KU LEUVEN

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

Citation/Reference	Lanneer W., Moonen M., Tsiaflakis P., and Maes J. **Linear and Nonlinear Precoding Based Dynamic Spectrum Management for Downstream Vectored G.fast Transmission** Published in IEEE Transactions on Communications, vol. 65, no. 3, pp. 1247-1259, March 2017.
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/TCOMM.2016.2641952
Journal homepage	http://www.comsoc.org/TC
Author contact	wouter.lanneer@esat.kuleuven.be *+ 32 16 32 79 75*
IR	https://lirias.kuleuven.be/handle/123456789/566899

(article begins on next page)

Linear and Nonlinear Precoding Based Dynamic Spectrum Management for Downstream Vectored G.fast Transmission

Wouter Lanneer, *Student Member, IEEE,* Paschalis Tsiaflakis, *Member, IEEE,*
Jochen Maes, *Senior Member, IEEE,* and Marc Moonen, *Fellow, IEEE*

***Abstract*—In the G.fast digital subscriber line (DSL) frequency range (up to 106 or 212 MHz), where crosstalk channels may even become larger than direct channels, linear zero-forcing (ZF) precoding is no longer near-optimal for downstream (DS) vectored transmission. To improve performance, we develop a novel low-complexity algorithm for both linear and nonlinear precoding based dynamic spectrum management (DSM) that maximizes the weighted sum-rate under realistic per-line total power and per-tone spectral mask constraints. It applies to DS scenarios with a single copper line at each customer site [i.e. broadcast channel (BC) scenarios], as well as to DS scenarios with multiple copper lines at some or all customer sites (i.e. the so-called multiple-input-multiple-output (MIMO)-BC scenarios). The algorithm alternates between precoder and equalizer optimization, where the former relies on a Lagrange multiplier based transformation of the DS dual decomposition approach formulation into its dual upstream (US) formulation, together with a low-complexity iterative fixed-point formula to solve the resulting US problem. Simulations with measured G.fast channel data of a very high crosstalk cable binder are provided revealing a significantly improved performance of this algorithm over ZF techniques for various scenarios, and in addition, a faster convergence rate compared to the state-of-the-art WMMSE algorithm.**

***Index Terms*—DSL, G.fast, dynamic spectrum management (DSM), optimal spectrum balancing (OSB), precoding, broadcast channel (BC), MIMO-BC**

I. Introduction

In DSL systems, the main source of performance degradation is traditionally crosstalk interference between different copper lines in the same cable bundle. The crosstalk interference problem has been tackled with the introduction of vectoring (also known as dynamic spectrum management (DSM) level 3) for VDSL2 [2]. Vectoring removes the crosstalk by employing signal coordination solely at the access node, resulting in single-sided precoding techniques for downstream (DS) transmission which corresponds to a so-called broadcast channel (BC) scenario. Since in the VDSL2 frequency range (below 30 MHz) the channel matrix typically has a diagonally dominant structure, the linear zero-forcing (ZF) precoder is near-optimal for DS vectored transmission [3].

However G.fast, a new standard approved by the International Telecommunication Union (ITU), already exploits a much broader spectrum up to 106 MHz, while a 212 MHz profile is under definition [4]. At these high frequencies, the diagonally dominant structure of the channel matrix is no longer valid, and crosstalk channels may even become larger than direct channels [5]. As a result, the linear ZF precoder is not near-optimal anymore as it will suffer from increased per-line transmit power penalties due to the large precompensation signals needed for crosstalk cancellation [6]. This makes nonlinear precoding (NLP) in particular interesting for G.fast, as NLP sequentially encodes the user transmit signals in order to “pre-subtract" the crosstalk from previously encoded users without transmit power penalties.

At the same time, the availability of multiple copper lines for data transmission to one end-user, as a result of the historical installation of multiple phone lines at most customer sites, offers the opportunity to significantly improve performance by using bonding and phantom mode transmission [7]. Bonding is used to combine multiple copper lines into one big data pipe to the end-user. On top of that, phantom mode transmission can be used to create, for instance, a third (phantom or virtual) channel over two physical copper lines by exploiting the difference between the common mode voltages of the two lines[1]. Moreover, these techniques allow for receiver signal coordination (equalization) at the customer sites, in addition to the transmitter signal coordination (precoding) at the access node, which corresponds to a so-called multiple-input-multiple-output (MIMO)-BC scenario.

To improve the performance of DS vectored G.fast transmission with multiple lines available to each user, we study linear precoding (LP) and NLP based DSM to maximize the achievable weighted sum-rate under realistic per-line total

Copyright (c) 2016 IEEE. Personal use of this material is permitted, but republication/redistribution requires IEEE permission.

This research work was carried out at the ESAT Laboratory of KU Leuven in the frame of VLAIO O&O Project nr. HBC.2016.0055 ‘The 5th Generation Broadband’, Research Project FWO nr. G.0912.13 ‘Cross-layer optimization with real-time adaptive dynamic spectrum management for 4th generation broadband access networks’, the Interuniversity Attractive Poles Programme initiated by the Belgian Science Policy Office: IUAP P7/23 ‘Belgian network on stochastic modeling analysis design and optimization of communication systems’, 2012-2017, and the KU Leuven Research Council CoE PFV/10/002 (OPTEC). The scientific responsibility is assumed by its authors. This paper was preliminarily presented at the IEEE Global Communications Conference, San Diego, USA, in December 2015 [1].

W. Lanneer and M. Moonen are with the KU Leuven, Dept. of Electrical Engineering (ESAT), STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, Leuven, Belgium (e-mail: {wouter.lanneer, marc.moonen}@esat.kuleuven.be).

P. Tsiaflakis and J. Maes are with the Copper team of Nokia Bell Labs, Antwerp, Belgium (e-mail: {paschalis.tsiaflakis, jochen.maes}@nokia-bell-labs.com).

[1]Note that in this paper the term ‘line’ will be used for physical DSL lines as well as for virtual phantom modes.

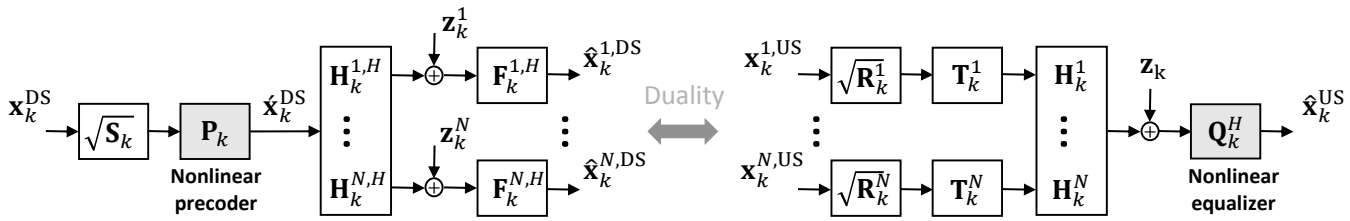

Fig. 1. DS (left-hand side) MIMO-BC and its dual US (right-hand side) MIMO-MAC system model for transmission on tone k. The nonlinear parts are bypassed for linear transmission. Switching between both domains is done by interchanging the role of precoders and equalizers and hermitian transposing the channel matrix.

power and per-tone spectral mask constraints. These DS rate maximization problems are typically non-convex and therefore difficult to solve in a globally optimal manner.

In a wireless communication and single carrier context, DS rate maximization problems under per-line power constraints have been shown to be equivalent with their more easily solvable dual upstream (US) problems under a sum-power constraint by incorporating an unknown noise covariance matrix [8]. This leads to a minmax optimization of the weighted sum-rate function over transmit covariance matrices and the unknown noise covariance matrix.

A simpler, but equivalent, approach where only the transmit covariance matrices have to be optimized, is provided by the broadcast channel optimal spectrum balancing (BC-OSB) [9]. BC-OSB has been proposed for NLP based DSM in a DSL BC scenario under per-line total power constraints. It consists of an iterative dual decomposition approach where in each iteration the Lagrange dual function is transformed based on the Lagrange multipliers before exploiting US-DS duality. The resulting dual US Lagrangian is then maximized by per-tone discrete exhaustive searches, which however have a large computational complexity. Thus although the BC-OSB is globally optimal, it is only tractable for small scenarios, for example with up to four users.

A different approach for solving DS rate maximization problems which also applies to MIMO-BC scenarios is to adopt a transformation into an equivalent weighted minimum mean square error (WMMSE) minimization problem [10]. The WMMSE problem can then be solved by iteratively updating the weight matrices, the MMSE precoders and the MMSE equalizers, which provably converges to a locally optimal stationary point of both problems. Such an algorithm for a LP based DSL system with per-line total power constraints is the discrete multi-tone (DMT)-WMMSE [11], [12]. However, these WMMSE-based algorithms typically suffer from slow convergence rates.

In this paper, we focus on extending the BC-OSB approach for MIMO-BC scenarios employing both LP and NLP, and also on including per-tone spectral mask constraints on top of the per-line total power constraints. This is enabled by alternating optimization of the precoders and equalizers, where for the former, we propose a low-complexity method to maximize the US Lagrangian by means of an iterative fixed-point formula. The resulting algorithm will be referred to as the broadcast channel distributed spectrum balancing (BC-DSB) algorithm, as it is similar to the DSB algorithm of [13] for interference channels. Despite that BC-DSB is only provably convergent under certain conditions, in our simulations we have observed that it always converges to a solution outperforming those of ZF techniques. In addition, for BC scenarios BC-DSB has a significantly reduced computational complexity compared to BC-OSB such that also larger scenarios can be simulated, and will be shown to exhibit faster convergence than the state-of-the-art DMT-WMMSE algorithm.

This paper is organized as follows. Section II introduces the MIMO-BC system model and corresponding rate maximization problem. Section III and IV present BC-DSB for NLP and LP respectively. Section V discusses convergence properties and computational complexity of BC-DSB. Section VI compares the performance of BC-DSB with ZF techniques, BC-OSB and DMT-WMMSE in a G.fast context. Finally, section VII concludes the paper.

II. System Model and Problem Statement

We consider DS transmission in a MIMO-BC scenario with N interfering users. Each user n has A_n G.fast lines (see left-hand side of Fig. 1), meaning there are a total of $L = \sum_n A_n$ lines. We refer to the i-th line of user n as line (n, i). The transmission uses DMT modulation with K subcarriers or tones spaced by Δ_f Hz. Besides full transmitter signal coordination at the access node, there is also receiver signal coordination possible at the customer sites. Both LP and NLP is considered. Assuming perfect DMT synchronization, the linear part of the DS transmission for user n on tone k can be modeled as

$$\mathbf{y}_k^{n,\mathrm{DS}} = \mathbf{H}_k^{n,H} \acute{\mathbf{x}}_k^{\mathrm{DS}} + \mathbf{z}_k^n, \tag{1}$$

where $\acute{\mathbf{x}}_k^{\mathrm{DS}}$ is the L-vector of transmit signals of the access node on tone k, having a covariance matrix defined as $\mathbf{C}_k \triangleq \mathbb{E}\{\acute{\mathbf{x}}_k^{\mathrm{DS}} \acute{\mathbf{x}}_k^{\mathrm{DS},H}\}$. $\mathbf{y}_k^{n,\mathrm{DS}}$ is the A_n-vector of received signals of user n on tone k. $\mathbf{z}_k^n$ is the A_n-vector of additive noise signals of user n on tone k, which we assume to be uncorrelated and pre-whitened, i.e., $\mathbb{E}\{\mathbf{z}_k^n \mathbf{z}_k^{n,H}\} = I_{A_n}$. $\mathbf{H}_k^{n,H}$ is the $A_n \times L$ DS channel matrix between the access node and user n. The total $L \times L$ channel matrix is $\mathbf{H}_k^H \triangleq [\mathbf{H}_k^1, \cdots, \mathbf{H}_k^N]^H$ on tone k. The diagonal elements of $\mathbf{H}_k^H$ are the direct channels, and the off-diagonal elements are the crosstalk channels. We highlight that although $\mathbf{H}_k^H$ typically has a diagonally dominant structure (i.e. $|h_k^{l,j}| \ll |h_k^{l,l}|, j \neq l$) below 30 MHz, recent measurements show that this structure is not valid anymore for higher frequencies of G.fast where the crosstalk channels may even become larger than the direct channels [5]. Furthermore, we

assume perfect channel state information is available. In DSL systems the channel characteristics vary slowly with time such that the vectoring control entity at the access node is indeed able to estimate and track the channel characteristics by sending pilot symbols interleaved with the data symbols [4].

