Joint Level 2 and 3 Dynamic Spectrum Management for Downstream VDSL

Amir R. Forouzan†, Marc Moonen†, Jochen Maes‡, and Mamoun Guenach‡

† Dept. of Electrical Engineering (ESAT-SISTA), Katholieke Universiteit Leuven, Leuven, Belgium
‡ Bell Labs, Alcatel-Lucent, Antwerp, Belgium

Abstract

In this paper, we investigate joint level 2 and 3 dynamic spectrum management (joint DSM 2/3) for downstream (DS) DSL. We consider a DS scenario in which users are divided into a few separate groups, where vector encoding based signal coordination can be applied in each group and spectrum coordination is possible for all users. This can be seen as a mixed interference/broadcast channel (IF/BC) scenario. We investigate several candidates for vector encoding the signals inside the groups, including the linear zero-forcing (ZF) pre-compensator, ZF Tomlinson-Harashima pre-coder (THP), optimal linear pre-compensator (OLP), and THP with optimal transmit filters. The calculation of the optimal transmit filters (for both the linear pre-compensator and THP) is a non-convex problem. To resolve this problem, we develop a generalized duality between the vector broadcast and multiple-access channel (MAC) for scenarios in which partial signal coordination is available among users, such as the IF/BC and its dual IF/MAC. In order to achieve the highest data rates, optimal spectrum balancing (OSB) is applied to all users on top of the vector encoding inside the groups. Simulation results show that the grouped THP with optimal transmit filters achieves considerably higher bit rates than the other schemes. The resulting joint DSM 2/3 algorithm is referred to as the IF/BC-OSB algorithm, and encompasses the OSB (i.e., IF-OSB) and the earlier developed BC-OSB algorithm as special cases. A simplified version of IF/BC-OSB with near-optimal performance is also proposed.

Index Terms—Digital subscriber line (DSL), dynamic spectrum management (DSM), broadcast channel (BC), interference channel, BC-MAC duality, dirty paper coding (DPC), Tomlinson-Harashima pre-coder (THP), resource allocation, vectoring.

I. Introduction

Far-end crosstalk (FEXT) is one of the dominant impairments in very-high-speed digital subscriber line (VDSL) networks. Promising crosstalk mitigation and canceling techniques have been proposed in the last decade to counteract crosstalk based on spectrum coordination and signal coordination. These techniques are jointly referred to as dynamic spectrum management (DSM) techniques. When the bit rate of the users in the network are configured centrally but the computation of the transmit spectrum of the users is left uncoordinated, we obtain a level 1 DSM scheme, also known as DSM 1. When the transmit spectrum of the users is controlled by a spectrum management center (SMC), we obtain a DSM 2 scheme. DSM 2 schemes may achieve higher rates for the users compared to DSM 1 schemes. Finally, when the signals of the users are modulated and/or demodulated jointly using so-called vectoring schemes, we obtain DSM 3. DSM 3 is capable of achieving considerably higher bit rates than DSM 2. In state-of-the-art DSM 3 schemes, also the transmit spectra and bit rate of the users may be managed by an SMC.

DSM 3 requires spectrum coordination as well as signal coordination among all users which is usually achieved at the line termination (LT) end (e.g. at the central office (CO) or a remote terminal (RT)) only. When signal coordination among all users is possible at the LT, DSM 3 schemes can attain rates as if there is no crosstalk or even slightly higher rates by exploiting the power of the desired signal received through the crosstalk channels. However, sometimes due to the excessive complexity or physical limitations (e.g. as in joint CO/RT deployments), DSM 3 is not possible for all users. Hence, the users may be divided into separate groups, where signal coordination is possible inside each group but only spectrum coordination is possible for the users belonging to different groups. For these cases, so-called joint DSM 2/3 schemes can be used to maximize the achievable bit rates.

In this paper, we investigate joint DSM 2/3 for downstream (DS) VDSL. The joint DSM 2/3 problem for upstream (US) VDSL has been studied in [1]. The results show a considerable bit

This research work was carried out at the ESAT Laboratory of Katholieke Universiteit Leuven, in the frame of

- Concerted Research Action GOA-MaNet,
- The Belgian Programme on Interuniversity Attraction Poles initiated by the Belgian Federal Science Policy Office IUAP P6/04 (DYSCO, 'Dynamical systems, control and optimization', 2007-2011),
- K.U.Leuven Research Council CoE PFV/10/002 'Optimization in Engineering' (OPTEC),
- Research Project FWO nr.G.0235.07('Design and evaluation of DSL systems with common mode signal exploitation'), and
- IWT Project 'PHANTER: PHysical layer and Access Node TEchnology Revolutions: enabling the next generation broadband network'.

The scientific responsibility is assumed by its authors.

rate boost for the users when optimal joint DSM 2/3 schemes are used. Although, the joint DSM 2/3 in US and DS are closely related in practice, in general the vectoring schemes for them are totally different because in US the vectoring is applied at the receiver side while in DS it is applied at the transmitter side. Moreover, the DS problem is more difficult because, unlike in the US problem, the calculation of the optimal filters is a non-convex problem.

The Joint DSM 2/3 scenario can be viewed as a generalization of the so-called interference (IF) channel and vector broadcast channel (BC) and so will be referred to as an IF/BC. In the special case, where each group consists of a single user, the scenario is reduced to an IF channel. On the other hand, when all users are in the same group, the problem is reduced to a vector BC. The optimal strategy to achieve the capacity region of the Gaussian IF channel is unknown in the general case, although, near-optimal solutions are available [2, 3]. These techniques involve decoding the signal of unintended users using the interference (i.e. crosstalk) signal. As the IF/BC is a generalization of the IF channel, the capacity region is also unknown in this case. Nevertheless, the problem is different to some extent for VDSL channels compared to ideal Gaussian channels due to practical considerations. Particularly, crosstalk is always treated as (Gaussian) noise in VDSL channels. That is, decoding the signal of the unintended users is not an available option. Other important differences include the VDSL transceiver gap to Shannon capacity, and an upper bound on per-tone QAM constellation size, i.e., the number of bits that can be loaded to each tone. Under such practical considerations, the achievable rate region (RR) for a VDSL IF channel is obtained by the optimal spectrum balancing (OSB) algorithm [4].

On the other hand, it has been shown that the capacity region of the vector BC is achievable using Costa's dirty paper coding (DPC) [5] scheme. First, it was shown that DPC achieves the sum capacity of the BC with two [6] or multiple [7–9] transmitters under a sum transmit power constraint. The solutions in [7, 8] are based on a duality theory between the BC and multiple-access channel (MAC), which states that the same rates can be achieved under equal sum power constraints over a BC and a MAC with their channel matrices being the transpose of each other. Later it was shown that the entire capacity region can be achieved by DPC as well [10].

All these results consider a constraint on the sum transmit power, summed over all transmitters (i.e., lines in the DSL context). However, for VDSL systems, we are more interested in a per-line power constrained solution, because the transmit power on each line is bounded by the nominal power spectral density (PSD) mask constraints. Moreover, the total transmit power on each line over all tones can be bounded by analog front-end limitations or standard

regulations. This problem has been addressed in [11] for single-carrier vector BCs, where the BC-MAC duality is generalized using the Lagrange dual optimization technique. The outcome is that a vector BC problem with per-transmitter power constraint can be solved in a dual MAC with uncertain noise powers. The uncertain noise powers are in fact the Lagrange multipliers corresponding to the per-transmitter power constraints. The solution is found by selecting appropriate values for the noise powers which satisfy the constraints. In [12], optimal spectrum and signal coordination for DS DSL has been considered and a BC-OSB algorithm has been proposed. The BC-OSB algorithm uses the Lagrange dual optimization approach to enforce the regulatory PSD mask and per-line total power constraints. The Lagrange multipliers are used to scale the dual MAC. The optimality of this technique can be shown using the results in [11].

In [13], the linear zero-forcing (ZF) pre-compensator (ZFP) has been proposed to cancel crosstalk in DS DSL and it is has been shown that it is near-optimal in DSM 3 due to the row-wise diagonal dominance (RWDD) property [14] of the DS DSL channel. By using the linear ZFP in each group and applying OSB over all users, the grouped ZFP-OSB algorithm has been derived for the joint DSM 2/3 scenarios [15]. As we will show in this paper, grouped ZFP-OSB does not necessarily achieve the highest possible rates even under the aforementioned practical assumptions, because the ZF solution is selfish and does not take into account the interests of the users in other groups. A related solution is based on the non-linear vector DPC (e.g., the vector Tomlinson-Harashima pre-coder (THP) [14]) with ZF transmit filters to encode the signals inside the groups. The ZF-THP is more complicated than the linear ZFP and achieves higher rates than the linear ZFP's in the general case, e.g., in the wireless channel. However, as it is shown in [13], it achieves almost the same rates as the linear ZFP in DS DSL channels when the PSD of the users is set by static spectrum management (SSM). The main benefit of the linear/nonlinear grouped ZFP techniques is their small implementation complexity, as their transmit filters can be calculated directly from the channel transfer matrix and are independent of the transmit power and service requirements of the users.

In this paper, we consider the design of grouped linear/nonlinear vector encoding with optimal transmit filters for the joint DSM 2/3 scenario in DS DSL. The optimal transmit filter for each user takes into account the overall benefit of all users. In order to calculate the optimal transmit filters, we develop a generalized duality theory between the MAC and BC with partial signal coordination and with per-line total power and nominal PSD mask constraints[1]. This solution can be considered as a generalization of the BC-MAC duality theory with per-transmitter

[1]Note that the current BC-MAC duality theory is applicable only when full signal coordination is available among the users.

power constraints in [11]. Using this theory, the optimal transmit filters for the grouped linear pre-compensator or the grouped non-linear THP can be calculated. Our simulation results show that by using the THP with optimal transmit filters for each group together with OSB over all users, considerably higher rates are achieved compared to the other available schemes. We refer to the resulting algorithm as the IF/BC-OSB algorithm. The IF/BC-OSB algorithm encompasses the OSB (i.e. IF-OSB) [4] and the BC-OSB algorithm [12] as special cases.

Finally, when the number of users is large, the IF/BC-OSB algorithm is impractical because of a high computational complexity. To resolve this complexity problem, we propose a simplified algorithm with a much lower computational complexity. We discuss the potential rate loss by using this sub-optimal algorithm and we show that the simplified algorithm is nearly optimal using analytical and simulation results.

The paper is organized as follows. In Sec. II, we describe the grouped DS DSL transmission model. Then, we develop our generalized BC-MAC duality theory and we propose an optimal joint DSM 2/3 algorithm for DS DSL accordingly in Sec. III. We also propose a simplified version of the algorithm with a near-optimal performance and reduced computational complexity. Simulation results are provided in Sec. IV. Finally the paper is concluded in Sec. V.

