-1} \, \underline{\mathbf{U}}_n^{\mathrm{T}}. \tag{37}$$

In (37), the subscript 'ub' indicates that the resulting $\hat{X}_n^{\mathbb{R}}$ is an unbiased estimator of $X_n^{\mathbb{R}}[k]$, as is evidenced by substituting (37) into (35).

$$\hat{X}_{n,\mathrm{ub}}^{\mathbb{R}}[k] = X_n^{\mathbb{R}}[k] + \mathbf{P}_n^{\mathrm{T}} \, (\underline{\mathbf{U}}_n^{\mathrm{T}} \, \underline{\mathbf{G}}_n)^{-1} \, \underline{\mathbf{U}}_n^{\mathrm{T}} \, \boldsymbol{z}_n^{\mathbb{R}}[k] \tag{38}$$

After applying the unbiased WL PTEQ, the axes of the received PAM symbols are aligned with the real and imaginary axis in $\hat{X}_{n,\mathrm{ub}}[k]$. The rotation matrix $\mathbf{P}_n$ will be chosen such that the eigenvectors of the covariance matrix of the interference-plus-noise in $\hat{X}_{n,\mathrm{ub}}^{\mathbb{R}}[k]$ are aligned with the real and imaginary axis as well. With the eigenvalue decomposition of this interference-plus-noise covariance matrix given as

$$(\underline{\mathbf{U}}_n^{\mathrm{T}} \, \underline{\mathbf{G}}_n)^{-1} \, \underline{\mathbf{U}}_n^{\mathrm{T}} \, \boldsymbol{\Psi}_{\boldsymbol{z}_n^{\mathbb{R}} \boldsymbol{z}_n^{\mathbb{R}}} \, \underline{\mathbf{U}}_n \, (\underline{\mathbf{U}}_n^{\mathrm{T}} \, \underline{\mathbf{G}}_n)^{-\mathrm{T}} = \mathbf{W}_n \, \boldsymbol{\Lambda}_n \, \mathbf{W}_n^{\mathrm{T}}, \tag{39}$$

$\mathbf{P}_n$ will thus be chosen as $\mathbf{P}_n = \mathbf{W}_n$.

With this choice for the rotation matrix $\mathbf{P}_n$, the expression for the achievable bit loading as in (29) can be further simplified. If $\mathbf{P}_n = \mathbf{W}_n$, then the SINRs of the real and imaginary part of $\hat{X}_{n,\mathrm{ub}}[k]$ are given by $\frac{S_n}{2}$ multiplied with the inverse of $\boldsymbol{\Lambda}_n$'s diagonal elements. Assuming w.l.o.g. that $\boldsymbol{\Lambda}_n$ is invertible, the achievable bit loading as in (29) becomes

$$b_n = \frac{1}{2} \, \log_2 \det \left(\mathbf{I} + \Gamma^{-1} \frac{S_n}{2} \, \boldsymbol{\Lambda}_n^{-1} \right), \tag{40}$$

which can equivalently be expressed as

$$b_n = \frac{1}{2} \, \log_2 \frac{\det \left(\underline{\mathbf{U}}_n^{\mathrm{T}} \, \left(\boldsymbol{\Psi}_{\boldsymbol{z}_n^{\mathbb{R}} \boldsymbol{z}_n^{\mathbb{R}}} + \Gamma^{-1} \frac{S_n}{2} \, \underline{\mathbf{G}}_n \, \underline{\mathbf{G}}_n^{\mathrm{T}} \right) \, \underline{\mathbf{U}}_n \right)}{\det \left(\underline{\mathbf{U}}_n^{\mathrm{T}} \, \boldsymbol{\Psi}_{\boldsymbol{z}_n^{\mathbb{R}} \boldsymbol{z}_n^{\mathbb{R}}} \, \underline{\mathbf{U}}_n \right)}. \tag{41}$$

With this expression for the achievable bit loading, the following expression for the optimal WL PTEQ filter coefficients $\underline{\mathbf{U}}_n$ can be obtained.[14]

$$\underline{\mathbf{U}}_n = \boldsymbol{\Psi}_{\boldsymbol{z}_n^{\mathbb{R}} \boldsymbol{z}_n^{\mathbb{R}}}^{-1} \, \underline{\mathbf{G}}_n \tag{42}$$

After plugging (42) back into (41), the following expression for the achievable bit loading is obtained.

$$b_n = \frac{1}{2} \, \log_2 \det \left(\mathbf{I} + \frac{S_n}{2} \Gamma^{-1} \, \boldsymbol{\Psi}_{\boldsymbol{z}_n^{\mathbb{R}} \boldsymbol{z}_n^{\mathbb{R}}}^{-1} \, \underline{\mathbf{G}}_n \, \underline{\mathbf{G}}_n^{\mathrm{T}} \right). \tag{43}$$

In addition to simplifying the expression for the achievable bit loading, (42) can also be employed to simplify the calculation of the rotation matrix $\mathbf{P}_n$. Substituting (42) into (39) yields

$$\left(\underline{\mathbf{G}}_n^{\mathrm{T}} \, \boldsymbol{\Psi}_{\boldsymbol{z}_n^{\mathbb{R}} \boldsymbol{z}_n^{\mathbb{R}}}^{-1} \, \underline{\mathbf{G}}_n \right)^{-1} = \mathbf{W}_n \, \boldsymbol{\Lambda}_n \, \mathbf{W}_n^{\mathrm{T}}, \tag{44}$$

such that $\mathbf{W}_n = \mathbf{P}_n$ can equivalently be determined from the following eigenvalue decomposition.

$$\underline{\mathbf{G}}_n^{\mathrm{T}} \, \boldsymbol{\Psi}_{\boldsymbol{z}_n^{\mathbb{R}} \boldsymbol{z}_n^{\mathbb{R}}}^{-1} \, \underline{\mathbf{G}}_n = \mathbf{W}_n \, \boldsymbol{\Lambda}_n^{-1} \, \mathbf{W}_n^{\mathrm{T}} \tag{45}$$

This expression makes clear that the receiver needs no additional information to determine the rotation matrix $\mathbf{P}_n$: the interference-plus-noise covariance matrix $\boldsymbol{\Psi}_{\boldsymbol{z}_n^{\mathbb{R}} \boldsymbol{z}_n^{\mathbb{R}}}$ and direct channels $\underline{\mathbf{G}}_n$ need to be determined in order to calculate the optimal WL PTEQ filter coefficients using (34), regardless of whether or not a symbol rotation is applied. The transmitter-side symbol rotation does require a feedback channel between receiver and transmitter, through which the desired rotation matrix $\mathbf{P}_n$ is to be communicated. The existence of such a feedback channel is however already implicitly assumed, as the transmitter should be made aware of the achieved SINR to determine the bit loading. The desired rotation matrix $\mathbf{P}_n$ can be communicated using a single real scalar θ_n per subcarrier, as any 2×2 rotation matrix can w.l.o.g. be written as

$$\mathbf{P}_n = \begin{bmatrix} \cos\theta_n & -\sin\theta_n \\ \sin\theta_n & \cos\theta_n \end{bmatrix}. \tag{46}$$

With the rotation matrix chosen as $\mathbf{P}_n = \mathbf{W}_n$, the interference-plus-noise in the real and imaginary part of $\hat{X}_n$ is no longer correlated. As a result, the proposed scheme of constructing the 2-D symbols X_n from two 1-D PAM symbols can be combined with the adapted version of Wei's 4-D 16-state TCM encoding that is used in G.fast [2]. This TCM encoding requires the constellations of its constituent

[13] It is noted that this type of transmitter-side symbol rotation is compatible with the G.fast standard [2], as the standard specifies that a complex gain scaling should be applied to the transmitted symbols upon updating the vectoring matrix in order to absorb "FEQ shock".

