Zero-Forcing and MMSE Precoding for G.fast

Conference Paper · December 2015

DOI: 10.1109/GLOCOM.2014.7417707

Technische Universität München

Zero-Forcing and MMSE Precoding for G.fast

Globecom 2015 - Symposium on Selected Areas in Communications (GC15 SAC)

Rainer Strobel, Andreas Barthelme, Wolfgang Utschick

Department of Electrical Engineering and Information Technology
Associate Institute for Signal Processing
Univ.-Prof. Dr.-Ing. Wolfgang Utschick

Zero-Forcing and MMSE Precoding for G.fast

Rainer Strobel*‡, Andreas Barthelme*, Wolfgang Utschick*
*Fachgebiet Methoden der Signalverarbeitung, Technische Universität München, 80290 München, Germany
{rainer.strobel,a.barthelme,utschick}@tum.de
‡Lantiq GmbH & Co. KG, 85579 Neubiberg, Germany, rainer.strobel@lantiq.com

***Abstract*—Hybrid copper/fiber networks bridge the gap between the fiber link at the distribution point and the customer by using copper wires over the last meters. The G.fast transmission technology has been designed to be used in such a fiber to the distribution point (FTTdp) network. Crosstalk management using MIMO precoding in downlink and MIMO equalization in uplink is a key to the required performance of FTTdp.**

Currently, there are mostly linear and nonlinear zero-forcing (ZF) methods discussed for precoding in G.fast. This paper presents a spectrum optimization algorithm for both zero-forcing precoding methods.

Minimum mean squared error (MMSE) precoding shows significant advantages under certain channel conditions. A performance comparison of MMSE precoding for G.fast with linear and nonlinear zero-forcing methods indicates that MMSE precoding outperforms standard linear and nonlinear ZF precoding. Applying the proposed spectrum optimization brings all three methods to a similar performance.

I. Introduction

To address increasing consumer demand on higher data rates, ITU has recently finished the next generation copper access technology standard called G.fast [1]. For G.fast, mostly linear and nonlinear zero-forcing (ZF) precoding methods are discussed [2], [3]. Spectrum optimization for linear zero-forcing precoding is investigated in [4],

This paper extends the approach of [4] to nonlinear precoding and demonstrates a different optimization algorithm based on quadratic programming. In addition to zero-forcing precoding, weighted MMSE precoding [5] for sum-rate maximization is derived for the G.fast application which requires to decompose the optimization problem over the subcarriers of the multi-carrier system. Results for weighted MMSE precoding are compared to the optimized zero-forcing methods.

II. System Model

The key features of G.fast in the FTTdp network are summarized in [3] using a frequency spectrum from 2 MHz to 106 or 212 MHz with DMT multi-carrier modulation and synchronized time division duplexing. Crosstalk coupling between different lines of a binder becomes very strong for this spectrum. Therefore, linear or nonlinear crosstalk mitigation methods are required to achieve the desired performance.

This work has been founded by the research project "FlexDP - Flexible Breitband Distribution Points", funded by the Bayerische Forschungsstiftung.

A. Architecture

The FTTdp network consists of a fiber link to a remotely powered distribution point unit (DPU) which is connected to the customer premises equipment (CPE) via traditional copper wires. This work investigates the downlink transmission from a multi-port DPU, to a small number of customers, e.g. 24. They share the same cable bundle for some distance from the DP to the CPEs. We focus on MIMO transceiver optimization. Linear zero-forcing, nonlinear zero-forcing and linear MMSE precoding are compared in terms of achievable data rates vs. line length.

Every subscriber uses a set of subcarriers $k = 1,\dots,K$ for data transmission over the crosstalk channel described by the matrix $\boldsymbol{H}^{(k)} \in \mathbb{C}^{L\times L}$ for subcarrier k and a system with L lines. The linear precoder matrix $\boldsymbol{P}^{(k)} \in \mathbb{C}^{L\times L}$ at the DPU is used to pre-compensate crosstalk between the lines.

The gain-scaling for ZF precoding is collected in diagonal matrices $\boldsymbol{S}^{(k)} = \mathrm{diag}(s_1^{(k)},\dots,s_L^{(k)}) \in \mathbb{R}^{L\times L}$. The operation $\mathrm{diag}(.)$ transforms a vector into a diagonal matrix and a diagonal matrix into a vector. The receivers also apply the diagonal equalizer $\boldsymbol{G}^{(k)} = \mathrm{diag}(g_1^{(k)},\dots,g_L^{(k)}) \in \mathbb{C}^{L\times L}$ to compensate direct channel distortion. The block diagram in Fig. 1 holds for linear as well as for nonlinear precoding. For linear precoding, the nonlinear operation and modulo blocks are bypassed.

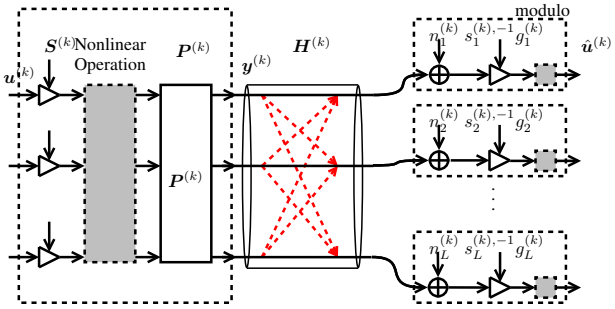

Fig. 1. Downstream system model representing one subcarrier

The signal processing steps for linear precoding are

$$\boldsymbol{y}^{(k)} = \boldsymbol{P}^{(k)}\boldsymbol{S}^{(k)}\boldsymbol{u}^{(k)} \tag{1}$$

$$\hat{\boldsymbol{u}}^{(k)} = \boldsymbol{S}^{(k),-1}\boldsymbol{G}^{(k)}\left(\boldsymbol{H}^{(k)}\boldsymbol{y}^{(k)} + \boldsymbol{n}^{(k)}\right) \tag{2}$$

for each subcarrier k. The transmit and receive signal vectors are $\boldsymbol{u}^{(k)} \in \mathbb{C}^L$ and $\hat{\boldsymbol{u}}^{(k)} \in \mathbb{C}^L$, respectively. $\boldsymbol{y}^{(k)} \in \mathbb{C}^L$ is the

DP output signal. The transmit signals are assumed to be statistically independent, zero-mean, unit power QAM (quadrature amplitude modulation) signals. The receivers experience additive white Gaussian noise (AWGN) $\boldsymbol{n}^{(k)} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0},\sigma^2\boldsymbol{I})$. For linear zero-forcing precoding, the zero-forcing condition

$$\boldsymbol{G}^{(k)}\boldsymbol{H}^{(k)}\boldsymbol{P}^{(k)} = \boldsymbol{I} \tag{3}$$

must be satisfied.