II. Grouped DS DSL Transmission Model

A. Transmission Model

We consider discrete multi-tone (DMT) transmission with K tones. We assume that N managed users form G groups with N_g users in group g ($g = 1 \dots G$). We assume that the users in each group are coordinated at the transmit signal level and all managed users are coordinated with each other at the transmit spectrum (i.e. PSD) level. All of the users are assumed to be DMT frame synchronized at the receiver side. Therefore, the transmission over tone k can be modeled as

$$\boldsymbol{y}^k = \left[\boldsymbol{H}^k\right]^{\mathrm{H}} \boldsymbol{x}^k + \boldsymbol{z}^k, \tag{2.1}$$

where $[\cdot]^{\mathrm{H}}$ denotes the conjugate transpose operation[2], $\boldsymbol{x}^k \equiv \left(\left[\boldsymbol{x}_1^k\right]^{\mathrm{T}}, \dots, \left[\boldsymbol{x}_G^k\right]^{\mathrm{T}}\right)^{\mathrm{T}}$, $\boldsymbol{y}^k \equiv \left(\left[\boldsymbol{y}_1^k\right]^{\mathrm{T}}, \dots, \left[\boldsymbol{y}_G^k\right]^{\mathrm{T}}\right)^{\mathrm{T}}$, and $\boldsymbol{z}^k \equiv \left(\left[\boldsymbol{z}_1^k\right]^{\mathrm{T}}, \dots, \left[\boldsymbol{z}_G^k\right]^{\mathrm{T}}\right)^{\mathrm{T}}$ are transmitted, received, and noise vector, respectively, and

$$\boldsymbol{H}^k \equiv \begin{bmatrix} \boldsymbol{H}_{11}^k & \cdots & \boldsymbol{H}_{1G}^k \\ \vdots & \ddots & \vdots \\ \boldsymbol{H}_{G1}^k & \cdots & \boldsymbol{H}_{GG}^k \end{bmatrix}. \tag{2.2}$$

[2]Following the literature, we define the IF/BC matrix as $\left[\boldsymbol{H}^k\right]^{\mathrm{H}}$. As we will see later, the dual IF/MAC matrix will then be $\boldsymbol{H}^k$ which simplifies the notation.

We indicate the n-th user in group g by (g,n) and we refer to it as user (g,n). Similarly, we refer to the line connecting this user to the DSL network as line (g,n). The sub-vectors $\boldsymbol{x}_g^k \equiv \left(x_{(g,1)}^k, \ldots, x_{(g,N_g)}^k\right)^{\mathrm{T}}$, $\boldsymbol{y}_g^k \equiv \left(y_{(g,1)}^k, \ldots, y_{(g,N_g)}^k\right)^{\mathrm{T}}$, and $\boldsymbol{z}_g^k \equiv \left(z_{(g,1)}^k, \ldots, z_{(g,N_g)}^k\right)^{\mathrm{T}}$ are the transmitted, received, and noise vector for group g, respectively, and the sub-matrix $\boldsymbol{H}_{gg'}^k$ is (the conjugate transpose of) the crosstalk channel matrix from group g to group g' on tone k with the n-th column denoted by $\boldsymbol{h}_{(gg',n)}^k$. We assume that the channel is normalized such that $\mathrm{E}\left\{\left|z_{(g,n)}^k\right|^2\right\} = 1$, where $\mathrm{E}\{\cdot\}$ denotes the expectation operator. The correlation between the elements of $\boldsymbol{z}^k$ can be nonzero (e.g., in the presence of unmanaged users) but is ignored as it cannot be exploited [12]. Each group g forms the following vector BC

$$\boldsymbol{y}_g^k = \left[\boldsymbol{H}_{gg}^k\right]^{\mathrm{H}} \boldsymbol{x}_g^k + \boldsymbol{\zeta}_g^k, \tag{2.3}$$

where $\boldsymbol{x}_g^k = \sum_{n=1}^{N_g} u_{(g,n)}^k \boldsymbol{q}_{(g,n)}^k$ is the transmitted vector in group g, $u_{(g,n)}^k$ and $\boldsymbol{q}_{(g,n)}^k$ are the data symbol and the transmit filter for user (g,n) on tone k, and

$$\boldsymbol{\zeta}_g^k = \boldsymbol{z}_g^k + \sum_{g' \neq g} \left[\boldsymbol{H}_{g'g}^k\right]^{\mathrm{H}} \boldsymbol{x}_{g'}^k \tag{2.4}$$

is the sum noise and crosstalk for group g. Finally, the transmit power for user (g,n) on tone k is $s_{(g,n)}^k \left|\left|\boldsymbol{q}_{(g,n)}^k\right|\right|^2$, where $s_{(g,n)}^k \equiv \mathrm{E}\left\{\left|u_{(g,n)}^k\right|^2\right\}$ and $||\cdot||$ denotes the ℓ_2 norm.

Equation (2.3) can be arranged into the following matrix form

$$\boldsymbol{y}^k = \left[\hat{\boldsymbol{H}}^k\right]^{\mathrm{H}} \boldsymbol{u}^k + \boldsymbol{z}^k, \tag{2.5}$$

where $\boldsymbol{u}^k \equiv \left(\left[\boldsymbol{u}_1^k\right]^{\mathrm{T}}, \ldots, \left[\boldsymbol{u}_G^k\right]^{\mathrm{T}}\right)^{\mathrm{T}}$, $\boldsymbol{u}_g^k \equiv \left(u_{(g,1)}^k, \ldots, u_{(g,N_g)}^k\right)^{\mathrm{T}}$,

$$\hat{\boldsymbol{H}}^k \equiv \left[\boldsymbol{Q}^k\right]^{\mathrm{H}} \boldsymbol{H}^k, \tag{2.6}$$

$$\boldsymbol{Q}^k \equiv \begin{bmatrix} \boldsymbol{Q}_1^k & 0 & \cdots & 0 \\ 0 & \boldsymbol{Q}_2^k & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \boldsymbol{Q}_G^k \end{bmatrix}, \tag{2.7}$$

and $\boldsymbol{Q}_g^k \equiv \left[\boldsymbol{q}_{(g,1)}^k \middle| \boldsymbol{q}_{(g,2)}^k \middle| \cdots \middle| \boldsymbol{q}_{(g,N_g)}^k\right]$.

B. Rate, Power, and SNR

In DSL, the number of bits that can be loaded on tone k of user (g,n) is given by

$$b_{(g,n)}^k = \min\left\{b_{\max}, \left\lfloor \log_2\left(1 + \tfrac{1}{\Gamma}\mathrm{SNR}_{(g,n)}^k\right)\right\rfloor\right\}, \tag{2.8}$$

where $b_{\max}$ is the maximum number of bits that can be loaded on a tone, $\lfloor\cdot\rfloor$ denotes the floor function, Γ is the SNR gap, and $\mathrm{SNR}_{(g,n)}^k$ is the SNR of user (g,n) on tone k. The SNR depends

on the structure of the vector encoders for each BC group. For linear encoders the SNR is given by

$$\mathrm{SNR}_{(g,n)}^{k} = \frac{s_{(g,n)}^{k} \left|\hat{h}_{(g,n)(g,n)}^{k}\right|^2}{1 + \sum_{n'=1;n'\neq n}^{N_g} s_{(g,n')}^{k} \left|\hat{h}_{(g,n')(g,n)}^{k}\right|^2 + \sum_{g'=1;g'\neq g}^{G} \sum_{n'=1}^{N_{g'}} s_{(g',n')}^{k} \left|\hat{h}_{(g',n')(g,n)}^{k}\right|^2}, \tag{2.9}$$

where by definition $\hat{h}_{(g',n')(g,n)}^{k}$ is the element of $\hat{\boldsymbol{H}}^{k}$ located on the row and the column corresponding to the lines (g',n') and (g,n), respectively. For non-linear vector DPC encoders (e.g., the vector THP [14]), the SNR depends on the transmit filters (i.e., the matrix $\boldsymbol{Q}^{k}$) as well as the encoding order. A vector DPC encodes the users one by one, such that the crosstalk originating from the previously encoded users is canceled. Assuming the encoding order is from the first user to the last user in each group, the SNR is given by

$$\mathrm{SNR}_{(g,n)}^{k} = \frac{s_{(g,n)}^{k} \left|\hat{h}_{(g,n)(g,n)}^{k}\right|^2}{1 + \sum_{n'=n+1}^{N_g} s_{(g,n')}^{k} \left|\hat{h}_{(g,n')(g,n)}^{k}\right|^2 + \sum_{g'=1;g'\neq g}^{G} \sum_{n'=1}^{N_{g'}} s_{(g',n')}^{k} \left|\hat{h}_{(g',n')(g,n)}^{k}\right|^2}. \tag{2.10}$$

The total bit rate of user (g,n) and transmit power on line (g,n) are respectively given by

$$R_{(g,n)} = f_{\mathrm{s}} \sum_{k=1}^{K} b_{(g,n)}^{k}, \tag{2.11}$$

and

$$P_{(g,n)} = \Delta_{\mathrm{f}} \sum_{k=1}^{K} p_{(g,n)}^{k}, \tag{2.12}$$

where f_{s} is the DMT symbol rate, Δ_{f} is the tone spacing,

$$p_{(g,n)}^{k} \equiv \left[\left[\boldsymbol{Q}_{g}^{k}\right]^{\mathrm{H}} \boldsymbol{S}_{g}^{k} \boldsymbol{Q}_{g}^{k} \right]_{n,n} \tag{2.13}$$

is the transmit power on tone k of line (g,n), $\boldsymbol{S}_{g}^{k} \equiv \mathrm{diag}\left\{\boldsymbol{s}_{g}^{k}\right\}$, $\boldsymbol{s}_{g}^{k} \equiv \left(s_{(g,1)}^{k}, \ldots, s_{(g,N_g)}^{k}\right)^{\mathrm{T}}$, and $\mathrm{diag}\left\{\boldsymbol{a}\right\}$ denotes a diagonal matrix with diagonal elements equal to the elements of vector $\boldsymbol{a}$. Note that the transmit power on line (g,n) is not equal to the transmit power of user (g,n) in the general case.

C. Problem Definition

The joint DSM 2/3 problem is defined as follows. Find the optimal encoder structure and the required transmit spectra (or equivalently $s_{(g,n)}^{k}$), in order to maximize

$$\sum_{g=1}^{G} \sum_{n=1}^{N_G} \omega_{(g,n)} R_{(g,n)} / f_{\mathrm{s}}, \tag{2.14}$$

subject to the following per-line total transmit power constraints

$$P_{(g,n)} \leq P_{(g,n)}^{\max}, \qquad \forall (g,n); 1 \leq g \leq G, 1 \leq n \leq N_g, \tag{2.15}$$

and/or the following per-line nominal PSD mask constraints

$$p_{(g,n)}^{k} \leq p_{(g,n)}^{k,\text{mask}}, \qquad \forall k, (g,n); 1 \leq k \leq K, 1 \leq g \leq G, 1 \leq n \leq N_g, \tag{2.16}$$

where the variable $\omega_{(g,n)} \geq 0$ is the (bit rate) weight factor for user (g,n), $P_{(g,n)}^{\max}$ is the total transmit power budget for line (g,n), and $p_{(g,n)}^{k,\text{mask}}$ is the nominal PSD mask for line (g,n) on tone k.