[14] The optimality of this expression is established in Section V-C.

2-D symbols X_n to be partitioned into four subsets that — after scaling the real and imaginary part of X_n such that they experience the same interference-plus-noise power — maximize the minimum squared distance among them [25]. It is noted that this combining of two 1-D symbols into a single 2-D symbol which is then partitioned into four subsets is similar to how, in G.fast, the 2-PAM symbols of single-bit tones are combined into 2-D symbols for TCM encoding [2].

This section is concluded with a comment regarding bit loading discretization, a subject that has not been touched upon so far. In G.fast, cross-constellations enable loading a fractional number of bits per dimension. The proposed scheme of constructing 2-D symbols X_n from two 1-D PAM symbols however requires an integer number of bits to be loaded per dimension. The discretization of the bit loading in (29) thus involves rounding down the bit loading of the real and imaginary part of X_n to the nearest integer separately, whereas the discretization in (17) involves only a single rounding down operation. As a result, the bit loading discretization loss is expected to be higher for the proposed scheme. It is however possible to scale any rectangular constellation — i.e. any 2-D constellation consisting of two orthogonal 1-D PAM constellations — such that an additional 1/2 bit can be loaded per dimension using the technique from [26, Section II.F]. The resulting family of constellations will have essentially the same properties as the original family of rectangular constellations [26, Section II.F], such that the same value for the SNR gap Γ will result in the same BER being achieved. It is noted that the technique to add an additional 1/2 bit per dimension to a rectangular constellation as in [26, Section II.F] will result in a family of constellations with the same properties only if the real and imaginary part of interference-plus-noise in $\hat{X}_n$ are uncorrelated — a condition that is satisfied when $\mathbf{P}_n = \mathbf{W}_n$.

C. Symbol Rotation Optimality

This section shows that the proposed choice for $\mathbf{P}_n$ and $\underline{\mathbf{U}}_n$ maximizes the bit loading on subcarrier n. It is noted that the result is trivial when $\Gamma = 0\,\mathrm{dB}$. Indeed, for $\Gamma = 0\,\mathrm{dB}$ equation (43) corresponds to the capacity of a real MIMO channel with proper $X_n^{\mathbb{R}}[k]$, an orthogonal transmit beamforming matrix $\mathbf{P}_n$, and which treats the ISI/ICI as noise [27]. As such, a higher bit loading cannot be achieved, and the proposed $\mathbf{P}_n$ and $\underline{\mathbf{U}}_n$ from Section V-B are optimal. In order to show that the proposed $\mathbf{P}_n$ and $\underline{\mathbf{U}}_n$ from Section V-B are optimal in general, it is proven here that (43) upper bounds the achievable bit loading when $\Gamma \geq 0\,\mathrm{dB}$ as well. The desired result is stated in Proposition 1.

Proposition 1: Consider the real MIMO channel which is described by

$$\hat{\boldsymbol{X}} = \mathbf{U}^{\mathsf{T}} \left(\mathbf{G}\,\mathbf{P}\,\boldsymbol{X} + \boldsymbol{Z} \right), \tag{47}$$

where $\boldsymbol{X}$ is a real 2×1-vector with $\mathbb{E}\{\boldsymbol{X}\boldsymbol{X}^{\mathsf{T}}\} = S\,\mathbf{I}$, $\mathbf{P} = [\boldsymbol{p}_0, \boldsymbol{p}_1]$ is an orthogonal transmit beamforming matrix, $\mathbf{U} = [\boldsymbol{u}_0, \boldsymbol{u}_1]$ is a receive beamforming matrix, and where $\boldsymbol{Z}$ is a vector of zero-mean additive Gaussian noise which has a non-singular covariance matrix $\boldsymbol{\Psi} = \mathbb{E}\{\boldsymbol{Z}\boldsymbol{Z}^{\mathsf{T}}\}$. If $\Gamma \geq 0\,\mathrm{dB}$, then the achievable bit loading in this channel — which is considered to be given by

$$\frac{1}{2}\sum_{i=0}^{1} \log_2 \left(1 + \Gamma^{-1} \frac{S\,\boldsymbol{u}_i^{\mathsf{T}}\,\mathbf{G}\boldsymbol{p}_i\,\boldsymbol{p}_i^{\mathsf{T}}\mathbf{G}^{\mathsf{T}}\,\boldsymbol{u}_i}{\boldsymbol{u}_i^{\mathsf{T}}\,\left(\boldsymbol{\Psi} + S\,\mathbf{G}\,\boldsymbol{p}_{-i}\,\boldsymbol{p}_{-i}^{\mathsf{T}}\mathbf{G}^{\mathsf{T}}\right)\,\boldsymbol{u}_i} \right) \tag{48}$$

— cannot exceed

$$\frac{1}{2}\log_2 \det\left(\mathbf{I} + S\Gamma^{-1}\,\boldsymbol{\Psi}^{-1}\,\mathbf{G}\,\mathbf{G}^{\mathsf{T}}\right) \tag{49}$$

for any $\mathbf{U}$ and orthogonal $\mathbf{P}$.