Without loss of generality, we specify the transmit signal vector as

$$\acute{\mathbf{x}}_k^{\mathrm{DS}} = \mathbf{P}_k \sqrt{\mathbf{S}_k}\, \mathbf{x}_k^{\mathrm{DS}}, \quad \forall k. \tag{2}$$

In this equation $\mathbf{x}_k^{\mathrm{DS}} \triangleq [\mathbf{x}_k^{1,\mathrm{DS}}, \cdots, \mathbf{x}_k^{N,\mathrm{DS}}]^T$ denotes the L-vector of symbols intended for all users on tone k with covariance matrix $\mathbb{E}\{\mathbf{x}_k^{\mathrm{DS}} \mathbf{x}_k^{\mathrm{DS},H}\} = \mathbf{I}_L$. $\mathbf{P}_k \triangleq [\mathbf{P}_k^1, \cdots, \mathbf{P}_k^N]$ is the precoder matrix on tone k, where $\mathbf{P}_k^n \triangleq [\mathbf{p}_k^{(n,1)}, \cdots, \mathbf{p}_k^{(n,A_n)}]$ contains the precoder vectors of user n, with $\mathbf{p}_k^{(n,i)}$ being an L-vector with unity l_2-norm and representing the linear processing corresponding to the data symbol of line (n,i). $\sqrt{\mathbf{S}_k} \triangleq \mathrm{blockdiag}\left\{\sqrt{\mathbf{S}_k^1}, \cdots, \sqrt{\mathbf{S}_k^N}\right\}$ is the diagonal scaling matrix on tone k, where $\mathbf{S}_k^n \triangleq \mathrm{diag}\left\{s_k^{(n,1)}, \cdots, s_k^{(n,A_n)}\right\}$ contains the symbol transmit powers of user n. Combining (2) and the definition of transmit covariance matrix $\mathbf{C}_k$ results in

$$\mathbf{C}_k = \mathbf{P}_k \mathbf{S}_k \mathbf{P}_k^H. \tag{3}$$

Expressions (2) and (3) can be used both for LP and NLP as will be shown next.

A. Nonlinear Precoding

For NLP, we use the theoretical concept of dirty paper coding (DPC) [14] which is a successive interference subtraction technique that is sum-capacity-achieving for the MIMO-BC scenario [15]. It can be seen as the dual of the sum-capacity-achieving minimum mean squared error generalized decision feedback equalizer (MMSE-GDFE) for the US so-called multiple access channel (MAC) scenario [16]. DPC can be implemented in practice with Tomlinsom-Harashima precoding (THP) [17], [18] which is a well-known technique in the DSL community. However, note that there is a small performance gap between the DPC concept and the THP implementation due to the necessary modulo operations resulting in some power penalties [19].

To implement DPC, a subtraction or encoding order of the user is required. Without loss of generality, we assume that the encoding order is given by the user index. This means that user 1, which is encoded first, sees all other users as interference; while user N, which is encoded last, has the interference from all other users subtracted. For this NLP based transmission, the capacity for user n on tone k is [15]

$$c_k^{n,\mathrm{DS}} = \log_2 \frac{\left| I_{A_n} + \mathbf{H}_k^{n,H} \left(\sum_{m \geq n} \mathbf{C}_k^m \right) \mathbf{H}_k^n \right|}{\left| I_{A_n} + \mathbf{H}_k^{n,H} \left(\sum_{m > n} \mathbf{C}_k^m \right) \mathbf{H}_k^n \right|}, \tag{4}$$

where $\mathbf{C}_k^n = \mathbf{P}_k^n \mathbf{S}_k^n \mathbf{P}_k^{n,H}$ and $\mathbf{C}_k = \sum_n \mathbf{C}_k^n$.

To investigate the encoding and decoding process of each user in more detail, we introduce a block diagonal equalizer matrix $\mathbf{F}_k^H \triangleq \mathrm{blockdiag}\left\{\mathbf{F}_k^{1,H}, \cdots, \mathbf{F}_k^{N,H}\right\}$. This yields an estimated data signal vector of user n on tone k given by

$$\hat{\mathbf{x}}_k^{n,\mathrm{DS}} = \mathbf{F}_k^{n,H} \mathbf{y}_k^n, \tag{5}$$

and a signal-to-interference-and-noise-ratio (SNR) of line (n,i) on tone k without intra-user successive interference subtraction defined as

$$\mathrm{SNR}_{k,\mathrm{DS}}^{(n,i)} = \frac{s_k^{(n,i)} |\mathbf{f}_k^{(n,i),H} \mathbf{H}_k^{n,H} \mathbf{p}_k^{(n,i)}|^2}{1 + \sum_{(m,j) \succ (n,i)} s_k^{(m,j)} |\mathbf{f}_k^{(n,i),H} \mathbf{H}_k^{n,H} \mathbf{p}_k^{(m,j)}|^2}, \tag{6}$$

where $(m,j) \succ (n,i)$ denotes the condition that either $m > n$, or $m = n$ and $j \neq i$, and $\mathbf{f}_k^{(n,i)}$ is an A_n-vector having unity l_2-norm and representing the equalizer for line (n,i) on tone k. Then, the achievable bit rate for user n on tone k can be expressed as

$$b_k^{n,\mathrm{DS}} = \sum_i \log_2 \left(1 + \frac{1}{\Gamma} \mathrm{SNR}_{k,\mathrm{DS}}^{(n,i)} \right), \tag{7}$$

where Γ denotes the capacity gap for practical QAM implementations, and is a function of the desired BER, coding gain, and noise margin [20]. Throughout the paper, we always refer to the case of a zero capacity gap ($\Gamma = 0$ dB) when making a statement about "global optimality" or "capacity-achieving" for the case of nonlinear encoding and decoding. However, we remark that even for nonzero capacity gap ($\Gamma \neq 0$ dB) BC-DSB can exploit the DPC transmission structure to obtain a set of achievable user rates (not necessarily capacity-achieving) which outperform ZF schemes, as will be shown in section VI. Note that users have a total data rate in bits per second defined as $R_n = f_s \sum_k b_k^{n,\mathrm{DS}}$, where f_s is the DMT symbol rate.

Furthermore, when the SNR (6) is maximized using MMSE equalizers $\{\mathbf{F}_k^{n,H}\}$ and diagonal per-user MSE matrices $\{\mathbf{E}_k^{n,\mathrm{DS}} \triangleq \mathbb{E}[(\hat{\mathbf{x}}_k^{n,\mathrm{DS}} - \mathbf{x}_k^{n,\mathrm{DS}})(\hat{\mathbf{x}}_k^{n,\mathrm{DS}} - \mathbf{x}_k^{n,\mathrm{DS}})^H], \forall n\}$ are enforced, it is known that (see [21]) the achievable bit rates (7) attain capacity equal to (4). Enforcing diagonal per-user MSE matrices (which is explained in section III-C) results in intra-user interference-free transmission meaning $\mathbf{F}_k^{n,H} \mathbf{H}_k^{n,H} \mathbf{P}_k^n$ is a $A_n \times A_n$ diagonal matrix for each user n. This relaxes the capacity-achieving requirement of applying intra-user successive interference subtraction where all lines are sequentially encoded (like in [16]) to applying only inter-user successive interference subtraction where all users are sequentially encoded, reducing in this way the implementation complexity considerably.

B. Linear Precoding

Since LP requires less implementation complexity than NLP, it offers a valuable alternative. For the LP based transmission, the capacity of user n on tone k is given as

$$c_k^{n,\mathrm{DS}} = \log_2 \frac{\left| I_{A_n} + \mathbf{H}_k^{n,H} \left(\sum_m \mathbf{C}_k^m \right) \mathbf{H}_k^n \right|}{\left| I_{A_n} + \mathbf{H}_k^{n,H} \left(\sum_{m \neq n} \mathbf{C}_k^m \right) \mathbf{H}_k^n \right|}, \tag{8}$$

while the corresponding achievable SNR of line (n,i) is

$$\mathrm{SNR}_{k,\mathrm{DS}}^{(n,i)} = \frac{s_k^{(n,i)} |\mathbf{f}_k^{(n,i),H} \mathbf{H}_k^{n,H} \mathbf{p}_k^{(n,i)}|^2}{1 + \sum_{(m,j) \neq (n,i)} s_k^{(m,j)} |\mathbf{f}_k^{(n,i),H} \mathbf{H}_k^{n,H} \mathbf{p}_k^{(m,j)}|^2}. \tag{9}$$

Also for LP, a diagonal per-user MSE matrix requirement ensures that lines of the same user do not self-interfere [21]. Then, the achievable bit rate for user n and tone k, given by (7), attains capacity (8) for the specific case of MMSE equalizers $\{\mathbf{F}_k^{n,H}\}$ and a zero capacity gap ($\Gamma = 0$ dB). While the algorithm developed in this paper is applicable to both LP and NLP, we mainly focus on the NLP case, and adopt the line SNR definition given in (6) unless stated otherwise.

C. Problem Statement

We focus on finding the precoders $\{\mathbf{P}_k\}$, symbol transmit powers $\{\mathbf{S}_k\}$, and equalizers $\{\mathbf{F}_k^H\}$ that maximize the achievable weighted sum-rate under per-line total power and per-tone spectral mask constraints

$$\begin{aligned}\underset{\{\mathbf{P}_k\},\{\mathbf{S}_k\succeq\mathbf{0}\},\{\mathbf{F}_k^H\}}{\text{maximize}} \quad & \sum_k\sum_n w_n b_k^{n,\mathrm{DS}} \\ \text{s.t.} \quad & \sum_k \left[\mathbf{P}_k\mathbf{S}_k\mathbf{P}_k^H\right]_{(n,i),(n,i)} \le P^{\mathrm{line}}, \quad \forall(n,i) \\ & \left[\mathbf{P}_k\mathbf{S}_k\mathbf{P}_k^H\right]_{(n,i),(n,i)} \le P_k^{\mathrm{mask}}, \quad \forall(n,i),k \end{aligned} \tag{10}$$

where $[\cdot]_{l,l}$ denotes the (l,l)-entry of a matrix, and w_n is the weight for user n. Since the optimal user encoding order is defined by the user weights, we assume w.l.o.g. $w_1 \ge \cdots \ge w_N$ (i.e. the user with the largest weight is to be encoded first [9], [22]). P_k^{mask} is the spectral mask for every line and for tone k, which is typically kept low in G.fast in order not to generate too much interference into other technologies. We remark that when P^{line} is very high compared to the spectral masks, enforcing the latter will automatically lead to the former being satisfied as well. These power constraints are particularly important for G.fast transmission due to the lack of a diagonally dominant structure of the channel matrix. The strong crosstalk then results in large precompensation signals which (especially for LP) may increase the per-line transmit powers, which should therefore be accounted for by means of DSM. Although not explicitly taken into account in (10), BC-DSB as developed in the next sections can be simply extended to comply with the practical constraint on the maximal bit loading (i.e. a bitcap equal to 12 bits for G.fast) as will be shown.