III. IF/BC Design and Optimization

The joint DSM 2/3 design consists of the encoder design inside the groups (the DSM 3 part) and spectrum balancing among all of the users in all groups (the DSM 2 part). A joint DSM 2/3 algorithm, namely, the *grouped ZFP-OSB* algorithm has been proposed in [15]. In this algorithm, a ZFP is used in each group on each tone. Then, the OSB algorithm is applied jointly over all users in all groups to calculate the transmit PSDs of the users maximizing the weighted sum rate in (2.14) under the total power and PSD mask constraints in (2.15) and (2.16). A similar structure is the *grouped ZF-THP-OSB*. The grouped ZF-THP-OSB is similar to the grouped ZFP-OSB, and is obtained by using a ZF-THP instead of the ZFP inside each group[3]. The structure of the ZF-THP for DSL has been explained in [14]. The QR decomposition of the channel matrix is used to implement the THP encoder efficiently in an iterative manner.

The transmit filters for the grouped ZFP and the grouped ZF-THP are obtained from the channel matrix inverse in each group. However, calculation of the optimal transmit filters for either the grouped linear pre-coder or the grouped non-linear THP is a non-convex problem. In this section, we solve this problem by using dual optimization techniques and generalizing the MAC/BC duality theory with per-line power constraints [11].

A. Dual Optimization

The joint DSM 2/3 problem defined by (2.14)-(2.16) is a non-convex problem with its complexity growing exponentially in the number of tones K (it is also exponential in the number of users N).

[3]The performance of a ZF-THP can vary depending on the order of encoding in the THP structure. In order to keep the structure of the grouped ZF-THP-OSB simple, we assume that the encoding order in each group g is determined by the order of the weight factors $\omega_{(g,1)}$ to $\omega_{(g,N_g)}$ where the user with the largest weight factor is encoded first and the user with the smallest weight factor is encoded last. This strategy is known to be optimal when the capacity gap is zero and optimal filters are used.

By applying dual optimization techniques[4], the complexity is reduced to a linear function of K. Applying dual optimization techniques [16,17], we obtain the following K per-tone optimization problems

$$\text{maximize } \mathcal{L}_k, \quad \text{for} \quad k = 1, \ldots, K, \tag{3.1}$$

where $\mathcal{L}_k$ is the Lagrangian on tone k defined by

$$\mathcal{L}_k \equiv \sum_{(g,n)} \omega_{(g,n)} b^k_{(g,n)} - \sum_{(g,n)} \left(\theta_{(g,n)} + \lambda^k_{(g,n)}\right) p^k_{(g,n)}, \tag{3.2}$$

where $\theta_{(g,n)} \geq 0$ and $\lambda^k_{(g,n)} \geq 0$ are the dual variables associated with the constraints in (2.15) and (2.16), respectively. These variables will be referred to as the Lagrange multiplier, and the per-tone Lagrange multiplier associated with line (g,n), respectively. The maximization in (3.1) is carried out over the structure of the encoders and the transmit powers of the users. For a linear encoder, the structure is described by the transmit filters $\boldsymbol{q}^k_{(g,n)}$. For a non-linear THP, it is described by the transmit filters as well as the order of encoding in each group.

B. Generalized BC-MAC Duality (with Partial Signal Coordination and Per-Line Power Constraints)

In this section, we generalize the well-known duality between (fully coordinated) BC and MAC scenarios to BC and MAC scenarios with partial signal coordination among users and per-line power constraints. Among the BC and MAC scenarios with partial signal coordination, the IF/BC and its dual IF/MAC will be studied in more detail due to its application in joint DSM 2/3 for DSL. The original BC-MAC duality states that with the same sum transmit power, the same set of multiuser SNRs (or bit rate) is achievable in a (flat single-carrier) BC and in its dual MAC, where the dual MAC matrix is simply the conjugate transpose of the primal BC matrix [7,8]. This result was later generalized in [11], to the case where per-transmitter power constraints are defined in the BC rather than a sum transmit power constraint. This form of duality is of a much higher interest in the DSL context as usually the relevant constraints in DSL are indeed the per-line total transmit power and nominal PSD mask constraints, as described by (2.15) and (2.16).

First we consider linear vector encoders (see Sec. III.C for the case of non-linear encoders). Our goal here is to calculate the optimal transmit filters $\boldsymbol{q}^k_{(g,n)}$ and the transmit power of the users (or alternatively $s^k_{(g,n)}$) maximizing $\mathcal{L}_k$ in (3.2). For ease of notation, we drop the tone index k and ignore the group index g and use a single index (e.g., i $(1 \leq i \leq N)$) to refer to the

[4]We will discuss the optimality of the dual optimization techniques for solving problem (2.14)-(2.16) in the appendix.

users. With these simplifications, (2.9) is written as

$$\mathrm{SNR}_i^{\mathrm{BC}} = \frac{s_i \left|\hat{h}_{ii}\right|^2}{1 + \sum_{j \neq i} s_j \left|\hat{h}_{ji}\right|^2}, \tag{3.3}$$

where $\hat{h}_{ji} \equiv \left[\hat{\boldsymbol{H}}\right]_{j,i}$, and $\hat{\boldsymbol{H}} \equiv \boldsymbol{Q}^{\mathrm{H}}\boldsymbol{H}$. From this, the required transmit powers $\boldsymbol{s} \equiv (s_1, s_2, \ldots, s_N)^{\mathrm{T}}$ to reach the SNR vector $(\gamma_1, \gamma_2, \ldots, \gamma_N)$ is obtained by [4]

$$\boldsymbol{s} = \boldsymbol{X}^{-\mathrm{T}}\mathbf{1}, \tag{3.4}$$

where $\mathbf{1}$ is the all one column vector of size N, which is the noise power vector in the normalized channel and

$$\boldsymbol{X} \equiv \begin{bmatrix} \gamma_1^{-1}\left|\hat{h}_{11}\right|^2 & -\left|\hat{h}_{12}\right|^2 & \cdots & -\left|\hat{h}_{1N}\right|^2 \\ -\left|\hat{h}_{21}\right|^2 & \gamma_2^{-1}\left|\hat{h}_{22}\right|^2 & & \vdots \\ \vdots & & \ddots & -\left|\hat{h}_{(N-1)N}\right|^2 \\ -\left|\hat{h}_{N1}\right|^2 & \cdots & -\left|\hat{h}_{N(N-1)}\right|^2 & \gamma_N^{-1}\left|\hat{h}_{NN}\right|^2 \end{bmatrix}. \tag{3.5}$$

Matrix $\boldsymbol{X}$ is ill-defined when any of the SNRs is zero. In this case, it is easy to see that the transmit power corresponding to users with zero SNR is zero. In order to calculate the power for the other users, the rows and columns of $\boldsymbol{X}$ corresponding to users with zero SNR are removed and then the so reduced (3.4) is used.

Now consider the following dual MAC

$$\tilde{\boldsymbol{y}} = \boldsymbol{H}\tilde{\boldsymbol{x}} + \tilde{\boldsymbol{z}}, \tag{3.6}$$

with noise covariance matrix

$$\mathrm{E}\left\{\tilde{\boldsymbol{z}}\tilde{\boldsymbol{z}}^{\mathrm{H}}\right\} = \boldsymbol{\Lambda} \equiv \mathrm{diag}\left\{\lambda_1, \lambda_2, \ldots, \lambda_N\right\}, \tag{3.7}$$

for given λ_i's and assume that the reception filters are $\boldsymbol{Q}^{\mathrm{H}}$. That is, the decision variables are obtained by

$$\hat{\boldsymbol{y}} = \boldsymbol{Q}^{\mathrm{H}}\boldsymbol{H}\tilde{\boldsymbol{x}} + \boldsymbol{Q}^{\mathrm{H}}\tilde{\boldsymbol{z}}. \tag{3.8}$$

The SNR for user i is calculated by

$$\mathrm{SNR}_i^{\mathrm{MAC}} = \frac{\tilde{s}_i \left|\hat{h}_{ii}\right|^2}{\sigma_i^2 + \sum_{j \neq i} \tilde{s}_j \left|\hat{h}_{ij}\right|^2}, \tag{3.9}$$

where $\tilde{s}_i$ is the transmit power for user i, and $\sigma_i^2 \equiv \left[\mathrm{E}\left\{\boldsymbol{Q}^{\mathrm{H}}\tilde{\boldsymbol{z}}\tilde{\boldsymbol{z}}^{\mathrm{H}}\boldsymbol{Q}\right\}\right]_{i,i} = \left[\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{\Lambda}\,\boldsymbol{Q}\right]_{i,i}$. Then the required transmit powers $\tilde{\boldsymbol{s}} \equiv (\tilde{s}_1, \tilde{s}_2, \ldots, \tilde{s}_N)^{\mathrm{T}}$ to reach the SNR vector $(\gamma_1, \gamma_2, \ldots, \gamma_N)$ is obtained by

$$\tilde{\boldsymbol{s}} = \boldsymbol{X}^{-1}\boldsymbol{\sigma}, \tag{3.10}$$

where $\boldsymbol{\sigma} \equiv (\sigma_1^2, \sigma_2^2, \ldots, \sigma_N^2)^{\mathrm{T}}$. It can be seen that when $\mathrm{SNR}_i^{\mathrm{MAC}} = \mathrm{SNR}_i^{\mathrm{BC}}$ the following relationship exists between the transmit powers in the dual MAC and the per-line transmit powers in the primal BC

$$\boxed{\sum_i \tilde{s}_i = \sum_i \lambda_i p_i}. \tag{3.11}$$

The proof is straightforward using (3.4) and (3.10) as follows

$$\begin{aligned} \textstyle\sum_i \tilde{s}_i &= \mathbf{1}^{\mathrm{T}}\tilde{\boldsymbol{s}} \overset{(3.10)}{=} \mathbf{1}^{\mathrm{T}}\boldsymbol{X}^{-1}\boldsymbol{\sigma} \overset{(3.4)}{=} \boldsymbol{s}^{\mathrm{T}}\boldsymbol{\sigma} = \mathrm{trace}\left\{\boldsymbol{S}\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{\Lambda}\boldsymbol{Q}\right\} \\ &= \mathrm{trace}\left\{\boldsymbol{\Lambda}\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{S}\boldsymbol{Q}\right\} = (\lambda_1, \lambda_2, \ldots, \lambda_N) \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_N \end{bmatrix} = \textstyle\sum_i \lambda_i p_i, \end{aligned}$$

