Katholieke Universiteit Leuven

Departement Elektrotechniek

ESAT-SISTA/TR 11-89

Dynamic Spectrum Management in DSL with Asynchronous Crosstalk [1]

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

May 2011

[1]This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/sista/rmoraes/reports/11-89.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 Katholieke Universiteit Leuven, in the frame of K.U.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), 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.

Dynamic Spectrum Management in DSL with Asynchronous Crosstalk

Rodrigo B. Moraes*, Paschalis Tsiaflakis and Marc Moonen

Abstract

In this paper we focus on discrete multitone (DMT) dynamic spectrum management (DSM) in digital subscriber lines (DSL) networks with asynchronous crosstalk. DSM aims to optimally allocate per-user transmit spectra so that the effect of multi-user crosstalk is minimized and the capabilities of the network maximized. Most DSM solutions so far address an idealized situation, one in which all users' DMT blocks in the network are perfectly synchronized and crosstalk is dealt with on a per-tone basis. We focus on the case in which the DMT blocks of the various users are offset amongst each other, which leads to inter-carrier interference (ICI). ICI significantly impacts the performance of the system and complicates the DSM optimization problem considerably, as the per-tone decoupling used for the synchronous case is no longer directly applicable. Our contribution is twofold: First, we derive a new and accurate model for the effect of the ICI; Second, we provide a novel DMT DSM algorithm that outperforms the existing state-of-the-art.

Index Terms

A preliminary version of this paper appeared at the IEEE International Conference on Audio, Speech and Signal Processing (ICASSP), Prague, Czech Republic, in May 2011 [1].

This research work was carried out at the ESAT Laboratory of Katholieke Universiteit Leuven, in the frame of K.U.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), 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/SISTA), Katholieke Universiteit 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).

SPC-TDLS, SPC-MULT, SPC-CRDS.

I. Introduction

Digital Subscriber Line (DSL) is today one of the main technologies for broadband access. For about two decades now, there has been a strong activity in the research community to deal with DSL's main problems and to try to expand its lifetime as much as possible. One such area of research is focused on the optimal allocation of per-user transmit spectra so that the impact of multi-user *crosstalk*, the main source of performance degradation for DSL, is minimized and the capabilities of the network are maximized. This is referred to as dynamic spectrum management (DSM).[1]

Work on DSM has progressed significantly in the past decade, see e.g. [3]–[13]. The optimal spectrum balancing (OSB) of Cendrillon *et al*. [4] provides a provably optimal spectrum allocation algorithm that is efficient when the number of users is small. Subsequent works focused on more practical issues, such as making the implementation distributed across the network and reducing computational complexity so as to enable solutions for large scale DSL scenarios. Today, some near optimal, semi-centralized, trustworthy and low complexity solutions exist, e.g. [10]–[13].

Most of this previous work considers a synchronous discrete multitone (DMT) model, one in which all users have their DMT blocks perfectly synchronized (i.e. all users' DMT blocks are aligned in time). This leads to crosstalk decoupled across tones, i.e. crosstalk that can be dealt with on a per-tone basis. This assumption simplifies the DSM optimization problem significantly. However, the synchronous DMT model may not be very realistic in practice. There are some proposals to overcome the asynchronicity of the DMT blocks by adding a cyclic suffix [14], but it must be said that the conditions for synchronous DMT transmission may not always be easy to attain. Situations where interfering users belong to different service providers or where transmitters are not co-located are specially troublesome. In this paper we therefore focus on the *asynchronous* DMT DSM problem [6], [7], [15], [16].

Although similar at a first glance, the synchronous and asynchronous DMT DSM problems are quite different, the latter being even more challenging. The consequence of the time offset between the DMT blocks from different users is inter-carrier interference (ICI). With ICI, the crosstalk decoupling is broken. Hence, a tone of an interferer user affects not only the corresponding tone of a victim user, but all neighboring tones too. DSM solutions that explore a per-tone decoupling, like the OSB and all the other good synchronous solutions, are no longer directly applicable.

[1]While this work deals only with spectrum coordination, literature also refers to DSM as a signal-level coordination paradigm. See e.g [2].

In this paper, we present two novel results for the asynchronous DMT DSM problem. First, we derive an accurate model for the ICI, one that takes into account all peculiarities of DMT transmission; Second, we provide a novel asynchronous DMT DSM algorithm, one that is based on partially decoupling the problem with the aid of *virtual lines* (VL) [7], [12]. Our approach can be interpreted as partially transforming the asynchronous DMT DSM problem into a synchronous one. For example, in the case when crosstalk is small, we transform one asynchronous victim user into a batch of K_n synchronous virtual victim users—K_n being the number of tones in which the asynchronous victim is active. Our algorithm is shown to outperform existing state-of-the-art algorithms.

This paper is organized as following: Section II presents the problem of interest and a brief summary of the previous work; Section III derives the novel model for the ICI; Section IV introduces the novel asynchronous DMT DSM algorithm; Section V contains experimental results; and finally Section VI presents final remarks.