For nonlinear precoding, the nonlinear operation block in Fig. 1 contains the feedback structure and a modulo operation $\mathrm{mod}\,(.)$ [6]. The signal processing steps for the generation of the transmit signal $\boldsymbol{y}^{(k)}$ are given by

$$\boldsymbol{u}_{\text{back}}^{(k)} = \boldsymbol{P}_{\text{b}}^{(k)}\boldsymbol{u}_{\text{mod}}^{(k)}, \tag{4}$$

$$\boldsymbol{u}_{\text{mod}}^{(k)} = \mod\left(\boldsymbol{u}_{\text{back}}^{(k)} + \boldsymbol{S}^{(k)}\boldsymbol{u}^{(k)}\right), \tag{5}$$

$$\boldsymbol{y}^{(k)} = \boldsymbol{P}^{(k)}\boldsymbol{u}_{\text{mod}}^{(k)}. \tag{6}$$

The nonlinear zero-forcing precoding coefficients are determined by using the QR decomposition of the channel matrix

$$\boldsymbol{Q}^{(k)}\boldsymbol{R}^{(k)} = \boldsymbol{H}^{(k),\text{T}} \tag{7}$$

into the unitary matrix $\boldsymbol{Q}^{(k)}$ and the upper triangular matrix $\boldsymbol{R}^{(k)}$. The forward matrix is given by

$$\boldsymbol{P}^{(k)} = \boldsymbol{Q}^{(k),*} \tag{8}$$

and the lower triangular feedback matrix $\boldsymbol{P}_{\text{b}}^{(k)} \in \mathbb{C}^{L\times L}$ is

$$\boldsymbol{P}_{\text{b}}^{(k)} = \left(\boldsymbol{I}_L - \mathrm{diag}\left(\mathrm{diag}\left(\boldsymbol{R}^{(k)}\right)\right)^{-1}\boldsymbol{R}^{(k),\text{T}}\right). \tag{9}$$

The equalizer at receive side, which satisfies the zero-forcing condition, is obtained by

$$\boldsymbol{G}^{(k)} = \mathrm{diag}\left(\mathrm{diag}\left(\boldsymbol{R}^{(k)}\right)\right)^{-1}. \tag{10}$$

The system of Fig. 1 can also be used for illustrating the MMSE precoding method [7]. Linear MMSE precoding performs the signal processing steps of Eq. (1) and (2). Therefore, the precoding complexity is equal to linear zero-forcing precoding which is an advantage compared to TH precoding.

For transmit power optimization, the real-valued transmit gains $s_l^{(k)}$ are represented by the squared values $x_l^{(k)} = \left|s_l^{(k)}\right|^2$. The channel matrix can be rearranged into one row vector $\boldsymbol{h}_l^{(k)} \in \mathbb{C}^L$ per subscriber $\boldsymbol{H}^{(k)} = \left[\boldsymbol{h}_1^{(k)},\ldots,\boldsymbol{h}_L^{(k)}\right]^{\text{T}}$. Similarly, the precoder matrix is split into one column vector $\boldsymbol{p}_l^{(k)} \in \mathbb{C}^L$ per subscriber, i.e., $\boldsymbol{P}^{(k)} = [\boldsymbol{p}_1^{(k)},\ldots,\boldsymbol{p}_L^{(k)}]$.

B. Data Rate Computation

For a realistic estimation of the achievable data rates of the G.fast links, some implementation limitations are considered in the data rate calculation.

With the precoder and equalizer satisfying the zero-forcing constraint (3), the signal-to-noise-ratio $SNR_l^{(k)}$ is given by

$$SNR_l^{(k)} = \frac{|\boldsymbol{h}_l^{(k),\text{T}}\boldsymbol{p}_l^{(k)}|^2 x_l^{(k)}}{\sigma_l^2} \tag{11}$$

such that the achievable number of bits per channel use for line l and subcarrier k reads as

$$b_l^{(k)} = \log_2\left(1 + \frac{SNR_l^{(k)}}{\Gamma}\right). \tag{12}$$

with the SNR gap Γ to accommodate coding gain, non-Gaussian modulation and to achieve the desired bit error rate.

The QAM modulation and the analog and mixed-signal components of the system do no allow arbitrary constellation sizes which gives an upper bound on the number of bits per channel use $b_l^{(k)} \leq b_{\text{max}}$. G.fast allows constellations up to 4096-QAM, which corresponds to $b_{\text{max}} = 12$ bit. The actual number of bits $\hat{b}_l^{(k)}$ per line l and subcarrier k is given by

$$\hat{b}_l^{(k)} = \min\left(\left\lfloor b_l^{(k)}\right\rfloor, b_{\text{max}}\right). \tag{13}$$

The corresponding line rate is

$$R_l = \frac{\eta}{t_{\text{sym}}}\sum_{k=1}^{K}\hat{b}_l^{(k)} - R_{\text{oh}}, \tag{14}$$

with t_{sym} including the cyclic prefix and windowing [1], an efficiency factor η and the overhead channel rate R_{oh} which account for the losses due to communication protocol, framing overhead and management communication.

Eq. (13) shows that the actual data rate is a discrete function. In [8], it is shown that the floor rounding operation can be neglected when the SNR gap Γ is selected accordingly. Therefore, (12) without rounding and the upper bound is used as an optimization objective, while reported data rates are calculated using (13).