where $\boldsymbol{S} \equiv \mathrm{diag}\{\boldsymbol{s}\}$ and $p_n = \left[\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{S}\boldsymbol{Q}\right]_{n,n}$. It should be noted that for any achievable set of SNRs in the primal BC, the calculated powers in the dual MAC are non-negative as long as λ_1 to λ_N are non-negative. To prove this, we first rewrite (3.4) as follows

$$\boldsymbol{s} = \left(\boldsymbol{I} - \boldsymbol{A}^{\mathrm{T}}\right)^{-1}\boldsymbol{b}, \tag{3.12}$$

where $\boldsymbol{I}$ is the identity matrix, $\boldsymbol{b} \equiv \left(\frac{\gamma_1}{\left|\hat{h}_{11}\right|^2}, \ldots, \frac{\gamma_N}{\left|\hat{h}_{NN}\right|^2}\right)^{\mathrm{T}}$ and $\boldsymbol{A} \equiv \boldsymbol{I} - \boldsymbol{X}\,\mathrm{diag}\,(\boldsymbol{b})$. Note that $\boldsymbol{A}$ is a non-negative matrix with diagonal elements equal to zero. If a positive solution to (3.12) exists, the Perron-Frobenius norm of $\boldsymbol{A}$ is smaller than 1, meaning that the inverse of $\boldsymbol{I} - \boldsymbol{A}^{\mathrm{T}}$ is an all positive matrix [7, 18, 19]. Thus $\boldsymbol{X}^{-1}$ is also an all positive matrix. Therefore, the solutions to (3.10) is positive for any positive set of λ_1 to λ_N.

Formula (3.11) holds for any partial signal coordination among the users. In fact the admitted signal coordination determines the non-zero elements of $\boldsymbol{Q}$, where $[\boldsymbol{Q}]_{i,j}$ is set to zero if users i and j are not coordinated at the signal level. For BCs with full signal coordination, there are no zeros in $\boldsymbol{Q}$. For the IF/BC scenario, $\boldsymbol{Q}$ consists of square sub-matrices $\boldsymbol{Q}_g$ $(1 \leq g \leq G)$ on its diagonal and the rest of its elements are zero, as shown in (2.7). In the general case, there is no restriction on the location of the zeros of $\boldsymbol{Q}$, other than there should be at least one non-zero element in each column of $\boldsymbol{Q}$. In DSL, normally the diagonal elements of $\boldsymbol{Q}$ are non-zero. An example, for which the zeros of $\boldsymbol{Q}$ could be irregularly located, is the partial crosstalk cancellation of DSL [20–22].

Now consider the maximization problem in (3.1)-(3.2). Consider the dual MAC defined by $\boldsymbol{H}^k$ and assume that its noise power for line (g,n) is $\theta_{(g,n)} + \lambda^k_{(g,n)}$. It is noteworthy that there is a one to one relationship between the primal BC bit loadings $b^k_{(g,n)}$ and the minimum SNRs needed to load them, i.e., $\gamma^k_{(g,n)} = \Gamma\left(2^{b^k_{(g,n)}} - 1\right)$. A similar formula holds for the dual MAC bit loadings $\tilde{b}^k_{(g,n)}$. Therefore, if use the same SNRs $\gamma^k_{(g,n)}$ in the primal BC and dual MAC, we can

load the same number of bits in the primal BC and dual MAC, i.e., $b^k_{(g,n)} = \tilde{b}^k_{(g,n)}$. Using (3.11), the Lagrangian in (3.2) can then be written as

$$\mathcal{L}_k = \tilde{\mathcal{L}}_k \equiv \sum_{(g,n)} \omega_{(g,n)} \tilde{b}^k_{(g,n)} - \sum_{(g,n)} \tilde{s}^k_{(g,n)}. \tag{3.13}$$

Thus, instead of solving (3.1), which is a non-convex problem, we can maximize (3.13). Maximizing (3.13) is not a convex problem in the dual IF/MAC channel either. However, for each set of bit loadings, the required transmit powers in the dual IF/MAC and the optimal reception filters can be calculated easily (Alg. 2 in [1]). Therefore, by running an exhaustive search over the space of bit loadings, the optimal solution in the dual channel can be found. It is also possible to run the exhaustive search over the space of transmit powers in the dual channel as explained in [1]. The optimal transmit filters in the dual MAC are the minimum mean squared error (MMSE) filters.

In the general case, the MMSE filters can be calculated as follows. Let $\boldsymbol{q}_i$ denote the i-th column of $\boldsymbol{Q}$ and assume M_i denotes the number of non-zero elements of $\boldsymbol{q}_i$. Assume the $M_i \times N$ matrix $\widehat{\boldsymbol{H}}_i$ is obtained by removing the rows of $\boldsymbol{H}$ corresponding to the zero elements of $\boldsymbol{q}_i$ and let $\widehat{\boldsymbol{h}}_i$ denote its i-th column. Similarly, assume the $M_i \times M_i$ matrix $\widehat{\boldsymbol{\Lambda}}_i$ is obtained by removing the rows and the columns of $\boldsymbol{\Lambda}$ corresponding to the zero elements of $\boldsymbol{q}_i$. Then the non-zero elements of $\boldsymbol{q}_i$ are calculated by

$$\widehat{\boldsymbol{q}}_i = \boldsymbol{C}_i^{-1}\, \widehat{\boldsymbol{h}}_i, \tag{3.14}$$

where $\boldsymbol{C}_i \equiv \widehat{\boldsymbol{H}}_i\, \tilde{\boldsymbol{S}}_i\, \widehat{\boldsymbol{H}}_i^{\mathrm{H}} + \widehat{\boldsymbol{\Lambda}}_i$ is the sum noise and interference covariance matrix for user i and $\tilde{\boldsymbol{S}} \equiv \mathrm{diag}\{\tilde{s}_1, \ldots, \tilde{s}_{i-1}, 0, \tilde{s}_{i+1}, \ldots, \tilde{s}_N\}$. The SNR is obtained as

$$\mathrm{SNR}_i = \tilde{s}_i\, \widehat{\boldsymbol{q}}_i^{\mathrm{H}}\, \boldsymbol{h}_i. \tag{3.15}$$

In joint DSM 2/3 scenarios, the MMSE filter for user (g,n) is calculated from (3.14) by

$$\tilde{\boldsymbol{q}}^k_{(g,n)} = \left(\boldsymbol{H}^k_g \tilde{\boldsymbol{S}}^k_{(g,n)} \left[\boldsymbol{H}^k_g\right]^{\mathrm{H}} + \boldsymbol{\Theta}_g + \boldsymbol{\Lambda}^k_g\right)^{-1} \boldsymbol{h}^k_{(gg,n)} \tag{3.16}$$

and the SNR is

$$\mathrm{SNR}^k_{(g,n)} = \tilde{s}^k_{(g,n)} \tilde{\boldsymbol{q}}^k_{(g,n)} \boldsymbol{h}^k_{(gg,n)}, \tag{3.17}$$

where $\boldsymbol{H}^k_g \equiv \left[\boldsymbol{H}^k_{g1} \middle| \boldsymbol{H}^k_{g2} \middle| \cdots \middle| \boldsymbol{H}^k_{gG}\right]$, $\boldsymbol{h}^k_{(gg,n)}$ is the n-th column of $\boldsymbol{H}^k_{gg}$, $\tilde{\boldsymbol{S}}^k \equiv \mathrm{diag}\left\{\tilde{s}^k_{(1,1)}, \ldots, \tilde{s}^k_{(g,n-1)}, 0, \tilde{s}^k_{(g,n+1)}, \ldots, \tilde{s}^k_{(G,N_G)}\right\}$, $\boldsymbol{\Theta}_g \equiv \mathrm{diag}\left\{\theta_{(g,1)}, \ldots, \theta_{(g,N_g)}\right\}$, and $\boldsymbol{\Lambda}^k_g \equiv \mathrm{diag}\left\{\lambda^k_{(g,1)}, \ldots, \lambda^k_{(g,N_g)}\right\}$. At the termination of the exhaustive search the optimal solution found is used to calculate the required per-user and per-line transmit powers in the primal BC using (3.4) and (2.13).

C. Non-Linear Encoder and Decoder Structures

So far we have studied the general duality between the BC and MAC assuming linear encoders and decoders. By using the non-linear vector DPC encoder, the contribution of the crosstalk for the users that are already encoded in each group is removed in the SNR (compare (2.10) to (2.9)), the effect of which is that the corresponding entries in matrix $\boldsymbol{X}$ (defined in (3.5)) will be zero. Consequently

$$\boldsymbol{s} = \overline{\boldsymbol{X}}^{-\mathrm{T}}\mathbf{1}, \tag{3.18}$$

where

$$\left[\overline{\boldsymbol{X}}\right]_{i,j} \equiv \begin{cases} 0, & \text{if } i \text{ and } j \text{ are coordinated at the signal level and } i \text{ is encoded before } j \\ [\boldsymbol{X}]_{i,j}, & \text{otherwise.} \end{cases} \tag{3.19}$$

The BC-MAC duality will not hold unless we use the *dual* decoder structure of the vector DPC encoder in the MAC. The dual of the vector DPC encoder is the generalized decision feedback equalizer (GDFE) [7]. Assuming the same combination of signal coordination among users, by using a GDFE structure in the dual MAC and setting the decoding order as the reverse of the encoding order in the primal BC, we will be able to restore the BC-MAC duality. The GDFE removes the crosstalk originating from the already decoded users, resulting in exactly the same zeros in $\boldsymbol{X}$ as those resulting from the DPC encoder operation in the primal BC. Therefore,

$$\tilde{\boldsymbol{s}} = \overline{\boldsymbol{X}}^{-1}\boldsymbol{\sigma}. \tag{3.20}$$

Equations (3.18) and (3.20) portray the duality between the BC using the DPC encoder and the MAC using the GDFE with partial signal coordination. Again, the MMSE filters maximize the SNR and are optimal. In the general case, the non-zero elements of the MMSE-GDFE filter for user i are calculated by

$$\widehat{\boldsymbol{q}}_i = \overline{\boldsymbol{C}}_i^{-1}\,\widehat{\boldsymbol{h}}_i, \tag{3.21}$$

where $\overline{\boldsymbol{C}}_i \equiv \widehat{\boldsymbol{H}}_i\,\overline{\boldsymbol{S}}_i\,\widehat{\boldsymbol{H}}_i^{\mathrm{H}} + \widehat{\boldsymbol{\Lambda}}_i$, $\overline{\boldsymbol{S}}_i \equiv \mathrm{diag}\left\{\overline{s}_{i,1},\ldots,\overline{s}_{i,i-1},0,\overline{s}_{i,i+1},\ldots,\overline{s}_{i,N}\right\}$, and

$$\overline{s}_{i,j} \equiv \begin{cases} 0, & \text{if } i \text{ and } j \text{ are coordinated at the signal level and } i \text{ is decoded after } j \\ \tilde{s}_j, & \text{otherwise.} \end{cases} \tag{3.22}$$

In joint DSM 2/3 scenarios, assuming that the decoding order is from the first user to the last user in each group, the MMSE-GDFE filter for user (g,n) is calculated by

$$\tilde{\boldsymbol{q}}^k_{(g,n)} = \left(\boldsymbol{H}^k_g\overline{\boldsymbol{S}}^k_{(g,n)}\left[\boldsymbol{H}^k_g\right]^{\mathrm{H}} + \boldsymbol{\Theta}_g + \boldsymbol{\Lambda}^k_g\right)^{-1}\boldsymbol{h}^k_{(gg,n)}, \tag{3.23}$$

where $\overline{\boldsymbol{S}}^k \equiv \mathrm{diag}\left\{\tilde{s}^k_{(1,1)},\ldots,\tilde{s}^k_{(g-1,N_{g-1})},\overbrace{0,\ldots,0}^{n\text{ times}},\tilde{s}^k_{(g,n+1)},\ldots,\tilde{s}^k_{(G,N_G)}\right\}$.

