Katholieke Universiteit Leuven

Departement Elektrotechniek

ESAT-SISTA/TR 12-26

DMT MIMO IC Rate Maximization in DSL with Combined Signal and Spectrum Coordination [1]

Rodrigo B. Moraes, Paschalis Tsiaflakis, Jochen Maes and Marc Moonen[2]

Published in IEEE Transactions on Signal Processing
April 2013

[1]This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/sista/rmoraes/reports/12-26.pdf*

[2]K.U.Leuven, Dept. of Electrical Engineering (ESAT), Research group SISTA, Kastelpark Arenberg 10, 3001 Leuven, Belgium, Tel. 32/16/32 17 09, Fax 32/16/32 19 70, WWW: *http://www.esat.kuleuven.ac.be/sista*. E-mail: *rodrigo.moraes@esat.kuleuven.ac.be*. This research work was carried out at the ESAT Laboratory of the KU Leuven, in the frame of the KU Leuven Research Council CoE EF/05/006 Optimization in Engineering (OPTEC) and PFV/10/002 (OPTEC); 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); IUAP P7/23 (Belgian network on Stochastic modeling, analysis, design and optimization of communication systems, BESTCOM, 2012-2017); 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.

DMT MIMO IC Rate Maximization in DSL with Combined Signal and Spectrum Coordination

Rodrigo B. Moraes, *Student Member, IEEE,* Paschalis Tsiaflakis, *Member, IEEE,* Jochen Maes, *Senior Member, IEEE* and Marc Moonen, *Fellow, IEEE*

Abstract

Theoretical research has demonstrated that the achievable gains in data rate with dynamic spectrum management, i.e. signal coordination or spectrum coordination, are substantial for digital subscriber line (DSL) networks. Work on these two fronts has progressed steadily and, more often than not, independently. In this paper, we combine the two types of coordination for a mixed DSL scenario, in which some of the infrastructure required for full two-sided signal coordination is available, but not all. This scenario, which is referred to as the discrete multitone multiple-input, multiple-output interference

Copyright (c) 2012 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

A preliminary version of this paper was presented at the European Signal Processing Conference (EUSIPCO), Barcelona, Spain, in September 2011.

This research work was carried out at the ESAT Laboratory of the KU Leuven, in the frame of the KU Leuven Research Council CoE EF/05/006 Optimization in Engineering (OPTEC) and PFV/10/002 (OPTEC); 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); IUAP P7/23 (Belgian network on Stochastic modeling, analysis, design and optimization of communication systems, BESTCOM, 2012-2017); 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.

R. B. Moraes, P. Tsiaflakis and M. Moonen are with the Department of Electrical Engineering (ESAT-SCD), KU Leuven, Kasteelpark Arenberg 10, B-3001 Leuven, Belgium (e-mail: rodrigo.moraes@esat.kuleuven.be; paschalis.tsiaflakis@esat.kuleuven.be; and marc.moonen@esat.kuleuven.be). P. Tsiaflakis is also a postdoctoral fellow funded by the Research Foundation—Flanders (FWO). J. Maes is with the Access Research Domain, Alcatel-Lucent Bell Labs, Antwerp, Belgium. (e-mail: jochen.maes@alcatel-lucent.com).

channel (DMT MIMO IC), consists of multiple interfering users, each operating a distinct subset of DSL lines as a MIMO system. Coordination is done both on the signal level (with per user MIMO techniques) and on the spectrum level (with multi-user power allocation). We propose two algorithms for the DMT MIMO IC weighted rate maximization problem. In the first algorithm, we profit from recent work showing the close relation between the weighted rate sum maximization problem and the weighted MMSE minimization problem. We show that with a simple extension, we can adapt the previous work to the scenario of interest. In the second algorithm, the signal and spectrum coordination parts are solved separately. For the signal coordination part, we obtain multiple independent single tone MIMO IC's, which allows us to leverage on the previous work on the topic. For the spectrum coordination part, one of the interesting results of our analysis is a generalization of the distributed spectrum balancing (DSB) power allocation formula for the DMT MIMO IC scenario. Simulation results demonstrate that both algorithms obtain significant gains when compared to pure spectrum coordination algorithms.

Index Terms

SPC-TDLS, SPC-MULT, MSP-CODR, MSP-CAPC.

I. Introduction

As the market demand for higher data rates for broadband communication steadily increases, it appears that the only way to satisfy it is to connect optical fiber directly to the customers. Fiber deployment, however, is difficult and expensive, and it is expected that the fiber network will experience slow (albeit steady) enlargement towards the customer. The time needed for full fiber-to-the-home connections to dominate the market is counted in decades rather than years [1].

In the meantime, fiber often reaches the cabinets on the streets or in the basement of large buildings. Today, the last couple of kilometers or hundreds of meters are bridged primarily with digital subscriber line (DSL) technology. DSL plays a major role and is expected to continue to do so for many years to come. DSL is specially attractive because of the ubiquitousness of the twisted pair copper network. On the negative side, there is multi-user interference, i.e. *crosstalk*. While crosstalk is a mild annoyance for telephone services, in DSL, with its greatly enlarged transmission bandwidth, it has been identified as the main source of performance degradation.

In this landscape, where competition is fierce and new challenges abound, service providers have identified two directions in which the DSL standards can evolve. First, new generations of DSL are being adopted. Perhaps the incumbent standard is still the asymmetrical DSL (ADSL), which is able to provide data rates up to 24 Mb/s on lines that are several kilometers long. But ADSL is rapidly losing ground

to the very-high-bit-rate DSL (VDSL). The success of VDSL is strongly related to the enlargement of the fiber network. Accordingly, VDSL is primarily designed to connect the customers from a fiber-fed cabinet, and hence to operate on much shorter loops. VDSL is able to provide data rates up to 100 Mb/s.

A second direction of evolution is dynamic spectrum management (DSM). DSM looks for improvements by means of spectrum or signal coordination techniques. In the past ten to fifteen years, theoretical research has shown that data rate improvements achievable with DSM can be formidable. The main objective is to avoid, cancel or even profit from crosstalk. This second direction is the focus of this paper.

Spectrum coordination (also known as DSM levels 1 and 2) applies to the interference channel (IC) scenario with multiple interfering users, each of them operating a single DSL line as a single-input, single-output (SISO) system. Spectrum coordination aims to allocate transmit power to the multiple users in the available spectrum so that crosstalk is avoided or minimized. Previous work on this topic includes [2]–[6]. Spectrum coordination does not deliver the same gains as signal coordination does, but it profits from simplified requirements on infrastructure and smaller online complexity. For spectrum coordination, users do not necessarily have to have physically co-located transmitters and receivers, and a number of algorithms optimize a network in which little or no message exchanges between users take place. DSM levels 1 and 2 are in use today in the forms of waterfilling-based bit loading, downstream power back-off and upstream power back-off.

Signal coordination (also known as DSM level 3 or vectoring) generally involves multiple-input, multiple-output (MIMO) processing, either two sided or single sided. Previous work includes [7]–[11]. With signal coordination, requirements on infrastructure are considerably higher. For signal coordination, users have to have physically co-located transmitters and/or receivers at the same access node, and knowledge of all signals and all channel gains involved is usually required. On the plus side, signal coordination is able to deliver substantial gains in comparison with spectrum coordination, eliminating most or all crosstalk and sometimes even using crosstalk for its benefit. Recent results [10], [12] have shown that with full two sided signal coordination, i.e. with transmit and receive signal coordination among all users in the binder, data rates up to 1 Gb/s can be provided on loops a couple of hundreds meters long.

Work on these two fronts has progressed steadily and, more often than not, independently. Recently, attention was given to mixed scenarios, in which some of the infrastructure for signal coordination is available, but not all [13]–[15]. In this paper we consider a DSL scenario where each user has a number of lines that can be operated jointly, but where signal coordination among all users in the binder is not

possible. In this scenario, coordination involves both the signal and the spectrum levels.

These mixed scenarios fit well in the framework of enlargement of the fiber network and may be encountered in real deployments in some years to come—if they are not already. As fiber progresses towards the customer, fewer DSL lines share the same binder. Whenever some of the lines terminate on different access nodes (e.g. when not co-located or when local loop unbundling is regulatory required), signal coordination among all the users is not possible, and so part of the coordination should be done on the spectrum level. Moreover, these mixed scenarios encompass the possible utilization of not only the direct mode of a DSL line, but also its phantom or common modes [12], [16], [17]. For example, one user with two copper pairs (terminating on the same equipment on both the customer and the service provider sides) can have three communication channels, corresponding to two direct modes (one for each DSL line) and the phantom mode. Henceforth we will only use the term 'transceiver'—we use it as a generalization of 'line'. We define a transceiver to be connected to a physical communications channel, which can be a differential or a common mode of a DSL line, or a phantom mode of two or more twisted wire pairs. In wireless, a similar terminology would be antenna.

This scenario is generally referred to as MIMO interference channel (MIMO IC). Because of the discrete multitone (DMT) modulation used in DSL, where tones are coupled through per-user power constraints, we refer to this scenario as the DMT MIMO IC. To maximize the weighted rate sum of the network in such a scenario, we need an approach where elements of both spectrum and signal coordination are present, i.e. crosstalk that cannot be canceled on the signal level should be avoided on the spectrum level.

In this paper, we propose two algorithms for weighted rate sum maximization for the DMT MIMO IC. The first algorithm builds on the recently suggested equivalence between the weighted rate sum maximization problem and the weighted minimum mean squared error (WMMSE) minimization problem [18] for the (single tone) MIMO IC [19]–[21]. We show that a simple adjustment on the single tone algorithm is sufficient to adapt it to the multitone case. The resulting multitone algorithm solves the signal and spectrum coordination parts of the problem simultaneously and is guaranteed to reach a stationary point.

The WMMSE minimization approach is surprisingly easy to adapt from the single tone to the multitone problem, but that is not necessarily true for other solutions. That is why, in the second proposed algorithm, we separate the signal and spectrum coordination parts of the problem and solve each of them separately. The advantage here is that many options are possible for solving these two separate parts. We end up with many independent single tone MIMO IC's (one for each tone). In this way, we can rely on previous

work on this topic and select any of the available solutions. For the spectrum coordination part, one of the interesting outcomes of our analysis is a generalization of the distributed spectrum balancing (DSB) [5], [6], [22] power allocation formula for the DMT MIMO IC scenario. Simulation results demonstrate that both algorithms obtain significant gains when compared to pure spectrum coordination algorithms.

This paper is organized as follows. In Section II, we present the system model and the problem statement. In Section III we describe the algorithm based on solving the equivalent WMMSE minimization problem and in Section IV we describe the algorithm based on separately solving the signal and spectrum coordination parts. Simulation results are presented in Section V and final remarks are given in Section VI.

We use lower-case boldface letters to denote vectors, upper-case boldface letters for matrices and calligraphic letters for sets (for example, $\mathbf{a}$, $\mathbf{A}$ and $\mathcal{A}$). We use $\mathbf{I}_A$ as the identity matrix of size A, $(\cdot)^{\mathrm{H}}$ as the Hermitian transpose, $(\cdot)^*$ as the complex conjugate, $\mathrm{E}[\cdot]$ as expectation, $\mathrm{tr}\{\cdot\}$ as trace, $|\cdot|$ as determinant and $\mathrm{diag}\{\mathbf{a}\}$ as the matrix with a vector $\mathbf{a}$ in the main diagonal.

II. Problem Statement

We consider an N user DSL system with discrete multitone (DMT) modulation with K Δ_f-spaced tones. We denote the set of users by $\mathcal{N} = \{1, \ldots, N\}$ and the set of tones by $\mathcal{K} = \{1, \ldots, K\}$. We let p_n^k be the transmit power of user n on tone k and we organize these values in the matrix $\mathbf{P} \in \mathbb{R}^{K \times N}$. The nth column of $\mathbf{P}$, denoted by $\mathbf{p}_n = \begin{bmatrix} p_n^1 & \ldots & p_n^K \end{bmatrix}^{\mathrm{T}}$, contains the power allocation of user n in all tones. The kth row of $\mathbf{P}$, $\mathbf{p}^k = \begin{bmatrix} p_1^k & \ldots & p_N^k \end{bmatrix}$, represents the power allocation of all users in tone k. User n has A_n transceivers and can adopt two sided MIMO processing among them. Throughout this paper, we focus on a linear design for both transmitters and receivers and treat interference as noise. All channel gains are considered perfectly known (not such a tall order in DSL systems). Also, we consider the simplifying assumption of perfect DMT block synchronization between users.[1] Taking that into account, we obtain the received signal for user n on tone k as

$$\mathbf{y}_n^k = \mathbf{H}_{n,n}^k \mathbf{T}_n^k \mathbf{x}_n^k + \sum_{j \neq n} \mathbf{H}_{n,j}^k \mathbf{T}_j^k \mathbf{x}_j^k + \mathbf{z}_n^k. \tag{1}$$

Here $\mathbf{y}_n^k,\ \mathbf{x}_n^k \in \mathbb{C}^{A_n}$ are, respectively, the received and transmitted signal vector for user n on tone k; $\mathbf{H}_{n,j}^k \in \mathbb{C}^{A_n \times A_j},\ \mathbf{T}_n^k \in \mathbb{C}^{A_n \times A_n}$ are, respectively, the channel matrix from user j to user n on tone

[1]For the case when the DMT blocks of different users are offset in relation to each other, inter carrier interference (ICI) arises. ICI complicates the problem significantly. See e.g. [23].