III. BC-DSB-NLP

Optimization problem (10) is non-trivial and has to be solved in an iterative fashion. Albeit the DS weighted sum-rate functions are neither concave nor convex in $\mathbf{C}_k^n$ [15], the optimal precoders $\{\mathbf{P}_k\}$ and transmit powers $\{\mathbf{S}_k\}$ can be calculated in the dual US domain using US-DS duality theory provided that the equalizers $\{\mathbf{F}_k^H\}$ are known (see [1]). On the other hand, the optimal equalizers $\{\mathbf{F}_k^H\}$ can be calculated by a closed-form expression when the precoders $\{\mathbf{P}_k\}$ and transmit powers $\{\mathbf{S}_k\}$ are known. However, since their optimal solutions are a function of one another, the optimal precoders and equalizers cannot be calculated jointly. To overcome this difficulty, we alternate between precoder and transmit power versus equalizer optimization, meaning that the precoders $\{\mathbf{P}_k\}$ and transmit powers $\{\mathbf{S}_k\}$ are optimized for fixed equalizers $\{\mathbf{F}_k^H\}$, and the other way around, $\{\mathbf{F}_k^H\}$ is optimized for fixed $\{\mathbf{P}_k\}$ and $\{\mathbf{S}_k\}$.

A. Precoder and Transmit Power Optimization for Fixed Equalizers

For fixed DS equalizers, we can define the equivalent DS channel vector $\tilde{\mathbf{h}}_k^{(n,i),H}$ and noise signal $\tilde{z}_k^{(n,i)}$ for line (n,i) and tone k as

$$\tilde{\mathbf{h}}_k^{(n,i),H} = \mathbf{f}_k^{(n,i),H}\mathbf{H}_k^{n,H}, \tag{11}$$

$$\tilde{z}_k^{(n,i)} = \mathbf{f}_k^{(n,i),H}\mathbf{z}_k^n. \tag{12}$$

This turns each line (n,i) into a virtual single-line user, creating an equivalent BC scenario with only transmitter coordination. Note that still $\{\mathbb{E}\{|\tilde{z}_k^{(n,i)}|^2\} = 1\}$ due the unity normalized $\{\mathbf{f}_k^{(n,i)}\}$. To simplify notations, we will now use line index $l = \{1,\cdots,L\}$ instead of (n,i).

For this equivalent BC scenario, denoting the total equivalent channel as $\widetilde{\mathbf{H}}_k^{n,H} = [\tilde{\mathbf{h}}_k^1,\cdots,\tilde{\mathbf{h}}_k^L]^H$, the data transmission model (5) on tone k for all users can be simplified to

$$\hat{\mathbf{x}}_k^{\mathrm{DS}} = \widetilde{\mathbf{H}}_k^H\mathbf{P}_k\sqrt{\mathbf{S}_k}\mathbf{x}_k^{\mathrm{DS}} + \tilde{\mathbf{z}}_k, \tag{13}$$

for which (10) reduces to the following DS sub-problem[2]

$$\begin{aligned}\underset{\{\mathbf{P}_k\},\{\mathbf{S}_k\succeq\mathbf{0}\}}{\text{maximize}} \quad & \sum_k\sum_l \tilde{w}_l\tilde{b}_k^{l,\mathrm{DS}} \\ \text{s.t.} \quad & \sum_k \left[\mathbf{P}_k\mathbf{S}_k\mathbf{P}_k^H\right]_{l,l} \le P^{\mathrm{line}}, \quad \forall l \\ & \left[\mathbf{P}_k\mathbf{S}_k\mathbf{P}_k^H\right]_{l,l} \le P_k^{\mathrm{mask}}, \quad \forall l,k \end{aligned} \tag{14}$$

$$\text{with} \quad \tilde{b}_k^{l,\mathrm{DS}} = \log_2\left(1 + \frac{s_k^l|(\tilde{\mathbf{h}}_k^l)^H\mathbf{p}_k^l|^2}{\Gamma(1+\sum_{j>l}s_k^j|(\tilde{\mathbf{h}}_k^l)^H\mathbf{p}_k^j|^2)}\right). \tag{15}$$

It is noted that, although the achievable bit rates (15) correspond to successive interference subtraction between all virtual single-line users in the equivalent BC, intra-user successive interference subtraction in the original MIMO-BC becomes superfluous when enforcing the diagonal MSE requirement (see (32) and (33) in section III-C).

To solve (14) we formulate a dual decomposition approach with standard subgradient based updating of the Lagrange multipliers (similar to for instance [13]), relying on the "zero duality gap"-result for multi-carrier systems[3] [23]. The Lagrangian of (14) in this case is given as

$$\begin{aligned}\mathcal{L}^{\mathrm{DS}}(\mathbf{\Theta},\{\mathbf{\Lambda}_k\},\{\mathbf{P}_k\},\{\mathbf{S}_k\}) = & \\ \sum_k\sum_l \left(w_l\tilde{b}_k^{l,\mathrm{DS}} - s_k^l\mathrm{Tr}\{(\mathbf{\Theta}+\mathbf{\Lambda}_k)\mathbf{p}_k^l(\mathbf{p}_k^l)^H\}\right) & \\ + P^{\mathrm{line}}\mathrm{Tr}\{\mathbf{\Theta}\} + \sum_k P_k^{\mathrm{mask}}\mathrm{Tr}\{\mathbf{\Lambda}_k\}, & \end{aligned} \tag{16}$$

where $\mathbf{\Theta} = \mathrm{diag}\{\theta^1,\cdots,\theta^L\}$ and $\mathbf{\Lambda}_k = \mathrm{diag}\{\lambda_k^1,\cdots,\lambda_k^L\}$ are diagonal matrices containing the Lagrange multipliers corresponding to the per-line total power and spectral masks

[2] The virtual single-user line weight vector $[\tilde{w}_1,\cdots,\tilde{w}_L]^T$ is equal to $[\underbrace{w_1,\cdots,w_1}_{A_1\text{ times}},\cdots,\underbrace{w_N,\cdots,w_N}_{A_N\text{ times}}]^T$.

[3] We remark that, although this "zero duality gap"-result of [23] is not valid for the case with per-tone spectral mask constraints, the dual problem of (10) can be shown to be equivalent to an US dual problem corresponding to an US primal problem with one total sum-power constraint across all lines and tones, and with unknown noise covariance matrices for which [23] is valid (similar to the proof in [9]).

constraints for tone k, respectively. The Lagrange dual function is then defined as an unconstrained maximization of the Lagrangian for a given set of Lagrange multipliers

$$g^{\mathrm{DS}}(\mathbf{\Theta}, \{\mathbf{\Lambda}_k\}) = \underset{\{\mathbf{P}_k\}, \{\mathbf{S}_k \geq \mathbf{0}\}}{\text{maximize}} \mathcal{L}^{\mathrm{DS}}(\mathbf{\Theta}, \{\mathbf{\Lambda}_k\}, \{\mathbf{P}_k\}, \{\mathbf{S}_k\}) \quad (17)$$

which can be decoupled into K per-tone independent sub-problems. The DS dual optimization problem is

$$\underset{\mathbf{\Theta} \geq \mathbf{0}, \{\mathbf{\Lambda}_k \geq \mathbf{0}\}}{\text{minimize}} g^{\mathrm{DS}}(\mathbf{\Theta}, \{\mathbf{\Lambda}_k\}) \quad (18)$$

As the achievable bit rates (15) are neither convex nor concave in $\mathbf{C}_k^n$ [15], (18) is a difficult problem to solve. To overcome this difficulty, we use US-DS duality theory, which states that the same set of bit rates can be achieved in the dual US channel under the same total transmit power constraint across all lines and tones [15]. However, this also implies that US-DS duality transformations do not preserve the more realistic per-line total power and spectral mask constraints when converting directly an US solution of (18) to a DS solution. Therefore, we first transform the Lagrangian (16) into an equivalent version with a virtual sum power constraint before exploiting US-DS duality. In particular, this transformation consists in incorporating a virtual precoding matrix based on the Lagrange multipliers and rescaling the equivalent channel matrix for each tone [9]. The transmission model of tone k is then given as

$$\hat{\mathbf{x}}_k^{\mathrm{DS}} = \underbrace{\widetilde{\mathbf{H}}_k^H (\mathbf{\Theta} + \mathbf{\Lambda}_k)^{-1/2}}_{\widetilde{\mathbf{H}}_k'^H} \mathbf{P}_k' \sqrt{\mathbf{S}_k} \mathbf{x}_k^{\mathrm{DS}} + \tilde{\mathbf{z}}_k, \quad (19)$$

where $\mathbf{P}_k' = (\mathbf{\Theta} + \mathbf{\Lambda}_k)^{1/2} \mathbf{P}_k$. Now, the Lagrange dual function (17) is rewritten as

$$g^{\mathrm{DS}}(\mathbf{\Theta}, \{\mathbf{\Lambda}_k\}) = \underset{\{\mathbf{P}_k'\}, \{\mathbf{S}_k \geq 0\}}{\text{maximize}} \mathcal{L}'^{\mathrm{DS}}(\mathbf{\Theta}, \{\mathbf{\Lambda}_k\}, \{\mathbf{P}_k'\}, \{\mathbf{S}_k\}), \quad (20)$$

where the transformed Lagrangian is

$$\mathcal{L}'^{\mathrm{DS}}(\mathbf{\Theta}, \{\mathbf{\Lambda}_k\}, \{\mathbf{P}_k'\}, \{\mathbf{S}_k\}) = \sum_k \sum_l \left[w_l \log_2 \left(1 + \frac{s_k^l |(\tilde{\mathbf{h}}_k'^l)^H \mathbf{p}_k'^l|^2}{\Gamma(1 + \sum_{j>l} s_k^j |(\tilde{\mathbf{h}}_k'^l)^H \mathbf{p}_k'^j|^2)} \right) - s_k^l \mathrm{Tr}\{\mathbf{p}_k'^l (\mathbf{p}_k'^l)^H\} \right] + P_{\mathrm{v}}, \quad (21)$$

and where $\tilde{\mathbf{h}}_k'^l = (\mathbf{\Theta} + \mathbf{\Lambda}_k)^{-1/2} \tilde{\mathbf{h}}_k^l$ and $\mathbf{p}_k'^l = (\mathbf{\Theta} + \mathbf{\Lambda}_k)^{1/2} \mathbf{p}_k^l$ are denoting the transformed equivalent channel and precoder vector for line l, respectively. Furthermore, for a given set of Lagrange multipliers, the per-line power and spectral mask constraints appear as a virtual sum power constraint $P_{\mathrm{v}} = P^{\mathrm{line}} \mathrm{Tr}\{\mathbf{\Theta}\} + \sum_k P_k^{\mathrm{mask}} \mathrm{Tr}\{\mathbf{\Lambda}_k\}$. This transformation hides the Lagrange multipliers into the equivalent channel $\{\tilde{\mathbf{h}}_k'^l\}$ and precoder $\{\mathbf{p}_k'^l\}$ vectors. As a result, the transformed Lagrangian (21) can be interpreted as the Lagrangian of a DS problem with a virtual sum power constraint and with its virtual Lagrange multiplier equal to one. Therefore, we can now apply US-DS duality theory to transform (20) into an equivalent US Lagrange dual function for a fixed set of Lagrange multipliers $\{\mathbf{\Theta} + \mathbf{\Lambda}_k\}$ and under the same virtual sum power constraint.