Proof: From (48), it is readily seen that the optimal $\boldsymbol{u}_i$ maximizes the following generalized Rayleigh quotient

$$r(\boldsymbol{u}_i) = \frac{S\,\boldsymbol{u}_i^{\mathsf{T}}\,\mathbf{G}\boldsymbol{p}_i\,\boldsymbol{p}_i^{\mathsf{T}}\mathbf{G}^{\mathsf{T}}\,\boldsymbol{u}_i}{\boldsymbol{u}_i^{\mathsf{T}}\,\left(\boldsymbol{\Psi} + S\,\mathbf{G}\,\boldsymbol{p}_{-i}\,\boldsymbol{p}_{-i}^{\mathsf{T}}\mathbf{G}^{\mathsf{T}}\right)\,\boldsymbol{u}_i}$$

and is therefore given as [28, Chapter 7.9]

$$\boldsymbol{u}_i^{\mathrm{opt}} = \left(\boldsymbol{\Psi} + S\,\mathbf{G}\,\boldsymbol{p}_{-i}\,\boldsymbol{p}_{-i}^{\mathsf{T}}\mathbf{G}^{\mathsf{T}}\right)^{-1}\mathbf{G}\,\boldsymbol{p}_i.$$

Substituting $\boldsymbol{u}_i^{\mathrm{opt}}$ back into (48) yields

$$\frac{1}{2}\sum_{i=0}^{1} \log_2 \left(1 + \Gamma^{-1}\boldsymbol{p}_i^{\mathsf{T}}\,\mathbf{G}^{\mathsf{T}}\left(\boldsymbol{\Psi} + S\,\mathbf{G}\,\boldsymbol{p}_{-i}\,\boldsymbol{p}_{-i}^{\mathsf{T}}\mathbf{G}^{\mathsf{T}}\right)^{-1}\mathbf{G}\,\boldsymbol{p}_i\,S\right)$$

which — as it is assumed that $\Gamma \geq 0\,\mathrm{dB}$ — is upper bounded by equation (50) at the top of the next page. Some algebraic manipulations reveal that (50) is equivalent to

$$\frac{1}{2}\log_2 \frac{\det\left(\mathbf{I} + S\Gamma^{-1}\,\boldsymbol{\Psi}^{-1}\,\mathbf{G}\,\mathbf{G}^{\mathsf{T}}\right)^2}{\prod_{i=0}^{1}\boldsymbol{p}_i^{\mathsf{T}}\,\left(\mathbf{I} + S\Gamma^{-1}\,\mathbf{G}^{\mathsf{T}}\boldsymbol{\Psi}^{-1}\,\mathbf{G}\right)\,\boldsymbol{p}_i}. \tag{51}$$

Applying Hadamard's inequality to the denominator in (51) reveals that the achieved bit loading is upper bounded by

$$\frac{1}{2}\log_2 \frac{\det\left(\mathbf{I} + S\Gamma^{-1}\,\boldsymbol{\Psi}^{-1}\,\mathbf{G}\,\mathbf{G}^{\mathsf{T}}\right)^2}{\det\left(\mathbf{I} + S\Gamma^{-1}\,\boldsymbol{\Psi}^{-1}\,\mathbf{G}\,\mathbf{G}^{\mathsf{T}}\right)}, \tag{52}$$

which is equivalent to (49). ■

VI. Simulation Results

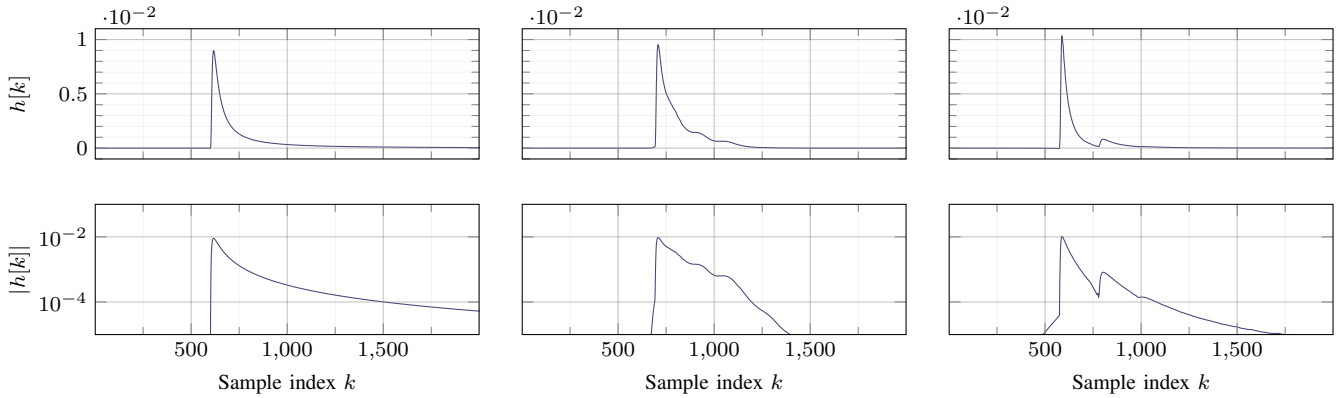
The performance gains that can be achieved by explicitly taking interference-plus-noise impropriety into account are assessed in the context of single-user long-reach G.fast systems. Three long-reach DSL scenarios are considered, with CIRs — which are given in Fig. 3 — corresponding to twisted pair cables of length $600\,\mathrm{m}$. One CIR is synthesized using the KM1 model [29], while the other two are derived from measured frequency responses. These measured CIRs are artificially extended to $600\,\mathrm{m}$ using the following equation:

$$H_1(f) = \left(H_0(f)\right)^{\frac{l_1}{l_0}}, \tag{53}$$

where l_0 is the original length and $l_1 = 600\,\mathrm{m}$.

The simulation parameters are chosen conform the 106b-profile of G.fast [2], and can be found in Table IV. Note that Table IV defines the frequency domain noise power σ_{f}^2 instead of the time domain noise power σ_{t}^2. With the DFT operation in

$$\frac{1}{2}\sum_{i=0}^{1}\log_2\left(1+\Gamma^{-1}\boldsymbol{p}_i^{\mathrm{T}}\,\mathbf{G}^{\mathrm{T}}\left(\boldsymbol{\Psi}+S\Gamma^{-1}\,\mathbf{G}\,\boldsymbol{p}_{-i}\,\boldsymbol{p}_{-i}^{\mathrm{T}}\mathbf{G}^{\mathrm{T}}\right)^{-1}\mathbf{G}\,\boldsymbol{p}_i\,S\right) \tag{50}$$



(a) KM1 (b) ISP1 (c) ISP2

Fig. 3: The three CIRs that are considered in the simulations. The CIR labeled "KM1" was synthesized using the KM1 model with k-parameters corresponding to a homogeneous and perfectly terminated Cat5 cable [29, Table II]. The CIRs labeled "ISP1" and "ISP2" were obtained by measurement of the transfer function of twisted-pair cables belonging to two different undisclosed Internet service providers. These measured CIRs were artificially extended to $600\,\mathrm{m}$.