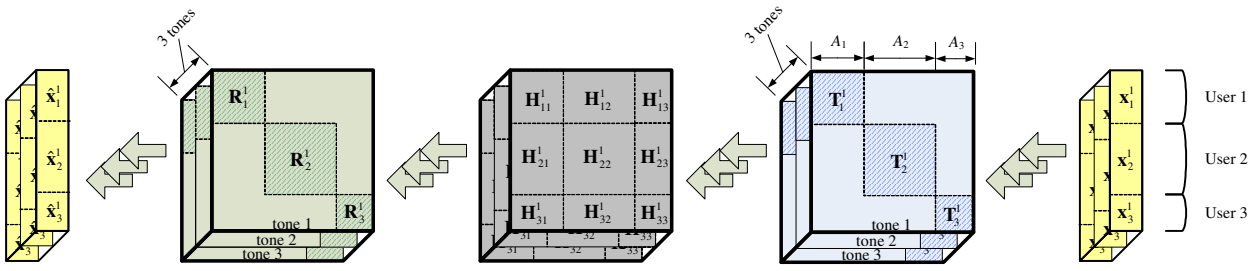

Fig. 1. Illustration of the DMT MIMO IC for a scenario with 3 users and 3 tones. Each user n has A_n transceivers. For simplicity, the zero mean Gaussian noise is not shown.

k and the transmit matrix for user n on tone k. In (1), we assume $\mathrm{E}\left[\mathbf{x}_n^k(\mathbf{x}_n^k)^{\mathrm{H}}\right] = \mathbf{I}_{A_n}$, and hence $\mathrm{tr}\{\mathbf{T}_n^k(\mathbf{T}_n^k)^{\mathrm{H}}\} = p_n^k$. Without loss of generality, the noise vector $\mathbf{z}_n^k$ is assumed to be spatially white with covariance matrix $\mathrm{E}\left[\mathbf{z}_n^k(\mathbf{z}_n^k)^{\mathrm{H}}\right] = \mathbf{I}_{A_n}$. The estimated signal vector for user n on tone k is given by

$$\hat{\mathbf{x}}_n^k = \mathbf{R}_n^k \mathbf{y}_n^k, \tag{2}$$

where $\mathbf{R}_n^k$ is the receive matrix for user n on tone k. We illustrate this situation in Fig. 1 for a system with three users and three tones. The receive matrix used in (2) is the linear MMSE (LMMSE) matrix. It has been shown that in a MIMO IC scenario the LMMSE receiver provides an optimal linear receiver given a set of linear transmit matrices [18], [19]. For a given set of transmit matrices, we write

$$\mathbf{R}_n^k = (\mathbf{T}_n^k)^{\mathrm{H}}(\mathbf{H}_{n,n}^k)^{\mathrm{H}}\left(\mathbf{M}_n^k + \mathbf{H}_{n,n}^k\mathbf{T}_n^k(\mathbf{T}_n^k)^{\mathrm{H}}(\mathbf{H}_{n,n}^k)^{\mathrm{H}}\right)^{-1}, \tag{3}$$

where

$$\mathbf{M}_n^k = \sum_{j\neq n}\mathbf{H}_{n,j}^k\mathbf{T}_j^k(\mathbf{T}_j^k)^{\mathrm{H}}(\mathbf{H}_{n,j}^k)^{\mathrm{H}} + \mathbf{I}_{A_n}$$

is the noise plus interference covariance matrix for user n on tone k. We remark that, although we do not write it explicitly, this matrix may be normalized by a capacity gap Γ.

With the LMMSE receiver and assuming Gaussian signaling, the achievable data rate for user n on tone k is given by

$$b_n^k = \log\left|(\mathbf{T}_n^k)^{\mathrm{H}}(\mathbf{H}_{n,n}^k)^{\mathrm{H}}(\mathbf{M}_n^k)^{-1}\mathbf{H}_{n,n}^k\mathbf{T}_n^k + \mathbf{I}_{A_n}\right|. \tag{4}$$

We use $\log(\cdot)$ as the natural logarithm. Users have their total data rate in bits per second defined as $r_n = {}^{f_s}/_{\log(2)}\sum_{k\in\mathcal{K}} b_n^k$, where f_s is the symbol rate. In this paper, we ignore the practical constraint of discrete bit loading.

We denote the set of all matrices $\mathbf{T}_n^k$ as $\mathbb{T} = \left\{\mathbf{T}_n^k \mid n \in \mathcal{N},\ k \in \mathcal{K}\right\}$. The problem we would like to solve is the weighted rate sum (WRS) maximization, which can be written as

$$\begin{aligned} \max_{\mathbb{T}} \quad & \sum_{n\in\mathcal{N}} \sum_{k\in\mathcal{K}} u_n b_n^k \\ \text{subject to} \quad & \sum_{k\in\mathcal{K}} \mathrm{tr}\left\{\mathbf{T}_n^k (\mathbf{T}_n^k)^{\mathrm{H}}\right\} \leq P_n^{\max}\ \forall n \end{aligned} \tag{5}$$

Here, $P_n^{\max}$ is the power budget and u_n is the weight for user n.[2] Note that we use a per-user total power constraint, not a per-transceiver total power constraint. For the sake of appropriately emphasizing the main problem, we also do not use a per-tone power constraint, i.e. we do not consider a spectral mask. We remark that the design of the receive matrices is the easy part of the problem and that makes these matrices not to be included in the optimization variables in (5). We refer to (5) as the DMT MIMO IC problem.

The optimization in (5) is non-concave with respect to $\mathbf{T}_n^k$ and hence it is not trivial. It comprises K distinct N-user MIMO ICs, in which, for all tones, user n has A_n transceivers. The challenge in (5) is twofold. First, we should design the matrices $\mathbf{T}_n^k$ for all users and tones so that the resulting signal vectors—the columns of $\mathbf{T}_n^k$—are easy to identify in the intended receiver (user n) and easy to mitigate in the unintended receivers (users $j \neq n$). The second challenge is about power allocation. Since $\mathrm{tr}\left\{\mathbf{T}_n^k (\mathbf{T}_n^k)^{\mathrm{H}}\right\} = p_n^k$, we should appropriately choose how much power each tone of each user gets, i.e. we should choose $\mathbf{P}$. The design of the matrices $\mathbf{T}_n^k$ corresponds to the signal coordination part of the problem. The design of $\mathbf{P}$, i.e. the power allocation, is the spectrum coordination part. Albeit the spectrum coordination part is not explicitly shown in (5), it is crucially important. Note that the per-user power constraint couples the optimization through the tones, which complicates the problem significantly.

We remark that special cases of (5) are well-known in the literature.

- The special case when $A_n = 1\ \forall n$ corresponds to the SISO case, i.e. the pure spectrum coordination problem (DSM levels 1 or 2). The spectrum coordination problem can also be obtained by restricting the transmit matrices $\mathbf{T}_n^k$ to be diagonal for all n, k. Ref. [2] proposes an algorithm with an approximately optimal solution [26], and several other papers have worked on practical and low complexity solutions for an efficient implementation, e.g. [3]–[6].

[2]We remark that, with similar arguments as [24], [25], it can be shown that, as the tone spacing Δ_f tends to zero, the vector of data rates $\mathbf{r} = \{r_n\},\ r_n = \int b_n(f)df$ forms a convex set. That in turn allows us to map the whole border of the rate region by changing the weights u_n. Problem (5), however, is discretized, with a finite tone spacing. Theoretically, this may imply that some points in the border of the rate region cannot be found. However, the tone spacing for DSL impound to be small enough so that, for practical purposes, the WRS approach can characterize the border of the rate region almost fully. See e.g. [2].

- The special case when $N = 1$ corresponds to full two-sided signal coordination [10], [27]. For this case, the optimal solution comprises two steps: first, set $\mathbf{R}_1^k = (\mathbf{U}_1^k)^{\mathrm{H}}$ and $\mathbf{T}_1^k = {}^1/\sqrt{A_1}\mathbf{V}_1^k$ for all tones, where $\mathbf{U}_1^k$ and $\mathbf{V}_1^k$ are, respectively, the matrices of left and right singular vectors of $\mathbf{H}_1^k$, i.e. $\mathbf{H}_1^k = \mathbf{U}_1^k \mathrm{diag}\left\{\begin{bmatrix}\gamma_1^k(1) & \dots & \gamma_1^k(A_1)\end{bmatrix}\right\}(\mathbf{V}_1^k)^{\mathrm{H}}$; and second, consider the noise to channel ratio to be ${}^1/\gamma_1^k(i)^2$ and allocate power with a waterfilling algorithm.
- For the special case when $K = 1$, several solutions are also available, e.g. [19]–[21], [28]–[31], but they are at best guaranteed to converge to a local optimum.

Although there is some related work in the wireless communication context, to the best of our knowledge the DMT MIMO IC problem in the more general form of (5) has not been analyzed in full in the literature. In the wireless context, for example [32], [33] use signal and spectrum coordination algorithms jointly for a multicell broadcast channel problem (in [32] scheduling is also considered), but simplifying assumptions are used. In [32] users are restricted to one data stream. The power constraints are also simplified. In [33], users are restricted to have one receive antenna. These limitations clearly do not apply to our scenario. In [34], [35], power is evenly distributed through the sub-carriers. The DSL case, however, is quite different because the channel is highly frequency selective. This makes power allocation across frequency to be a fundamental feature of high performance systems.

III. Algorithm 1: DMT WMMSE minimization

A. WMMSE vs. WRS

A recent paper [18] develops an interesting solution for the design of the transmit matrices for the WRS problem in a MIMO broadcast channel. Instead of directly focusing on the WRS problem, a relation is established to the WMMSE problem, which is simpler [36]. The original WRS problem is then solved through the WMMSE problem. Ref. [18] presents the conditions for the two problems to have the same stationary points and proposes an iterative algorithm that is efficient and that provides good results.

Following the same idea, [19]–[21] extend the algorithm to the MIMO IC. These three references plus [18] focus on the single tone problem ($K = 1$), i.e. they focus on only one layer in Fig. 1.

In this section, we show that the solution for the single tone case is straightforwardly extended to the multitone case. We begin with the definition of the MSE matrix $\mathbf{E}_n^k$ for user n, tone k, after the LMMSE receive filter $\mathbf{R}_n^k$, i.e.

$$\begin{aligned}\mathbf{E}_n^k &= \mathrm{E}\left[(\mathbf{R}_n^k\mathbf{y}_n^k - \mathbf{x}_n^k)(\mathbf{R}_n^k\mathbf{y}_n^k - \mathbf{x}_n^k)^{\mathrm{H}}\right] \\ &= \left((\mathbf{T}_n^k)^{\mathrm{H}}(\mathbf{H}_{n,n}^k)^{\mathrm{H}}(\mathbf{M}_n^k)^{-1}\mathbf{H}_{n,n}^k\mathbf{T}_n^k + \mathbf{I}_{A_n}\right)^{-1} \end{aligned} \quad (6)$$

As is well known, we can rewrite (4) as $b_n^k = \log\left(\left|(\mathbf{E}_n^k)^{-1}\right|\right)$.

The DMT WMMSE minimization problem is then given by

$$\begin{aligned} \max_{\mathbb{T}} \quad & \sum_{n\in\mathcal{N}}\sum_{k\in\mathcal{K}} -\mathrm{tr}\{\mathbf{W}_n^k\mathbf{E}_n^k\} \\ \text{subject to} \quad & \sum_{k\in\mathcal{K}} \mathrm{tr}\{\mathbf{T}_n^k(\mathbf{T}_n^k)^{\mathrm{H}}\} \le P_n^{\max}\ \forall n. \end{aligned} \tag{7}$$

Here $\mathbf{W}_n^k \in \mathbf{C}^{A_n\times A_n}$ is a weighting matrix. Note that (7) is concave in $\mathbf{T}_n^k$ when $\mathbf{W}_n^k$ is fixed and vice-versa.

We now write the Lagrangean of (5) and (7) as

$$\begin{aligned} L_{\mathrm{wsr}}(\mathbb{T},\boldsymbol{\lambda}) &= \sum_{n\in\mathcal{N}}\sum_{k\in\mathcal{K}} u_n b_n^k - \\ &\qquad \sum_{n\in\mathcal{N}} \lambda_n\Big(\sum_{k\in\mathcal{K}} \mathrm{tr}\{\mathbf{T}_n^k(\mathbf{T}_n^k)^{\mathrm{H}}\} - P_n^{\max}\Big) \\ L_{\mathrm{wmmse}}(\mathbb{T},\boldsymbol{\lambda}) &= -\sum_{n\in\mathcal{N}}\sum_{k\in\mathcal{K}} \mathrm{tr}\{\mathbf{W}_n^k\mathbf{E}_n^k\} \\ &\qquad -\sum_{n\in\mathcal{N}} \lambda_n\Big(\sum_{k} \mathrm{tr}\{\mathbf{T}_n^k(\mathbf{T}_n^k)^{\mathrm{H}}\} - P_n^{\max}\Big), \end{aligned}$$

Next, we find the equations for the stationary conditions by calculating $\nabla_{\mathbf{T}_n^k} L_{\mathrm{wsr}}(\mathbb{T},\boldsymbol{\lambda}) = \mathbf{0}$ and $\nabla_{\mathbf{T}_n^k} L_{\mathrm{wmmse}}(\mathbb{T},\boldsymbol{\lambda}) = \mathbf{0}$ and compare the two. Just like in [20], [21], it is observed that if the weighting matrices are set as

$$\mathbf{W}_n^k = u_n(\mathbf{E}_n^k)^{-1} \tag{8}$$

then if $\mathbf{T}$ is a stationary point of (5), it is also a stationary point of (7). Since the WMMSE is easier to solve, that is the one we focus on. The solution of (7) for $\mathbf{T}_n^k$ is derived while keeping $\mathbf{W}_n^k$ and $\mathbf{R}_n^k$ fixed and it is similar to that for the MIMO broadcast channel case [18], [36], i.e.

$$\begin{aligned} \mathbf{T}_n^k = \big(\sum_{j\in\mathcal{N}} (\mathbf{H}_{j,n}^k)^{\mathrm{H}}(\mathbf{R}_j^k)^{\mathrm{H}}\mathbf{W}_j^k\mathbf{R}_j^k\mathbf{H}_{j,n}^k + \lambda_n\mathbf{I}_{A_n}\big)^{-1} \\ \times (\mathbf{H}_{n,n}^k)^{\mathrm{H}}(\mathbf{R}_n^k)^{\mathrm{H}}\mathbf{W}_n^k, \end{aligned} \tag{9}$$

Because $\mathbf{W}_n^k$ and $\mathbf{R}_n^k$ are fixed, the whole problem has to solved iteratively. In (9), the Lagrange multipliers λ_n should be adjusted so as to satisfy the power constraints, i.e. for every user $\sum_{k\in\mathcal{K}} \mathrm{tr}\{\mathbf{T}_n^k(\mathbf{T}_n^k)^{\mathrm{H}}\} \le P_n^{\max}$.