In order to use this US-DS duality, we first introduce the dual US so-called MIMO-MAC system of (5) which is obtained by switching the role of precoders and equalizers and by Hermitian transposing the channel matrix (see right-hand side of Fig. 1). The fixed DS equalizers now correspond to fixed US precoders, simplifying the scenario to a MAC with only receiver coordination between the L single-line virtual users. Hence, the dual US transmission model of (19) is given as[4]

$$\hat{\mathbf{x}}_k^{\mathrm{US}} = \widetilde{\mathbf{Q}}_k^H \widetilde{\mathbf{H}}_k \sqrt{\widetilde{\mathbf{R}}_k} \mathbf{x}_k^{\mathrm{US}} + \widetilde{\mathbf{Q}}_k^H \underbrace{(\mathbf{\Theta} + \mathbf{\Lambda}_k)^{1/2} \mathbf{z}_k}_{\mathbf{z}_k'}. \quad (22)$$

We observe that this model has a noise covariance matrix equal to $\mathbb{E}\{\mathbf{z}_k' (\mathbf{z}_k')^H\} = (\mathbf{\Theta} + \mathbf{\Lambda}_k)$, as in [8]. We remark that this formulation is preferred over scaling the channel and equalizer matrices since it is numerically more robust whenever some of the Lagrange multipliers converge to zero. $\sqrt{\widetilde{\mathbf{R}}_k} \triangleq \mathrm{diag}\{\sqrt{\tilde{r}_k^1}, \cdots, \sqrt{\tilde{r}_k^L}\}$ is the diagonal scaling matrix, where $\tilde{r}_k^l$ represents the US transmit power on line l for tone k. $\widetilde{\mathbf{Q}}_k^H$ is the $L \times L$ US equalizer matrix on tone k. Assuming the theoretically optimal receiver with successive interference cancellation, which is the MMSE-GDFE [16], the US achievable bit rate of line l on tone k is

$$\tilde{b}_k^{l,\mathrm{US}} = \log_2 \left(1 + \frac{1}{\Gamma} \tilde{r}_k^l (\tilde{\mathbf{h}}_k^l)^H \left(\mathbf{\Theta} + \mathbf{\Lambda}_k + \sum_{j<l} \tilde{r}_k^j \tilde{\mathbf{h}}_k^j (\tilde{\mathbf{h}}_k^j)^H \right)^{-1} \tilde{\mathbf{h}}_k^l \right), \quad (23)$$

where the decoding order is the reverse of the encoding order (i.e. virtual single-line user 1 is decoded last, virtual single-line user L is decoded first) as US-DS duality dictates. The US bit rate functions are only dependent of the US transmit powers $\{\tilde{r}_k^l\}$ (and are independent of the US equalizers $\{\widetilde{\mathbf{Q}}_k^H\}$), and moreover, are concave in $\{\tilde{r}_k^l\}$ [15], meaning that the globally optimal power allocation in the dual US system can be efficiently computed.

Based on this US transmission model (22), we define the equivalent US Lagrange dual function of (20) as

$$g'^{\mathrm{US}}(\mathbf{\Theta}, \{\mathbf{\Lambda}_k\}) = \underset{\{\mathbf{R}_k \geq \mathbf{0}\}}{\text{maximize}} \mathcal{L}'^{\mathrm{US}}(\mathbf{\Theta}, \{\mathbf{\Lambda}_k\}, \{\mathbf{R}_k\}), \quad (24)$$

where the Lagrangian is

$$\mathcal{L}'^{\mathrm{US}}(\mathbf{\Theta}, \{\mathbf{\Lambda}_k\}, \{\mathbf{R}_k\}) = \sum_k \sum_l \left[w_l \tilde{b}_k^{l,\mathrm{US}} - \tilde{r}_k^l \right] + P_{\mathrm{v}}. \quad (25)$$

This means that for a given set of Lagrange multipliers $\{\mathbf{\Theta} + \mathbf{\Lambda}_k\}$, the original DS Lagrange dual function (17) of sub-problem (14) can be solved by solving the equivalent US function (24), resulting in the optimal US transmit powers and MMSE equalizers, and that then the corresponding DS transmit powers and precoders can be obtained using US-DS duality transformations (as will be explained in the next

[4] In this model, $\widetilde{\mathbf{R}}_k$ and $\widetilde{\mathbf{Q}}_k^H$ are used for the US transmit powers and equalizer on tone k, as later a unity rotation will be applied to the US precoders and equalizers when enforcing the diagonal MSE requirement resulting in the final transmit powers $\{\mathbf{R}_k\}$ and equalizers $\{\mathbf{Q}_k^H\}$ (see (32) and (33)).

sections). Hence, updating the Lagrange multipliers such that the DS per-line total power and spectral mask constraints are enforced, will eventually solve (14) optimally. I.e., it will result in optimal DS transmit powers and precoders for a fixed set of DS equalizers.

B. Maximizing the Lagrangian of (24)

In order to maximize the Lagrangian of (24), we propose a low-complexity iterative method, based on a fixed-point reformulation of the KKT stationary condition of (24) (similar to [13], [24]) which is both necessary and sufficient for NLP. Thus, extracting $\tilde{r}_k^l$ to one side of equation

$$\sum_j \frac{\partial w_j \tilde{b}_k^{j,\mathrm{US}}}{\partial \tilde{r}_k^l} - 1 = 0, \quad \forall l, k, \tag{26}$$

and enforcing positivity results in a fixed-point US transmit power update formula for line l and tone k given as

$$\tilde{r}_k^l = \left[\frac{w_l/\log(2)}{1 + \sum_{j>l} \frac{(w_j \tilde{r}_k^j)/\log(2)}{\Gamma + \tilde{r}_k^j (\tilde{\mathbf{h}}_k^j)^H (\mathbf{X}_k^j)^{-1} \tilde{\mathbf{h}}_k^j} |(\tilde{\mathbf{h}}_k^j)^H (\mathbf{X}_k^j)^{-1} \tilde{\mathbf{h}}_k^l|^2} - \frac{\Gamma}{(\tilde{\mathbf{h}}_k^l)^H (\mathbf{X}_k^l)^{-1} \tilde{\mathbf{h}}_k^l} \right]^+ \tag{27}$$

where $[x]^+ \triangleq \max(x, 0)$ and

$$\mathbf{X}_k^l = \left(\mathbf{\Theta} + \mathbf{\Lambda}_k + \sum_{j<l} \tilde{r}_k^j \tilde{\mathbf{h}}_k^j (\tilde{\mathbf{h}}_k^j)^H \right). \tag{28}$$

We observe that formula (27) is equivalent to the formula in [24] with the noise covariance matrix replaced by $(\mathbf{\Theta} + \mathbf{\Lambda}_k)$. Formula (27) with iterative Gauss-Seidel (sequential) updates over all lines is guaranteed to converge to a globally optimal stationary point, since it can be shown to be equivalent to a series of closed-form solutions of convex approximations of (24) (see [25] for more details). Moreover, this type of fixed-point iterations has been demonstrated in [13], [24] to work well with low complexity, and converges to good accuracy in only a few iterations. Furthermore, we highlight that (27) corresponds to a continuous transmit power and bit loading, unlike the BC-OSB algorithm that can only do exhaustive searches over a discrete power or bit loading.

Once the optimal US transmit powers are available, the corresponding MMSE-GDFE equalizer for each user n and tone k is computed as

$$\widetilde{\mathbf{Q}}_k^{n,H} = \sqrt{\widetilde{\mathbf{R}}_k^n} \widetilde{\mathbf{H}}_k^{n,H} \left(\mathbf{M}_k^n \right)^{-1} \tag{29}$$

where

$$\mathbf{M}_k^n = (\mathbf{\Theta} + \mathbf{\Lambda}_k) + \sum_{m \le n} \widetilde{\mathbf{H}}_k^m \widetilde{\mathbf{R}}_k^m \widetilde{\mathbf{H}}_k^{m,H},$$
$$\widetilde{\mathbf{R}}_k^n = \mathrm{diag}\left\{ \tilde{r}_k^{(n,1)}, \cdots, \tilde{r}_k^{(n,A_n)} \right\}. \tag{30}$$

C. US-DS Duality Transformations

The US transmit powers and equalizers can be transformed to DS transmit powers and precoders by combining a per-user US MSE matrix diagonalization step with per-line SNR duality transformations. The former guarantees the optimality of the decoding process in the US system without intra-user successive interference cancellation, by ensuring that lines of the same user do not self-interfere [21]. Note that the SNR duality transformations also result in DS diagonal MSE matrices and, thus, optimality of the DS encoding process.

Clearly, the expression of the US MSE matrix of user n

$$\begin{aligned} \mathbf{E}_k^{n,\mathrm{US}} &= \mathbb{E}[(\hat{\mathbf{x}}_k^{n,\mathrm{US}} - \mathbf{x}_k^{n,\mathrm{US}})(\hat{\mathbf{x}}_k^{n,\mathrm{US}} - \mathbf{x}_k^{n,\mathrm{US}})^H] \\ &= I_{A_n} + \widetilde{\mathbf{Q}}_k^{n,H} \mathbf{M}_k^n \widetilde{\mathbf{Q}}_k^n \\ &\quad - \widetilde{\mathbf{Q}}_k^{n,H} \widetilde{\mathbf{H}}_k^n \sqrt{\widetilde{\mathbf{R}}_k^n} - \left(\widetilde{\mathbf{Q}}_k^{n,H} \widetilde{\mathbf{H}}_k^n \sqrt{\widetilde{\mathbf{R}}_k^n} \right)^H \\ &= I_{A_n} - \widetilde{\mathbf{Q}}_k^{n,H} \widetilde{\mathbf{H}}_k^n \sqrt{\widetilde{\mathbf{R}}_k^n}, \end{aligned} \tag{31}$$

where we used (29) in the third line, shows that $\mathbf{E}_k^{n,\mathrm{US}}$ is diagonalized by updating the precoders and equalizers of the US system for user n and tone k as follows [21]

$$\begin{aligned} \mathbf{T}_k^n &= \mathbf{F}_k^n \sqrt{\widetilde{\mathbf{R}}_k^n} \mathbf{W}_k^n, \\ \mathbf{Q}_k^{n,H} &= \mathbf{W}_k^{n,H} \widetilde{\mathbf{Q}}_k^{n,H}, \end{aligned} \tag{32}$$

where $\mathbf{W}_k^n$ is calculated as the unitary eigenvector matrix of $\widetilde{\mathbf{Q}}_k^{n,H} \widetilde{\mathbf{H}}_k^n \sqrt{\widetilde{\mathbf{R}}_k^n}$. In order to remain consistent with previous notation (see right-hand side of Fig. (1)), we then set

$$\begin{aligned} \mathbf{R}_k^n &\leftarrow \mathrm{diag}\left\{ \|\mathbf{t}_k^{(n,1)}\|_2^2, \cdots, \|\mathbf{t}_k^{(n,A_n)}\|_2^2 \right\} \\ \mathbf{T}_k^n &\leftarrow \mathbf{T}_k^n \mathrm{diag}\left\{ \|\mathbf{t}_k^{(n,1)}\|_2^2, \cdots, \|\mathbf{t}_k^{(n,A_n)}\|_2^2 \right\}^{-1/2} \\ \mathbf{Q}_k^{n,H} &\leftarrow \mathrm{diag}\left\{ \|\mathbf{q}_k^{(n,1)}\|_2^2, \cdots, \|\mathbf{q}_k^{(n,A_n)}\|_2^2 \right\}^{-1/2} \mathbf{Q}_k^{n,H}, \forall k, n, \end{aligned} \tag{33}$$

where $\mathbf{T}_k^n$ and $\mathbf{Q}_k^{n,H}$ are the final US $A_n \times A_n$ precoder and $A_n \times L$ equalizer for user n with unity l_2-norm column and row vectors respectively. Since the $\{\mathbf{W}_k^n\}$ are unitary, the diagonalization step neither increases the US virtual sum transmit power nor decreases the weighted sum rate, meaning the maximized value of Lagrangian (25) for a fixed set of Lagrange multipliers remains unmodified.