C. Spectral Constraints

Following [4], three power constraints are of interest. The first is the per-line spectral mask constraint which is defined by the regulator, e.g. in [9], in terms of a per-line power spectral density (PSD) limit. It can be translated into an equivalent power limit $p_{\text{mask},l}^{(k)}$ per line l and per carrier k.

$$\mathrm{diag}\left(\boldsymbol{P}^{(k)}\mathrm{diag}(\boldsymbol{x}^{(k)})\boldsymbol{P}^{(k),\text{H}}\right) \leq \boldsymbol{p}_{\text{mask}}^{(k)}. \tag{15}$$

The second constraint comes from the limited transmit amplifier capabilities. It is a limit on the per-line sum-power $\boldsymbol{p}_{\text{sum}}$

$$\sum_{k=1}^{K}\mathrm{diag}\left(\boldsymbol{P}^{(k)}\mathrm{diag}(\boldsymbol{x}^{(k)})\boldsymbol{P}^{(k),\text{H}}\right) \leq \boldsymbol{p}_{\text{sum}}. \tag{16}$$

The third constraint reflects the limited capabilities of the modulator, which does not allow constellation sizes larger than $2^{b_{\text{max}}}$. This can be represented by an equivalent power limit $p_{\text{bmax},l}^{(k)}$. For zero-forcing precoding as well as for the high SNR region of MMSE precoding, it can be derived from Eq. (11) according to

$$x_l^{(k)} \leq p_{\text{bmax},l}^{(k)} = \frac{(2^{b_{\text{max}}}-1)\sigma^2\Gamma}{\left|\boldsymbol{h}_l^{(k),\text{T}}\boldsymbol{p}_l^{(k)}\right|^2}. \tag{17}$$

The zero-forcing approximation for the bit loading constraint in MMSE precoding is acceptable, because the constraint is only active for high SNR where it converges to the ZF solution.

III. Spectrum Optimization for ZF Precoding

For ZF precoding, the structure of the precoders $\boldsymbol{P}^{(k)}$ depends on the channel, only, and is independent of the transmit signal power. Spectrum optimization for zero-forcing precoding is a convex problem. A solution for linear precoding is shown in [4]. This section presents a novel algorithm which is based on quadratic programming and converges faster. It is formulated for linear and nonlinear precoding.

A. Linear ZF Precoding Constraint Set and Objective

The per-carrier constraints (15) and (17) are represented in terms of the linear inequality $\boldsymbol{A}^{(k)}\boldsymbol{x}^{(k)} \leq \boldsymbol{d}^{(k)}$,

$$\boldsymbol{A}^{(k)} = \begin{pmatrix} \tilde{\boldsymbol{A}}^{(k)} \\ \boldsymbol{I}_L \\ -\boldsymbol{I}_L \end{pmatrix} \quad \boldsymbol{d}^{(k)} = \begin{pmatrix} \boldsymbol{p}_{\text{mask}}^{(k)} \\ \boldsymbol{p}_{\text{bmax}}^{(k)} \\ \boldsymbol{0}_L \end{pmatrix} \tag{18}$$

with $\tilde{\boldsymbol{A}}^{(k)} = \boldsymbol{P}^{(k)} \odot \boldsymbol{P}^{(k),*}$ where $\odot$ is the Hadamard product and * the complex conjugate.

With the constraint set including Eq. (17) and under the assumption of linear zero-forcing precoding, the bit loading to be used for the objective function is given by

$$R_{\text{LZF},l} = \sum_{k=1}^{K} \log_2 \left(1 + \left|\boldsymbol{h}_l^{(k),\text{T}} \boldsymbol{p}_l^{(k)}\right|^2 x_l^{(k)} (\Gamma\sigma^2)^{-1}\right). \tag{19}$$

B. Nonlinear ZF Precoding Constraint Set and Objective

In general, the constraint set for linear and nonlinear precoding is the same. However, in practical implementations for Tomlinson Harashima (TH) precoding [6], the nonlinear operation causes a non-negligible power increase for small QAM constellation sizes, which shall be approximated in a diagonal gain matrix $\boldsymbol{P}_{\text{m}}^{(k)}$. With $\boldsymbol{u}_{\text{mod in}}^{(k)} = \boldsymbol{S}^{(k),-1}\boldsymbol{u}_{\text{back}}^{(k)} + \boldsymbol{u}^{(k)}$, it is obtained as

$$\left[\boldsymbol{P}_{\text{m}}^{(k)}\right]_{ll} = \sqrt{\text{E}\left[\boldsymbol{u}_{\text{mod}}^{(k)}\boldsymbol{u}_{\text{mod}}^{(k),\text{H}}\right]_{ll} \text{E}\left[\boldsymbol{u}_{\text{mod in}}^{(k)}\boldsymbol{u}_{\text{mod in}}^{(k),\text{H}}\right]_{ll}^{-1}}. \tag{20}$$

The constraint set is as (18), but with

$$\tilde{\boldsymbol{A}}^{(k)} = \left(\boldsymbol{P}^{(k)}\boldsymbol{P}_m^{(k)}\right) \odot \left(\boldsymbol{P}^{(k)}\boldsymbol{P}_m^{(k)}\right)^*. \tag{21}$$

An upper bound on the rate for TH precoding can be calculated with respect to the receive equalizer [6] which gives

$$R_{\text{TH},l}^{(k)} = \sum_{k}^{K} \log_2 \left(1 + x_l^{(k)} (\Gamma\sigma^2 |g_l^{(k)}|^2)^{-1}\right). \tag{22}$$

Encoding order optimization is not discussed, here. For simulations, a fixed encoding order is used.

C. Column Norm Scaling

There is a simplified power allocation approach, which allows to satisfy the described power constraints while neglecting the maximum bit loading constraint. In a first step, the limit PSD is modified to match the sum-power constraint

$$\hat{p}_{\text{mask}}^{(k)} = \min(p_{\text{mask}}^{(k)}, \mu_{\text{sum}}) \tag{23}$$

where μ_{sum} is selected to satisfy the constraint $\sum_{k=1}^{K} \hat{p}_{\text{mask}}^{(k)} \leq p_{\text{sum}}$, and in a second step, the transmit power value is selected for all lines such that the PSD value is satisfied

$$x_l^{(k)} = \frac{\hat{p}_{\text{mask}}^{(k)}}{\max\left(\text{diag}\left(\boldsymbol{P}^{(k)}\boldsymbol{P}^{(k),\text{H}}\right)\right)} \tag{24}$$

where the precoders are scaled in advance with the inverse of the column norm as described in [10]. This is used as a reference for linear ZF and TH precoding without spectrum optimization.