TABLE IV: Simulation parameters

Parameter	Value	Parameter	Value
P_{tot}	$8\,\mathrm{dBm}$	f_{s}	$212\,\mathrm{MHz}$
N	4096	Γ	$9.54\,\mathrm{dB}$
σ_{f}^2	$-145\,\mathrm{dBm/Hz}$	S_n^{mask}	See [1]
ν_{s}	64		

the DMT demodulation defined as in (3), the relation between the two is as follows.

$$\sigma_{\mathrm{t}}^2=\frac{1}{N}\sigma_{\mathrm{f}}^2 \tag{54}$$

A raised cosine window is used at the transmitter. More specifically, the diagonal elements of the windowing matrix $\mathbf{C}$ are defined as follows.

$$[\mathbf{C}]_{n,n}\triangleq\begin{cases}\sin^2\left(\frac{\pi(n+1)}{2\nu_{\mathrm{s}}}\right) & \text{for } 0\le n\le\nu_{\mathrm{s}}-1\\ 1 & \text{for } \nu_{\mathrm{s}}\le n\le s-1\\ \sin^2\left(\frac{\pi(n+1-s+\nu_{\mathrm{s}})}{2\nu_{\mathrm{s}}}\right) & \text{for } s\le n\le s+\nu_{\mathrm{s}}-1\end{cases} \tag{55}$$

The symbol powers S_n are chosen such that the G.fast aggregate transmit power (ATP) constraint and spectral mask constraint are satisfied, which are respectively given as

$$\frac{f_{\mathrm{s}}}{N}\sum_{n=1}^{N/2-1}S_n\le P_{\mathrm{tot}}, \tag{56a}$$

$$S_n\le S_n^{\mathrm{mask}}. \tag{56b}$$

As the LR extension of G.fast is expected to not use the first 18 subcarriers, $S_n=0\ \forall n\in\{0,\ldots,17\}$. The employed symbol powers S_n are determined using a known heuristic procedure. As a first step of this heuristic procedure, the available transmit power is distributed evenly over all subcarriers by setting $S_n=\frac{P_{\mathrm{tot}}N}{f_{\mathrm{s}}(N/2-18)}$, $\forall n\in\{18,\ldots,N/2-1\}$. Then, the resulting SINR is determined for each subcarriers using (15).[15] Finally, the symbol power is set to $S_n=0$ on all subcarriers after the highest subcarrier n for which $\gamma_n\ge\Gamma$, and the reclaimed transmit power is distributed evenly over the remaining subcarriers while making sure not to violate the spectral mask constraints.

In the ensuing simulations, SINR and bit loading values will be presented that are obtained with (15) and (17), as well as with (26) and (29). In all bit loading calculations — irrespective of which equation is considered — the value of the SNR gap Γ from Table IV is employed. The considered value of Γ includes a coding gain for the employed TCM encoding of G.fast, as well as a noise margin. It is emphasized that, with the considered value of Γ, achieving the target error performance requires the following.

- If (17) is employed to determine the bit loading, then the received signal $\hat{X}_n[k]$ must be free of interference from $\hat{X}_{n^*}[k]$ and the interference-plus-noise in $\hat{X}_n[k]$ must be proper.
- If (29) is employed to determine the bit loading, then the interference-plus-noise in the estimators of the real and imaginary parts of $\hat{X}_n[k]$ must be uncorrelated.

Plots will show the continuous bit loading, i.e. no bit loading discretization is executed.

A. Examples Fig. 2 and Table II

The scatter plots of Fig. 2 are obtained by Monte-Carlo simulation of a G.fast system. The transmitted symbols $X_n[k]$ are randomly drawn from a scaled 16-QAM constellation, where

[15] In practice, the SINR can be estimated from the training sequence transmitted during system initialization.

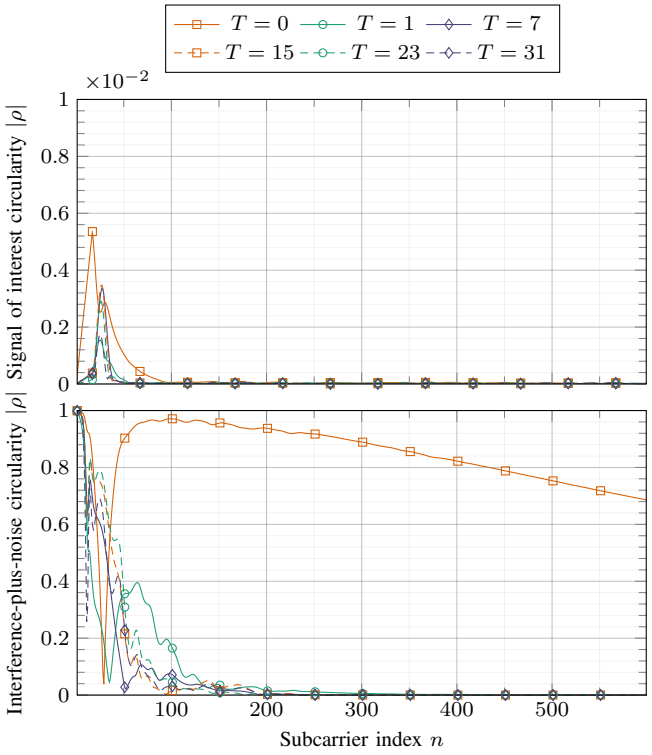


Fig. 4: Circularity coefficients of the received direct signal (top) and the received interference-plus-noise (bottom) in a DMT receiver using PTEQ. The same system as in Section VI-A is considered. The y-axes of the two plots have a different rage: the y-axis of the top plot ranges from 0 to 0.01, while the y-axis of the bottom plot ranges from 0 to 1.

the scaling factor is chosen such that the symbol power on each subcarrier n corresponds to the value of S_n obtained through the previously described heuristic procedure. The number of transmitted DMT symbols is 1000. The considered CP length is $\nu_{\mathrm{p}} = 128$, which is the lowest value allowed by the G.fast standard [2]. The delay δ is chosen by maximizing the energy in $\left[h[\delta], h[\delta+1], \ldots, h[\delta+\nu_{\mathrm{p}}]\right]$ as in [16]. Simulations employed CIR ISP1.

Fig. 2 consists of three subplots, each containing a left scatter plot depicting the received signal $\boldsymbol{y}_n[k]$ for a single subcarrier ($n = 100$), and a right scatter plot depicting the interference-plus-noise component of the same signal. The depicted signals are rotated prior to plotting, such that the received 16-QAM symbols are aligned with the real and imaginary axes. To obtain the scatter plots in Fig. 2b, the transmitted symbols $X_n[k]$ are rotated prior to transmission, where the rotation matrices $\mathbf{P}_n$ are chosen to be optimal for a DMT system with WL PTEQ and $T = 0$. Finally, Fig. 2c was obtained from Fig. 2a by subtracting the interference-plus-noise term from the received signal, and adding artificially generated circularly symmetric Gaussian noise with the same power as the previously subtracted interference-plus-noise.

B. PTEQ Output Impropriety

To further illustrate the need for paying attention to impropriety in DMT systems with a too-short CP, Fig. 4 shows the circularity coefficient $|\rho|$ for the received signal of interest and for the received interference-plus-noise in a DMT system using PTEQ. The employed definition of $|\rho|$ can be found in the appendix. The range of possible values for $|\rho|$ is $0 \leq |\rho| \leq 1$, with $\rho = 0$ when the considered signal is proper and $|\rho| = 1$ when the considered signal is maximally improper.