The SNR duality transformations ensure that each line SNR, and therefore also line bit rate, is preserved between the US and DS system on each tone k. The US SNR for line (n, i) and tone k is

$$\mathrm{SNR}_{k,\mathrm{US}}^{(n,i)} = \frac{r_k^{(n,i)} |\hat{h}_k^{(n,i),(n,i)}|^2}{\mathbf{q}_k^{(n,i),H} (\mathbf{\Theta} + \mathbf{\Lambda}_k) \mathbf{q}_k^{(n,i)} + \sum_{m<n} \sum_j r_k^{(m,j)} |\hat{h}_k^{(n,i),(m,j)}|^2}, \tag{34}$$

where $\widehat{\mathbf{H}}_k \triangleq \mathbf{Q}_k^H \mathbf{H}_k \mathbf{T}_k$ and $\hat{h}_k^{l,j} \triangleq \left[\widehat{\mathbf{H}}_k \right]_{l,j}$. Defining the relation between the DS and US precoding and respectively US and DS equalizers as

$$\mathbf{p}_k^{(n,i)} = \mathbf{q}_k^{(n,i)}, \quad \text{and} \quad \mathbf{f}_k^{(n,i)} = \mathbf{t}_k^{(n,i)} \quad \forall (n,i), k, \tag{35}$$

the DS SNR (6) is equivalent to

$$\mathrm{SNR}_{k,\mathrm{DS}}^{(n,i)} = \frac{s_k^{(n,i)}|\hat{h}_k^{(n,i),(n,i)}|^2}{1+\sum\limits_{m>n}\sum\limits_{j} s_k^{(m,j)}|\hat{h}_k^{(m,j),(n,i)}|^2}. \quad (36)$$

Equating the US and DS SNRs for all lines and tone k, generates a linear system of equations [26]

$$\mathbf{Z}_k \cdot [s_k^1, \cdots, s_k^L]^T = \mathbf{1}, \forall k \quad (37)$$

with

$$\mathbf{Z}_k = \begin{bmatrix} \mathbf{Z}_k^{1,1} & \cdots & \mathbf{Z}_k^{1,N} \\ \vdots & \ddots & \vdots \\ \mathbf{Z}_k^{N,1} & \cdots & \mathbf{Z}_k^{N,N} \end{bmatrix}, \quad (38)$$

where the diagonal and off-diagonal sub-matrices are respectively given as

$$\mathbf{Z}_k^{n,n} = \mathrm{diag}\left\{ \frac{|\hat{h}_k^{(n,1),(n,1)}|^2}{\mathrm{SNR}_{k,\mathrm{US}}^{(n,1)}}, \cdots, \frac{|\hat{h}_k^{(n,A_n),(n,A_n)}|^2}{\mathrm{SNR}_{k,\mathrm{US}}^{(n,A_n)}} \right\}$$

$$\text{and } [\mathbf{Z}_k^{n,p}]_{i,u} = \begin{cases} -|\hat{h}_k^{(p,u),(n,i)}|^2 & \forall (u,i) \text{ if } p > n \\ 0 & \forall (u,i) \text{ if } p < n. \end{cases}$$

This linear system of equations is always solvable yielding valid solutions $s_k^{(n,i)} \geq 0$, because $\mathbf{Z}_k$ has non-negative diagonal and non-positive off-diagonal entries, and $\mathbf{R}_k\mathbf{Z}_k$ is column diagonally dominant for each tone k [26], [27]. We remark that $r_k^{(n,i)} = 0$ results in a zero row and column in $\mathbf{Z}_k$ that have to be removed, together with the respective $s_k^{(n,i)}$ and $r_k^{(n,i)}$, before solving (37), and then $\mathbf{p}_k^{(n,i)} = \mathbf{0}$ and $s_k^{(n,i)} = 0$ are chosen. Furthermore, summing up the rows of (37) multiplied at both sides by $\mathbf{R}_k$, we obtain

$$\sum_{(n,i)} s_k^{(n,i)} \underbrace{\mathbf{q}_k^{(n,i),H}(\mathbf{\Theta}+\mathbf{\Lambda}_k)\mathbf{q}_k^{(n,i)}}_{\|\mathbf{p}_k'^{(n,i)}\|^2} = \sum_{(n,i)} r_k^{(n,i)}, \quad (39)$$

meaning that the same virtual sum power is transmitted in the DS and dual US system.

D. Equalizer Optimization for Fixed Precoders and Transmit Powers

When the transmit powers and precoders of the DS system are fixed, the optimal DS equalizers $\{\mathbf{F}_k^{n,H}\}$ are straightforwardly calculated by the well known MMSE closed-form expression for each user n, ignoring the signals of users $m < n$ and considering the signals of users $m > n$ as background noise, i.e.

$$\widetilde{\mathbf{F}}_k^{n,H} = \sqrt{\mathbf{S}_k^n}\mathbf{P}_k^{n,H}\mathbf{H}_k^n \times \left(\mathbf{I}_{A_n} + \mathbf{H}_k^{n,H}\left(\sum_{m\geq n}\mathbf{P}_k^m\mathbf{S}_k^m\mathbf{P}_k^{m,H}\right)\mathbf{H}_k^n\right)^{-1}, \quad (40)$$

with $\mathbf{F}_k^{n,H} = \mathrm{diag}\left\{\|\tilde{\mathbf{f}}_k^{(n,1)}\|_2, \cdots, \|\tilde{\mathbf{f}}_k^{(n,A_n)}\|_2\right\}^{-1}\widetilde{\mathbf{F}}_k^{n,H}$ having unity normalized rows. Furthermore, although it is not strictly necessary, we then also diagonalize the DS MSE matrices $\{\mathbf{E}_k^{n,\mathrm{DS}}\}$, similar to the diagonalization of $\{\mathbf{E}_k^{n,\mathrm{US}}\}$ in (32) and (33), to speed up convergence.

E. Alternating Optimization

The final algorithm solving the original problem (10) alternates through several steps of precoder and transmit power optimization on the one hand, and equalizer optimization on the other hand, as developed in the previous sections. A complete algorithm description is given in Algorithm 1 and will be referred to as BC-DSB-NLP.

Initialization of BC-DSB consists of choosing the DS equalizers $\{\mathbf{F}_k^{n,H}\}$ as the complex conjugate of the left-singular vectors of the DS channels $\{\mathbf{H}_k^{n,H}\}$. This initialization has shown to work well in the simulations. Furthermore, the predefined step size parameters ϵ_1 and ϵ_2, necessary for the subgradient updates (see line 5 of Algorithm. 1), are typically set to $10^7/P_k^{\mathrm{mask}}$ and $10^7/P^{\mathrm{line}}$ resp. in our simulations. The Lagrange multiplier search is stopped when the following two stopping criterions, based on the KKT conditions of (14), are satisfied

$$\left([\mathbf{P}_k\mathbf{S}_k\mathbf{P}_k^H]_{l,l} = P_k^{\mathrm{mask}}\right) \text{ or } \left([\mathbf{P}_k\mathbf{S}_k\mathbf{P}_k^H]_{l,l} < P_k^{\mathrm{mask}} \text{ and } \lambda_k^l = 0\right), \forall l,k, \quad (41)$$

$$\text{and } \left(\sum_k [\mathbf{P}_k\mathbf{S}_k\mathbf{P}_k^H]_{l,l} = P^{\mathrm{line}}\right) \text{ or } \left(\sum_k [\mathbf{P}_k\mathbf{S}_k\mathbf{P}_k^H]_{l,l} < P^{\mathrm{line}} \text{ and } \theta^l = 0\right), \forall l. \quad (42)$$

Remark 1: When all $\{A_n\}$ are equal to one, the MIMO-BC problem (10) reduces to a BC problem, for which only 1 iteration is necessary for convergence and the optimal equalizer matrices $\{\mathbf{F}_k^{n,H}\}$ reduce to identity matrices.

Remark 2: The BC-DSB can be simply extended to incorporate a constraint on the bit loading (i.e. a bitcap b^{cap}). It consists in particular in bounding (by $\Gamma(2^{b^{\mathrm{cap}}}-1)$) the US SNR (34) for each tone k and line (n,i) before performing the per-line SNR duality transformations, and hence, resulting in capped line bit rates. Even though this SNR upper bounding is not a truly optimal approach, it has shown to work very well with low complexity, as will be shown in the simulation section.

Remark 3: When $(\theta^l + \lambda_k^l) = 0$, the matrices $\{\mathbf{X}_k^l\}$ in (28) for tone k may become rank-deficient in the iterative fixed-point update formula for particular sets of transmit powers, and then, BC-DSB breaks down. Fortunately, this problem can be avoided by simply lower bounding $\{\theta^l\}$ by a small positive constant [26]. The per-line Lagrange multipliers $\{\theta^l\}$ go to zero when the per-tone spectral mask and bit loading constraints are more restrictive than the per-line total transmit power constraints. The latter are then automatically satisfied. In the same manner, the per-tone Lagrange multipliers $\{\lambda_k^l\}$ go to zero when the bit loading constraints are too restrictive, which occurs for low frequency tones in particular.

Remark 4: The iterative fixed-point power update formula (27) implemented in a Gauss-Seidel (sequential) manner is guaranteed to increase the US Lagrangian (25) in each iteration, converging eventually to a stationary point [25]. However, the downside is that $\{\mathbf{X}_k^l\}$ of all the lines (which are considered as virtual single-line users) have to be recomputed whenever one line transmit power is updated. To

reduce the computational complexity, the iterative fixed-point power update formula can also be implemented in a Jacobi (simultaneous) manner. In particular, the $\{\mathbf{X}_k^l\}$ of all lines are only re-computed when all the line transmit powers have been updated. Although convergence of the Jacobi update is not analytically proven, it has shown to converge fast in most of our simulations. However, especially for high frequency tones with high levels of crosstalk, the Jacobi update sometimes fails to converge and then the power update formula keeps jumping between multiple extrema of (25). In our simulations, we have avoided this by switching to Gauss-Seidel updating after a certain number of power update iterations. A similar remark about Gauss-Seidel and Jacobi updating was made by the authors in [28].

IV. BC-DSB-LP

LP based DSM for DS transmission can also be handled by BC-DSB. The only difference when using LP instead of NLP is that (9) is used for the DS SNR instead of (6), which results in a different US fixed-point power update formula for line l and tone k

$$\tilde{r}_k^l = \left[\frac{w_l/\log(2)}{1 + \sum_{j \neq l} \frac{(w_j \tilde{r}_k^j)/\log(2)}{\Gamma + \tilde{r}_k^j (\tilde{\mathbf{h}}_k^j)^H (\mathbf{X}_k^j)^{-1} \tilde{\mathbf{h}}_k^j} |(\tilde{\mathbf{h}}_k^j)^H (\mathbf{X}_k^j)^{-1} \tilde{\mathbf{h}}_k^l|^2} - \frac{\Gamma}{(\tilde{\mathbf{h}}_k^l)^H (\mathbf{X}_k^l)^{-1} \tilde{\mathbf{h}}_k^l} \right]^+ \tag{43}$$

where

$$\mathbf{X}_k^l = \left(\mathbf{\Theta} + \mathbf{\Lambda}_k + \sum_{j \neq l} \tilde{r}_k^j \tilde{\mathbf{h}}_k^j (\tilde{\mathbf{h}}_k^j)^H \right),$$

and a linear MMSE equalizer closed form expression for user n in the US system

$$\widetilde{\mathbf{Q}}_k^{n,H} = \sqrt{\widetilde{\mathbf{R}}_k^n} \widetilde{\mathbf{H}}_k^{n,H} \left((\mathbf{\Theta} + \mathbf{\Lambda}_k) + \sum_m \widetilde{\mathbf{H}}_k^m \widetilde{\mathbf{R}}_k^m (\widetilde{\mathbf{H}}_k^m)^H \right)^{-1}. \tag{44}$$

In addition, this results in a linear expression for both the US (34) and DS (36) SNR when using the per-line SNR duality transformations, for which the elements of the off-diagonal sub-matrices ($p \neq n$) of $\mathbf{Z}_k$ are given by $\{[\mathbf{Z}_k^{n,p}]_{i,u} = -|\hat{h}_k^{(p,u),(n,i)}|^2 | \forall u, i\}$; and in a DS linear MMSE equalizer expression for user n, treating the signals of all other users $m \neq n$ as background noise, i.e.,

$$\widetilde{\mathbf{F}}_k^{n,H} = \sqrt{\mathbf{S}_k^n} \mathbf{P}_k^{n,H} \mathbf{H}_k^n \times \left(\mathbf{I}_{A_n} + \mathbf{H}_k^{n,H} \left(\sum_m \mathbf{P}_k^m \mathbf{S}_k^m \mathbf{P}_k^{m,H} \right) \mathbf{H}_k^n \right)^{-1}. \tag{45}$$

A complete algorithmic description is included in Algorithm 1 and will be referred to as BC-DSB-LP.