D. Successive Quadratic Optimization

The objectives (19) and (22) are concave in $\boldsymbol{x}^{(k)}$ and the constraint set is linear. Therefore, the spectrum optimization problem for linear and nonlinear precoding, which reads as

$$\max_{\boldsymbol{x}^{(1)},\ldots,\boldsymbol{x}^{(K)}} \sum_{l=1}^{L} R_l \;\; \forall k : \boldsymbol{A}^{(k)}\boldsymbol{x}^{(k)} \leq \boldsymbol{d}^{(k)} \; \forall \; k = 1,\ldots,K$$
$$\sum_{k=1}^{K} \tilde{\boldsymbol{A}}^{(k)}\boldsymbol{x}^{(k)} \leq \boldsymbol{p}_{\text{sum}}, \tag{25}$$

is a convex problem. The optimization problem of Eq. (25) is coupled over all lines and all subcarriers which makes it difficult to solve the optimization problem directly for the typical system size of G.fast FTTdp. As in [4], Lagrange duality is used to separate the problem into a per-carrier subproblem and a sum-power allocation problem. Starting with the Lagrange function $\Phi(\boldsymbol{x},\boldsymbol{\mu}_{\text{sum}})$

$$\Phi(\boldsymbol{x},\boldsymbol{\mu}_{\text{sum}}) = \sum_{l=1}^{L} -R_l + \boldsymbol{\mu}_{\text{sum}}^{\text{T}} \left(\sum_{k=1}^{K} \tilde{\boldsymbol{A}}^{(k)}\boldsymbol{x}^{(k)} - \boldsymbol{p}_{\text{sum}}\right), \tag{26}$$

the per-carrier problem as stated in (26) is convex in $\boldsymbol{x}$ and it is twice differentiable. The first derivative is given by

$$\frac{\partial\Phi(\boldsymbol{x},\boldsymbol{\mu}_{\text{sum}})}{\partial x_l} = -\frac{1}{x_l + \Gamma|g_l|^2\sigma^2} + \tilde{\boldsymbol{A}}^{(k),\text{T}}\boldsymbol{\mu}_{\text{sum}}, \tag{27}$$

and the second derivative is obtained as

$$\frac{\partial^2\Phi(\boldsymbol{x},\boldsymbol{\mu}_{\text{sum}})}{\partial x_v \partial x_d} = \begin{cases} -\dfrac{1}{(\Gamma|g_v|^2\sigma_v^2)^2} & \text{for } v = d \\ 0 & \text{for } v \neq d. \end{cases} \tag{28}$$

Therefore, Newtons method can be used to solve the per-carrier problem. The vector $\nabla\Phi$ of first derivatives and the matrix $\nabla^2\Phi$ of second derivatives can be used to approximate the objective function.

In each step, the problem

$$\min_{\boldsymbol{x}^{(k)}} f_q(\boldsymbol{x},\boldsymbol{x}_0), \;\; \text{s.t. } \boldsymbol{A}^{(k)}\boldsymbol{x}^{(k)} \leq \boldsymbol{d}^{(k)} \tag{29}$$

is solved, where the objective function $f_q(\boldsymbol{x},\boldsymbol{x}_0)$ is given by

$$\Phi(\boldsymbol{x},\boldsymbol{\mu}_{\text{sum}}) \approx f_q(\boldsymbol{x},\boldsymbol{x}_0) = \Phi(\boldsymbol{x}_0,\boldsymbol{\mu}_{\text{sum}}) + \boldsymbol{c}_q^{\text{T}}\boldsymbol{x} + \boldsymbol{x}^{\text{T}}\boldsymbol{H}_q\boldsymbol{x} \tag{30}$$

with $\boldsymbol{c}_q = \nabla\Phi + \nabla^2\Phi\boldsymbol{x}_0$ and $\boldsymbol{H}_q = \frac{1}{2}\nabla^2\Phi$. The solution of Eq. (29) is used as a starting point for the next iteration. A small number of quadratic approximation steps is required to come very close to the optimal solution for the per-carrier problem.

The sum-power allocation problem is solved using a projected gradient approach. The gradient step for $\boldsymbol{\mu}_{\text{sum}}$ is

$$\boldsymbol{\mu}_{\text{sum grad}}^{[t]} = \boldsymbol{\mu}_{\text{sum}}^{[t]} + \alpha_\mu \left(\sum_{k=1}^{K} \tilde{\boldsymbol{A}}^{(k)} \boldsymbol{x}^{(k)} - \boldsymbol{p}_{\text{sum}} \right) \tag{31}$$

followed by the projection step

$$\mu_{\text{sum},l}^{[t+1]} = \max \left(\mu_{\text{sum grad},l}^{[t]}, 0 \right). \tag{32}$$

The proposed method is summarized in Algorithm 1.

Algorithm 1 Linear or nonlinear zero-forcing precoder

```
Initialize μ_sum = 0_L
repeat
  for k = 1 to K do
    Start with all lines active
    repeat
      Calculate ZF precoder (33)
      Initialize x^(k)
      Calculate constraint set (18) or (21)
      repeat
        Calculate quadratic approximation of (30) using
        (27) and (28) and solve (29)
      until Convergence of x^(k)
      Update selected active channels, H_active^(k)
    until All active carriers meet SNR lower bound
  end for
  Update power allocation using (31) and (32)
until Per-line sum-power is converged
```

Eq. (13) indicates that there is not only an upper bound on the bit loading, but also a lower bound, because tones where the SNR is insufficient to load one bit, will transmit zero bits. To use this knowledge, the row vectors $\boldsymbol{h}_{\text{active},l}^{(k)}$ corresponding to the lines of carrier k transporting at least one bit are collected in a reduced channel matrix $\boldsymbol{H}_{\text{active}}^{(k)} \in \mathbb{C}^{L_{\text{active}} \times L}$ with L_{active} active lines. Performing the pseudo-inverse $(.)^\dagger$ on the reduced channel matrix

$$\boldsymbol{P}_{\text{active}}^{(k)} = \boldsymbol{H}_{\text{active}}^{(k),\dagger} \tag{33}$$

gives the precoder matrix $\boldsymbol{P}_{\text{active}}^{(k)} \in \mathbb{C}^{L \times L_{\text{active}}}$ [4]. For TH precoding, QR decomposition of $\boldsymbol{H}_{\text{active}}^{(k)}$ is performed.