Algorithm 1: DMT-WMMSE

1 Initialize $\mathbf{T}_n^k = \mathbf{V}_n^k \ \forall n, k$, normalize s.t. $\sum_k \mathrm{tr}\{\mathbf{T}_n^k(\mathbf{T}_n^k)^{\mathrm{H}}\} = P_n^{\max}$;
2 **repeat**
3 Calculate $\mathbf{R}_n^k$ with (3) $\forall n, k$;
4 Calculate $\mathbf{W}_n^k$ with (8) $\forall n, k$;
5 **for** $n = 1, \ldots, N$ **do**
6 **repeat**
7 Calculate $\mathbf{T}_n^k$ with (9) $\forall k$;
8 **if** $\sum_k \mathrm{tr}\{\mathbf{T}_n^k(\mathbf{T}_n^k)^{\mathrm{H}}\} > P_n^{\max}$ **then**
9 increase λ_n;
10 **else**
11 decrease λ_n;
12 **until** $\left|\sum_k \mathrm{tr}\{\mathbf{T}_n^k(\mathbf{T}_n^k)^{\mathrm{H}}\} - P_n^{\max}\right| / P_n^{\max} < \epsilon_1$ *or* $\lambda_n < \epsilon_2$
13 **until** *until convergence*

B. Algorithm

We now describe the first proposed algorithm. Because of the DMT aspect of the problem, we refer to it as DMT WMMSE minimization (DMT-WMMSE). The full procedure is shown as Algorithm 1.

The initialization of the $\mathbf{T}_n^k$'s is done in line 1. Here we initialize these matrices with their corresponding matrices of right singular values of the direct channel matrix $\mathbf{H}_{n,n}^k$, i.e. $\mathbf{T}_n^k = \mathbf{V}_n^k$. We normalize the result to make sure the power constraints are satisfied.

Like [18], we fix two sets of variables and calculate the other one for all tones and users. We first calculate $\mathbf{R}_n^k$ with (3) for all users and tones in line 3 of Algorithm 1, then $\mathbf{W}_n^k$ with (8) for all users and tones in line 4 and then $\mathbf{T}_n^k$ with (9) for all users and tones in line 7. For the transmit matrices, we need to adjust the Lagrange multiplier λ_n so that the power budget is respected. This is done with a simple bisection search. In line 12, the constants ϵ_1 and ϵ_2 are very small positive numbers.

We remark that this algorithm works for any values of A_n, even when all A_n are equal to one. In other words, Algorithm 1 can also be used for the pure spectrum coordination case, i.e. DSM levels 1 and 2. Note that the power allocation is "altruistic" (as opposed to a "selfish" power allocation, e.g. [30], [37]). The allocated power on tone k of user n depends on the magnitude of the crosstalk channels from user n to all its potential victims. So, there is a penalty for user n to allocate power on tone k if there exist potential for large interference to the other users. We also remark that it is possible to calculate (9) for every tone simultaneously by organizing all matrices in a block diagonal structure.

C. Convergence and Complexity

Similar demonstrations of convergence and of the fact the algorithm reaches a stationary point given in [18], [20] also apply to Algorithm 1. In regard to computational cost, the most intensive part of the DMT-WMMSE algorithm is the calculation of the transmit matrices. For each user and each tone, the algorithm has computational complexity of $O(NA_n^3)$—N because the sum in (9) has N terms and A_n^3 because of the multiplications and matrix inverse— and hence a total computational complexity of $O(KN^2 \max_n\{A_n\}^3)$.

IV. Algorithm 2: WMMSE-GDSB

A. Solving the problem in two parts

First, eq. (5) can be rewritten in a slightly different way:

$$\begin{aligned} \max_{\{\mathbf{P},\mathbb{T}\}} \quad & \sum_{n\in\mathcal{N}}\sum_{k\in\mathcal{K}} u_n b_n^k \\ \text{subject to} \quad & \mathrm{tr}\big\{\mathbf{T}_n^k(\mathbf{T}_n^k)^{\mathrm{H}}\big\} = p_n^k \ \forall k,n \\ & \sum_{k\in\mathcal{K}} p_n^k \le P_n^{\max} \ \forall n \end{aligned} \tag{10}$$

The optimization problems are equivalent, but (10) emphasizes more clearly the signal and spectrum coordination parts of the problem. Note that, by fixing $\mathbf{P}$ in (10), we are left with K independent N-user single tone MIMO IC's—thus not coupled across tones anymore. By the same token, by decomposing $\mathbf{T}_n^k$ as $\mathbf{T}_n^k = \sqrt{p_n^k}\overline{\mathbf{T}}_n^k$ such that $\mathrm{tr}\big\{\overline{\mathbf{T}}_n^k(\overline{\mathbf{T}}_n^k)^{\mathrm{H}}\big\} = 1$ and fixing $\overline{\mathbf{T}}_n^k \ \forall n,k$, we are left with a pure spectrum coordination problem. The algorithm proposed in this section solves each of those two parts separately and independently.

Consider the Lagrangian of (10),

$$\begin{aligned} L(\mathbf{P},\mathbb{T},\boldsymbol{\lambda},\boldsymbol{\mu}) = \sum_{n\in\mathcal{N}}\sum_{k\in\mathcal{K}} u_n b_n^k - \sum_{n\in\mathcal{N}} \lambda_n\Big(\sum_{k\in\mathcal{K}} p_n^k - P_n^{\max}\Big) \\ - \sum_{n\in\mathcal{N}}\sum_{k\in\mathcal{K}} \mu_n^k\left(\mathrm{tr}\big\{\mathbf{T}_n^k(\mathbf{T}_n^k)^{\mathrm{H}}\big\} - p_n^k\right). \end{aligned} \tag{11}$$

The KKT condition of (10) states that, if $\{\mathbf{P},\mathbb{T}\}$ is a stationary point of (10), there exist $\boldsymbol{\mu} = \begin{bmatrix}\mu_1^1 & \dots & \mu_N^K\end{bmatrix}^T \in$

$\mathbb{R}^{KN}$, and $\boldsymbol{\lambda} = [\lambda_1 \quad \dots \quad \lambda_N]^T \in \mathbb{R}^N$, such that

$$\nabla_{\mathbf{T}_n^k} L(\mathbf{P}, \mathbb{T}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \mathbf{0} \ \forall n, k \tag{12}$$

$$\nabla_{p_n^k} L(\mathbf{P}, \mathbb{T}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = 0 \ \forall n, k \tag{13}$$

$$\mathrm{tr}\{\mathbf{T}_n^k (\mathbf{T}_n^k)^{\mathrm{H}}\} - p_n^k = 0 \ \forall n, k$$

$$P_n^{\max} - \sum_{k \in \mathcal{K}} p_n^k \geq 0 \ \forall n$$

$$\lambda_n (P_n^{\max} - \sum_{k \in \mathcal{K}} p_n^k) = 0 \ \forall n$$

$$\lambda_n \geq 0 \ \forall n$$

Eqs. (12) and (13) are the stationarity conditions related respectively to the signal and to the spectrum coordination parts of the problem. Our approach is to solve (12) and (13) in two steps. First, we fix $\mathbf{P}$ and $\boldsymbol{\mu}$ and optimize for $\mathbf{T}_n^k$ and $\boldsymbol{\lambda}$ such that $\mathrm{tr}\{\mathbf{T}_n^k (\mathbf{T}_n^k)^{\mathrm{H}}\} = p_n^k \ \forall n, k$. Second, we fix $\overline{\mathbf{T}}_n^k \ \forall n, k$ and $\boldsymbol{\lambda}$, with $\mathrm{tr}\{\overline{\mathbf{T}}_n^k (\overline{\mathbf{T}}_n^k)^{\mathrm{H}}\} = 1$, and optimize $\mathbf{P}$ and $\boldsymbol{\mu}$ such that the power constraints are satisfied. We apply this process iteratively until convergence.

We now focus on how to solve each of these steps.

B. Solving for $\mathbf{T}_n^k$

When solving for the transmit matrices, we can reduce (11) to

$$L(\mathbb{T}, \boldsymbol{\mu}) = \sum_{k \in \mathcal{K}} L^k(\mathbb{T}^k, \boldsymbol{\mu}^k),$$

where

$$L^k(\mathbb{T}^k, \boldsymbol{\mu}^k) = \sum_{n \in \mathcal{N}} u_n b_n^k - \sum_{n \in \mathcal{N}} \mu_n^k \left(\mathrm{tr}\{\mathbf{T}_n^k (\mathbf{T}_n^k)^{\mathrm{H}}\} - p_n^k \right). \tag{14}$$

Here $\mathbb{T}^k$ denotes the set of transmit matrices for all users and for one given tone k, i.e $\mathbb{T}^k = \{\mathbf{T}_n^k \mid n \in \mathcal{N}\}$ and $\boldsymbol{\mu}^k = \begin{bmatrix} \mu_1^k & \cdots & \mu_N^k \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^N$. As already mentioned, this implies that we have to solve K independent single tone MIMO IC problems with each user having a power budget of p_n^k. Previous work has dealt with the single tone MIMO IC problem several times. In the next paragraphs, we discuss (by no means exhaustively) some of the previously proposed solutions.

Recent work has proposed the concept of interference alignment [28]. In this reference it is shown that, when background noise tends to zero and interference is dominant, interference alignment is an optimal strategy. Ref. [29] deals with the case when background noise has to be taken into account,

and interference alignment is combined with an SNR maximizing algorithm. Some other direction of research focuses on game theory [30], [37]. These algorithms cast the single tone MIMO IC problem as a non-cooperative game and converge to a Nash equilibrium point. Other references include [19]–[21], [31].

However, in this paper we opt for the WMMSE strategy for the MIMO IC in [19]–[21]. This is so for a number of reasons. The WMMSE strategy has reasonable computational complexity and is guaranteed to converge to a stationary point. Convergence to a stationary point does not apply to the game theoretic strategies [30], [37]. The interference alignment strategy approaches optimality when background noise tends to zero, which is not necessarily the case for DSL. Ref. [29] would be more suitable for DSL, but it is also not guaranteed to converge to a stationary point. The algorithm in [31], though also able to reach a stationary point, is more computationally complex than the WMMSE approach.

Although we choose the WMMSE strategy, we emphasize that any other of the cited methods can be applied within our framework. We detail the implementation of the WMMSE approach in Section IV-E

C. Solving for $\mathbf{P}$

For the spectrum coordination part of the problem, we can reduce (11) to

$$L(\mathbf{P}, \boldsymbol{\lambda}) = \sum_{n \in \mathcal{N}} \lambda_n P_n^{\max} + \sum_{k \in \mathcal{K}} L^k(\mathbf{p}^k, \boldsymbol{\lambda}),$$

where

$$L^k(\mathbf{p}^k, \boldsymbol{\lambda}) = \sum_{n \in \mathcal{N}} u_n b_n^k - \sum_{n \in \mathcal{N}} \lambda_n p_n^k. \tag{15}$$

Note that here, unlike (14), the problem is not decoupled across tones. In this part of the algorithm, $\overline{\mathbf{T}}_n^k$ is fixed for all n, k. We remind that we have decomposed $\mathbf{T}_n^k$ as $\overline{\mathbf{T}}_n^k \triangleq {}^1/\sqrt{p_n^k}\mathbf{T}_n^k$ so that $\mathrm{tr}\{\overline{\mathbf{T}}_n^k(\overline{\mathbf{T}}_n^k)^{\mathrm{H}}\} = 1$. We can rewrite b_n^k in (15) as

$$b_n^k = \log\left|p_n^k(\overline{\mathbf{T}}_n^k)^{\mathrm{H}}(\mathbf{H}_{n,n}^k)^{\mathrm{H}}(\mathbf{M}_n^k)^{-1}\mathbf{H}_{n,n}^k\overline{\mathbf{T}}_n^k + \mathbf{I}_{A_n}\right|,$$

where $\mathbf{M}_n^k = \sum_{j \neq n} p_j^k \mathbf{H}_{n,j}^k \overline{\mathbf{T}}_j^k (\overline{\mathbf{T}}_j^k)^{\mathrm{H}} (\mathbf{H}_{n,j}^k)^{\mathrm{H}} + \mathbf{I}_{A_n}$

Solving (15) is non-concave and thus non trivial. To overcome this, we solve it in a simplified, per-user

fashion. Consider (15) as a function of the power allocation of one given user, say n,

$$\tilde{L}^k(p_n^k,\lambda_n) = u_n b_n^k + \sum_{j\neq n} u_j \Big(b_j^k \Big|_{\mathbf{p}^{k,(i-1)}} + (p_n^k - p_n^{k,(i-1)}) \frac{\partial b_j^k}{\partial p_n^k}\Big|_{\mathbf{p}^{k,(i-1)}} \Big) - \lambda_n p_n^k.$$

Here, we approximate $b_j^k,\ j\neq n$ by the first order Taylor expansion around the point $\mathbf{p}^{k,(i-1)}$, which represents the power allocation on tone k in iteration $i-1$. It can be shown that b_n^k is a concave function of p_n^k and that b_j^k is a convex function of p_n^k ([38], page 74). Because of the linearization of the convex part, $\tilde{L}^k(p_n^k,\lambda_n)$ is concave in p_n^k.