Remark 5: The US bit-rate functions with linear MMSE equalizers are now a non-concave function in $\{\tilde{r}_k^l\}$. As a result, the KKT stationary condition remains only necessary and is not sufficient anymore, and therefore (43) will only converge to a locally optimal stationary point.

Remark 6: For LP, the per-user US MSE matrix diagonalization step also guarantees intra-user interference-free transmission, and as a result, now even increases the maximized value of the US Lagrangian (25) [21].

Algorithm 1 BC-DSB-NLP or BC-DSB-LP

initialize $\mathbf{F}_k^{n,H}$ with $\|\mathbf{f}_k^{(n,i)}\|^2 = 1, \forall (n,i), k$
repeat
  Calculate $\tilde{h}_k^{(n,i)}$ using (11), $\forall (n,i), k$
  repeat
    Set/update
    $\forall l, k : \lambda_k^l = \left[\lambda_k^l + \epsilon_1 \left([\mathbf{P}_k \mathbf{S}_k \mathbf{P}_k^H]_{l,l} - P_k^{\text{mask}} \right) \right]^+$
    $\forall l : \theta_l = \left[\theta^l + \epsilon_2 \left(\sum_k [\mathbf{P}_k \mathbf{S}_k \mathbf{P}_k^H]_{l,l} - P^{\text{line}} \right) \right]^+$
    repeat
      for $l = 1 \ldots L$ **do**
        Calculate $\tilde{r}_k^l$ with (27) or (43) $\forall k$
      end for
    until *convergence*
    Calculate $\widetilde{\mathbf{Q}}_k^{n,H}$ with (29) or (44), $\forall n, k$
    Diagonalize $\mathbf{E}_k^{n,\text{US}}$ using (32) and (33), $\forall n, k$
    Calculate $\mathbf{S}_k$ with (37), $\forall k$
    Calculate $\mathbf{P}_k$ with (35), $\forall k$
  until (41) and (42) are satisfied
  Calculate $\mathbf{F}_k^{n,H}$ with (40) or (45), $\forall n, k$
  Diagonalize $\mathbf{E}_k^{n,\text{DS}}$ similar to (32) and (33), $\forall n, k$
until *convergence*

V. BC-DSB complexity and Convergence

A. Convergence

Convergence of the precoder optimization part of BC-DSB is difficult to prove due to the non-convexity of the US Lagrange dual function (24) for which the proposed iterative fixed-point power update formula can only reach a locally optimal stationary point. This non-globally optimal evaluation of the Lagrange dual function means that convergence of sub-gradient Lagrange multiplier updating is no longer guaranteed [23].

Only for the case of NLP and zero capacity gap ($\Gamma = 0$ dB), the BC-DSB algorithm is provably convergent. Indeed, for this case, (24) becomes convex [15], guaranteeing convergence of the precoder optimization part to its globally optimal solution. Convergence of the total BC-DSB algorithm can then be proven, similar to the proof in [29], by arguing that every step of the algorithm increases the objective function of (10); and also that this objective function is at the same time upper bounded by the per-line total power and spectral mask constraints. Hence, after convergence of BC-DSB, a limit point of (10) must be reached.

We remark, however, that extensive simulations have shown that BC-DSB consistently produces a monotonic increasing objective value both for NLP and LP and nonzero capacity gap.

B. Complexity

The complexity of the proposed BC-DSB algorithm is analyzed and compared to the complexity of BC-OSB and

DMT-WMMSE (summarized in Table I) for the case where $\{A_n = A\}$ so that $L = NA$.

Precoder and transmit power optimization for fixed equalizers is the most computationally intensive part of the BC-DSB algorithm. It implies an iterative Lagrange multiplier search where, for each set of Lagrange multipliers, the US Lagrange dual function (24) has to be solved, after which its solution is to be transformed to the DS solution.

To solve (24), the iterative fixed-point power update formula requires the computation of all L matrices $\{\mathbf{X}_k^l\}$ in (28) and their inverses, leading to a total complexity of $O(L^4)$ per tone. Luckily, this complexity can be reduced to $O(L^3)$ by recognizing that for NLP each $\mathbf{X}_k^l$ is a rank-1 update of $\mathbf{X}_k^{l-1}$ and using the matrix inversion lemma (similar to [30]), i.e.,

$$\begin{aligned}(\mathbf{X}_k^l)^{-1} &= \left(\mathbf{X}_k^{l-1} + \tilde{r}_k^{l-1}\tilde{\mathbf{h}}_k^{l-1}(\tilde{\mathbf{h}}_k^{l-1})^H\right)^{-1} \\ &= (\mathbf{X}_k^{l-1})^{-1} - \frac{\tilde{r}_k^{l-1}(\mathbf{X}_k^{l-1})^{-1}\tilde{\mathbf{h}}_k^{l-1}(\tilde{\mathbf{h}}_k^{l-1})^H(\mathbf{X}_k^{l-1})^{-1}}{1+\tilde{r}_k^{l-1}(\tilde{\mathbf{h}}_k^{l-1})^H(\mathbf{X}_k^{l-1})^{-1}\tilde{\mathbf{h}}_k^{l-1}}\end{aligned}$$

while for LP $\mathbf{X}_k^l$ is a rank-1 update of following matrix

$$\mathbf{B}_k \triangleq \left(\mathbf{\Theta} + \mathbf{\Lambda}_k + \sum_j \tilde{r}_k^j \tilde{\mathbf{h}}_k^j (\tilde{\mathbf{h}}_k^j)^H\right), \tag{46}$$

yielding

$$\begin{aligned}(\mathbf{X}_k^l)^{-1} &= \left(\mathbf{B}_k - \tilde{r}_k^l \tilde{\mathbf{h}}_k^l (\tilde{\mathbf{h}}_k^l)^H\right)^{-1} \\ &= \mathbf{B}_k^{-1} + \frac{\tilde{r}_k^l \mathbf{B}_k^{-1}\tilde{\mathbf{h}}_k^l(\tilde{\mathbf{h}}_k^l)^H\mathbf{B}_k^{-1}}{1-\tilde{r}_k^l(\tilde{\mathbf{h}}_k^{l-1})^H\mathbf{B}_k^{-1}\tilde{\mathbf{h}}_k^l}.\end{aligned}$$

Thus, the complexity of solving (24) is $O(KL^4)$ when using the Gauss-Seidel update (see Remark 4), and ignoring the few iterations by the fixed-point formula. On the contrary, when using the Jacobi update, the complexity is $O(KL^3)$.

The computation of the US linear or nonlinear MMSE equalizers $\{\widetilde{\mathbf{Q}}_k^{n,H}\}$ has a complexity of $O(KL^3)$. Although expression (29) for computing the nonlinear $\{\widetilde{\mathbf{Q}}_k^{n,H}\}$ has a complexity of $O(NL^3)$ due to the N matrix inversions of $\mathbf{M}_k^n$ of size $L\times L$ on tone k, the matrix inversion lemma can again be exploited, i.e.,

$$\begin{aligned}(\mathbf{M}_k^n)^{-1} =& \left(\mathbf{M}_k^{n-1} + \widetilde{\mathbf{H}}_k^n \widetilde{\mathbf{R}}_k^n \widetilde{\mathbf{H}}_k^{n,H}\right)^{-1} \\ =&(\mathbf{M}_k^{n-1})^{-1} - (\mathbf{M}_k^{n-1})^{-1}\widetilde{\mathbf{H}}_k^n \left((\widetilde{\mathbf{R}}_k^n)^{-1}\right. \\ &\left.+\widetilde{\mathbf{H}}_k^{n,H}(\mathbf{M}_k^{n-1})^{-1}\widetilde{\mathbf{H}}_k^n\right)^{-1}\widetilde{\mathbf{H}}_k^{n,H}(\mathbf{M}_k^{n-1})^{-1}\end{aligned} \tag{47}$$

in which the complexity of $\widetilde{\mathbf{H}}_k^{n,H}(\mathbf{M}_k^{n-1})^{-1}\widetilde{\mathbf{H}}_k^n$ dominates, meaning that $(\mathbf{M}_k^n)^{-1}$ can be computed in $O(N^2A^3)$ operations given $(\mathbf{M}_k^{n-1})^{-1}$. On the other hand, the $L\times L$ matrix inversion in expression (44) for the linear $\widetilde{\mathbf{Q}}_k^{n,H}$ is the same for all users and, hence, can be re-used. Furthermore, the complexity of transforming the US transmit powers and MMSE equalizers to the DS domain for all tones involves solving K linear system of equations and is always upper bounded by $O(KL^3)$.

The overall complexity of BC-DSB is summarized in Table I in which we considered the complexity for the case of NLP and LP the same as can be inferred from above analysis. In this table, α and β denote respectively the number of outer loop iterations needed, and the number of Lagrange multiplier iterations needed to satisfy all transmit power constraints. Typically setting α to 10-300 is sufficient for convergence in our simulations, while β drops to a small number (typically) after a few outer loop iterations.

TABLE I
COMPLEXITY COMPARISON

Algorithm	Computational complexity
BC-DSB with Gaus-Seidel update	$O(\alpha\beta KL^4)$
BC-DSB with Jacobi update	$O(\alpha\beta KL^3)$
BC-OSB [9]	$O(\alpha\beta K(b^{\max}+1)^L L^3)$
DMT-WMMSE [11], [12]	$O(\alpha\beta KL^3)$

Clearly, Table I reveals that, especially when the number of lines are large, BC-DSB is significantly less complex than BC-OSB; and furthermore, that the complexity of one outer loop iteration of BC-DSB with Jacobi update and the DMT-WMMSE algorithm is similar. Thus, the total complexity is mainly determined by the speed of convergence. Note that although Jacobi updating is not guaranteed to converge, we avoid this by switching to Gauss-Seidel updating after a certain number of power update iterations on some tones (see Remark 4)).

VI. SIMULATION RESULTS

In this section, we evaluate the performance of BC-DSB for the DS G.fast 106 MHz and 212 MHz profile. We consider for both profiles standard compliant G.fast parameters. The per-line total transmit power is 4 dBm. The capacity gap Γ is set to 10.25 dB and the noise PSD to -140 dBm/Hz. The tone spacing is 51.75 kHz and the symbol rate is 48 KHz. The bandwidth of the G.fast profiles starts at 2.2 MHz and ends at 106 MHz and 212 MHz corresponding to a total of about $K = 2000$ and $K = 4000$ tones respectively. Spectral masks are obtained from [31]. The channel matrices have been obtained by measurements of a cable binder consisting out of 10 lines of 80 m. We remark that the observed crosstalk levels in this particular cable binder are rather high compared to other G.fast measurements. For all simulations, we only consider the unweighted sum-rate case (i.e. $\{w_n = 1|\forall n\}$).