IV. MMSE Precoding

For MMSE Precoding, the precoder coefficients and the spectrum allocation are jointly optimized within $\boldsymbol{P}^{(k)}$ and $\boldsymbol{S}^{(k)} = \boldsymbol{I}$ is used.

A. Weighted MMSE Minimization with SNR Gap

Sum-rate maximization can be solved as a weighted MSE minimization, as shown in [5]. The receiver error vector is $\boldsymbol{e}^{(k)} = \hat{\boldsymbol{u}}^{(k)} - \boldsymbol{u}^{(k)}$ The per-line MSE ξ_l is for line l is then given by

$$\xi_l^{(k)} = \mathrm{E}\left[|e_l^{(k)}|^2 \right] = |g_l^{(k)}|^2 \sigma^2 + |g_l^{(k)}|^2 \boldsymbol{h}_l^{(k),\mathrm{T}} \boldsymbol{p}_l^{(k)} \boldsymbol{p}_l^{(k),\mathrm{H}} \boldsymbol{h}_l^{(k),*} - 2\Re\{ g_l^{(k)} \boldsymbol{h}_l^{(k),\mathrm{T}} \boldsymbol{p}_l^{(k)} \} + \sum_{d \neq l} |g_l^{(k)} \boldsymbol{h}_l^{(k),\mathrm{T}} \boldsymbol{p}_d^{(k)}|^2 + 1 \tag{34}$$

where $\boldsymbol{\xi}^{(k)}$ is the vector of MSEs of all lines for carrier k.

Introducing the SNR gap Γ, we have to check for $1 + \frac{SNR}{\Gamma} = \frac{1}{\xi_\Gamma}$, which leads to

$$1 + \frac{SNR_l^{(k)}}{\Gamma} = \frac{\left| \boldsymbol{h}_l^{(k),\mathrm{T}} \boldsymbol{p}_l^{(k)} \right|^2 + \Gamma \sum_{d \neq l} \left| \boldsymbol{h}_l^{(k),\mathrm{T}} \boldsymbol{p}_d^{(k)} \right|^2 + \Gamma \sigma^2}{\Gamma \sum_{d \neq l} \left| \boldsymbol{h}_l^{(k),\mathrm{T}} \boldsymbol{p}_d^{(k)} \right|^2 + \Gamma \sigma^2}. \tag{35}$$

and indicates that the equivalent MSE $\boldsymbol{\xi}_\Gamma^{(k)}$ to be optimized in presence of the SNR gap is

$$\xi_{\Gamma,l}^{(k)} = 1 - \frac{|\boldsymbol{h}_l^{(k),\mathrm{T}} \boldsymbol{p}_l^{(k)}|^2}{|\boldsymbol{h}_l^{(k),\mathrm{T}} \boldsymbol{p}_l^{(k)}|^2 + \Gamma \sum_{d \neq l} |\boldsymbol{h}_l^{(k),\mathrm{T}} \boldsymbol{p}_d^{(k)}|^2 + \Gamma \sigma^2}. \tag{36}$$

The optimal receiver with an SNR gap unequal 1 is given by

$$g_{\Gamma,l}^{(k),\star} = \frac{\left(\boldsymbol{h}_l^{(k),\mathrm{T}} \boldsymbol{p}_l^{(k)} \right)^*}{|\boldsymbol{h}_l^{(k),\mathrm{T}} \boldsymbol{p}_l^{(k)}|^2 + \Gamma \left(\sum_{d \neq l} |\boldsymbol{h}_l^{(k),\mathrm{T}} \boldsymbol{p}_d^{(k)}|^2 + \sigma^2 \right)}. \tag{37}$$

Following [5], sum-rate optimization is approximated by sequential weighted MMSE minimizations with weights which are the inverse of the MSEs $\boldsymbol{\xi}_\Gamma$,

$$w_l^{(k)} = \frac{1}{\xi_{\Gamma\, l}^{(k)}}. \tag{38}$$

It must be noted that the upper bound on the bit loading results in an upper bound on the weights $w_l^{(k)}$. The weighted MMSE optimization problem for given weights is

$$\min_{\substack{\boldsymbol{P}^{(1)},\ldots,\boldsymbol{P}^{(K)} \\ \boldsymbol{G}^{(1)},\ldots,\boldsymbol{G}^{(K)}}} \sum_{k=1}^{K} \boldsymbol{w}^{(k),\mathrm{T}} \boldsymbol{\xi}_\Gamma^{(k)} \tag{39}$$

$$\operatorname{diag}\left(\boldsymbol{P}^{(k)} \boldsymbol{P}^{(k),\mathrm{H}} \right) \leq \boldsymbol{p}_{\text{mask}}^{(k)} \; \forall \; k = 1,\ldots,K$$

$$\boldsymbol{1}_L \leq \boldsymbol{p}_{\text{bmax}}^{(k)} \; \forall \; k = 1,\ldots,K; \sum_{k=1}^{K} \operatorname{diag}\left(\boldsymbol{P}^{(k)} \boldsymbol{P}^{(k),\mathrm{H}} \right) \leq \boldsymbol{p}_{\text{sum}}.$$

While the optimal receive equalizers for fixed precoders can be calculated according to Eq. (37), finding the precoders for fixed equalizers is a non-convex optimization problem. The optimal precoders can be found in the dual uplink applying uplink-downlink duality. Strong duality holds between uplink and downlink [11] for the given constraint set, when the ZF approximation for the bit loading constraint (17) is used.