The next step is to calculate the derivative of $\tilde{L}^k(p_n^k,\lambda_n)$ in p_n^k, set the result to zero and solve for p_n^k. In order to do that, we take into account that [39]

$$\frac{\partial \log|\mathbf{A}|}{\partial \mathbf{X}^*_{m,n}} = \mathrm{tr}\Big\{ \mathbf{A}^{-1}\frac{\partial \mathbf{A}}{\partial \mathbf{X}^*_{m,n}}\Big\}, \tag{16}$$

$$\frac{\partial \mathbf{A}^{-1}}{\partial \mathbf{X}^*_{m,n}} = -\mathbf{A}^{-1}\frac{\partial \mathbf{A}}{\partial \mathbf{X}^*_{m,n}}\mathbf{A}^{-1}, \tag{17}$$

We proceed in two steps. First, using (16), we write

$$\frac{\partial b_n^k}{\partial p_n^k} = \mathrm{tr}\Big\{ \big(p_n^k \mathbf{S}_n^k + \mathbf{I}_{A_n}\big)^{-1}\mathbf{S}_n^k\Big\}. \tag{18}$$

Here, we define $\mathbf{S}_n^k \triangleq (\overline{\mathbf{T}}_n^k)^{\mathrm{H}}(\mathbf{H}_{n,n}^k)^{\mathrm{H}}(\mathbf{M}_n^k)^{-1}\mathbf{H}_{n,n}^k\overline{\mathbf{T}}_n^k$. Second, observing (16), (17) and using the chain rule, we write

$$\begin{aligned}\frac{\partial b_j^k}{\partial p_n^k}\Big|_{\mathbf{p}^{k,(i-1)}} &= \mathrm{tr}\Big\{ \mathbf{E}_j^k(\mathbf{T}_j^k)^{\mathrm{H}}(\mathbf{H}_{j,j}^k)^{\mathrm{H}}\frac{\partial(\mathbf{M}_j^k)^{-1}}{\partial p_n^k}\mathbf{H}_{j,j}^k\mathbf{T}_j^k\Big\} \\ &= -\mathrm{tr}\Big\{ (\overline{\mathbf{T}}_n^k)^{\mathrm{H}}(\mathbf{H}_{j,n}^k)^{\mathrm{H}}(\mathbf{M}_j^k)^{-1}\mathbf{H}_{j,j}^k\mathbf{T}_j^k\mathbf{E}_j^k \\ &\qquad \times (\mathbf{T}_j^k)^{\mathrm{H}}(\mathbf{H}_{j,j}^k)^{\mathrm{H}}(\mathbf{M}_j^k)^{-1}\mathbf{H}_{j,n}^k\overline{\mathbf{T}}_n^k\Big\},\end{aligned} \tag{19}$$

where $\mathbf{E}_j^k$ is given by (6). Finally, by combining (18) and (19), we obtain the resulting equation for power allocation:

$$\frac{\partial \tilde{L}(p_n^k,\lambda_n)}{\partial p_n^k} = u_n \mathrm{tr}\Big\{ \big(p_n^k \mathbf{S}_n^k + \mathbf{I}_{A_n}\big)^{-1}\mathbf{S}_n^k\Big\} - \tau_n^k - \lambda_n = 0. \tag{20}$$

Here, we define

$$\tau_n^k \triangleq -\sum_{j\neq n} u_j \frac{\partial b_j^k}{\partial p_n^k}\Big|_{\mathbf{p}^{k,(i-1)}}, \tag{21}$$

where the partial derivative is given by (19).

For the solution, for every user we should find p_n^k such that (20) holds for every tone. There should also be a search for the Lagrange multiplier λ_n so that the power budget is respected. Although not clear at a first glance, (20) is a type of waterfilling formula with frequency selectivity. The frequency selectivity is due to the term τ_n^k, which represents how much damage is inflicted to other users if user n allocates power on tone k. This variable should be large if there is potential for large interference to other users. We remark that the distributed spectrum balancing (DSB) algorithm has a similar power allocation formula [5], [6], which is a special case of (20) for the SISO case. For this case, (20) can be rearranged in a closed-form waterfilling formula where τ_n^k distorts the waterlevel so that damage to other users is accounted for. Eq. (20), however, is more complicated but, as we see next, can be solved efficiently.

The fact that (20) is a generalization of the DSB power allocation formula motivates our choice for the name *generalized DSB* (GDSB) for the spectrum coordination part of the algorithm.

D. Solving (20)

Let us rewrite (20) and define the function $f(p)$ as

$$f(p) = up^{-1}\mathrm{tr}\big\{(\mathbf{S} + p^{-1}\mathbf{I})^{-1}\mathbf{S}\big\} - \nu, \tag{22}$$

$\mathbf{S} \in \mathbb{H}^{A\times A}$, where $\mathbb{H}^{A\times A}$ represents hermitian matrices of size $A \times A$; and $\nu \in \mathbb{R}$, $\nu \geq 0$. For our case $\mathbf{S}$ is actually not only hermitian, but also positive semi-definite (it is a covariance matrix). However, in order to get some insight, the analysis in this section mostly focuses on the case of $\mathbf{S}$ being hermitian, except for Corollary 1. For convenience, in this section we ignore the subscripts and superscripts of (20). These will be recovered when suitable.

As there is no closed form solution for the roots of (20), one particularly straightforward approach is to use a root-finding algorithm, e.g. the Newton method [40]. However, such a method would be very inefficient, given that it should be repeated for all users and all tones of the DSL network. Plus, there is the power budget: for a given user, such a method would have to solve the problem for every tone multiple times until the right Lagrange multiplier λ_n in (20) is found and the power budget is respected [in (22), $\nu = \lambda_n + \tau_n^k$].

We can find the roots (22) efficiently with a different strategy. In the following, we use basic linear algebra concepts to show how to obtain the coefficients $\boldsymbol{\alpha} = [\alpha_1 \quad \dots \quad \alpha_{A+1}]^T \in \mathbb{R}^{A+1}$ of a polynomial $f'(p) = \sum_{i=1}^{A+1} \alpha_i p^{A+1-i}$ whose roots are the same as those of the original (22).

We consider the eigenvalue decomposition of $\mathbf{S}$, $\mathbf{S} = \mathbf{Q}\boldsymbol{\Xi}\mathbf{Q}^{\mathrm{H}}$. Because $\mathbf{S}$ is hermitian, $\mathbf{Q}$ is orthonormal and all eigenvalues $\xi_i,\ i = 1, \dots, A$ are real. Thus $\boldsymbol{\Xi} = \mathrm{diag}\left\{\begin{bmatrix}\xi_1 & \dots & \xi_A\end{bmatrix}\right\}$ is a real matrix. We rewrite

(22) as

$$f(p) = up^{-1}\mathrm{tr}\big\{(\mathbf{Q}\boldsymbol{\Xi}\mathbf{Q}^{\mathrm{H}} + p^{-1}\mathbf{Q}\mathbf{Q}^{\mathrm{H}})^{-1}\mathbf{Q}\boldsymbol{\Xi}\mathbf{Q}^{\mathrm{H}}\big\} - \nu$$
$$= up^{-1}\mathrm{tr}\big\{(\boldsymbol{\Xi} + p^{-1}\mathbf{I})^{-1}\boldsymbol{\Xi}\big\} - \nu = u\sum_{i=1}^{A}\frac{\xi_i}{p\xi_i + 1} - \nu. \quad (23)$$

1) Analysis: First, we do a brief analysis of $f(p)$ and its behavior. Clearly, as seen in (23), we can write $f(p) : \mathbb{R} \to \mathbb{R}$. Also, it is easy to see that

$$\lim_{p\to-\infty} f(p) = \lim_{p\to+\infty} f(p) = -\nu. \quad (24)$$

Proposition 1. If, $\nu > 0$, the number of roots to f(p) is equal to the number of distinct non-zero eigenvalues of $\mathbf{S}$.

Proof. To prove this statement, we need two things. First, we focus on the behavior of $f(p)$ when $p \to -1/\xi_i$. Second, we check the derivative of $f(p)$

We assume that $\mathbf{S}$ has D distinct non-zero eigenvalues. Note that eigenvalues equal to zero do not contribute to $f(p)$. When $p = -1/\xi_i$, then $p\xi_i + 1 = 0$ and (23) is undefined. Consider $\lim_{p\to\{-1/\xi_i\}^-} f(p)$ and $\lim_{p\to\{-1/\xi_i\}^+} f(p)$, $i = 1, \ldots, D$. Here $p \to \{P\}^-$ and $p \to \{P\}^+$ denote p approaching P from the left and from the right, respectively. It is clear that these two limits will be either ∞ or $-\infty$. At this point we cannot guarantee, however, whether they are both equal or whether they differ in sign.

Calculating the derivative of $f(p)$, we obtain

$$\frac{\partial f(p)}{\partial p} = -u\sum_{i=1}^{A}\frac{(\xi_i)^2}{(p\xi_i + 1)^2}. \quad (25)$$

Define $\mathcal{P} = \{p \in \mathbb{R} \mid p \neq -1/\xi_i,\ i = 1, \ldots, D\}$. Then, (25) implies that in $\mathcal{P}$ the derivative $\partial f(p)/\partial p$ is negative and so $f(p)$ is strictly decreasing. That leads us to conclude that

$$\lim_{p\to[-1/\xi_i]^-} f(p) = -\infty \quad (26)$$

$$\lim_{p\to[-1/\xi_i]^+} f(p) = \infty \quad (27)$$

For $p \in \mathcal{P}$, $f(p)$ is continuous. The continuity and the strictly decreasing behavior of $f(p)$ in $\mathcal{P}$ plus (24), (26) and (27) lead us to complete the proof. □

As (23) has D distinct solutions, an additional pertinent question is how to handle these different solutions. For example, of the D roots of $f(p)$, which one should be picked? We remind that these roots should be the power allocation p_n^k for a given user and tone in (10). For a given user with K tones, if

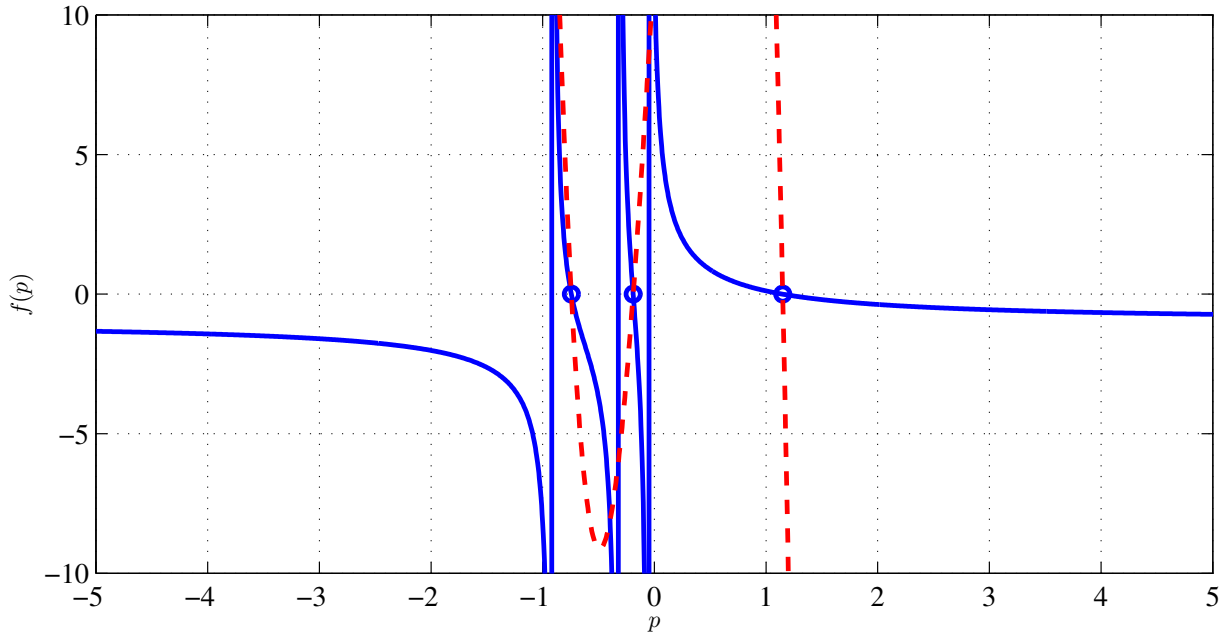

Fig. 2. Example of the behavior of the function $f(p) = u\mathrm{tr}\{(p\mathbf{S}+\mathbf{I})^{-1}\mathbf{S}\} - \nu$ for a randomly generated, hermitian and positive semi-definite $\mathbf{S}$ and $\nu = 0.75$. The solid line represents $f(p)$ and the dotted line represents the polynomial $f'(p)$. Their roots are the same. When $\mathbf{S}$ is positive semi-definite, at most one root is non-negative.

each tone has D solutions, is it necessary to combine these KD solutions? Would all combinations be equivalent? Fortunately, because of the following corollary we need not worry about this issue.

Corollary 1. If $\nu > 0$ and $\mathbf{S}$ is positive semi-definite, there can be at most one non-negative root to $f(p)$.

Proof. If $\mathbf{S}$ is positive semi-definite, then $\xi_i \geq 0$, $i = 1, \ldots, D$. That implies that $f(p)$ approaches infinity only for $p < 0$. Order the values of p for which $f(p)$ approaches infinity in ascending order as $\{-1/\xi_1,\ -1/\xi_2,\ \ldots,\ -1/\xi_D\}$. There are D roots to $f(p)$. Because of (24) and the strictly decreasing behavior of $f(p)$, there is one root between $-1/\xi_i$ and $-1/\xi_{i+1}$, $i = 1, \ldots, D-1$ and one root between $-1/\xi_D$ and $+\infty$. This means that, for $p \geq 0$, $f(p)$ crosses zero at most once. □

As already mentioned, for our problem $\mathbf{S}$ is a covariance matrix, so it is always positive semi-definite. For the solution of the spectrum coordination problem, what happens in practice is that either only one root of $f(p)$ is positive or all roots are negative. Since a negative p_n^k has no physical meaning, for the former case we need to discard the negative roots of $f(p)$. For the latter case, the optimal solution is $p_n^k = 0$ and user n allocates no power on tone k. Thus Corollary 1 guarantees there will always be only

one feasible solution.

In Fig. 2, the solid line illustrates the behavior of $f(p)$ for a hermitian, positive semi-definite and randomly generated $\mathbf{S} \in \mathbb{H}^{3\times 3}$ and $\nu = 0.75$. Note that there is only one non-negative root.