D. IF/BC-OSB Algorithm

A pseudocode representation of the optimal joint DSM 2/3 algorithm in DS direction, referred to as the IF/BC-OSB algorithm, is given in Alg. 1. The IF/BC-OSB algorithm finds the optimal transmit powers based on OSB over all users as well as the optimal encoding order and transmit filters for vector THPs per group. The algorithm consists of six nested loops. In the first loop, the Lagrange multipliers corresponding to the per-line total power constraints ($\theta_{(g,n)}$) are optimized. In the second loop, the tone index is selected in order to solve the corresponding per-tone optimization problem. In the third loop, the per-tone Lagrange multipliers corresponding to the nominal PSD mask constraints ($\lambda^k_{(g,n)}$) are optimized. In the fourth loop, an exhaustive search is carried out to find the optimal value of the transmit powers in the dual IF/MAC channel ($\tilde{s}^k_{(g,n)}$). The two inner-most loops test all decoding orders in the dual IF/MAC channel to find those resulting in the largest per-tone Lagrangian. The outer loop of the two selects the group index g and then the optimal reception filters ($\tilde{\boldsymbol{q}}^k_{(g,n)}$) and the corresponding bit loadings ($\tilde{b}^k_{(g,n)}$) are calculated in the inner loop based on the parameter values set in the outer loops.

Note that, given the transmit powers in the dual IF/MAC, the optimal decoding order in each group is independent of the decoding order in other groups. This allows us to optimize the decoding order of each group independently of the other groups, which results in a considerable computational complexity reduction. To implement this, the weighted sum rate in each group is optimized over all decoding orders and the optimal solution is stored in $\mathcal{W}^{\max}_g$. Then the Lagrangian is calculated by $\mathcal{L}_k = \sum_g \mathcal{W}^k_{g,\max} - \sum_n \tilde{s}^k_{(g,n)}$. This Lagrangian is partially optimized as the decoding orders in the groups are optimized. The global optimum is calculated by searching through the transmit powers in an outer loop. The optimal solution is later used to calculate the per-line transmit powers in the primal IF/BC, i.e., $p^k_{(g,n)}$.

When the sum of the Lagrange multipliers for a user ($\theta_{(g,n)}+\lambda^k_{(g,n)}$) is zero, the noise covariance matrix will be rank-deficient during the exhaustive search on the transmit powers for some particular sets of transmit powers and decoding orders. As a result the reception filters could not be calculated using (3.23). In order to avoid this problem, we can simply assume $\theta_{(g,n)}$ (or $\theta_{(g,n)} + \lambda^k_{(g,n)}$) are lower bounded by a sufficiently small positive constant.

E. Computational Complexity Reduction

The computational complexity of Alg. 1 is $\mathcal{O}\left(\beta_1 K \beta_2 \alpha^N \sum_{g=1}^{G} N_g! N_g^3\right)$, where β_1 is the number of iterations required for optimizing $\left\{\theta_{(g,n)}; \forall(g,n)\right\}$, K is the number of iterations of the second loop, β_2 is the number of iterations required for optimizing $\left\{\lambda^k_{(g,n)}; \forall(g,n)\right\}$, α^N is the size of

the dual transmit power search space assuming α distinct power levels for each user, $N_g!$ is the number of decoding orders in group g, and N_g^3 is the computational complexity of calculating the transmit filters $\tilde{\boldsymbol{q}}_{(g,n)}^k$ and bit loadings $\tilde{b}_{(g,n)}^k$ in group g. Note that when the GDFE is used, the calculation of the transmit filters for each user in group g involves a matrix inverse operation, each time with complexity $\mathcal{O}\left(N_g^3\right)$. However, using the matrix inversion lemma, it is possible to calculate all filters in a group with $\mathcal{O}\left(N_g^3\right)$ complexity in total [23]. The number of iterations required for optimizing the Lagrange multipliers, i.e., β_1 and β_2, depends on the employed algorithm. Using the iterative facet dividing algorithm [24], we have $\beta_1, \beta_2 = \mathcal{O}\left(N \log \frac{1}{\epsilon}\right)$, where ϵ is the desired precision.

The overall computational complexity of Alg. 1 is large when the number of users is large. In order to reduce the complexity we can apply (all or some of) the following three simplifying steps.

First, we can use a computationally efficient spectrum balancing algorithm such as the iterative spectrum balancing (ISB) instead of the OSB [4]. By doing so, the α^N factor, which is exponential in the number of users, is replaced by a (degree one to three) polynomial factor in N and α.

Second, instead of testing all decoding orders, the decoding order in the dual IF/MAC can be set according to the order of weights $\omega_{(g,n)}$ such that the user with the smallest weight is decoded first and the user with the largest weight is decoded last. Despite being optimal for ideal systems with zero capacity gap, this approach is not optimal in the general case. However, in [1] this has been shown to be near-optimal for IF/MAC channels in the presence of spectrum balancing. This will replace the factor $\sum_{g=1}^{G} N_g! N_g^3$ in the computational complexity by $\sum_{g=1}^{G} N_g^3$. Alternatively, one can use the computationally efficient algorithms proposed in [25] and [23] for finding the optimal decoding order, which are near-optimal in the general case.

Third, we can dismiss the optimization of $\lambda_{(g,n)}^k$ and set them to a sufficiently small constant (e.g., one) for all users (g, n). This will reduce the computational complexity with only a very small loss in the achieved bit rates. Aside from that, this improves the robustness of the algorithm against numerical errors because finding the Lagrange multipliers to meet the PSD mask constraints tightly proves to be a heavily error-prone process in typical DSL channels. We explain this for a special case in the following:

Consider the normalized 2×2 BC defined by $\boldsymbol{H} \equiv \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}$ and assume we aim to minimize the sum power $\lambda_1 p_1 + \lambda_2 p_2$ subject to $\mathrm{SNR}_1 \geq \gamma_1$ and $\mathrm{SNR}_2 \geq \gamma_2$. By solving the equations in Sec. III analytically, we have calculated the per-line transmit powers for a DPC

encoder with encoding order $\{1, 2\}$ as a function of λ_1 and λ_2 as follows:

$$p_1 = \gamma_1 \left| \frac{\lambda_2 a_1 h_{11} + \gamma_2 \left(h_{11} |h_{22}|^2 - h_{12} h_{21} h_{22}^* \right)}{a_1 a_2 (a_4 h_{11} + a_3 h_{21})} \right|^2 + \gamma_2 \frac{|h_{12}|^2}{\lambda_1^2 a_1^2} \left(1 + \gamma_1 \left| \frac{a_4 h_{12} + a_3 h_{22}}{a_4 h_{11} + a_3 h_{21}} \right|^2 \right) \tag{3.24}$$

$$p_2 = \gamma_1 \left| \frac{\lambda_1 a_1 h_{21} + \gamma_2 \left(h_{21} |h_{12}|^2 - h_{11} h_{22} h_{12}^* \right)}{a_1 a_2 (a_4 h_{11} + a_3 h_{21})} \right|^2 + \gamma_2 \frac{|h_{22}|^2}{\lambda_2^2 a_1^2} \left(1 + \gamma_1 \left| \frac{a_4 h_{12} + a_3 h_{22}}{a_4 h_{11} + a_3 h_{21}} \right|^2 \right) \tag{3.25}$$

where $a_1 \equiv \frac{|h_{12}|^2}{\lambda_1} + \frac{|h_{22}|^2}{\lambda_2}$, $a_2 \equiv \lambda_1 \lambda_2 (1 + \gamma_2)$, $a_3 \equiv a_5^{-1} (\lambda_1 h_{21}^* a_1 + \gamma_2 h_{21}^* |h_{12}|^2 - \gamma_2 h_{12} h_{11}^* h_{22}^*)$, $a_4 \equiv a_5^{-1} (\lambda_2 h_{11}^* a_1 + \gamma_2 h_{11}^* |h_{22}|^2 - \gamma_2 h_{22} h_{12}^* h_{21}^*)$, $a_5 \equiv \lambda_1 \lambda_2 a_1 + \lambda_1 \gamma_2 |h_{22}|^2 + \lambda_2 \gamma_2 |h_{12}|^2$, and $(\cdot)^*$ denotes complex conjugate operation. The limiting powers are obtained by

$$\begin{array}{ll} \lim\limits_{\lambda_1 \to 0^+} p_1 = \frac{\gamma_1(\gamma_2+1)}{|h_{11}|^2} + \frac{\gamma_2}{|h_{12}|^2}, & \lim\limits_{\lambda_1 \to 0^+} p_2 = 0, \\ \lim\limits_{\lambda_2 \to 0^+} p_1 = 0, & \lim\limits_{\lambda_2 \to 0^+} p_2 = \frac{\gamma_1(\gamma_2+1)}{|h_{21}|^2} + \frac{\gamma_2}{|h_{22}|^2}. \end{array} \tag{3.26}$$

For a two-user DSL scenario with loop lengths 400 and 800 m, the normalized channel matrix at the carrier frequency of 6 MHz is given by $|\boldsymbol{H}|^2 = -\begin{bmatrix} 25.8 & 61.4 \\ 81.1 & 48.4 \end{bmatrix} + 170$ dB. Figure 1 shows p_1 vs. p_2 in dB and linear scales to reach $\gamma_1 = \gamma_2 = 39.55$ dB which is the required SNR to load 10 bits assuming that $\Gamma = 9.45$ dB. As it can be seen, although p_1 and p_2 fluctuate vividly with respect to each other in the dB scale; the curve looks like two vertical and horizontal line sections with a common end in the linear scale. Geometrically, the process of minimizing $\lambda_1 p_1 + \lambda_2 p_2$ can be interpreted as finding the tangency point of a line with slope $-\lambda_1/\lambda_2$ with the curve in the linear scale. Since the curve has a slope that is either extremely close to 0 or extremely close to infinite (corresponding to the two line sections), the process is highly sensitive to numerical errors. Moreover, when any of the Lagrange multipliers tends to zero, the noise correlation matrix in the dual channel will be rank deficient during the exhaustive search over power-loadings and decoding orders. This makes it difficult to calculate the MMSE filters (and all subsequent parameters such as transmit powers in the dual and primal channels) accurately.