B. Uplink-Downlink Duality

After calculating the equalizer using Eq. (37), precoder optimization is performed in the dual uplink. The dual uplink system model in comparison to downlink is shown in Fig. 2.

The downlink equalizers $\boldsymbol{G}^{(k)}$ are transformed into uplink precoders $\boldsymbol{P}_{\text{u}}^{(k)} = \boldsymbol{G}^{(k)}$. The uplink channel is defined as $\boldsymbol{H}_{\text{u}}^{(k)} = \boldsymbol{H}^{(k),\mathrm{H}}$. We define an effective dual uplink channel $\bar{\boldsymbol{H}}_{\text{u}}^{(k)} = \boldsymbol{H}_{\text{u}}^{(k)} \boldsymbol{P}_{\text{u}}^{(k)} = [\bar{\boldsymbol{h}}_{\text{u},1}^{(k)}, \ldots, \bar{\boldsymbol{h}}_{\text{u},L}^{(k)}]$. The dual uplink model is used to calculate the dual uplink equalizer $\boldsymbol{G}_{\text{u}}^{(k)}$.

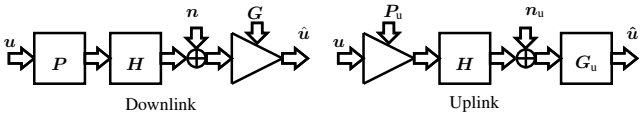

Fig. 2. Uplink Downlink Duality

The dual uplink MSE including the SNR gap Γ is given by

$$\begin{aligned}\xi_{\Gamma\mathrm{u},l}^{(k)} =& w_l \bar{\boldsymbol{g}}_{\mathrm{u},l}^{(k),\mathrm{T}} \boldsymbol{Q}^{(k)} \bar{\boldsymbol{g}}_{\mathrm{u},l}^{(k),*} + w_l \bar{\boldsymbol{g}}_{\mathrm{u},l}^{(k),\mathrm{T}} \bar{\boldsymbol{h}}_{\mathrm{u},l}^{(k)} \bar{\boldsymbol{h}}_{\mathrm{u},l}^{(k),\mathrm{H}} \bar{\boldsymbol{g}}_{\mathrm{u},l}^{(k),*} \\ & - 2\Re\{w_l \bar{\boldsymbol{g}}_{\mathrm{u},l}^{(k),\mathrm{T}} \bar{\boldsymbol{h}}_{\mathrm{u},l}^{(k)}\} + \sum_{d\neq l} \Gamma w_d |\bar{\boldsymbol{g}}_{\mathrm{u},l}^{(k),\mathrm{T}} \bar{\boldsymbol{h}}_{\mathrm{u},d}^{(k)}|^2 + 1\end{aligned} \tag{40}$$

where $\bar{\boldsymbol{g}}_{\mathrm{u},l}^{(k),\mathrm{T}}$ is the lth row of the dual uplink equalizer matrix $\boldsymbol{G}_{\mathrm{u}}^{(k)} = [\bar{\boldsymbol{g}}_{\mathrm{u},1}^{(k)}, \ldots, \bar{\boldsymbol{g}}_{\mathrm{u},L}^{(k)}]^{\mathrm{T}}$. The columns of $\boldsymbol{G}_{\mathrm{u}}^{(k)}$ are denoted as $\boldsymbol{g}_{\mathrm{u},l}$. $\boldsymbol{Q}^{(k)}$ is the uplink noise covariance matrix. For uplink-downlink duality, $\sum \xi_{\Gamma,l} = \sum \xi_{\Gamma\mathrm{u},l}$ must hold. Comparing (34) and (40), this gives $\Gamma \mathrm{tr}(\boldsymbol{G}\boldsymbol{G}^{\mathrm{H}}\sigma^2) = \mathrm{tr}(\boldsymbol{G}_u \boldsymbol{Q} \boldsymbol{G}_u^{\mathrm{H}})$.

The per-carrier dual uplink optimization problem is

$$\max_{\boldsymbol{\mu}_{\mathrm{bmax}}^{(k)}, \boldsymbol{\mu}_{\mathrm{mask}}^{(k)}} \min_{\boldsymbol{G}_{\mathrm{u}}^{(k)}} \sum_{l=1}^{L} \xi_{\Gamma\mathrm{u},l}^{(k)} \text{ s.t.} \sum_{l=1}^{L} (\mu_{\mathrm{mask},l}^{(k)} p_{\mathrm{mask},l}^{(k)} + \mu_{\mathrm{bmax},l}^{(k)} \sigma^2) \leq 1 \tag{41}$$

and with the substitution $\boldsymbol{R}_l^{(k)} = \bar{\boldsymbol{h}}_l^{(k)} \bar{\boldsymbol{h}}_l^{(k),\mathrm{H}} w_l^{(k)} + \sum_{d\neq l} \Gamma \bar{\boldsymbol{h}}_d^{(k)} \bar{\boldsymbol{h}}_d^{(k),\mathrm{H}} w_d^{(k)}$, the lth column of $\boldsymbol{G}_{\mathrm{u}}^{(k)}$, $\boldsymbol{g}_{\mathrm{u},l}^{(k)}$ is

$$\boldsymbol{g}_{\mathrm{u},l}^{(k)} = (\boldsymbol{R}_l^{(k)}/\bar{\sigma}^{(k)} + \boldsymbol{Q}_l^{(k)} \bar{\sigma}^{(k)})^{-1} \bar{\boldsymbol{h}}_l^{(k)} w_l^{(k)} \tag{42}$$

with the effective noise variance $\bar{\sigma}^{2,(k)} = \mathrm{tr}\left(\sigma^2 \mathrm{diag}\left(\boldsymbol{w}^{(k)}\right) \boldsymbol{G}^{(k)} \boldsymbol{G}^{(k),\mathrm{H}}\right)$.