A. Performance Comparison With ZF Techniques

In this section, the performance of BC-DSB is compared with ZF techniques. In the aforementioned cable binder, we vary the amount of receiver signal coordination, corresponding to scenarios with a different number of users N and lines per user $\{A_n\}$. Performance in these scenarios is measured in terms of the achieved unweighted sum-rate (shown in Table II and Table III in case of a 12 bitcap). The considered ZF technique is the singular value decomposition based block-diagonalization (SVD-BD) method for the MIMO-BC scenario extended for per-line power constraints in [32] which can be used for LP as well as for NLP. This method is combined with DSM by means of a bounded gain scaling optimization such that all per-line total power and per-tone spectral mask and bit

TABLE II
SUM-RATE COMPARISON BETWEEN VARIOUS PRECODING SCHEMES FOR THE G.FAST 106 MHZ AND 212 MHZ PROFILE IN 10×80 M BINDER AND DIFFERENT CASES OF RECEIVER SIGNAL COORDINATION.

			MMSE precoding				ZF precoding			
			BC-DSB-NLP		**BC-DSB-LP**		**SVD-BD-NLP**		**SVD-BD-LP**	
	Receiver signal coordination		[Gbps]	[%]	[Gbps]	[%]	[Gbps]	[%]	[Gbps]	[%]
212 MHz	100%	$N = 1, A = 10$	13.096	100.0	13.096	100.0	13.096	100.0	13.096	100.0
	50%	$N = 2, \{A_n = 5, \forall n\}$	13.038	99.6	12.725	97.2	13.025	99.5	12.550	95.8
	Phantom	$N = 4, \{A_n\} = \{2, 3, 2, 3\}$	12.977	99.1	12.368	94.5	12.940	98.8	11.939	91.2
	Bonded	$N = 5, \{A_n = 2, \forall n\}$	12.946	98.9	12.079	92.3	12.914	98.6	11.501	87.8
	None	$N = 10, \{A_n = 1, \forall n\}$	12.919	98.6	11.774	89.9	12.891	98.4	10.995	84.0
106 MHz	100%	$N = 1, A = 10$	11.180	100.0	11.180	100.0	11.180	100.0	11.180	100.0
	50%	$N = 2, \{A_n = 5, \forall n\}$	11.180	100.0	10.994	98.3	11.179	99.99	10.992	98.3
	Phantom	$N = 4, \{A_n\} = \{2, 3, 2, 3\}$	11.179	99.99	10.808	96.7	11.178	99.98	10.805	96.6
	Bonded	$N = 5, \{A_n = 2, \forall n\}$	11.178	99.98	10.627	95.1	11.178	99.98	10.623	95.0
	None	$N = 10, \{A_n = 1, \forall n\}$	11.178	99.98	10.399	93.0	11.178	99.98	10.387	92.9

TABLE III
SAME SCENARIO AS IN TABLE II WITH AN ADDITIONAL 12 BITCAP CONSTRAINT.

			MMSE precoding				ZF precoding			
			BC-DSB-NLP		**BC-DSB-LP**		**SVD-BD-NLP**		**SVD-BD-LP**	
	Receiver signal coordination		[Gbps]	[%]	[Gbps]	[%]	[Gbps]	[%]	[Gbps]	[%]
212 MHz	100%	$N = 1, A = 10$	12.591	100.0	12.591	100.0	12.591	100.0	12.591	100.0
	50%	$N = 2, \{A_n = 5, \forall n\}$	12.531	99.5	12.200	96.9	12.522	99.4	12.023	95.5
	Phantom	$N = 4, \{A_n\} = \{2, 3, 2, 3\}$	12.474	99.1	11.832	94.0	12.443	98.8	11.376	90.4
	Bonded	$N = 5, \{A_n = 2, \forall n\}$	12.449	98.9	11.535	91.6	12.417	98.6	10.870	86.3
	None	$N = 10, \{A_n = 1, \forall n\}$	12.423	98.7	11.196	88.9	12.396	98.4	10.196	81.1
106 MHz	100%	$N = 1, A = 10$	10.270	100.0	10.270	100.0	10.270	100	10.270	100
	50%	$N = 2, \{A_n = 5, \forall n\}$	10.276	100.1	10.094	98.3	10.274	100.1	10.093	98.3
	Phantom	$N = 4, \{A_n\} = \{2, 3, 2, 3\}$	10.277	100.1	9.871	96.2	10.277	100.1	9.869	96.1
	Bonded	$N = 5, \{A_n = 2, \forall n\}$	10.280	100.2	9.674	94.3	10.281	100.2	9.671	94.2
	None	$N = 10, \{A_n = 1, \forall n\}$	10.281	100.2	9.343	91.0	10.282	100.2	9.316	90.7

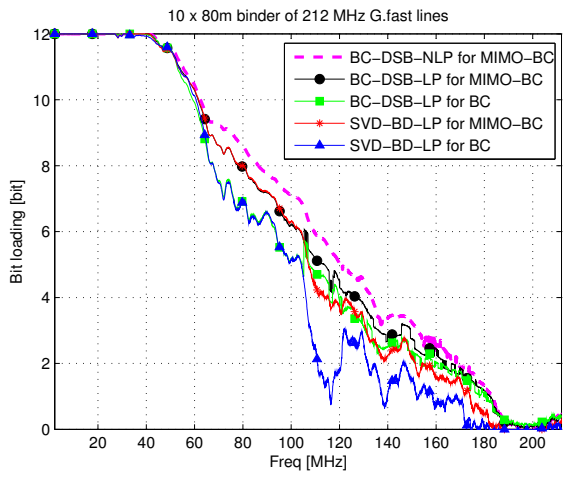

Fig. 2. The average line bit loading across all tones for the "phantom" scenario of Table II shows MMSE precoding outperforming ZF precoding both for the BC and MIMO-BC scenario at high frequencies due to the strong crosstalk at these frequencies.

loading constraints are satisfied. Note that SVD-BD-LP can be considered as the generalization of the linear ZF precoder [3] with multiple lines per user.

For the scenario with full two-sided signal coordination (case "100%" in Table II), all solutions reduce to the SVD-solution [33] which provides an upper bound on the achievable sum-rate in this binder (at least when the bit loading is not constrained). We also highlight the difference with the VDSL2 scenario where linear ZF precoding is near-optimal due to the diagonally dominant structure of the channel matrix at low (i.e. below 30 MHz) frequencies [3].

BC-DSB-LP outperforms SVD-BD-LP on all scenarios for the 212 MHz profile (see Table II) due to the very strong crosstalk channels and high direct channel attenuation at high frequencies resulting in low SNRs, such that MMSE outperforms ZF precoding for the BC and MIMO-BC. For instance, for the 212 MHz profile and "phantom" scenario (illustrated in Fig. 2), the bit loadings provided by BC-DSB-LP for the BC and MIMO-BC scenario are very similar to those of SVD-BD-LP respectively up to 100 MHz, after which the BC-DSB-LP solutions start to improve over the latter. This improved performance over SVD-BD-LP is most significant (by about 7%) in the scenario without receiver signal coordination (case "none" in Table II). However, in scenarios with more receiver signal coordination, this performance gap gets notably

reduced. Also, for the 106 MHz profile, the performance gain of BC-DSB-LP over SVD-BD is negligible.

The successive interference subtraction of NLP enhances the crosstalk environment resulting only in a small loss compared to the performance upper bound of the “100%" scenario, even without receiver signal coordination. This means that BC-DSB-NLP is expected to operate very close to the globally optimal solution notwithstanding the nonzero capacity gap ($\Gamma > 0$dB). Furthermore, this enhanced crosstalk environment for NLP decreases the performance gap between ZF and MMSE precoding, resulting only in a small gain of BC-DSB-NLP over SVD-BD-NLP.

Although these NLP results show an improved performance, they may be too optimistic for practical implementations like THP. For instance, the power penalties due to the modulo operations resulting in an increased transmit power and capacity gap have been ignored [19]. Furthermore, we have assumed perfect channel state information and ideal signal processing, even though NLP schemes have recently been shown to be sensitive to channel estimation errors and other non-idealities [34]. When taking these effects into account, the performance gap between NLP and LP will be reduced. Additional drawbacks of NLP are the significantly increased implementation complexity and the performance variation across users. Users encoded last achieve lower rate gains than users encoded first.

In case of a bitcap (Table III), the performance gap between BC-DSB and the SVD-BD schemes for the 212 MHz profile even increases. This is explained by the excess power of bitcapped tones at low frequencies that now is shifted to tones at higher frequencies. For example, for the BC scenario the gap between BC-DSB-LP and SVD-BD-LP increases from 7% to 9.6%. Furthermore, this bitcap operation generates a somewhat unintuitive result of NLP achieving a higher sum-rate than the SVD-solution for the 106 MHz profile. I.e., for the bitcap operation and this particular measured channel, the performance variation across lines of NLP is more beneficial than the one following the singular values of the SVD-solution on some tones.

B. Comparison With BC-OSB and DMT-WMMSE

Although the globally optimal BC-OSB optimizes the more realistic discrete bit loading (i.e. it uses an exhaustive discrete bit loading search to solve the US Lagrange dual function (24)), BC-DSB is more practical for large scenarios due to its significantly reduced complexity (see also section V-B). For instance, for 4 users, BC-OSB already takes more than a week [22], compared to only a few minutes for BC-DSB. Furthermore, the continuous bit loading of BC-DSB results in slightly better (<1%) line rates compared to the discrete bit loading of BC-OSB (illustrated in Fig. 3 for a 2-user scenario). In addition, this continuous nature of BC-DSB allows the transmit powers to satisfy the total per-line transmit power and spectral mask constraints very precisely, whereas the accuracy of BC-OSB is determined by the bit loading quantization step.

Extensive simulations have shown that BC-DSB-LP almost always needs significantly less iterations to converge than DMT-WMMSE. This is illustrated in Fig. 4 for the “50%", “Phantom", and “Bonded" scenario of Table II. While BC-DSB always convergence within a few tens of iterations, the DMT-WMMSE needs more than 3000 iterations. This is probably due to the reason that BC-DSB explicitly maximizes the sum-rate each iteration, while DMT-WMMSE aims to minimize the crosstalk interference between users. In addition, the slow convergence of WMMSE-based algorithms at high SNR is well-known. In our simulations, the DMT-WMMSE proposed in [11], [12] is extended with additional per-tone Lagrange multipliers for the spectral mask constraints on top of the per-line Lagrange multipliers, and a per-user MSE matrix diagonalization step. The DMT-WMMSE is initialized with the SVD-BD precoders and a flat power-allocation satisfying the power constraints.

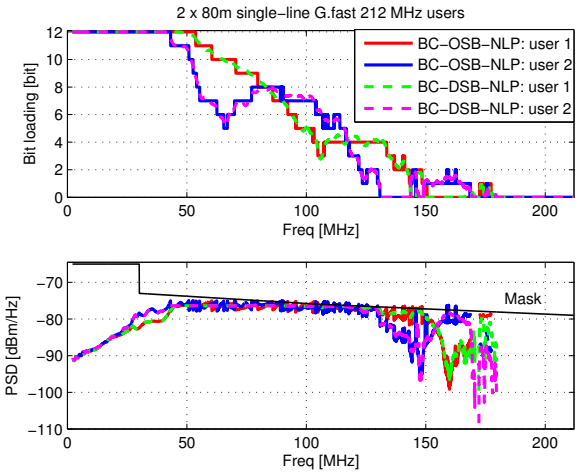

Fig. 3. Although BC-DSB optimizes the less realistic continuous bit loading, its computational complexity is significantly reduced w.r.t. BC-OSB.