Fortunately, as we will see in Sec. IV, assuming that the Lagrange multipliers corresponding to the PSD mask constraints ($\lambda_{(g,n)}^k$) are equal to a sufficiently small positive constant (e.g., one) is nearly optimal for practical DSL. Note that by removing the loop for optimizing $\lambda_{(g,n)}^k$, the per-line transmit powers may not necessarily satisfy the nominal PSD mask constraints. Therefore, before calculation of the per-tone Lagrangian $\mathcal{L}_k$, we should check whether $p_{(g,n)}^k \leq p_{(g,n)}^{k,\,\text{mask}}$ for all (g, n).

Note that this discussion only applies to the per-tone power regions and determination of $\lambda_{(g,n)}^k$. A similar problem could also happen in determining the Lagrange multipliers corresponding to the per-line total power constraints, i.e., $\theta_{(g,n)}$ as shown in [24]. However, the frequency sharing solution considered in [4] or the generalized IFDA proposed in [24] can be used to reach

the optimal solution with the desired precision without necessarily finding the optimal Lagrange variables precisely.

Taking all of the aforementioned points into account, we obtain the simplified algorithm listed in Alg. 2.

IV. Simulation Results

For a three-user VDSL2 scenario depicted in Fig. 2, we have simulated the proposed algorithms (Alg. 1 and Alg. 2) as well as several other algorithms or schemes, including grouped ZF-OSB, grouped ZF-THP-OSB, grouped optimal linear pre-coder (OLP)-OSB, OSB, and static spectrum management (SSM). These schemes are listed in Table I. The scenario consists of two groups of VDSL2 E17 loops [26]. The first group consists of an 800 m loop and a 400 m loop and the second group consists of a single 400 m loop. Simulation parameters are given in Table II.

Figure 3 shows the average bit rate of the lines in the two groups versus each other. To obtain this figure, we have executed the algorithms for $\omega_{(1,1)} = \omega_{(1,2)} = w$ and $\omega_{(2,1)} = 1 - w$, where w is swept from 0 to 1 in 0.05 steps. The performance of sub-optimal spectrum balancing algorithms such as ISB has been comprehensively studied in the literature. Therefore, in this figure, we have simulated the simplified IF/BC-OSB algorithm (and all other algorithms involving spectrum balancing) using the OSB, so that we could study the effect of our other simplifications (i.e., setting the encoding order according to the order of weights and fixing $\lambda^k_{(g,n)}$ to 1) on the optimality of the algorithm more precisely.

As it can be seen, the IF/BC-OSB algorithm and its simplified version achieve considerably higher bit rates compared to the grouped ZFP-OSB, grouped ZF-THP-OSB, and grouped OLP-OSB algorithms. This is a paradoxical result as the ZFP is known to be near-optimal in DSL [13], and ZF-THP and OLP are obviously superior to ZFP. Moreover, the same spectrum balancing algorithm (i.e., OSB) is used for all these algorithms. The combination of these near-optimal encoding schemes with OSB, however, is sub-optimal in IF/BC scenarios. This can be explained as follows. For any set of transmit powers, the ZFP or ZF-THP in the first group acts (near-) optimally if we consider the benefit of the users in the same group only. However, the resulting crosstalk power to the lines in the second group has not been taken into account and could reduce the bit rate of the line in that group. On the other hand, the IF/BC-OSB algorithm optimizes the transmit powers and transmit filters by taking into account the benefit of the users in one group as well as the users outside the group. Therefore, by proper design of the transmit filters, the crosstalk could be effectively canceled not only for the users in one group but also

for the users belonging to the other groups.

For the grouped OLP-OSB algorithm, the transmit filters are also calculated by taking into account the benefit of all users. However, the performance of the algorithm is still inferior to IF/MAC-OSB. This is justified by a deeper investigation into the structures of OLP and optimal THP. In the OLP structure, each line in group 1 receives crosstalk from the other user in group 1 and the user in group 2. The users are cooperative and try to cancel the crosstalk imposed by each user in the group to other users in the group as well as those outside the group. However, the available number of filter taps for each user (the degrees of freedom) is equal to the number of users in the group and hence smaller than the total number of users. This prevents the users from canceling the crosstalk effectively for all users particularly the crosstalk imposed to the users outside the group. However, the optimal THP structure in IF/BC-OSB pre-cancels the crosstalk for previously encoded users in each group, leaving a few degrees of freedom for each user to pre-cancel the crosstalk for the users outside the group. For example, assuming that for the simulated scenario the first user in group 1 (i.e., user (1,1)) is encoded first and the second user in group 1 (i.e., user (1, 2)) is encoded second in the THP structure, user (1, 2) can choose its filter taps to perfectly cancel the crosstalk from the users in group 1 into group 2. Therefore, the user in group 2 (i.e., user (2,1)) requires much less power to achieve a desirable bit rate. This in turn reduces the crosstalk received by the users in group 1 from user (2, 1), resulting in an increase in their achievable bit rates as well.

As it can be seen in Fig. 3, the performance of the simplified IF/BC-OSB algorithm is almost exactly the same as for IF/BC-OSB. The difference between the achieved bit rate is less than 0.3%. To understand this, let us consider two extreme IF/BC scenarios, namely the IF channel and the BC. In IF channels, when a desired SNR vector is reachable, a vector of minimal power exists which is optimal for all positive values of $\lambda_i, i = 1, \ldots, N$, meaning that changing λ_i has absoloutly no effect on the transmit powers. In BCs, unlike in IF channels, it is possible to reach a particular SNR vector by different sets of transmit powers. However, if the power constraints are not met at the minimum sum power solution (i.e., the one obtained by setting $\lambda_1 = \lambda_2 = \ldots = \lambda_N$) for a particular bit loading vector, it is unlikely to find a feasible solution to reach the bit loading vector by changing $\lambda_i, i = 1, \ldots, N$. This is because in DSL BCs, most of the desired signal for a user arrives through the direct channel and a significantly smaller portion of it arrives through the crosstalk channel, as the crosstalk channel is much weaker than the direct channel. Consider a 2×2 DSL BC and assume for a particular bit loading, the transmit power on line 2 corresponding to the minimum sum power solution is larger than the nominal

PSD mask. If we decrease the power on line 2 while the SNR for user 2 is kept at the same value, the power share of user 2 on line 1 should be increased significantly to compensate for the SNR loss. This would most likely cause the power on line 1 to go above the nominal PSD mask. Here, we provide some numerical results for channel $\boldsymbol{H}$ in Sec. III.E to clarify this point.

The required transmit powers on lines 1 and 2 to load 10 bits for both users is plotted in Fig. 4 as a function of λ_1/λ_2. The nominal PSD mask for VDSL2 at 6 MHz is approximately -56 dBm [26]. As it can be seen, p_1 is smaller and p_2 is larger than the nominal PSD mask at $\lambda_1/\lambda_2 = 1$. By decreasing λ_1/λ_2, p_2 will decrease slowly until it goes below the nominal PSD mask at $\lambda_1/\lambda_2 = 0.08183$. However, at the same time, p_1 increases with a much faster rate such that it goes above the nominal PSD mask at $\lambda_1/\lambda_2 = 0.3157$, i.e., before p_2 goes below the nominal PSD mask. As a result, there is no value of λ_1/λ_2 for which both lines meet the nominal PSD mask constraint. In fact, by decreasing λ_1/λ_2, the (relative) penalty for the power on line two in the Lagrangian, i.e., $\lambda_2 p_2$ increases. Therefore, by modifying the transmit filter for user 2, some of its power is sent on the first line to be received at the user's receiver front-end through the crosstalk channel. However, since the crosstalk channel is much weaker that the direct channel, the required increase in the power of line 1 to retain the same SNR for loading 10 bits is much larger than the decrease in the power on line 2. As a result p_1 goes above the nominal PSD mask before p_2 goes below the nominal PSD mask.

Since the degree of coordination among users in IF/BC is something between those for IF and BC, we expect that in most IF/BC scenarios, the simplified algorithm achieves bit rate close to the optimal solution as well.

V. Conclusion

In this paper, we have studied optimal joint DSM 2/3 in DS DSL. We have considered THP for signal encoding inside the groups and the OSB algorithm for spectrum balancing among all users. Calculation of the optimal transmit filters for the THP is a non-convex problem. To resolve this problem, we have extended the BC-MAC duality to scenarios with partial signal coordination and per-line power constraints. Based on this, we have proposed the IF/BC-OSB algorithm. We have also proposed a simplified version of the algorithm with practical complexity and near-optimal performance. We have compared the performance of the proposed algorithm with several other joint DSM 2/3 algorithms, namely, grouped ZF-OSB, grouped ZF-THP-OSB, and grouped OLP-OSB as well as the OSB algorithm and the SSM. Our simulation results have shown that the proposed algorithm performs significantly better than the OSB and

SSM. Moreover, it outperforms the other joint DSM 2/3 algorithms with a considerable margin. Our investigation have shown that this can be attributed to the grouped optimal THP structure which is capable of pre-canceling the crosstalk generated by the users in a group to the users inside the same group as well as part of the crosstalk to the users in other groups.

Appendix

A. A Brief Discussion on Optimality of IF/BC-OSB

The proposed IF/BC-OSB algorithm uses dual optimization techniques to solve the IF/BC problem (2.14)-(2.16). In order to show that this solution is optimal, we have to show that the duality gap is zero. The application of the dual optimization techniques to solve the IF/BC problem can be separated into two main steps. The first step corresponds to applying the dual technique on the per-line total power constraints in (2.15). The second step corresponds to applying it on the nominal PSD mask constraints in (2.16). To show the optimality, we discuss that the duality gap remains zero in both steps.