2) Eq. (27) in polynomial form: Clearly, (23) can be written as a polynomial and thus be solved efficiently. In this subsection, we consider some examples. First, consider $\mathbf{S} \in \mathbb{H}^{3\times 3}$. Set $f(p) = 0$. Re-arranging, we obtain

$$f'(p) = u\sum_{i=1}^{3}\xi_i \sum_{\substack{j=1\\ j\neq i}}^{3}(p\xi_i + 1) - \nu\prod_{i=1}^{3}(p\xi_i + 1) = 0.$$

Re-arranging again, we finally get $f'(p_n^k)$ to be (here we re-introduce the subscripts and superscripts)

$$(p_n^k)^3\big(-\nu_n^k|\mathbf{S}_n^k|\big) + (p_n^k)^2\big(3u_n|\mathbf{S}_n^k| - \nu_n^k(\xi_1\xi_2 + \xi_1\xi_3 + \xi_2\xi_3)\big) + p_n^k\Big(u_n\big(\sum_{i=1}^{3}\xi_i\sum_{\substack{j=1\\ j\neq i}}^{3}\xi_j\big) - \nu_n^k\mathrm{tr}\big\{\mathbf{S}_n^k\big\}\Big) + \big(u_n\mathrm{tr}\big\{\mathbf{S}_n^k\big\} - \nu_n^k\big) = 0. \quad (28)$$

We remind that $\nu_n^k = \tau_n^k + \lambda_n$. Eq. (28) is shown in the dotted line in Fig. 2 for the same hermitian, positive semi-definite and randomly generated $\mathbf{S}$ and $\nu = 0.75$. Note that the roots of the original $f(p)$ and of $f'(p)$ are the same.

As another example, if $\mathbf{S} \in \mathbb{H}^{2\times 2}$, then $f'(p)$ is written as

$$(p_n^k)^2\big(-\nu_n^k|\mathbf{S}_n^k|\big) + p_n^k\big(2u_n|\mathbf{S}_n^k| - \nu_n^k\mathrm{tr}\big\{\mathbf{S}_n^k\big\}\big) + \big(u_n\mathrm{tr}\big\{\mathbf{S}_n^k\big\} - \nu_n^k\big) = 0. \quad (29)$$

If S_n^k is a scalar then we are back at the SISO case, which was solved in [5], [6].

Solving this polynomial is much easier and computationally cheaper than, for example, the root-finding algorithm mentioned in Section IV-D. For example, for polynomials of degrees 2, 3 and 4 there are closed-form solutions to the roots. For polynomials of larger degree, finding the roots has computational complexity $O(A^2)$. However, we can do better than that. Because $\mathbf{S}$ in our problem is positive semi-definite, we are only interested in the largest (rightmost) root of $f(p)$ (see Corollary 1). We can thus use the *power method* ([41], Sec. 7.3.1). The power method obtains the largest eigenvalue in terms of absolute value of a given matrix. Consider the *companion matrix* formed with the coefficients of, say,

(28).

$$\begin{bmatrix} 0 & 0 & -\alpha_4/\alpha_1 \\ 1 & 0 & -\alpha_2/\alpha_1 \\ 0 & 1 & -\alpha_2/\alpha_1 \end{bmatrix}$$

Here $\alpha_1 = -\nu_n^k|\mathbf{S}_n^k|$, $\alpha_2 = 3u_n|\mathbf{S}_n^k| - \nu(\xi_1\xi_2 + \xi_1\xi_3 + \xi_2\xi_3)$, etc. It is well known that the eigenvalues of the companion matrix correspond to the roots of the polynomial $f'(p) = \sum_{i=1}^{A+1} \alpha_i p^{A+1-i}$. Hence, we only need to apply the power method and find the largest eigenvalue of the companion matrix. We can avoid obtaining a large negative eigenvalue with a simple procedure, which entails running the power method at most twice. First, run the power method. If a positive eigenvalue is found, then the procedure is successful and can finish. If a negative eigenvalue is found, then run the power method for a second time with a shift in $f'(p)$ to the right. The shift is denoted by $\delta \in \mathbb{R}$, $\delta > 0$. The polynomial in (28) becomes

$$\begin{aligned} &(p_n^k - \delta)^3\big(-\nu_n^k|\mathbf{S}_n^k|\big) \\ &+(p_n^k - \delta)^2\big(3u_n|\mathbf{S}_n^k| - \nu(\xi_1\xi_2 + \xi_1\xi_3 + \xi_2\xi_3)\big) \\ &+(p_n^k - \delta)\Big(u_n\big(\sum_{i=1}^{A}\xi_i\sum_{\substack{j=1\\ j\neq i}}^{A}\xi_j\big)\nu_n^k\mathrm{tr}\big\{\mathbf{S}_n^k\big\}\Big) \\ &+\big(u_n\mathrm{tr}\big\{\mathbf{S}_n^k\big\} - \nu_n^k\big) = 0. \end{aligned}$$

The shift should be equal to the absolute value of the negative number that was the result of the first run of the power method. The power method has computational complexity of $O(A)$.

E. Algorithm

We now proceed to writing an algorithm to solve (10). Since our method of choice for the signal coordination part is the WMMSE algorithm and our choice for the spectrum coordination part is the generalized DSB method (GDSB), we refer to the combination of the two as WMMSE-GDSB.

An algorithm description is provided in Algorithm 2. We initialize the algorithm with $\mathbf{T}_n^k$ equal to the right singular vector matrices in line 1. Also, we initialize the power allocation in line 2. In lines 4-14, we solve the spectrum coordination part and in lines 15-25 we solve the signal coordination part. We remark that it is not necessary to solve the spectrum coordination part first and the signal coordination part second. Inverting that order would work just as well. We briefly detail the two parts in the next paragraphs.

In lines 4 and 5, $\mathbf{S}_n^k$ and τ_n^k are calculated for every user and tone. Since these are calculated outside the 'for' loop in lines 6-14, power allocation is done in parallel and the calculation of τ_n^k takes into account the power allocation of the previous iteration. In line 9, for every user separately p_n^k is found by solving a polynomial—ξ_i is the ith eigenvalue of $\mathbf{S}_n^k$. Here we can either find all roots and keep the largest root or use the power method. Note that we should find an appropriate Lagrange multiplier so that the power budget is respected. We do this with a simple bisection search.

For the signal coordination part, the procedure is very similar to the one described in Algorithm 1. The variables $\mathbf{R}_n^k$, $\mathbf{W}_n^k$ and $\mathbf{T}_n^k$ should be calculated. As before, we fix two variables and solve for the third one. One difference between the signal coordination part in Algorithm 1 and in Algorithm 2 is that for the latter we should solve an equality constrained problem for every tone separately. That implies that the Lagrange multipliers, denoted as μ_n^k in Algorithm 2, can be negative. The loop in lines 19-25 will converge when the total power allocated for the matrix $\mathbf{T}_n^k$ is sufficiently close to p_n^k, calculated in the spectrum coordination part of the algorithm.

F. Convergence and complexity

First, we analyze the convergence of the spectrum coordination part of the algorithm under fixed $\overline{\mathbf{T}}_n^k \ \forall n, k$ when $\mathbf{T}_n^k \in \mathbb{C}^{2\times 2}$ and with a high SNR approximation.

Proposition 2. Consider $\overline{\mathbf{T}}_n^k$, u_n and τ_n^k fixed $\forall n, k$, and $A_n = 2 \ \forall n$. Define $\overline{\mathbf{H}}_{n,j}^k = \mathbf{H}_{n,j}^k \overline{\mathbf{T}}_j^k (\overline{\mathbf{T}}_j^k)^{\mathrm{H}} (\mathbf{H}_{n,j}^k)^{\mathrm{H}}$. Define $\mathbf{S}_n^k = \left[\begin{smallmatrix} a & b^* \\ b & a \end{smallmatrix}\right]$. Considering a high SNR approximation of the form

$$\frac{\left|b\right|^2}{(a^2 - \left|b\right|^2)^2} \approx 0,$$

if

$$(N-1) \max_{\substack{j,n,k \\ j \neq n}} \left\{ \frac{\mathrm{tr}\{\mathrm{adj}\{\overline{\mathbf{H}}_{n,j}^k\}\overline{\mathbf{H}}_{n,n}^k\}}{2\left|\overline{\mathbf{H}}_{n,n}^k\right|} \right\} < 1,$$

where $\mathrm{adj}\{\cdot\}$ stands for the matrix adjugate operator, then the spectrum coordination part of Algorithm 2 converges to a unique fixed point.

The details of the proof are left to the Appendix. It basically consists in showing that the sequence of iterations is a contraction mapping in the power updates if users operate with high SNR and if the crosstalk channels are relatively weak compared to the direct channels. We emphasize that Proposition 2 deals only with the spectrum coordination algorithm, i.e. the GDSB algorithm—the pseudo-code of the GDSB is the same as that in Algorithm 2 excluding lines 15-25. Although Proposition 2 deals only with

Algorithm 2: WMMSE-Generalized DSB

Initialize $\mathbf{T}_n^k = \mathbf{V}_n^k \ \forall n,k$, normalize s.t. $\mathrm{tr}\{\mathbf{T}_n^k(\mathbf{T}_n^k)^{\mathrm{H}}\} = P_n^{\max}$;
Initialize $\mathbf{P} = \mathbf{0}_{K\times N}$;
repeat
  Calculate τ_n^k with (21) $\forall n,k$;
  Calculate $\mathbf{S}_n^k \triangleq (\overline{\mathbf{T}}_n^k)^{\mathrm{H}}(\mathbf{H}_{n,n}^k)^{\mathrm{H}}(\mathbf{M}_n^k)^{-1}\mathbf{H}_{n,n}^k\overline{\mathbf{T}}_n^k \ \forall n,k$;
  for $n = 1,\ldots,N$ **do**
    repeat
      $\nu_n^k = \tau_n^k + \lambda_n$;
      $p_n^k = \max(\mathcal{P},0),\ \mathcal{P} = \{p | u_n \sum_{i=1}^{A_n}\xi_i \sum_{\substack{j=1\\ j\neq i}}^{A_n} - \nu_n^k \prod_{i=1}^{A_n}(p\xi_i+1) = 0\}\ \forall k$;
      if $\sum_k p_n^k > P_n^{\max}$ **then**
        increase λ_n;
      else
        decrease λ_n;
    until $\left|\sum_k p_n^k - P_n^{\max}\right|/P_n^{\max} < \epsilon_1$ *or* $\lambda_n < \epsilon_2$
  Calculate $\mathbf{R}_n^k$ with (3) $\forall n,k$;
  Calculate $\mathbf{W}_n^k$ with (8) $\forall n,k$;
  for $n = 1,\ldots,N$ **do**
    for $k = 1,\ldots,K$ **do**
      repeat
        Calculate $\mathbf{T}_n^k$ with (9);
        if $\mathrm{tr}\{\mathbf{T}_n^k(\mathbf{T}_n^k)^{\mathrm{H}}\} > p_n^k$ **then**
          increase μ_n^k;
        else
          decrease μ_n^k;
      until $\left|\mathrm{tr}\{\mathbf{T}_n^k(\mathbf{T}_n^k)^{\mathrm{H}}\} - p_n^k\right|/p_n^k < \epsilon_1$
until *until convergence*

the 2×2 case under a high SNR assumption, we observe convergence experimentally for a broad range of scenarios.

We make a small digression to remark that Proposition 2 is a generalization of the convergence result in [23]. In this reference, the SISO spectrum balancing problem is considered, whereas our result focuses on a MIMO spectrum balancing—both [23] and Proposition 2 assume a high SNR regime. The power allocation formula in [23] is the equivalent of (20) for the SISO case. To show that the result in [23] is a special case of ours, consider $\mathbf{H}_{n,n}^k = d_{n,n}^k\mathbf{I}_2 \ \forall n,k$. Also consider $\mathbf{H}_{n,j}^k = h_{n,j}^k\mathbf{I}_2 \ \forall n,j,k,\ n\neq j$. Because of the diagonal structure of $\mathbf{H}_{n,n}^k$, the optimal $\overline{\mathbf{T}}_n^k$ equals the identity matrix times $1/\sqrt{2}$. Since the optimal transmit matrix is diagonal, this corresponds to a pure spectrum coordination case (see Section

I). The condition in Proposition 2 is rewritten as

$$(N-1)\max_{\substack{j,n,k\\ j\neq n}}\left\{\frac{\mathrm{tr}\left\{|h_{n,j}^k|^2\mathbf{I}_2|d_{n,n}^k|^2\mathbf{I}_2\right\}}{2|d_{n,n}^k|^4}\right\}$$

$$=(N-1)\max_{\substack{j,n,k\\ j\neq n}}\frac{|h_{n,j}^k|^2}{|d_{n,n}^k|^2}<1,$$

which is exactly the result of [23].

Focusing back on Algorithm 2, we know that signal coordination part (the WMMSE algorithm) converges. By applying the signal and spectrum coordination parts in sequence and independently, we have consistently observed from the simulations a monotonically increasing objective function. Some results will be given in Section V. Also, these results also show that after convergence the WMMSE-GDSB algorithm, just like the DMT-WMMSE, reaches a stationary point. We will see, however, that the WMMSE-GDSB needs a significant smaller number of iterations to converge.

The computational complexity of the signal coordination part is $O(KN^2\max_n\{A_n\}^3)$. The computational complexity of the spectrum coordination part is also $O(KN^2\max_n\{A_n\}^3)$.

V. Numerical Results

In this section, we present some results for the two proposed algorithms. Two scenarios are considered: an ADSL downstream scenario and a VDSL upstream scenario. For all simulations, the values for the inter-channel spacing Δ_f and the symbol rate f_s are set to 4.3125 kHz and 4 kHz, respectively.

A. ADSL downstream

We simulate a system with cables of 0.5 mm (AWG 24) and an SNR gap of 12.8 dB. We consider the near-far downstream ADSL scenario depicted in Fig. 3, where two users, each with 2 transceivers (2 direct modes), share the system. Each user has a total power budget of 23.4 dBm. For each line, noise model ANSI A is adopted.

The near-far scenario is one of the main testing grounds for DSM algorithms and it encompasses a good deal of the complexity of the problem. User 1, connected to a central office (CO), shares a binder with a far user connected to a remote terminal (RT). Crosstalk from user 2 into user 1 can be prohibitively high, to the point where transmission for user 1 may be seriously compromised. A good solution is one in which user 2 protects user 1 from excessive crosstalk so that both users socially share the systems resources. Referring to Fig. 3, we set $l_1 = 4$ km and $l_2 = 3$ km and $d_2 = 3$ km.