The dual uplink reciver noise covariance matrix $\boldsymbol{Q}^{(k)}$ depends on the constraints and the corresponding Lagrangian multipliers and is given by

$$\boldsymbol{Q}_l^{(k)} = \mathrm{diag}(\boldsymbol{\mu}_{\mathrm{sum}} + \boldsymbol{\mu}_{\mathrm{mask}}^{(k)}) + \mu_{\mathrm{bmax}\,l}^{(k)} \boldsymbol{A}_{\mathrm{bmax}\,l}^{(k)}. \tag{43}$$

with where $\boldsymbol{A}_{\mathrm{bmax}\,l}^{(k)} = \frac{\boldsymbol{h}_l^{(k),*} \boldsymbol{h}_l^{(k),\mathrm{T}}}{(2^{b_{\max}}-1)\Gamma}$.

The Lagrangian multipliers are $\boldsymbol{\mu}_{\mathrm{mask}}^{(k)}$ for the spectral mask constraint (15), $\boldsymbol{\mu}_{\mathrm{bmax}}^{(k)}$ for the bit loading upper bound (17) and $\boldsymbol{\mu}_{\mathrm{sum}}$ for the per-line sum-power constraint (16). The values of the per-subcarrier Lagrange multipliers are derived by fixed-point iterations [12] using

$$\boldsymbol{\mu}_{\mathrm{mask}}^{(k),[t+1]} = \max\left(\boldsymbol{\mu}_{\mathrm{mask}}^{(k)[t]} \frac{\boldsymbol{p}_{\mathrm{tmp}}^{(k)}}{\boldsymbol{p}_{\mathrm{mask}}^{(k)}}, \varepsilon_{\mathrm{mask}}\right), \tag{44}$$

$$\mu_{\mathrm{bmax}\,l}^{(k),[t+1]} = \max\left(\mu_{\mathrm{bmax}\,l}^{(k),[t]} \frac{\beta \boldsymbol{g}_{\mathrm{u},l}^{(k),\mathrm{H}} \boldsymbol{A}_{\mathrm{bmax}\,l}^{(k)} \boldsymbol{g}_{\mathrm{u},l}^{(k)}}{\sigma^2}, \varepsilon_{\mathrm{bmax}}\right). \tag{45}$$

The dowlink transmit power vector $\boldsymbol{p}_{\mathrm{tmp}}$ is given by

$$\boldsymbol{p}_{\mathrm{tmp}}^{(k)} = \mathrm{diag}(\boldsymbol{G}_u^{(k)} \boldsymbol{G}_u^{(k),\mathrm{H}}) \beta \tag{46}$$

with the scaling factor β which is defined in (47) for uplink-downlink conversion. The lower bound $\varepsilon_{\mathrm{mask}}$ is required to guarantee that the matrix inversion for the equalizer calculation in (42) [12] is feasible and the lower bound $\varepsilon_{\mathrm{bmax}}$ is required to avoid that the fixed-point iterations gets stuck.

Algorithm 2 Weighted MMSE Precoder

```
Initialize P^(k) with the ZF precoder and μ_sum = 0_L
Calculate A^(k)_bmax,l for all k, l
repeat
  for k = 1 to K do
    Initialize μ^(k)_bmax and μ^(k)_mask
    repeat
      Calculate MSE (36) and update weights (38)
      repeat
        Update equalizer matrix G^(k) using (37)
        Downlink-uplink transform σ̄^2,(k), P_u^(k), H̄_u^(k)
        repeat
          Update noise covariance Q^(k) and R_l^(k)
          Update dual uplink equalizer (42)
          Evaluate (46) and (47)
          Update Lagrangians using (44) and (45)
        until Convergence of μ^(k)_bmax and μ^(k)_mask
        Uplink-downlink transform (48)
      until Convergence precoder and equalizer
    until Convergence of MSE weights
  end for
  Perform sum-power update using Eq. (49) and (32)
until Convergence of sum-power
```

The scale factor $\beta^{(k)}$ for uplink-downlink conversion is

$$\beta^{(k)} = \frac{\bar{\sigma}^2}{\mathrm{tr}(\boldsymbol{G}_u^{(k)} \boldsymbol{Q}_{\mathrm{sum}} \boldsymbol{G}_u^{(k),\mathrm{H}})} \tag{47}$$

where $\boldsymbol{Q}_{\mathrm{sum}} = \mathrm{diag}(\boldsymbol{\mu}_{\mathrm{sum}} + \boldsymbol{\mu}_{\mathrm{mask}}^{(k)}) + \sum_{l=1}^{L} \mu_{\mathrm{bmax},l}^{(k)} \boldsymbol{A}_{\mathrm{bmax},l}^{(k)}$ and the new downlink precoder is then

$$\boldsymbol{P}^{(k)} = \boldsymbol{G}_{\mathrm{u}}^{(k)} \sqrt{\beta}. \tag{48}$$

The per-line sum-power constraint is handled separately, because it couples the subcarriers. The Lagrangian multipliers for the per-line sum-power constraints $\boldsymbol{\mu}_{\mathrm{sum}}$ are derived using a gradient projection approach similar to the one used for ZF precoding, Eq. (31) and (32). For weighted MMSE precoding, the gradient step is given by

$$\boldsymbol{\mu}_{\mathrm{sum\ grad}}^{[t]} = \boldsymbol{\mu}_{\mathrm{sum}}^{[t]} + \alpha_\mu \left(\sum_{k=1}^{K} \mathrm{diag}\left(\boldsymbol{P}^{(k)} \boldsymbol{P}^{(k),\mathrm{H}}\right) - \boldsymbol{p}_{\mathrm{sum}}\right) \tag{49}$$

With multiple iterations of the uplink-downlink duality, the algorithm converges. A possible implementation for weighted MMSE precoding in G.fast is shown in Algorithm 2.

V. Simulations

The target reach of G.fast has been increased from 250m to 400m [1], which is covered in the shown simulations. The channel model from [13] with a 0.6mm PE loop according to

[14] is used to create channels with different line lengths from the DP to the customer premises (non-colocated scenario). The cable bundles have 24 pairs up to 400 m line length. Background noise is AWGN with −140 dBm/Hz power spectral density. The transmit PSD from the G.fast standard [9] for the 212 MHz profile with 4 dBm per-line sum-power is used.