In the received signal of interest — i.e. in the contribution of $X_n[k]$ to $\hat{X}_n[k]$ — impropriety can cause a loss of orthogonality and difference in scaling between the real and imaginary parts of $X_n[k]$. It is observed in Fig. 4 that the circularity coefficient of the received signal of interest is rather low, meaning that signal of interest is close to being proper. This is to be expected, as the power of the interference from tone n^* is outweighed by the power of the direct signal, even if the CP is too short.

Fig. 4 confirms that the interference-plus-noise is close to maximally improper when $T = 0$ (i.e. when no channel shortening is applied), as was already illustrated in Fig. 2. When a channel shortening filter is then applied, i.e. when $T \geq 1$, the interference-plus-noise at higher frequencies becomes mostly proper. On subcarriers with index $n \leq 100$ however, the interference-plus-noise remains improper. On subcarriers with index $n \leq 50$, the interference-plus-noise impropriety often even seems to be increased by the PTEQ. Moreover, there seems to be no straightforward relation between the level of interference-plus-noise impropriety and the PTEQ filter order T. Such a relation is not expected to exist, as the PTEQ filter design equation in (18) is based on maximizing the SINR, not on minimizing the impropriety.

C. SINR and Achievable Bit Loading

Fig. 5 displays the achieved SINR and achievable bit loading for the system considered in Section VI-A. Results labeled 'CL' are obtained using (15) and (17), whereas results labeled 'WL' are obtained using (26) and (29). For DMT systems with WL PTEQ, separate SINR curves are given for the real and imaginary part of the received signal. However, the SINR curves of the real and imaginary parts of the same received signal are assigned the same line style and the same label in the legend. SINR and bit loading values are provided for a CL receiver with 0-order PTEQ ($T = 0$ - CL, Assumed Proper), for a WL receiver with 0-order WL PTEQ without transmitter-side symbol rotation ($T = 0$ - WL), for a WL receiver with both 0-order WL PTEQ and transmitter-side symbol rotation ($T = 0$ - Rotated WL), and for a WL receiver with both 1-st-order WL PTEQ and transmitter-side symbol rotation ($T = 1$ - Rotated WL).

A first observation is that the bit loading curve labeled '$T = 0$ - WL' has distinct peaks and troughs. At the peaks — which are located at $n = 47$, $n = 185$, $n = 315$, and $n = 530$ — the axes of the received PAM symbols are well aligned with the eigenvectors of the residual interference-plus-noise covariance matrix. It is observed that no further performance

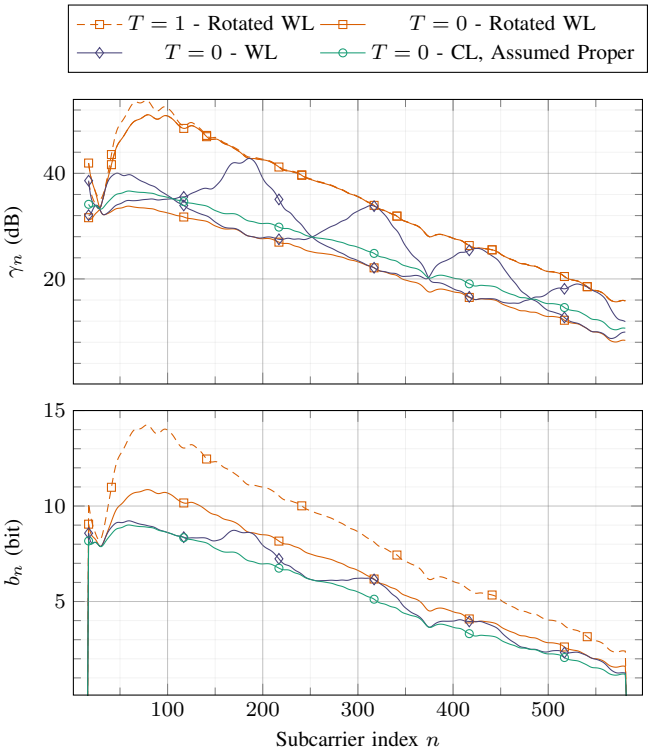


Fig. 5: Achieved SINR (top) and achievable bit loading (bottom) for the system considered in Section VI-A. A line mark offset was applied to the curves labeled '$T = 1$ - Rotated WL' to make the difference between this curve and the curve labeled '$T = 0$ - Rotated WL' more clear.

increase can be obtained by rotating the symbols prior to transmission. At the troughs of the same bit loading curve — which are located at $n = 104$, $n = 252$, $n = 375$, and $n = 482$ — the axes of the received PAM symbols are poorly aligned with the eigenvectors of the residual interference-plus-noise covariance matrix, and significant performance improvements can be obtained by optimally rotating the symbols prior to transmission. Note that this misalignment between the axes of the received PAM symbols and the eigenvectors of the residual interference-plus-noise covariance matrix is also illustrated by Fig. 2a, where the received signal of the considered system is plotted for a subcarrier that is near the trough at $n = 104$.

A second observation is that the two SINR curves labeled '$T = 1$ - Rotated WL' almost coincide with the upper SINR curve labeled '$T = 0$ - Rotated WL'. Increasing the PTEQ filter order from $T = 0$ to $T = 1$ thus improves the SINR for a single PAM signal, leaving the SINR of the other PAM signal on the same subcarrier mostly unaltered. On subcarriers $n \geq 100$, this observation indicates that the interference-plus-noise in the optimal unbiased WL PTEQ filter output is approximately proper for $T = 1$ — as was the case in the PTEQ output for $T = 1$ (see Fig. 4).

A final observation is that the bit loading curve labeled '$T = 0$, CL, Assumed proper' consistently lies below the other bit loading curves in Fig. 5. As such, it can be concluded that — in the considered DSL system — explicitly taking the interference-plus-noise impropriety into account increases the system performance.