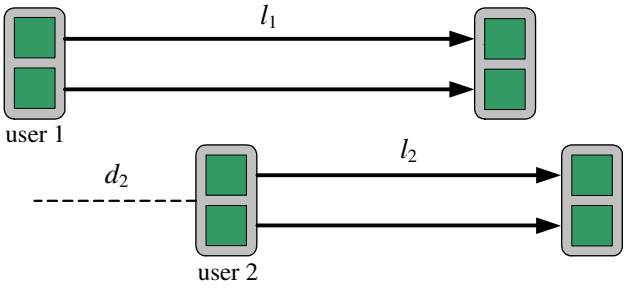

Fig. 3. Near-far downstream ADSL scenario.

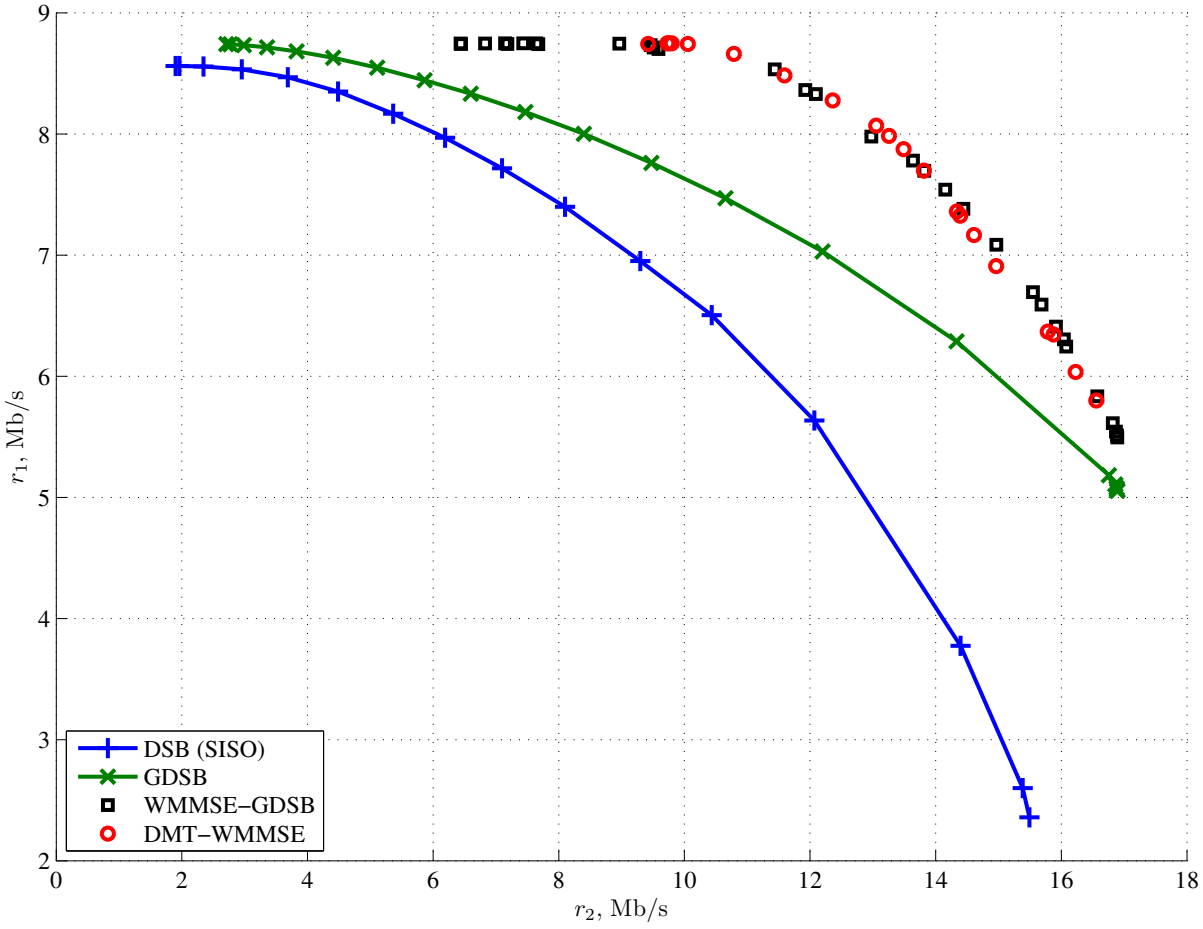

Fig. 4. Rate region for the downstream ADSL scenario for four different algorithms: DSB (SISO), GDSB, WMMSE-GDSB and DMT-WMMSE.

We simulate four different algorithms for this scenario, namely the DMT-WMMSE, the WMMSE-GDSB, the GDSB and the SISO DSB. The first two are the algorithms presented in Sections III and IV. As already mentioned, the GDSB corresponds to the spectrum coordination part in WMMSE-GDSB. This algorithm fixes the transmit matrices to the right singular vectors matrix of the direct channel matrix

for all users and iterates only the power allocation formula until convergence. The SISO DSB algorithm is the solution of [5], [6]. It is a pure spectrum coordination algorithm where $A_n = 1\ \forall n$. For SISO DSB, the number of users in the network is 4.

In Fig. 4 we depict the four corresponding rate regions. We see that the gains obtained with the combined signal and spectrum coordination for this scenario are quite significant. We also see that the DMT-WMMSE and the WMMSE-GDSB have very similar rate performance, while, not surprisingly, the GDSB is worse. The latter does not profit from signal coordination, which is the reason for its weaker performance. Referring to Fig. 4, when fixing a data rate of 7 Mb/s for user 1, the SISO DSB can provide approximately 9.3 Mb/s for user 2. For the GDSB, we achieve approximately 12.2 Mb/s for user 2. For both DMT-WMMSE and WMMSE-GDSB, we achieve about 15 Mb/s for user 2. Comparing these last two to the SISO DSB, the gain is approximately 80 %.

The power allocations for the GDSB and the WMMSE-GDSB are shown in Figs. 5 and 6. The two figures have user 1 with a data rate of 7 Mb/s. The difference between the two algorithms is the signal coordination part, and this reflects on the final power allocation. For GDSB there is no signal coordination, so excessive crosstalk to user 1 should be avoided by significant power reduction by user 2 in lower frequencies. WMMSE-GDSB can coordinate users on both the signal and spectrum levels. The signal coordination part can cancel some of the crosstalk. However, some of it may still remain, and hence should be avoided with spectrum coordination. This is reflected in the final power allocation. On lower frequencies, a lot of the crosstalk can be canceled, so the power reduction for user 2 on lower frequencies is significantly smaller.

Finally, convergence of the two algorithms proposed in this paper is depicted in Fig. 7. If the data rate performance of the two algorithms is very similar, in terms of convergence there is a big difference. In the example of the figure, the WMMSE-GDSB attains convergence in less than 200 iterations, while the DMT-WMMSE needs more than 3000 iterations. This observation holds true for a large range of scenarios.

The run-time complexity of one iteration of the DMT-WMMSE (one run of the “repeat” loop in Algorithm 1) takes a bit less time than one iteration of the WMMSE-GDSB (one run of the “repeat” loop in Algorithm 2). To give a practical example, in our implementation for the scenario in Fig. 3 one iteration of the DMT-WMMSE takes on average 0.3 seconds and one iteration of the WMMSE-GDSB takes on average 0.9 seconds. However, the fact that the latter needs much less iterations to converge more than makes up for this. Plus, in a changing environment, such faster convergence is of great importance. Applied either to wireless or DSL, a good algorithm should be able to quickly respond to a changing

network. In DSL, users turn on and off, in wireless the channel changes constantly.

B. VDSL upstream

In the second experiment, we simulate a VDSL upstream scenario. Cables of 0.4 mm (AWG 26) are used and an SNR gap of 9.45 dB is considered. Each user has a total power budget of 14.5 dBm. For each line, noise model ETSI A is adopted with a background noise level of -140 dBm/Hz. We use the FDD 998 frequency bandplan over POTS up to 12 MHz. The scenario has 4 users, two with long lines and two with short lines, each with two transceivers (two direct modes). The long lines are 1.2 km and the short lines are 0.6 km. For this scenario, the users with the short lines have to avoid excessive crosstalk to the users with the long lines.

We depict the rate regions of the SISO DSB, the GDSB and the WMMSE-GDSB in Fig. 8. We again conclude that the gains of adding signal coordination are very significant.

VI. Conclusion

Signal coordination and spectrum coordination have become crucial for mitigating the effects of multi-user interference in DSL networks. In this paper, we have combined the two strategies for the DMT MIMO IC, a scenario where only some of the infrastructure for full two-sided signal coordination is available. We have presented two distinct algorithms, the DMT-WMMSE and the WMMSE-GDSB. By means of numerical experiments, we have demonstrated that the addition of per-user MIMO signal coordination can have a significant impact on the achievable rates when compared to multi-user spectrum coordination only.

Appendix

Proof of Proposition 2

We analyze the N user case. In the following we focus on parallel updates, i.e. all users change their powers at the same time while taking into account the interference of the previous iteration. We write $\mathbf{S}_n^{k,(i+1)} = (\overline{\mathbf{T}}_n^k)^{\mathrm{H}}(\mathbf{H}_{n,n}^k)^{\mathrm{H}}(\mathbf{M}_n^{k,(i)})^{-1}\mathbf{H}_{n,n}^k\overline{\mathbf{T}}_n^k$ and $\mathbf{M}_n^{k,(i)} = \mathbf{I}+\sum_{j\neq n} p_j^{k,(i)}\mathbf{H}_{n,j}^k\overline{\mathbf{T}}_j^k(\overline{\mathbf{T}}_j^k)^{\mathrm{H}}(\mathbf{H}_{n,j}^k)^{\mathrm{H}}$, where the bracketed superscripts denote iteration. Also define $[x]^+ = \max\{x, 0\}$ and $[x]^- = \max\{-x, 0\}$. For the 2×2 case, the power allocation for user n, tone k on iteration i can be found in closed form. Taking into account the coefficients of the polynomial in (29), and defining $\alpha_1 = -\nu_n^{k,(i)}\left|\mathbf{S}_n^{k,(i)}\right|$, $\alpha_2 = 2u_n\left|\mathbf{S}_n^{k,(i)}\right| - \nu_n^{k,(i)}\mathrm{tr}\{\mathbf{S}_n^{k,(i)}\}$ and $\alpha_3 = u_n\mathrm{tr}\{\mathbf{S}_n^{k,(i)}\} - \nu_n^{k,(i)}$, the two roots are given by $(-\alpha_2 \pm \sqrt{\alpha_2^2 - 4\alpha_1\alpha_3})/(2\alpha_1)$.

Because of Corollary 1, we only need the largest root. If this roots is non-negative, we assign it to $p_n^{k,(i)}$, otherwise we set $p_n^{k,(i)}=0$. Thus, we can write

$$p_n^{k,(i)}=\left[\max\left\{\frac{u_n}{\nu_n^{k,(i)}}-\frac{\mathrm{tr}\{\mathbf{S}_n^{k,(i)}\}}{2|\mathbf{S}_n^{k,(i)}|}+\sqrt{\Delta};\right.\right.$$
$$\left.\left.\frac{u_n}{\nu_n^{k,(i)}}-\frac{\mathrm{tr}\{\mathbf{S}_n^{k,(i)}\}}{2|\mathbf{S}_n^{k,(i)}|}-\sqrt{\Delta}\right\}\right]^+,$$

where $\Delta=(u_n)^2/(\nu_n^{k,(i)})^2+\mathrm{tr}\{\mathbf{S}_n^{k,(i)}\}^2/4|\mathbf{S}_n^{k,(i)}|^2-1/|\mathbf{S}_n^{k,(i)}|$. Without loss of generality, we consider that the power budget for every user is tight. If that is not the case, we can use the concept of the "virtual tone" [23].

With the high SNR assumption, we can write

$$\frac{\mathrm{tr}\{\mathbf{S}_n^{k,(i)}\}^2}{4|\mathbf{S}_n^{k,(i)}|^2}-\frac{1}{|\mathbf{S}_n^{k,(i)}|}=\frac{|b|^2}{(a^2-|b|^2)^2}\approx 0$$
$$\Rightarrow p_n^{k,(i)}=\left[\frac{2u_n}{\nu_n^{k,(i)}}-\frac{\mathrm{tr}\{\mathbf{S}_n^{k,(i)}\}}{2|\mathbf{S}_n^{k,(i)}|}\right]^+. \tag{30}$$

This assumption translates into considering that the diagonal elements of SNR matrix $\mathbf{S}_n^{k,(i)}$ are much larger than the off diagonal elements. Now we write

$$\max_n\left\{\max\left\{\sum_k\left[p_n^{k,(i+1)}-p_n^{k,(i)}\right]^+;\sum_k\left[p_n^{k,(i+1)}-p_n^{k,(i)}\right]^-\right\}\right\}$$
$$=\max_n\left\{\max\left\{\sum_k\left[\left[\frac{2u_n}{\nu_n^{k,(i+1)}}-\frac{\mathrm{tr}\{\mathbf{S}_n^{k,(i+1)}\}}{2|\mathbf{S}_n^{k,(i+1)}|}\right]^+-\left[\frac{2u_n}{\nu_n^{k,(i)}}-\frac{\mathrm{tr}\{\mathbf{S}_n^{k,(i)}\}}{2|\mathbf{S}_n^{k,(i)}|}\right]^+\right]^+;\right.\right.$$
$$\left.\left.\sum_k\left[\left[\frac{2u_n}{\nu_n^{k,(i+1)}}-\frac{\mathrm{tr}\{\mathbf{S}_n^{k,(i+1)}\}}{2|\mathbf{S}_n^{k,(i+1)}|}\right]^+-\left[\frac{2u_n}{\nu_n^{k,(i)}}-\frac{\mathrm{tr}\{\mathbf{S}_n^{k,(i)}\}}{2|\mathbf{S}_n^{k,(i)}|}\right]^+\right]^-\right\}\right\} \tag{31}$$
$$\le\max_n\left\{\max\left\{\sum_k\left[\frac{2u_n}{\nu_n^{k,(i)}}-\frac{\mathrm{tr}\{\mathbf{S}_n^{k,(i+1)}\}}{2|\mathbf{S}_n^{k,(i+1)}|}-\frac{2u_n}{\nu_n^{k,(i)}}+\frac{\mathrm{tr}\{\mathbf{S}_n^{k,(i)}\}}{2|\mathbf{S}_n^{k,(i)}|}\right]^+;\right.\right.$$
$$\left.\left.\sum_k\left[\frac{2u_n}{\nu_n^{k,(i)}}-\frac{\mathrm{tr}\{\mathbf{S}_n^{k,(i+1)}\}}{2|\mathbf{S}_n^{k,(i+1)}|}-\frac{2u_n}{\nu_n^{k,(i)}}+\frac{\mathrm{tr}\{\mathbf{S}_n^{k,(i)}\}}{2|\mathbf{S}_n^{k,(i)}|}\right]^-\right\}\right\} \tag{32}$$