A. Performance Evaluation

Performance is reported in rate vs. reach curves. The rate-reach curves show average data rates at a specific line length range for various channel realizations. The method allows a performance comparison that depends on the line length.

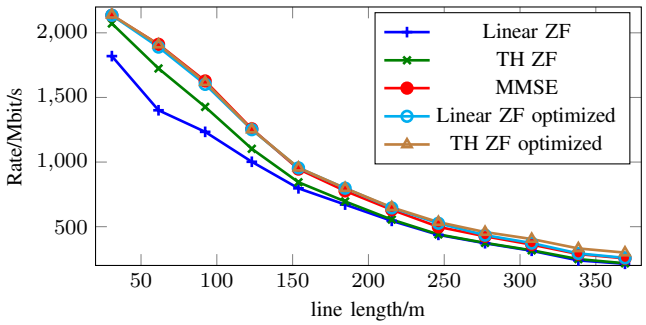

Fig. 3. Rate vs. reach curves for different precoding methods

TH ZF precoding is expected to achieve higher performance than linear ZF precoding because nonlinear precoding does not face the issue of power increase. MMSE precoding is expected to perform better than ZF precoding because of the joint precoder and spectrum optimization. Fig. 3 shows the rate vs. reach curves for the algorithms proposed algorithms.

With column norm scaling using Eq. (23) and (24), the results show TH precoding outperforming ZF precoding with 12% rate increase at 100 m and MMSE precoding giving a gain of 24% over linear ZF precoding.This is illustrated in Fig. 4 where the rate distribution at this line length is shown in terms of a cumulative density function.

Applying the spectrum optimization methods from Sec. III to both ZF methods brings all three methods into the same performance region. Only on very long lines, optimized TH precoding outperforms the other methods. This result indicates that selecting active lines per carrier and performing spectrum optimization with fixed precoders (33) works almost as good as precoder/spectrum optimization using MMSE precoding.

VI. Conclusion

The paper shows an algorithmic solution for MMSE precoding which is feasible for G.fast. A comparison with linear ZF and TH ZF precoding is presented, which demonstrates the performance gain of TH precoding over linear ZF precoding and the gain of MMSE precoding over ZF precoding.

Taking into account spectrum optimization changes the picture. The presented spectrum optimization algorithm for linear and nonlinear ZF precoding brings all three methods to a similar performance under the shown conditions of 212

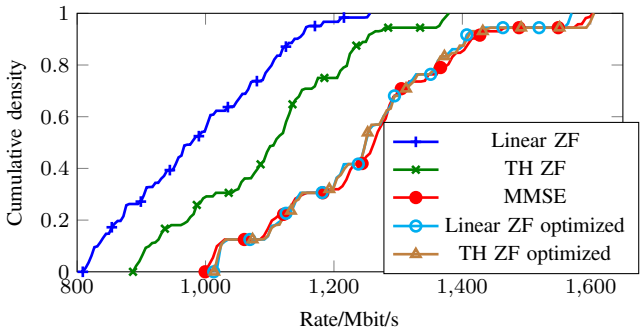

Fig. 4. Rate cumulative density function around 100m line length

MHz G.fast for G.fast-only FTTdp. However, Alg. 2 requires a very high algorithmic complexity to derive the precoders and TH precoding needs a more complex precoder hardware structure than linear methods.

Optimization of the TH encoding order is a topic of further research and a performance comparison in G.fast/VDSL coexistence scenarios may change the results.

References

[1] ITU-T Rec. G.9701, "Fast Access to Subscriber Terminals - Physical layer specification," 2015.

[2] F. C. Muller, C. Lu, P.-E. Eriksson, S. Host, and A. Klautau, "Optimizing power normalization for g. fast linear precoder by linear programming," in *2014 IEEE International Conference on Communications (ICC)*. IEEE, 2014, pp. 4160–4165.

[3] M. Timmers, M. Guenach, C. Nuzman, and J. Maes, "G.fast: evolving the copper access network," *IEEE Communications Magazine*, 2013.

[4] R. Strobel, M. Joham, and W. Utschick, "Achievable rates with implementation limitations for g.fast-based hybrid Copper/Fiber networks," in *IEEE ICC 2015 SAC - Access Networks and Systems (ICC'15 (01) SAC 7-ANS)*, London, United Kingdom, Jun. 2015.

[5] S. S. Christensen, R. Agarwal, E. Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 4792–4799, 2008.

[6] G. Ginis and J. Cioffi, "Vectored transmission for digital subscriber line systems," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 5, pp. 1085–1104, 2002.

[7] R. Hunger, W. Utschick, D. A. Schmidt, and M. Joham, "Alternating optimization for MMSE broadcast precoding," in *2006 IEEE International Conference on Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings.*, vol. 4. IEEE, 2006, pp. IV–IV.

[8] T. Lee and M. Peeters, "G.fast: Further simulation results for FEC evaluation," 2012, Contribution ITU-T SG15/Q4a 2012-07-4A-033.

[9] ITU-T Rec. G.9700, "Fast access to subscriber terminals (FAST) - Power spectral density specification," 2013.

[10] J. Maes and C. Nuzman, "Energy efficient discontinuous operation in vectored g. fast," in *IEEE International Conference on Communications (ICC)*. IEEE, 2014, pp. 3854–3858.

[11] A. Gründinger, M. Joham, J. P. González-Coma, L. Castedo, and W. Utschick, "Average Sum MSE Minimization in the Multi-User Downlink With Multiple Power Constraints," in *48th Asilomar Conference on Signals, Systems and Computers*, Nov 2014.

[12] T. E. Bogale and L. Vandendorpe, "Sum MSE optimization for downlink multiuser MIMO systems with per antenna power constraint: Downlink-uplink duality approach," in *2011 IEEE 22nd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*. IEEE, 2011, pp. 2035–2039.

[13] R. Strobel, R. Stolle, and W. Utschick, "Wideband Modeling of Twisted-Pair Cables for MIMO Applications," in *IEEE Globecom 2013 - Symposium on Selected Areas in Communications (GC13 SAC)*, 2013.

[14] Broadband Forum, "WT-285 Cable Models for Physical Layer Testing of G.fast Access Network," 2014.