By applying the dual optimization technique merely on the per-line total power constraints in (2.15), we obtain the following constrained optimization problem on each tone k

$$\text{maximize} \quad \sum_{(g,n)} \omega_{(g,n)} b^k_{(g,n)} - \sum_{(g,n)} \theta_{(g,n)} p^k_{(g,n)} \tag{A.1}$$

$$\text{subject to} \quad p^k_{(g,n)} \leq p^{k,\text{mask}}_{(g,n)}, \qquad \forall\ (g,n); 1 \leq g \leq G, 1 \leq n \leq N_g, \tag{A.2}$$

which will be transformed into (3.2), if we apply the dual optimization technique on the constraints in (A.2) too. The optimality of the first step comes from the general optimality of dual techniques for MIMO-OFDM problems when the number of tones is large [4, 17, 27]. This is a well-known property of dual techniques and for this reason we do not discuss it here. To prove the optimality of the second step, consider the fourth loop in Alg. 1, i.e., the exhaustive search over the transmit powers. The exhaustive search over the transmit powers can be replaced by an exhaustive search over the bit loadings. Then the order of this loop and the third loop, i.e., the search over the Lagrange multipliers $\lambda^k_{(g,n)}$, can be exchanged. The transmit powers in the dual channel can be calculated as a function of the bit loadings by Alg. 2 in [1]. Since changing the decoding order in each group changes the transmit power of the users and consequently the crosstalk power into other groups, the decoding order in all groups should be optimized together. Therefore, in order not to lose the optimality, the fifth and sixth loops in Alg. 1 should be merged into a single loop for checking all decoding orders in all groups.

By performing these changes, part of the algorithm inside the loop corresponding to the exhaustive search over the bit loadings space solves the dual of the following optimization problem for the bit loading vector $\left(b^k_{(1,1)}, \ldots, b^k_{(G,N_G)}\right)$

$$\begin{array}{ll} \text{minimize} & \sum_{(g,n)} \theta_{(g,n)} p^k_{(g,n)} \\ \text{subject to} & \mathrm{SNR}^k_{(g,n)} \geq \gamma^k_{(g,n)} \equiv \Gamma\left(2^{b^k_{(g,n)}} - 1\right) \\ \text{and} & p^k_{(g,n)} \leq p^{k,\mathrm{mask}}_{(g,n)}. \end{array} \tag{A.3}$$

Note that for a fixed bit loading vector, the term $\sum_{(g,n)} \omega_{(g,n)} b^k_{(g,n)}$ is constant and has been ignored in (A.3). The optimality of the solution can be shown if we show that the duality gap is zero for (A.3). In [11] the following BC optimization problem

$$\begin{array}{ll} \underset{\alpha,\ s_i,\ \boldsymbol{q}_i}{\text{minimize}} & \alpha \sum_{i=1}^{N} p_i \\ \text{subject to} & \mathrm{SNR}_i \geq \gamma_i \\ \text{and} & p_i \leq \alpha P_i^{\max} \end{array} \tag{A.4}$$

has been solved using dual techniques and a rigorous proof of the optimality is provided for the case where the users are fully coordinated at the signal level. The proof can be extended to show that the duality gap is zero for the dual of (A.3). This means that the second application of the dual techniques is also optimal. We omit the proof here for the sake of brevity.

References

[1] A. R. Forouzan, M. Moonen, J. Maes, and M. Guenach, "Joint level 2 and 3 dynamic spectrum management for upstream VDSL," *IEEE Trans. Commun.*, 2011, to be published. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5957380

[2] R. H. Etkin, D. N. C. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inform. Theory*, vol. 54, no. 12, pp. 5534–5562, Dec. 2008.

[3] A. S. Motahari and A. K. Khandani, "Capacity bounds for the Gaussian interference channel," *IEEE Trans. Inform. Theory*, vol. 55, no. 2, pp. 620–643, Feb. 2009.

[4] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen, "Optimal multiuser spectrum balancing for digital subscriber lines," *IEEE Trans. Commun.*, vol. 54, no. 5, pp. 922–933, May 2006.

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

[6] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inform. Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.

[7] P. Viswanath and D. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Inform. Theory*, vol. 49, no. 8, pp. 1912–1921, Aug. 2003.

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

[9] W. Yu and J. Cioffi, "Sum capacity of Gaussian vector broadcast channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 1875–1892, Sep. 2004.

[10] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inform. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.

[11] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Trans. Signal Processing*, vol. 55, no. 6, pp. 2646–2660, Jun. 2007.

[12] 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 Processing*, vol. 57, no. 2, pp. 690–697, Feb. 2009.

[13] 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.

[14] G. Ginis and J. M. Cioffi, "Vectored transmission for digital subscriber line systems," *IEEE J. Select. Areas Commun.*, vol. 20, no. 5, pp. 1085–1104, Jun. 2002.

[15] A. R. Forouzan and L. M. Garth, "Generalized iterative spectrum balancing and grouped vectoring for maximal throughput of digital subscriber lines," in *IEEE Global Telecom. Conf., GLOBECOM'05*, vol. 4, Jun. 2005, pp. 2359–2363.

[16] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.

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

[18] S. V. Hanly and D.-N. Tse, "Power control and capacity of spread spectrum wireless networks," *Automatica*, vol. 35, no. 12, pp. 1987–2012, Dec. 1999.

[19] E. Seneta, *Non-Negative Matrices and Markov Chains*, 2nd ed. New York, NY: Springer-Verlag, 1981.

[20] R. Cendrillon, M. Moonen, G. Ginis, K. V. Acker, T. Bostoen, and P. Vandaele, "Partial crosstalk precompensation for downstream VDSL," *Signal Processing*, vol. 84, no. 11, pp. 2005–2019, Nov. 2004.

[21] A. R. Forouzan and L. M. Garth, "Computationally efficient partial crosstalk cancellation in fast time-varying DSL crosstalk environments," *EURASIP J. Adv. Signal Processing*, vol. 2007, 2007, doi: 10.1155/2007/72041.

[22] M. Maesoumi, M.-A. Masnadi-Shirazi, and M. Dousti, "Computationally efficient joint partial crosstalk cancellation and multi-user power control algorithm in xDSL networks," *European Trans. on Telecomm.*, pp. 770–781, 2009.

[23] A. R. Forouzan, M. Moonen, J. Maes, and M. Guenach, "Efficient calculation of decoding order in non-ideal DSL multiple-access channels," in *17th IEEE Symp. on Commun. & Veh. Technol. in the Benelux, SCVT'10*, Enschede, The Netherlands, Nov. 2010, pp. 1–6.

[24] A. R. Forouzan and M. Moonen, "Lagrange multiplier optimization for optimal spectrum balancing of DSL with logarithmic complexity," in *IEEE Int'l. Conf. on Commun., ICC'11*, Kyoto, Japan, Jun. 2011.

[25] C.-Y. Chen, K. Seong, R. Zhang, and J. Cioffi, "Optimized resource allocation for upstream vectored DSL systems with zero-forcing generalized decision feedback equalizer," *IEEE J. Select. Topics Signal Processing*, vol. 1, no. 4, pp. 686–699, Dec. 2007.

[26] ITU-T G.993.2, "Very high speed digital subscriber line transceivers 2 (VDSL2)," Geneva, Switzerland, Feb. 2006.

[27] Z.-Q. Luo and S. Zhang, "Duality gap estimation and polynomial time approximation for optimal spectrum management," *IEEE Trans. Signal Processing*, vol. 57, no. 7, pp. 2675–2689, Jul. 2009.

[28] ITU-T G.993.2; Amendment 1, "Very high speed digital subscriber line transceivers 2 (VDSL2)," Geneva, Switzerland, Apr. 2007.

[29] P. Golden, H. Dedieu, and K. Jacobsen, *Fundamentals of DSL Technology.* Boca Raton, FL: Auerbach Publications, 2006.

Algorithm 1: The IF/BC-OSB Algorithm

repeat
- Set/update $\theta_{(g,n)}, \forall (g,n) (1 \leq g \leq G \text{ and } 1 \leq n \leq N_g)$;
- **for** $k = 1$ **to** K **do**
  - **repeat**
    - Set/update $\lambda^k_{(g,n)}$, $\forall\ (g,n)$;
    - $\mathcal{L}^{\max}_k \leftarrow 0$;
    - **repeat**
      - Set/update $\tilde{s}^k_{(g,n)}$, $\forall\ (g,n)$;
      - **for** $g = 1$ **to** G **do**
        - $\mathcal{W}^k_{g,\max} \leftarrow 0$;
        - **for** *all decoding orders in group* g **do**
          - `/* Solving the problem in the dual channel */`
          - Calculate $\tilde{\boldsymbol{q}}^k_{(g,n)},\ \forall\ n$ using (3.23);
          - Calculate $\tilde{b}^k_{(g,n)},\ \forall\ n$ using (3.17) and (2.8);
          - $\mathcal{W}^k_g \leftarrow \sum_n \omega_{(g,n)} \tilde{b}^k_{(g,n)}$;
          - **if** $\mathcal{W}^k_g > \mathcal{W}^k_{g,\max}$ **then**
            - $\mathcal{W}^k_{g,\max} \leftarrow \mathcal{W}^k_g$;
            - `/* Store/update the so far optimal solution for group` g `on tone` k `*/`
            - $\hat{\boldsymbol{q}}^k_{(g,n)} \leftarrow \tilde{\boldsymbol{q}}^k_{(g,n)}$;
            - $\hat{b}^k_{(g,n)} \leftarrow \tilde{b}^k_{(g,n)}$;
            - Store/update the optimal decoding order in the dual channel for group g on tone k;
      - $\mathcal{L}_k \leftarrow \sum_g \mathcal{W}^k_{g,\max} - \sum_{(g,n)} \tilde{s}^k_{(g,n)}$;
      - **if** $\mathcal{L}_k > \mathcal{L}^{\max}_k$ **then**
        - $\mathcal{L}^{\max}_k \leftarrow \mathcal{L}_k$;
        - `/* Store/update the so far optimal solution on tone` k `*/`
        - $\boldsymbol{q}^{k,\ \text{opt}}_{(g,n)} \leftarrow \hat{\boldsymbol{q}}^k_{(g,n)}$;
        - $b^{k,\ \text{opt}}_{(g,n)} \leftarrow \hat{b}^k_{(g,n)}$;
        - Encoding order $\leftarrow$ Reverse of the optimal decoding order in the dual channel;
      - Calculate $p^k_{(g,n)},\ \forall\ (g,n)$, using (3.4) and (2.13);
    - **until** *the entire transmit power space in the dual channel is searched*;
    - $P_{(g,n)} \leftarrow \Delta_{\text{f}} \sum_{k=1}^{K} p^{k,\ \text{opt}}_{(g,n)},\ \forall\ (g,n)$;
  - **until** $(p^k_{(g,n)} = p^{k,\text{mask}}_{(g,n)})$ or $(p^k_{(g,n)} < p^{k,\text{mask}}_{(g,n)}$ and $\lambda^k_{(g,n)} = 0),\ \forall\ (g,n)$;

until $(P_{(g,n)} = P^{\max}_{(g,n)})$ or $(P_{(g,n)} < P^{\max}_{(g,n)}$ and $\theta_{(g,n)} = 0),\ \forall\ (g,n)$;