D. Achievable Data Rate

Finally, the performance increase enabled by explicitly taking the interference-plus-noise covariance into account is assessed. The metric under consideration is the achievable data rate, which is calculated as

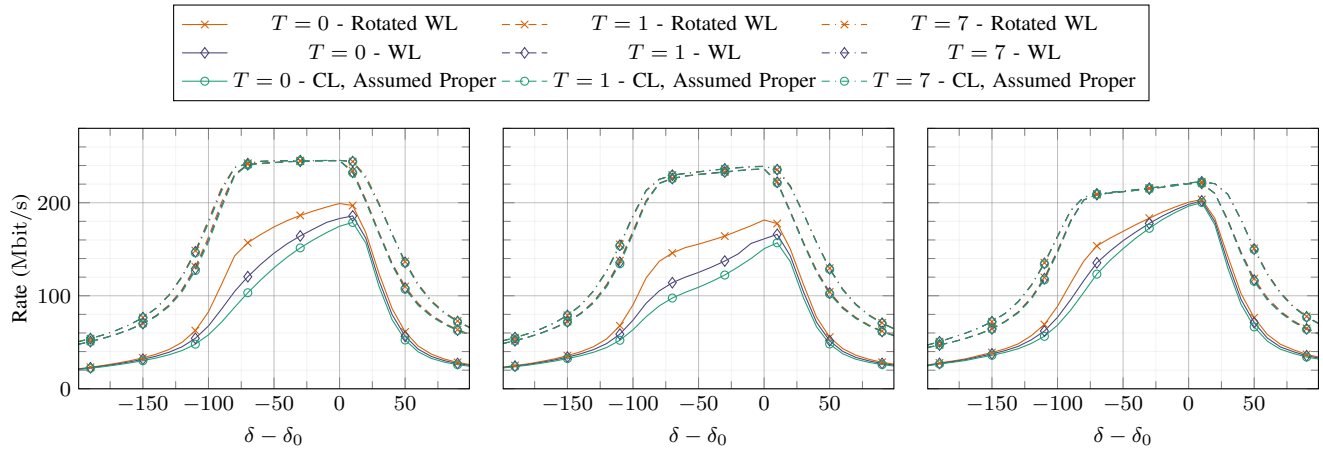
$$R = \frac{f_{\mathrm{s}}}{N + \nu_{\mathrm{p}}} \sum_{n=0}^{N/2} b_n. \tag{57}$$

Fig. 6 displays the achievable data rate as a function of the delay δ for G.fast systems using either PTEQ or WL PTEQ. The delay sweep has a granularity of 10 samples and ranges from $\delta_0 - 200$ to $\delta_0 + 100$, where δ_0 is chosen by maximizing the energy in $\left[h[\delta_0], \ldots, h[\delta_0 + \nu_{\mathrm{p}}]\right]$ as in [16]. The considered CP length is $\nu_{\mathrm{p}} = 128$. In addition, Fig. 7 displays the maximum achievable data rate as a function of the CP length ν_{p}. Each data point in Fig. 7 is obtained by executing a delay sweep ranging from $\delta_0 - 200$ to $\delta_0 + 100$ with a granularity of 10 samples, and determining the maximum achieved data rate over the considered values of δ.

The results in Fig. 6 and Fig. 7 indicate that explicitly taking the interference-plus-noise impropriety into account by using (26) and (29) yields better performance, especially for $T = 0$ — which corresponds to a DMT receiver without channel shortening and with (WL) FEQ. When $T \geq 1$ however, the results for DMT systems using PTEQ and WL PTEQ mostly coincide, indicating that the impact of interference-plus-noise impropriety on the data rate calculation is reduced by the PTEQ filter even without generalizing it to the WL PTEQ filter structure. It can be concluded that, when the employed channel shortening technique is able to eliminate most of the ISI/ICI, the impact of impropriety is reduced and the gap between the legacy SINR and bit loading expressions from [8]–[18] and the newly developed expressions disappears.

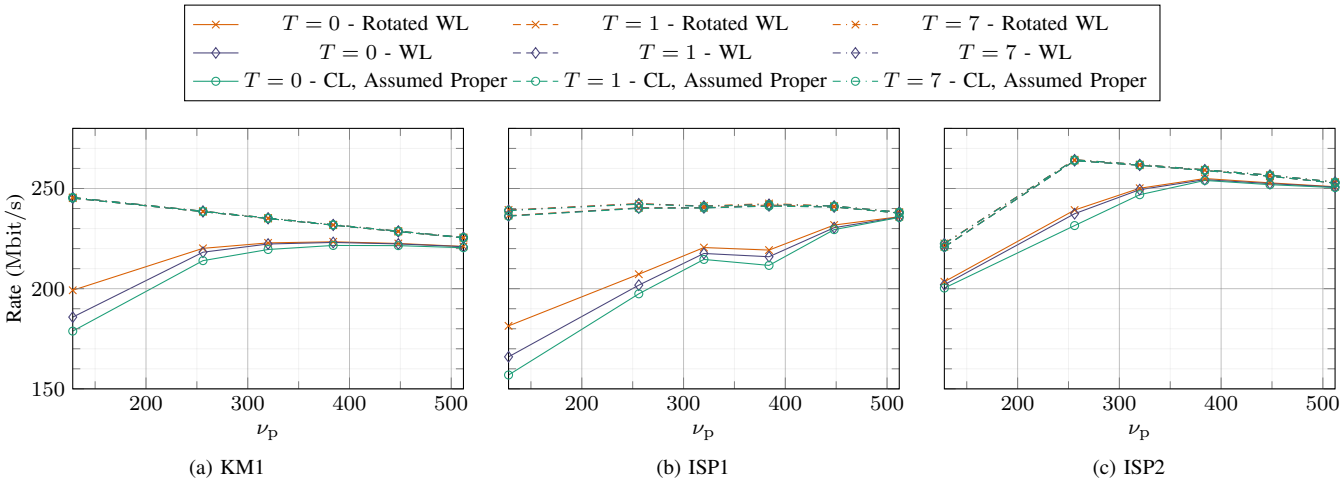
It is once more emphasized that, if regular QAM symbols are transmitted and (17) is used to determine the bit loading, then the target BER may not be achieved by a receiver employing a PTEQ filter due to the residual impropriety at the lower frequencies (see Section VI-B). This failure to achieve the target BER is even more likely if the receiver is employing a 0-order PTEQ filter, as the ISI/ICI may be improper on all active subcarriers. Failing to achieve the target BER can be counteracted by reducing the bit loading, resulting in an even lower achieved data rate for 0-order PTEQ than is indicated in Fig. 6 and Fig. 7 ($T = 0$ - CL, Assumed Proper). Even without this further reduction in data rate for 0-order PTEQ however, it is seen that the 0-order WL PTEQ filter with double PAM symbols achieves part of the data rate increase without the more significant increase in computational complexity that is required by 1-order PTEQ. As such, 0-order WL PTEQ provides an interesting trade-off between the more complex 1-order PTEQ and the worse-performing 0-order PTEQ.

Finally, the results in Fig. 6 and Fig. 7 also indicate that performance of the PTEQ saturates with a very short FIR



(a) KM1 (b) ISP1 (c) ISP2

Fig. 6: Achievable data rate as a function of the delay δ for G.fast systems using WL PTEQ with impulse responses given by Fig. 3. Delays are relative to a reference delay δ_0, which is chosen by maximizing the energy in $\left[h[\delta_0], \ldots, h[\delta_0 + \nu_\mathrm{p}]\right]$ as in [16]. The granularity of the delay sweep is 10 samples. The considered CP length is $\nu_\mathrm{p} = 128$.