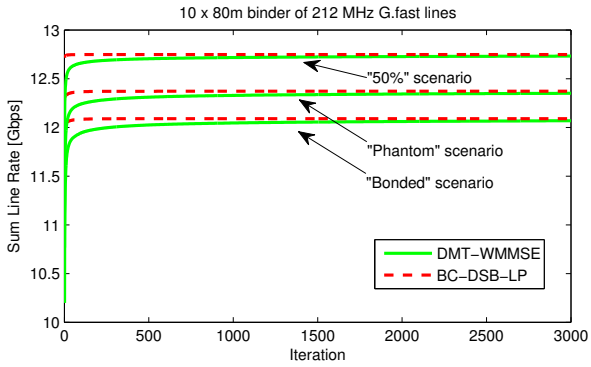

Fig. 4. BC-DSB is observed to exhibit faster convergence than DMT-WMMSE.

VII. Conclusion

Motivated by the strong crosstalk at high frequencies characterizing G.fast cable binders, we have investigated both linear and nonlinear precoding based DSM for weighted sum-rate maximization in the context of DS vectored G.fast transmission with multiple lines available to each user and under realistic per-line total power and per-tone spectral mask constraints. For this problem, we have developed a novel

algorithm called BC-DSB that alternates between precoder and transmit power versus equalizer optimization, and transforms the DS dual decomposition approach formulation of the former before exploiting US-DS duality. Furthermore, a low-complexity iterative fixed-point formula has been proposed to solve the resulting US problem, and SNR duality transformations have been presented to convert the US solutions back to the DS domain without any rate-loss due to intra-user interference. The computational complexity of the BC-DSB has been shown to be similar to the state-of-the-art DMT-WMMSE, while a faster convergence rate of BC-DSB has been experimentally observed. Finally, a performance comparison in a very high crosstalk cable binder up to 212 MHz has been provided revealing an improved performance of BC-DSB over ZF techniques for various scenarios.

REFERENCES

[1] W. Lanneer, M. Moonen, P. Tsiaflakis, and J. Maes, "Linear and non-linear precoding based dynamic spectrum management for downstream vectored G.fast transmission," in *Proc. IEEE Global Comm. Conf. (GLOBECOM)*, San Diego, USA, Dec. 2015, pp. 1–6.

[2] *Self-FEXT cancellation (vectoring) for use with VDSL2 transceivers*, Recommendation ITU-T G.993.5, April 2010.

[3] R. Cendrillon, G. Ginis, E. Van den Bogaert, and M. Moonen, "A Near-Optimal Linear Crosstalk Precoder for Downstream VDSL," *IEEE Trans. Commun.*, vol. 55, no. 5, pp. 860–863, May 2007.

[4] *Fast Access to Subscriber Terminals (FAST) - Physical Layer Specification*, Recommendation ITU-T G.9701, Nov. 2015.

[5] R. Strobel, R. Stolle, and W. Utschick, "Wideband modeling of twisted-pair cables for MIMO applications," in *IEEE Global Communications Conf. (GLOBECOM)*, 2013, pp. 2828–2833.

[6] M. Guenach, C. Nuzman, P. Tsiaflakis, and J. Maes, "Power optimization in vectored and non-vectored G.fast transmission," in *IEEE Global Communications Conf. (GLOBECOM)*, 2014, pp. 2229–2233.

[7] M. Peeters and S. Vanhastel, "The Copper Phantom," *OSP Mag.*, Dec. 2010. [Online]. Available: https://www.isemag.com/2017/01/the-copper-phantom/

[8] W. Yu and T. Lan, "Transmitter Optimization for the Multi-Antenna Downlink with Per-Antenna Power Constraints," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2646–2660, June 2007.

[9] V. Le Nir, M. Moonen, J. Verlinden, and M. Guenach, "Optimal power allocation for downstream xDSL with per-modem total power constraints: Broadcast Channel Optimal Spectrum Balancing (BC-OSB)," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 690–697, Feb. 2009.

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

[11] R. B. Moraes, P. Tsiaflakis, J. Maes, and M. Moonen, "DMT MIMO IC rate maximization in DSL with per-transceiver power constraints," *IEEE Trans. Signal Process.*, vol. 101, no. 0, pp. 87 – 98, 2014.

[12] R. Moraes, P. Tsiaflakis, J. Maes, and M. Moonen, "General Framework and Algorithm for Data Rate Maximization in DSL Networks," *IEEE Trans. Commun.*, vol. 62, no. 5, pp. 1691–1703, May 2014.

[13] P. Tsiaflakis, M. Diehl, and M. Moonen, "Distributed spectrum management algorithms for multiuser DSL networks," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 4825–4843, Oct. 2008.

[14] M. Costa, "Writing on dirty paper (Corresp.)," *IEEE Trans.Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[15] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.

[16] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. New York, NY, USA: Cambridge University Press, 2005.

[17] M. Tomlinson, "New automatic equaliser employing modulo arithmetic," *Electron. Lett.*, vol. 7, no. 5, pp. 138–139, March 1971.

[18] H. Harashima and H. Miyakawa, "Matched-transmission technique for channels with intersymbol interference," *IEEE Trans. Commun.*, vol. 20, no. 4, pp. 774–780, Aug. 1972.

[19] J. Neckebroek, M. Moeneclaey, W. Coomans, M. Guenach, P. Tsiaflakis, R. B. Moraes, and J. Maes, "Novel bitloading algorithms for coded G.fast DSL transmission with linear and nonlinear precoding," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2015, pp. 945–951.

[20] T. Starr, J. M. Cioffi, and P. J. Silverman, *Understanding Digital Subscriber Line Technology*. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1999.

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

[22] P. Tsiaflakis, J. Vangorp, J. Verlinden, and M. Moonen, "Multiple Access Channel Optimal Spectrum Balancing for Upstream DSL Transmission," *IEEE Commun. Lett.*, vol. 11, no. 4, pp. 398–300, April 2007.

[23] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310–1322, July 2006.

[24] P. Tsiaflakis, R. Moraes, and M. Moonen, "A low-complexity algorithm for joint spectrum and signal coordination in upstream DSL transmission," in *18th IEEE Symp. Communications and Vehicular Technology in the Benelux (SCVT)*, 2011, pp. 1–6.

[25] P. Tsiaflakis and F. Glineur, "A novel class of iterative approximation methods for DSL spectrum optimization," in *IEEE Int. Conf. Communications (ICC)*, 2012, pp. 3154–3159.

[26] A. Forouzan, M. Moonen, J. Maes, and M. Guenach, "Joint level 2 and 3 dynamic spectrum management for downstream DSL," *IEEE Trans. Commun.*, vol. 60, no. 10, pp. 3111–3122, Oct. 2012.

[27] R. Hunger, M. Joham, and W. Utschick, "On the MSE-duality of the broadcast channel and the multiple access channel," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 698–713, Feb. 2009.

[28] D. Nguyen and T. Le-Ngoc, "Sum-rate maximization in the multicell MIMO broadcast channel with interference coordination," *IEEE Trans. Signal Process.*, vol. 62, no. 6, pp. 1501–1513, March 2014.

[29] S. Shi, M. Schubert, and H. Boche, "Rate optimization for multiuser MIMO systems with linear processing," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 4020–4030, Aug. 2008.

[30] A. Forouzan, M. Moonen, J. Maes, and M. Guenach, "Efficient calculation of the MMSE-GDFE decoding order in non-ideal DSL multiple-access channels," in *17th IEEE Symp. Communications and Vehicular Technology Benelux (SCVT)*, Nov. 2010, pp. 1–6.

[31] *Fast Access to Subscriber Terminals (FAST) - Power Spectral Density Specification*, Recommendation ITU-T G.9700, April 2014.

[32] R. Zhang, "Cooperative multi-cell block diagonalization with per-base-station power constraints," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1435–1445, Dec. 2010.

[33] V. Le Nir, M. Moonen, J. Verlinden, and M. Guenach, "Full vectoring optimal power allocation in xDSL channels under per-modem power constraints and spectral mask constraints," *IEEE Trans. Commun.*, vol. 57, no. 1, pp. 194–202, Jan. 2009.

[34] J. Maes, C. Nuzman, and P. Tsiaflakis, "Sensitivity of nonlinear precoding to imperfect channel state information in G.fast," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2016.



Wouter Lanneer (S'15) received the M.Sc. degree in electrical engineering from the KU Leuven, Leuven, Belgium, in 2014, where he is currently pursuing the Ph.D. degree under the supervision of Prof. M. Moonen at the Electrical Engineering Department of KU Leuven. He is currently involved in joint projects with KU Leuven and Nokia Bell Labs, Antwerp, Belgium. His research concerns about signal processing and optimization for digital communication systems with an emphasis on DSL wireline access networks.



Paschalis Tsiaflakis (S'06-M'09) received the Master's degree in Electrical Engineering and the Ph.D. degree in Engineering Sciences from the KU Leuven, Leuven, Belgium, in 2004 and 2009, respectively.

From 2010 to 2013, he was a Post-Doctoral Research Fellow at the KU Leuven. He was a Visiting Researcher with Princeton University, University of California at Los Angeles, and Tsinghua University, and a Post-Doctoral Research Associate with the UniversitÃl' catholique de Louvain. Since 2013, he has been a Research Engineer with Nokia Bell Labs, where he drives innovation into next generation communication systems. He has published over 60 papers in internationally recognized journals and conferences, and has 10+ patents pending. His research interests include vectoring, resource allocation, MAC/PHY design for ultrabroadband access systems, algorithm design, optimization theory, and signal processing.

Dr. Tsiaflakis was a recipient of the Belgian Young ICT Personality Award sponsored by FITCE in 2010, two IEEE ICC best paper awards in 2013, the Best Multimedia Master Thesis Prize Award sponsored by PIMC in 2001, and was a top-12 finalist for the European ERCIM Cor Baayen Award 2010. He was also a recipient of the FWO Aspirant Grant in 2004, the PDMK postdoc grant in 2009, the Francqui Intercommunity Postdoc Grant in 2010, the FWO postdoc grant in 2011, and the FNRS postdoc grant in 2011.



Jochen Maes (SM'11) received the M.Sc. degree in physics and the Ph.D. degree in science from the KU Leuven, Leuven, Belgium. He was with Bell Labs in 2006, where he continuously shifts the limits of copper. He was the Head of Copper Access and Indoor Department with the Fixed Networks Laboratory, focused on leveraging legacy infrastructure through innovations in systems and transceiver design. He contributes to several ITU standards. His work was a recipient of the Broadband Infovision Awards (2010, 2014) and the Bell Labs Presidents Award (2011).



Marc Moonen (M'94-SM'06-F'07) is currently a Full Professor with Electrical Engineering Department, KU Leuven, where he is the head of the research team working in the area of numerical algorithms and signal processing for digital communications, wireless communications, DSL, and audio signal processing.

He is currently a member of the Editorial Board of EURASIP Journal on Advances in Signal Processing. He was the Chairman of the IEEE Benelux Signal Processing Chapter (1998-2002), a member of the IEEE Signal Processing Society Technical Committee on Signal Processing for Communications, and the President of EURASIP (2007-2008 and 2011-2012).

He was a recipient of the 1994 KU Leuven Research Council Award, the 1997 Alcatel Bell (Belgium) Award (with Piet Vandaele), the 2004 Alcatel Bell (Belgium) Award (with Raphael Cendrillon), and was a 1997 Laureate of the Belgium Royal Academy of Science. He received best paper awards from the IEEE TRANSACTIONS ON SIGNAL PROCESSING (with Geert Leus and with Daniele Giacobello) and from Signal Processing (with Simon Doclo).

He has served as an Editor-in-Chief of the EURASIP Journal on Applied Signal Processing (2003-2005), an Area Editor for feature articles in the IEEE Signal Processing Magazine (2012-2014), and has been a member of the Editorial Board of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II, the IEEE Signal Processing Magazine, Integration, the VLSI Journal, EURASIP Journal on Wireless Communications and Networking, and Signal Processing.