In (31), we substitute (30); in (32) we use $[a^+-b^+]^+\le[a-b]^+$ (and likewise for $[a^+-b^+]^-$) and

Corollary 1 of [23]. In (32), note that

$$\frac{\mathrm{tr}\{\mathbf{S}_n^{k,(i)}\}}{2|\mathbf{S}_n^{k,(i)}|} = \frac{\mathrm{tr}\{(\mathbf{M}_n^{k,(i-1)})^{-1}\overline{\mathbf{H}}_{n,n}^k\}}{2|(\mathbf{M}_n^{k,(i-1)})^{-1}\overline{\mathbf{H}}_{n,n}^k|}$$
$$= \frac{\mathrm{tr}\{\mathrm{adj}\{\mathbf{M}_n^{k,(i-1)}\}\overline{\mathbf{H}}_{n,n}^k\}}{2|\mathbf{M}_n^{k,(i-1)}||\mathbf{M}_n^{k,(i-1)}|^{-1}|\overline{\mathbf{H}}_{n,n}^k|} = \frac{\mathrm{tr}\{\mathrm{adj}\{\mathbf{M}_n^{k,(i-1)}\}\overline{\mathbf{H}}_{n,n}^k\}}{2|\overline{\mathbf{H}}_{n,n}^k|}.$$

Here we first use the definition of $\mathbf{S}_n^{k,(i)}$ and then the definition of $\overline{\mathbf{H}}_{n,j}^k$ (notice that, since all matrices are square, we can write $|\mathbf{S}_n^{k,(i)}| = |(\overline{\mathbf{T}}_n^k)^{\mathrm{H}}(\mathbf{H}_{n,n}^k)^{\mathrm{H}}(\mathbf{M}_n^{k,(i-1)})^{-1}\mathbf{H}_{n,n}^k\overline{\mathbf{T}}_n^k| = |(\mathbf{M}_n^{k,(i-1)})^{-1}\overline{\mathbf{H}}_{n,n}^k|$; then the adjugate formula for matrix inverses, i.e. $(\mathbf{M}_n^{k,(i-1)})^{-1} = \mathrm{adj}\{\mathbf{M}_n^{k,(i-1)}\}|\mathbf{M}_n^{k,(i-1)}|^{-1}$; and $|\mathbf{A}^{-1}\mathbf{B}| = |\mathbf{A}|^{-1}|\mathbf{B}|$. Consider $\overline{\mathbf{H}}_{n,j}^k = \left[\begin{smallmatrix} d_{n,j}^k & (e_{n,j}^k)^* \\ e_{n,j}^k & d_{n,j}^k \end{smallmatrix}\right]$ for all n, j. Then $\mathrm{adj}\{\mathbf{M}_n^{k,(i)}\}$ can be written as

$$\mathrm{adj}\{\mathbf{M}_n^{k,(i)}\} = \sum_{j\neq n}\begin{bmatrix} p_j^{k,(i)}d_{n,j}^k + 1 & -p_j^{k,(i)}(e_{n,j}^k)^* \\ -p_j^{k,(i)}e_{n,j}^k & p_j^{k,(i)}d_{n,j}^k + 1 \end{bmatrix}$$
$$= \sum_{j\neq n} p_j^{k,(i)} \underbrace{\begin{bmatrix} d_{n,j}^k & -(e_{n,j}^k)^* \\ -e_{n,j}^k & d_{n,j}^k \end{bmatrix}}_{=\mathrm{adj}\{\mathbf{H}_{n,j}^k\}} + \mathbf{I}. \quad (33)$$

Substituting (33) in (32), we obtain

$$\max_n\Big\{\max\Big\{\sum_k\Big[\sum_{j\neq n}\frac{\mathrm{tr}\{\mathrm{adj}\{\mathbf{H}_{n,j}^k\}\overline{\mathbf{H}}_{n,n}^k\}}{2|\overline{\mathbf{H}}_{n,n}^k|}\big(p_j^{k,(i-1)} - p_j^{k,(i)}\big)\Big]^+;$$
$$\sum_k\Big[\sum_{j\neq n}\frac{\mathrm{tr}\{\mathrm{adj}\{\mathbf{H}_{n,j}^k\}\overline{\mathbf{H}}_{n,n}^k\}}{2|\overline{\mathbf{H}}_{n,n}^k|}\big(p_j^{k,(i-1)} - p_j^{k,(i)}\big)\Big]^-\Big\}\Big\}$$
$$\le \max_n\Big\{\max\Big\{\sum_{\substack{k\\ j\neq n}}\frac{\mathrm{tr}\{\mathrm{adj}\{\mathbf{H}_{n,j}^k\}\overline{\mathbf{H}}_{n,n}^k\}}{2|\overline{\mathbf{H}}_{n,n}^k|}\big[p_j^{k,(i-1)} - p_j^{k,(i)}\big]^+;$$
$$\sum_{\substack{k\\ j\neq n}}\frac{\mathrm{tr}\{\mathrm{adj}\{\mathbf{H}_{n,j}^k\}\overline{\mathbf{H}}_{n,n}^k\}}{2|\overline{\mathbf{H}}_{n,n}^k|}\big[p_j^{k,(i-1)} - p_j^{k,(i)}\big]^-\Big\}\Big\} \quad (34)$$
$$\le (N-1)\max_{\substack{j,n,k\\ j\neq n}}\Big\{\frac{\mathrm{tr}\{\mathrm{adj}\{\mathbf{H}_{n,j}^k\}\overline{\mathbf{H}}_{n,n}^k\}}{2|\overline{\mathbf{H}}_{n,n}^k|}\Big\}\max_n\Big\{\max\Big\{\sum_k\big[p_n^{k,(i-1)} - p_n^{k,(i)}\big]^+;\sum_k\big[p_n^{k,(i-1)} - p_n^{k,(i)}\big]^-\Big\}\Big\} \quad (35)$$

In (34), we use $\sum_k[\sum_j a_j^k b_j^k]^+ \le \sum_k\sum_j a_j^k[b_j^k]^+$ for non-negative a_j^k (and likewise for $\sum_k[\sum_j a_j^k b_j^k]^-$).

In (35) we use

$$\sum_k \sum_j^{N-1} a_j^k b_j^k \leq \sum_j \max_k\{a_j^k\} \sum_k b_j^k$$

$$\leq (N-1) \max_{j,k}\{a_j^k\} \max_j\{\sum_k b_j^k\}$$

for non-negative a_j^k, b_j^k. In (35), the condition $(N-1)\max_{\substack{j,n,k \\ j\neq n}} \left\{ \frac{\mathrm{tr}\left\{\mathrm{adj}\left\{\mathbf{H}_{n,j}^k\right\}\overline{\mathbf{H}}_{n,n}^k\right\}}{2\left|\overline{\mathbf{H}}_{n,n}^k\right|} \right\} < 1$ completes the proof.

ACKNOWLEDGMENT

We thank Prof. Raf Vandebril, from the Department of Computer Science, KU Leuven, Belgium, for the helpful discussions.

REFERENCES

[1] P. Ödling, T. Magesacher, S. Höst, P. O. Börjesson, M. Berg, and E. Areizaga, "Fourth generation broadband concept," *IEEE Commun. Mag.*, vol. 47, no. 1, pp. 63–69, 2009.

[2] 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, 2006.

[3] R. B. Moraes, B. Dortschy, A. Klautau, and J. R. i Riu, "Semiblind spectrum balancing for DSL," *IEEE Trans. Signal Process.*, vol. 58, no. 7, pp. 3717–3727, 2010.

[4] J. Papandriopoulos and J. S. Evans, "SCALE: a low-complexity distributed protocol for spectrum balancing in multiuser DSL networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3711–3724, 2009.

[5] 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, 2008.

[6] W. Yu, "Multiuser water-filling in the presence of crosstalk," in *Inf. Theory and Appl. Workshop*, San Diego, USA, 2007.

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

[8] R. Cendrillon, G. Ginis, E. Van den Bogaert, and M. Moonen, "A near-optimal linear crosstalk canceler for upstream VDSL," *IEEE Trans. Signal Process.*, vol. 54, no. 8, pp. 3136–3146, 2006.

[9] ——, "A near-optimal linear crosstalk precoder for downstream VDSL," *IEEE Trans. Commun.*, vol. 55, no. 5, pp. 860–863, 2007.

[10] B. Lee, J. M. Cioffi, S. Jagannathan, and M. Mohseni, "Gigabit DSL," *IEEE Trans. Commun.*, vol. 55, no. 9, pp. 1689–1692, 2007.

[11] S. Jagannathan, V. Pourahmad, J. M. Cioffi, M. Ouzzif, and R. Tarafi, "Common-mode data transmission using the binder sheath in digital subscriber lines," *IEEE Trans. Commun.*, vol. 57, no. 3, pp. 831–840, 2009.

[12] J. M. Cioffi, S. Jagannathan, M. Mohseni, and G. Ginis, "CuPON: The copper alternative to PON 100 Gb/s DSL networks," *IEEE Commun. Mag.*, vol. 45, no. 6, pp. 132–139, 2007.

[13] A. Chowdhery and J. M. Cioffi, "Dynamic spectrum management for upstream mixtures of vectored & non-vectored DSL," in *IEEE Global Telecomm. Conf.*, Miami, USA, 2010.

[14] M. Guenach, J. Louveaux, L. Vandendorpe, P. Whiting, J. Maes, and M. Peeters, "On signal-to-noise ratio-assisted crosstalk channel estimation in downstream DSL systems," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 2327–2338, 2010.

[15] J. Maes, C. Nuzman, A. Van Wijngaarden, and D. Van Bruyssel, "Pilot-based crosstalk channel estimation for vector-enabled VDSL systems," in *Annual Conf. Inf. Sciences and Syst.*, Princeton, USA, 2010.

[16] T. Magesacher, P. Ödling, P. O. Börjesson, and T. Nordström, "Exploiting the common-mode signal in xDSL," in *European Signal Process. Conf.*, Vienna, Austria, 2004.

[17] T. Magesacher, P. Ödling, P. O. Börjesson, W. Henkel, T. Nordström, R. Zukunft, and S. Haar, "On the capacity of the copper cable channel using the common mode," in *IEEE Global Telecomm. Conf.*, Taipei, Taiwan, 2004.

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

[19] F. Negro, S. P. Shenoy, I. Ghauri, and D. T. M. Slock, "On the MIMO interference channel," in *Inf. Theory and Appl. Workshop*, San Diego, USA, 2010.

[20] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, 2011.

[21] J. Shin, J. Moon, , and J. Ahn, "Weighted-sum-rate-maximizing linear transceiver filters for the K-user MIMO interference channel," in *Global Telecomm. Conf.*, Houston, USA, 2011.

[22] P. Tsiaflakis, R. B. Moraes, and M. Moonen, "A low complexity algorithm for joint spectrum and signal coordination in upstream DSL transmission," in *18th IEEE Symp. on Commun. Veh. Technol. Benelux*, Ghent, Belgium, 2011.

[23] R. Cendrillon, J. Huang, M. Chiang, and M. Moonen, "Autonomous spectrum balancing for digital subscriber lines," *IEEE Trans. Signal Process.*, vol. 55, no. 8, pp. 4241–4257, 2007.

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

[25] Z.-Q. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE Trans. Signal Process.*, vol. 2, no. 1, pp. 57–73, 2009.

[26] ——, "Duality gap estimation and polynomial time approximation for optimal spectrum management," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2675 – 2689, 2009.

[27] I. E. Telatar, "Capacity of multi-antenna gaussian channels," *European Trans. Telecommun.*, vol. 10, pp. 585–595, 1999.

[28] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, 2008.

[29] K. Gomadam, V. R. Cadambe, and S. A. Jafar, "Approaching the capacity of wireless networks through distributed interference alignment," in IEEE *Global Telecomm. Conf.*, New Orleans, USA, 2008.

[30] G. Scutari, D. P. Palomar, and S. Barbarossa, "The MIMO iterative waterfilling algorithm," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1917–1935, 2009.

[31] S. J. Kim and G. B. Giannakis, "Optimal resource allocation for MIMO ad-hoc cognitive radio networks," in *Allerton Conf. Comm., Control and Comput.*, Urbana-Champaign, USA, 2008.

[32] W. Yu, T. Kwon, and C. Shin, "Multicell coordination via joint scheduling, beamforming and power spectrum adaptation," in *IEEE INFOCOM*, Shanghai, China, 2011.

[33] L. Venturino, N. Prasad, and X. Wang, "Coordinated linear beamforming in downlink multi-cell wireless networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, pp. 1451–1461, 2010.

[34] H. Bölcskei, D. Gesbert, and A. J. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 225–234, 2002.

[35] O. El Ayach, S. W. Peters, and R. W. Heath, Jr., "The feasibility of interference alignment over measured MIMO-OFDM channels," *IEEE Trans. Veh. Technol.*, vol. 59, no. 9, pp. 4309–4321, 2010.

[36] M. Joham, K. Kusume, M. H. Gzara, W. Utschick, and J. A. Nossek, "Transmit Wiener filter," Munich University of Technology, Tech. Rep., 2002.

[37] W. Yu, G. Ginis, and J. M. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE J. Sel. Areas of Commun.*, vol. 20, no. 5, pp. 1105–1115, 2002.

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

[39] K. B. Petersen and M. S. Pedersen, "The matrix cookbook," Technical University of Denmark, Tech. Rep., 2008.