(a) KM1 (b) ISP1 (c) ISP2

Fig. 7: Maximum data rate achieved for different CP lengths. The data points correspond to the maximum data rates that are achieved when sweeping the delay δ from $\delta_0 - 200$ to $\delta_0 + 100$ with a granularity of 10 samples, as in Fig. 6.

filter (only two taps), as moving from $T = 2$ to $T = 7$ provides negligible gain. This observation suggests that the employed channel impulse responses have few poles which are readily canceled with the PTEQ. Moreover, the PTEQ clearly enables reducing ν_p for all considered systems without reducing the achieved data rate, and even achieves performance improvements for KM1. A notable exception however are the results for ISP2, where a significant performance drop is observed when ν_p is reduced from $\nu_\mathrm{p} = 256$ to $\nu_\mathrm{p} = 128$. A possible explanation is as follows. It is seen in Fig. 3c that the CIR of ISP2 has a specific structure: the CIR looks like the sum of two impulse responses of almost identical systems characterized by only few poles, with a distance between the impulse response peeks of ca. 225 samples. The PTEQ is likely able to significantly shorten these impulse responses individually — as they are similar in shape to the CIR of KM1 — but seems to struggle with reducing the length of their sum to below $\nu_\mathrm{p} - \nu_\mathrm{s}$ for smaller values of ν_p.

VII. Conclusion

In this paper, new SINR and bit loading expressions have been presented for DMT systems with (P)TEQ, which are valid when the received signals are improper. The performance of such DMT systems has been further optimized by generalizing the PTEQ to a widely linear PTEQ, and by optimally rotating the symbols $X_n[k]$ prior to transmission. Finally, simulation results have demonstrated that explicitly taking signal impropriety into account leads to an increased achievable bit loading when residual ISI/ICI is present in the system. When the employed channel shortening technique is able to eliminate most of the ISI/ICI, the impact of impropriety is reduced and the gap between the state-of-the-art SINR and bit loading expressions from [8]–[18] and the newly developed expressions disappears.

In future work, power loading algorithms will be developed to determine optimal symbol powers S_n. Such algorithms can be based on filter-less bit loading expressions as in (43),

and may employ successive convex approximation (SCA) algorithms similar to those developed in [18]. Furthermore, it will be interesting to generalize the bit loading expression as in (43) so that it also applies to the case where the transmitted symbols are improper, as the use of improper transmitted symbols is indeed expected to yield further performance gains [22], [30]–[34]. Such a bit loading expression could then be used to design SCA algorithms optimizing the improper symbols' transmit covariance matrices as well.

APPENDIX
CIRCULARITY COEFFICIENT

The circularity coefficient is defined as the magnitude of the complex correlation coefficient ρ [21], which in turn is defined for scalar random variables of the form $x = x_r + jx_i$ as

$$\rho = \frac{\mathbb{E}\left\{x^2\right\}}{\mathbb{E}\left\{xx^*\right\}}. \tag{58}$$

The range of possible values for the circularity coefficient $|\rho|$ is $0 \leq |\rho| \leq 1$. Moreover, $\rho = 0$ when x is proper, and $|\rho| = 1$ when x is maximally improper. Given $\Psi_{x_r x_r} \triangleq \mathbb{E}\{x_r x_r\}$, $\Psi_{x_i x_i} \triangleq \mathbb{E}\{x_i x_i\}$, $\Psi_{x_r x_i} \triangleq \mathbb{E}\{x_r x_i\}$, the complex correlation coefficient ρ can equivalently be calculated as [21]

$$\rho = \frac{\Psi_{x_r x_r} - \Psi_{x_i x_i}}{\Psi_{x_r x_r} + \Psi_{x_i x_i}} + j\frac{2\Psi_{x_r x_i}}{\Psi_{x_r x_r} + \Psi_{x_i x_i}}. \tag{59}$$

Expressions will now be given that can be used to determine the complex correlation coefficient of the interference-plus-noise, as well as for the complex correlation coefficient of the direct signals. If x in (59) is given by $x^{\mathbb{R}} = \underline{\mathbf{U}}_n^{\mathsf{T}} \underline{\boldsymbol{z}}_n^{\mathbb{R}}$, then it can be seen from the composite real channel model in (25) that $\Psi_{x_r x_r}$, $\Psi_{x_i x_i}$, and $\Psi_{x_r x_i}$ can be calculated as follows.

$$\begin{bmatrix} \Psi_{x_r x_r} & \Psi_{x_r x_i} \\ \Psi_{x_i x_r} & \Psi_{x_i x_i} \end{bmatrix} = \underline{\mathbf{U}}_n^{\mathsf{T}} \boldsymbol{\Psi}_{\underline{\boldsymbol{z}}_n^{\mathbb{R}} \underline{\boldsymbol{z}}_n^{\mathbb{R}}} \underline{\mathbf{U}}_n \tag{60}$$

If, on the other hand, x in (59) is given by $x^{\mathbb{R}} = \underline{\mathbf{U}}_n^{\mathsf{T}} \underline{\mathbf{G}}_n X_n^{\mathbb{R}}[k]$, then — using $\mathbb{E}\{X_n^{\mathbb{R}}[k]\,(X_n^{\mathbb{R}}[k])^{\mathsf{T}} = \frac{S_n}{2}\mathbf{I}_2$ — it can be seen that $\Psi_{x_r x_r}$, $\Psi_{x_i x_i}$, and $\Psi_{x_r x_i}$ can be calculated as follows.

$$\begin{bmatrix} \Psi_{x_r x_r} & \Psi_{x_r x_i} \\ \Psi_{x_i x_r} & \Psi_{x_i x_i} \end{bmatrix} = \underline{\mathbf{U}}_n^{\mathsf{T}} \left(\frac{S_n}{2} \underline{\mathbf{G}}_n \underline{\mathbf{G}}_n^{\mathsf{T}} \right) \underline{\mathbf{U}}_n \tag{61}$$

REFERENCES

[1] ITU-T Study Group 15, "Fast access to subscriber terminals (G.fast) - Power spectral density specification," International Telecommunication Union, Tech. Rep. ITU-T G.9700, 2014.

[2] ——, "Fast access to subscriber terminals (G.fast) - Physical layer specification," International Telecommunication Union, Tech. Rep. ITU-T G.9701, 2015.

[3] R. Strobel and W. Utschick, "Coexistence of G.fast and VDSL in FTTDP and FTTC deployments," in *2015 23rd European Signal Processing Conference (EUSIPCO)*, Aug. 2015, pp. 1103–1107, iSSN: 2076-1465.

[4] F. Effenberger, "Future broadband access networks," *Proceedings of the IEEE*, vol. 104, no. 11, pp. 2078–2081, Nov. 2016.