Algorithm 2: The Simplified IF/BC-OSB Algorithm

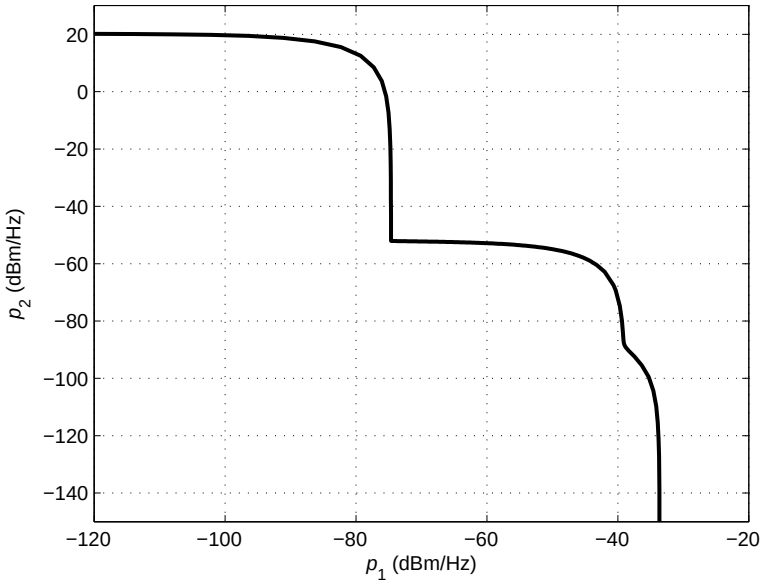
Set the decoding order in the groups according to the order of weight factors $\omega_{(g,n)}$;
$\lambda^k_{(g,n)} \leftarrow 1$, for all (g,n) $(1 \le g \le G$ and $1 \le n \le N_g)$;
repeat
  Set/update $\theta_{(g,n)},\ \forall (g,n)$;
  for $k = 1$ **to** K **do**
    $\mathcal{L}_k^{\max} \leftarrow 0$;
    repeat
      Set/update $\tilde{s}^k_{(g,n)}$, $\forall\ (g,n)$ (using a fast spectrum balancing algorithm);
      `/* Solving the problem in the dual channel */`
      Calculate $\tilde{\boldsymbol{q}}^k_{(g,n)},\ \forall\ (g,n)$ using (3.23);
      Calculate $\tilde{b}^k_{(g,n)},\ \forall\ (g,n)$ using (3.17) and (2.8);
      Calculate $p^k_{(g,n)},\ \forall\ (g,n)$, using (3.4) and (2.13);
      if $p^k_{(g,n)} \le p^{k,\text{mask}}_{(g,n)},\ \forall\ (g,n)$ **then**
        Calculate $\mathcal{L}_k$ using (3.13);
        if $\mathcal{L}_k > \mathcal{L}_k^{\max}$ **then**
          $\mathcal{L}_k^{\max} \leftarrow \mathcal{L}_k$;
          `/* Store/update the so far optimal solution on tone` k `*/`
          $\boldsymbol{q}^{k,\ \text{opt}}_{(g,n)} \leftarrow \tilde{\boldsymbol{q}}^k_{(g,n)}$;
          $b^{k,\ \text{opt}}_{(g,n)} \leftarrow \tilde{b}^k_{(g,n)}$;
          $p^{k,\ \text{opt}}_{(g,n)} \leftarrow p^k_{(g,n)}$;
    until *the transmit power space in the dual channel is searched*;
  $P_{(g,n)} \leftarrow \Delta_\text{f} \sum_{k=1}^K p^{k,\ \text{opt}}_{(g,n)},\ \forall\ (g,n)$;
until $P_{(g,n)} = P^{\max}_{(g,n)}$ or $(P_{(g,n)} < P^{\max}_{(g,n)}$ and $\theta_{(g,n)} = 0),\ \forall\ (g,n)$;

Table I: Simulated Algorithms

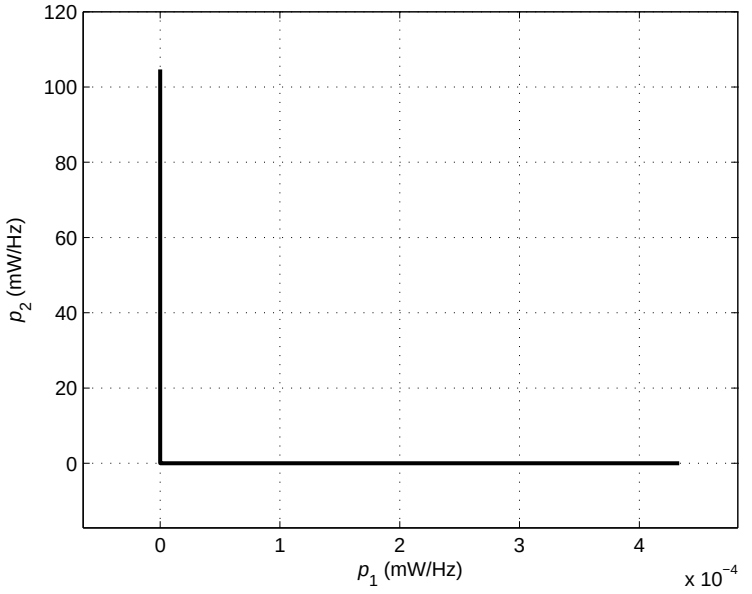
NAME	ENCODING	FILTER TAPS	SPECTRUM BALANCING
IF/BC-OSB	Non-Linear	Optimal	OSB
Simplified IF/BC-OSB	Non-Linear	Sub-Optimal	OSB (or its alternatives)
Grouped ZF-OSB	Linear	ZF	OSB
Grouped ZF-THP-OSB	Non-Linear	ZF	OSB
Grouped OLP-OSB	Linear	Optimal	OSB
OSB	–	–	OSB
SSM	–	–	–

Table II: Simulation Parameters

PARAMETER	VALUE
Bandplan and PSD mask	DS VDSL2E17 B7-9 bandplan [26, 28]
Cable type	26 AWG [29]
Noise	White noise, -140 dBm/Hz
Tone spacing, Δ_{f}	4.3125 kHz
Symbol rate, f_{s}	4 kHz
$b_{\max}$	15
$b_{\min}$	2
SNR Gap	9.45 dB



(a)



(b)

Figure 1: The power region for a 2 user BC scenario to load 10 bits for each user in (a) dBm and (b) milliwatts.

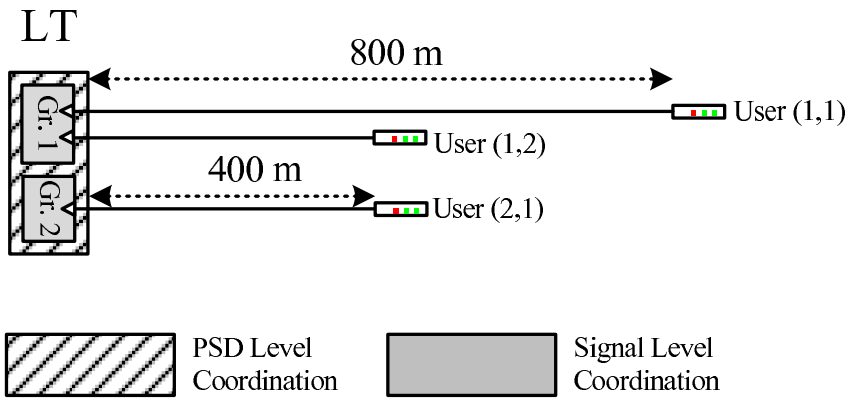


Figure 2: The simulated scenario.

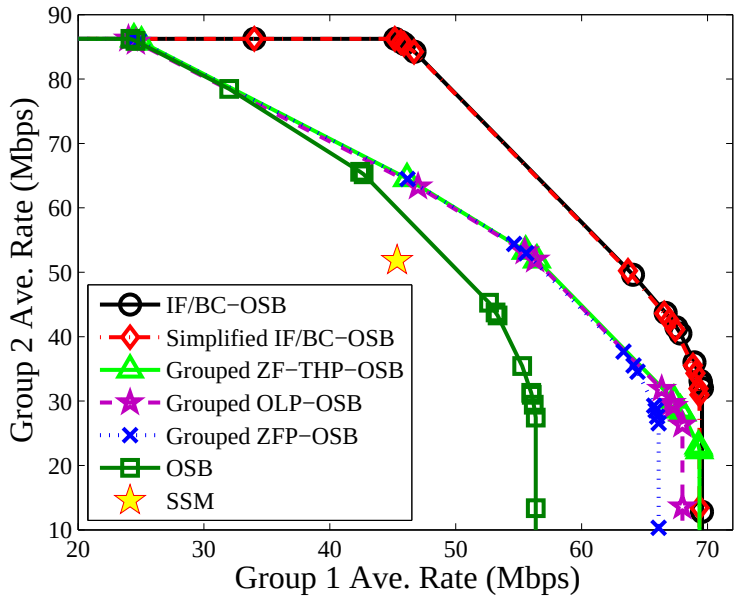


Figure 3: Average achievable bit rate of the users in group 1 vs. achievable bit rate of the user in group 2 for the scenario in Fig. 2 using different crosstalk mitigation schemes.

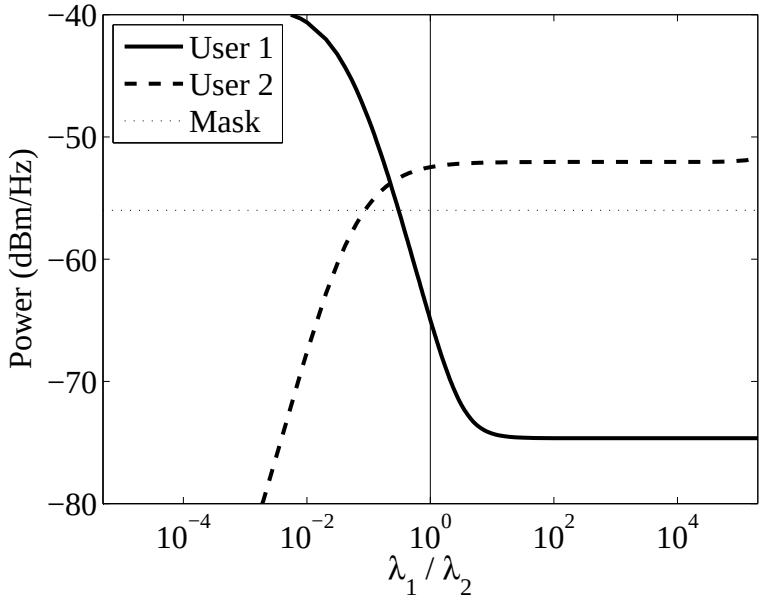


Figure 4: The transmit power on lines for a 2×2 DSL broadcast channel to load 10 bits for both users on a tone located at 6 MHz as a function of λ_1/λ_2.