[40] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computing*, 3rd ed. Cambridge University Press, 2007.

[41] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. The Johns Hopkins University Press, 1996.



Rodrigo B. Moraes (S'08) was born in Belém, Brazil, in 1982. He obtained the Bachelor degree at the Federal University of Pará, Belém, Brazil, in 2005 and the M.Sc. degree at the Pontifical Catholic University, Rio de Janeiro, Brazil, in 2009, both in Electrical Engineering. Since October 2009, he is working towards a Ph.D. degree at the department of Electrical Engineering at the Katholieke Universiteit Leuven (KU Leuven), Leuven, Belgium.

In 2006, he was a visiting researcher at Ericsson's Broadband Technologies laboratories. His research interests are in signal processing for communications.

Mr. Moraes has been awarded the FAPERJ *Nota Dez* Scholarship by state of Rio de Janeiro, Brazil, and the IEEE Travel Grants, both in 2008.



Paschalis Tsiaflakis (S'06-M'09) received the M.Eng. degree in information and communication technology from the Katholieke Hogeschool Limburg (Belgium) in 2001, and the M.S. degree and Ph.D. degree, both in electrical engineering, from the KU Leuven (Belgium) in 2004 and 2009, respectively.

He is currently a FWO Postdoctoral Research Fellow with the Department of Electrical Engineering, KU Leuven. He was a visiting scholar at Princeton University with Prof. Mung Chiang in 2007, a visiting postdoc at the University of California Los Angeles with Prof. Lieven Vandenberghe in 2010, and a postdoctoral fellow in the Center for Operations Research and Econometrics at the Université Catholique de Louvain (Belgium) with Prof. François Glineur in 2011. Since 2004, he has been involved in several research projects with industry through joint collaborations and consultancy. His research interests include signal processing and optimization for wireline and wireless digital communication systems.

Dr. Tsiaflakis received the 'Belgian Young ICT Personality 2010 Award' in 2010, the 'Best Multimedia Master Thesis Prize' in 2001, and was a top-12 finalist for the European ERCIM Cor Baayen Award in 2010. He also received a FWO Aspirant scholarship (2004-2008), a PDMK postdoc grant in 2009, a Francqui Intercommunity Postdoc Grant in 2011, and a FWO postdoc grant (2011-2014). He holds 1 granted patent, has 8 standard and broadband forum contributions, and has published over 20 IEEE journal/conference peer reviewed papers.



Jochen Maes (M'10-SM'11) is a senior researcher at the Access Research Domain of Bell Labs. He researches transceiver and system design for G.fast Digital Subscriber Line technology, delivering 1 Gb/s over the telephony access network. His previous work includes vectoring and phantom mode transmission, which received the Infovision Award for Broadband Innovation in 2010 and the Bell Labs President's Award in 2011. He received his Masters in Physics and Ph.D. in Science from the Katholieke Universiteit Leuven in 2000 and 2004 respectively.



Marc Moonen (M'94, SM'06, F'07) received the electrical engineering degree and the PhD degree in applied sciences from KU Leuven, Belgium, in 1986 and 1990 respectively. Since 2004 he is a Full Professor at the Electrical Engineering Department of KU Leuven, where he is heading a research team working in the area of numerical algorithms and signal processing for digital communications, wireless communications, DSL and audio signal processing.

He received 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 a journal best paper award from the IEEE Transactions on Signal Processing (with Geert Leus) and from Elsevier Signal Processing (with Simon Doclo).

He was chairman of the IEEE Benelux Signal Processing Chapter (1998-2002), and a member of the IEEE Signal Processing Society Technical Committee on Signal Processing for Communications, and is currently President of EURASIP (European Association for Signal Processing).

He has served as Editor-in-Chief for the EURASIP Journal on Applied Signal Processing (2003-2005), and has been a member of the editorial board of IEEE Transactions on Circuits and Systems II, IEEE Signal Processing Magazine, Integration-the VLSI Journal, EURASIP Journal on Wireless Communications and Networking, and Signal Processing. He is currently a member of the editorial board of EURASIP Journal on Applied Signal Processing and Area Editor for Feature Articles in IEEE Signal Processing Magazine.

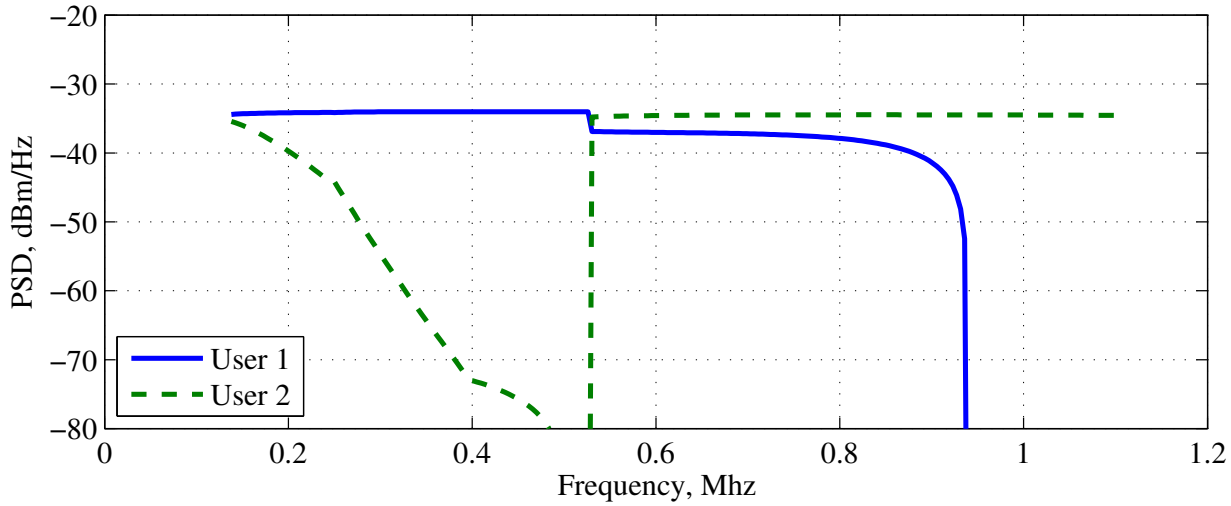

Fig. 5. Power allocation obtained with GDSB. Data rates are approximately $r_1 = 7$ Mb/s and $r_2 = 12.2$ Mb/s.

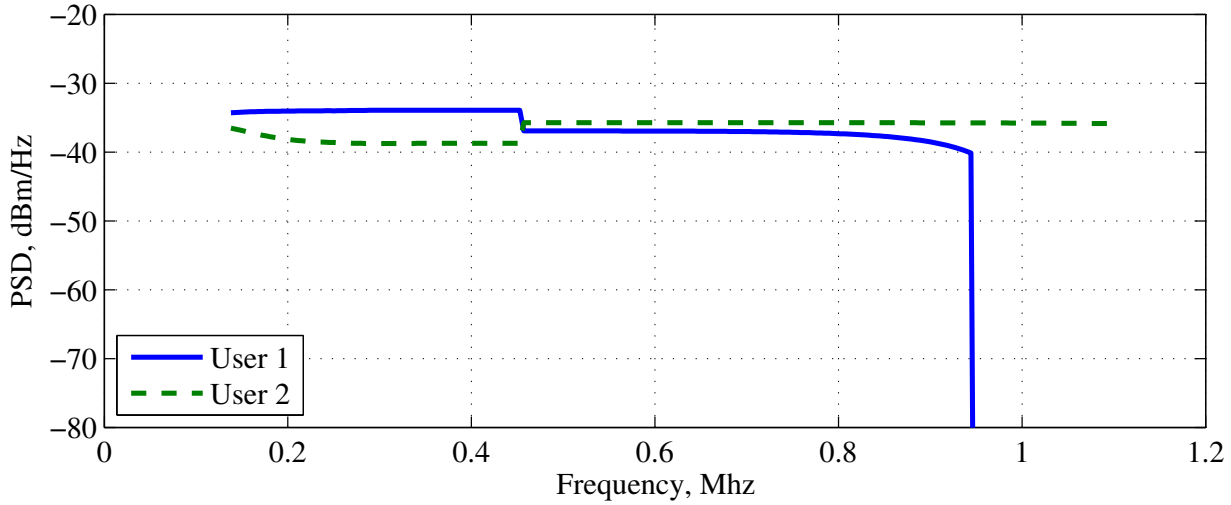

Fig. 6. Power allocation obtained with WMMSE-GDSB. Data rates are approximately $r_1 = 7$ Mb/s and $r_2 = 15$ Mb/s.

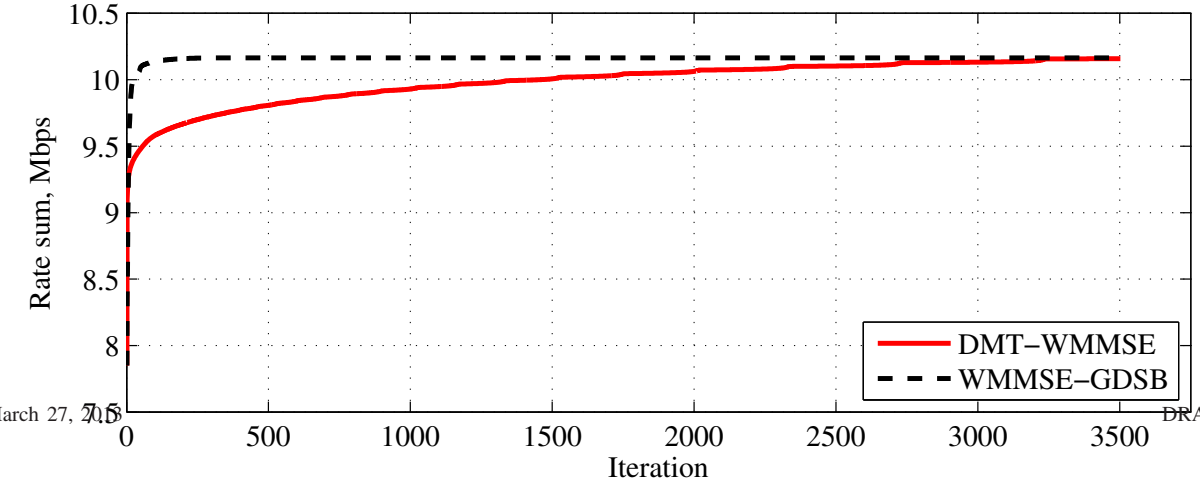

Fig. 7. Convergence speed for the scenario in Fig. 3 for the DMT-WMMSE and the WMMSE-GDSB.

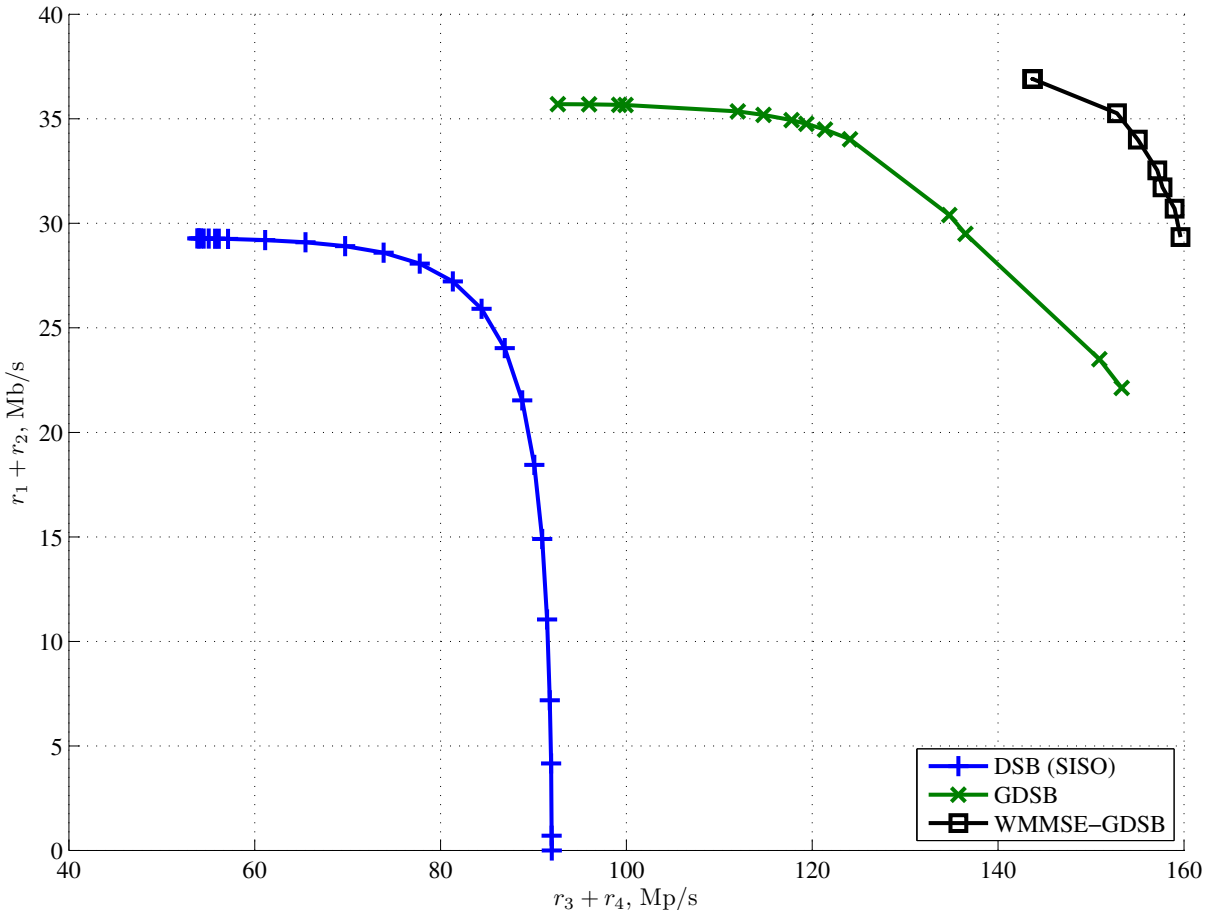

Fig. 8. Rate region for the upstream VDSL scenario for DSB (SISO) and WMMSE-GDSB.