[5] R. Strobel and V. Oksman, "Line grouping for crosstalk avoidance," US Patent US20 180 041 246A1, Feb., 2018.

[6] D. Hincapie, M. Leibiger, and E. Oswald, "Roll-out of giga-bit copper: Impact-analysis of deploying G.fast in FTTC access networks," in *Proc. 12th ITG-Symposium; Broadband Coverage in Germany*, Apr. 2018, pp. 1–6.

[7] F. Mazzenga and R. Giuliano, "Analysis of zero-forcing vectoring for VDSL2 and long reach VDSL technologies," *IEEE Transactions on Communications*, vol. 67, no. 11, pp. 7811–7823, Nov. 2019.

[8] J. Chow, J. Tu, and J. Cioffi, "A discrete multitone transceiver system for HDSL applications," *IEEE Journal on Selected Areas in Communications*, vol. 9, no. 6, pp. 895–908, Aug. 1991.

[9] N. Al-Dhahir and J. Cioffi, "Optimum finite-length equalization for multicarrier transceivers," *IEEE Transactions on Communications*, vol. 44, no. 1, pp. 56–64, Jan. 1996.

[10] P. Melsa, R. Younce, and C. Rohrs, "Impulse response shortening for discrete multitone transceivers," *IEEE Transactions on Communications*, vol. 44, no. 12, pp. 1662–1672, Dec. 1996.

[11] W. Henkel and T. Kessler, "Maximizing the channel capacity of multicarrier transmission by suitable adaptation of the time-domain equalizer," *IEEE Transactions on Communications*, vol. 48, no. 12, pp. 2000–2004, Dec. 2000.

[12] G. Arslan, B. Evans, and S. Kiaei, "Equalization for discrete multitone transceivers to maximize bit rate," *IEEE Transactions on Signal Processing*, vol. 49, no. 12, pp. 3123–3135, Dec. 2001.

[13] R. Martin, J. Balakrishnan, W. Sethares, and C. Johnson, "A blind adaptive TEQ for multicarrier systems," *IEEE Signal Processing Letters*, vol. 9, no. 11, pp. 341–343, Nov. 2002.

[14] K. Vanbleu, G. Ysebaert, G. Cuypers, M. Moonen, and K. V. Acker, "Bitrate-maximizing time-domain equalizer design for DMT-based systems," *IEEE Transactions on Communications*, vol. 52, no. 6, pp. 871–876, Jun. 2004.

[15] M. Sharma, M. Moonen, Y. Lefevre, and P. Tsiaflakis, "MIMO time domain equalizer design for long reach xDSL," ESAT-STADIUS, KU Leuven, Leuven, Belgium, Internal Report 20-70, 2020.

[16] K. V. Acker, G. Leus, M. Moonen, O. v. d. Wiel, and T. Pollet, "Per tone equalization for DMT-based systems," *IEEE Transactions on Communications*, vol. 49, no. 1, pp. 109–119, Jan. 2001.

[17] P. K. Pandey, M. Moonen, and L. Deneire, "Sparse approximation based resource allocation in DSL/DMT transceivers with per-tone equalization," *Signal Processing*, vol. 95, pp. 43–48, Feb. 2014.

[18] J. Verdyck, Y. Lefevre, P. Tsiaflakis, and M. Moonen, "Per-tone precoding and per-tone equalization for OFDM and DMT transmission systems: Duality, filter optimization, and resource allocation," ESAT-STADIUS, KU Leuven, Leuven, Belgium, Internal Report 20-78, 2020.

[19] G. Taubock, "Complex noise analysis of DMT," *IEEE Transactions on Signal Processing*, vol. 55, no. 12, pp. 5739–5754, Dec. 2007.

[20] R. Strobel, *Channel Modeling and Physical Layer Optimization in Copper Line Networks*, ser. Signals and Communication Technology. Springer International Publishing, 2019.

[21] T. Adali, P. J. Schreier, and L. L. Scharf, "Complex-valued signal processing: The proper way to deal with impropriety," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5101–5125, Nov. 2011.

[22] C. Hellings and W. Utschick, "Block-skew-circulant matrices in complex-valued signal processing," *IEEE Transactions on Signal Processing*, vol. 63, no. 8, pp. 2093–2107, Apr. 2015.

[23] D. E. Knuth, *The Art of Computer Programming, Volume 2: Seminumerical Algorithms*. Reading, Mass.: Addison-Wesley Pub, Jan. 1981.

[24] B. Picinbono and P. Chevalier, "Widely linear estimation with complex

data," *IEEE Transactions on Signal Processing*, vol. 43, no. 8, pp. 2030–2033, Aug. 1995.

[25] G. Ungerboeck, "Trellis-coded modulation in DSL systems," in *Fundamentals of DSL Technology*. CRC Press, Sep. 2005.

[26] G. Forney and L.-F. Wei, "Multidimensional constellations. I. Introduction, figures of merit, and generalized cross constellations," *IEEE Journal on Selected Areas in Communications*, vol. 7, no. 6, pp. 877–892, Aug. 1989.

[27] R. Hunger and M. Joham, "A general rate Duality of the MIMO multiple access channel and the MIMO broadcast channel," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM)*, Nov. 2008, pp. 1–5.

[28] X.-D. Zhang, *Matrix Analysis and Applications*. Cambridge University Press, Oct. 2017, google-Books-ID: YBs0DwAAQBAJ.

[29] D. Acatauassu, S. Höst, C. Lu, M. Berg, A. Klautau, and P. O. Börjesson, "Simple and causal twisted-pair channel models for G.fast systems," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, Dec. 2013, pp. 2834–2839.

[30] S. A. Jafar, V. Cadambe, and C. Wang, "Interference alignment with asymmetric complex signaling," in *Proc. 47th Annual Allerton Conference on Communication, Control, and Computing*, Allerton, Sep. 2009, pp. 991–996.

[31] C. Lameiro and I. Santamaría, "Degrees-of-freedom for the 4-user SISO interference channel with improper signaling," in *Proc. IEEE International Conference on Communications (ICC)*, Jun. 2013, pp. 3053–3057.

[32] Z. K. Ho and E. Jorswieck, "Improper Gaussian signaling on the two-user SISO interference channel," in *Proc. 8th International Symposium on Wireless Communication Systems*, Nov. 2011, pp. 824–828.

[33] Y. Zeng, C. M. Yetis, E. Gunawan, Y. L. Guan, and R. Zhang, "Transmit optimization with improper gaussian signaling for interference channels," *IEEE Transactions on Signal Processing*, vol. 61, no. 11, pp. 2899–2913, Jun. 2013.

[34] Y. Zeng, R. Zhang, E. Gunawan, and Y. L. Guan, "Optimized transmission with improper gaussian signaling in the K-User MISO interference channel," *IEEE Transactions on Wireless Communications*, vol. 12, no. 12, pp. 6303–6313, Dec. 2013.