II. Problem Statement and Previous Work

Consider an N user DMT system with K Δ_f-spaced tones and let $\mathbf{P} = \{p_n^k\} \in \mathbb{R}^{K\times N}$ be a matrix in which p_n^k is the transmit power of user n on tone k. The nth column $\mathbf{p}_n = \begin{bmatrix} p_n^1 & \cdots & p_n^K \end{bmatrix}^{\mathrm{T}}$ represents the power allocation of user n on all tones and the kth row $\mathbf{p}^k = \begin{bmatrix} p_1^k & \cdots & p_N^k \end{bmatrix}$ contains the power allocation of all users on a given tone. Let $\tilde{\sigma}_n^k$ be the background noise power observed by the user n on tone k, $h_{i,n}^k$ be the channel gain between transmitter i and receiver n at tone k and Γ be the SNR gap to capacity. The bit loading for user n on tone k in the asynchronous case is a function of the whole matrix $\mathbf{P}$—and that is in contrast with the synchronous case, in which bit loading is only a function of $\mathbf{p}^k$. The bit loading for user n on tone k is defined as

$$b_n^k = \log\left(1 + \frac{p_n^k}{\sigma_n^k + \mathrm{XT}_n^k}\right), \tag{1}$$

where

$$\mathrm{XT}_n^k = \sum_{i\neq n}^{N}\sum_{j=1}^{K} \mathrm{xt}_{i,n}^{j,k} \tag{2}$$

and

$$\mathrm{xt}_{i,n}^{j,k} = \alpha_{i,n}^{j,k} p_i^j \tag{3}$$

$$\alpha_{i,n}^{j,k} = \frac{\Gamma\gamma_{i,n}^{j,k}|h_{i,n}^j|^2}{|h_{n,n}^k|^2} \tag{4}$$

We use $\log$ to denote base two logarithm. Here $\sigma_n^k = \Gamma\tilde{\sigma}_n^k(|h_{n,n}^k|^2)^{-1}$. In (3) and (4), $\mathrm{xt}_{i,n}^{j,k}$ (lowercase), $\alpha_{i,n}^{j,k}$ and $\gamma_{i,n}^{j,k}$ are respectively the crosstalk, the normalized channel gain and the ICI coefficient specifically

from user i to user n, and from tone j to tone k. Notice that XT_n^k (uppercase) in (2) is the total crosstalk for user n on tone k. For the synchronous case, $\gamma_{i,n}^{j,k} = 1$ for $k = j$ and zero otherwise for all users and tones. The data rate for user n is given by

$$R_n = f_s \sum_k b_n^k,$$

where f_s is the symbol rate. Except for a brief treatment of discrete bit loading in Section IV-F, we consider continuous bit loading throughout this paper. We also do not consider spectral masks.

The DSM problem of interest is that of finding a matrix $\hat{\mathbf{P}}$ that maximizes data rates of all users in the network given a power budget for each user. This problem is called rate adaptive (RA) [17]. In mathematical form, it can be written as the maximization of the weighted rates i.e.

$$\begin{aligned} \hat{\mathbf{P}} &= \arg\max_{\mathbf{P}} \sum_n w_n R_n \\ &\text{subject to} \quad \sum_k p_n^k \leq P_n^{\max} \ \forall n \end{aligned} \tag{5}$$

The weight w_n can be interpreted as a priority given to user n. For convenience, it is assumed that the w_n's sum to 1. For a fixed set of weights, we achieve a point on the border of the rate region (RR) whose tangent line has a slope determined by the weights.

The asynchronous DMT DSM problem is quite challenging. The kind of per-tone decoupling used in OSB has no direct application, since the ICI coefficients re-couple tones and users much more strongly. Previous work includes an analysis of the convergence properties of an iterative waterfilling-like algorithm [16] and three alternative solutions. The first two solutions are the greedy bit adding/subtracting, by Chan and Yu [15], and the asynchronous autonomous spectrum balancing, by Cendrillon *et al.* [7]. Both solutions are local search procedures starting from the PSDs of one of the solutions for the synchronous case. They basically work by iteratively withdrawing some power (or bits) from the current power (or bit loading) allocation and then re-introducing these where best suited. These two solutions are computationally expensive and, because of the local searches having as starting point a solution to the synchronous problem, are generally sub-optimal. The third solution is the modified iterative water-filling (MIW), by Yu [6]. The MIW algorithm has been shown to outperform the other two solutions. This solution is based on solving the KKT stationarity conditions so that PSDs are found for every user and then updating an interference-dependent term that should be taken into account for the next iteration of power allocation. In [6], the author first presents an algorithm for the synchronous case (one that is equivalent to the distributed spectrum balancing [DSB], by Tsiaflakis *et al.* [10]) to then show that it can

be adapted to the asynchronous case.

For single-user systems, the modeling of the ICI and inter-symbol interference due to an insufficient cyclic prefix length is well studied in the literature (e.g. [18]). In the present paper, we focus on an ICI that emerges for another reason, namely the asynchronism between different users sharing the DSL network. This phenomenon was first modeled by Chan and Yu [15] and all subsequent works followed their model. Referring to Fig. 1, consider two non-synchronized users. The delay is η, $0 \leq \eta \leq 1$, indicating a fraction of the DMT block length. According to [15], the ICI coefficients as a function of η are given by

$$\gamma_{i,n}^{j,k} = \begin{cases} \frac{(\eta K)^2 + (K - \eta K)^2}{K^2}, & j = k; \\ \dfrac{2\sin^2\left(\pi(k-j)\eta\right)}{K^2 \sin^2\left(\pi/K(k-j)\right)}, & j = 1, \ldots, K,\ j \neq k. \end{cases} \tag{6}$$

The authors of [15] also consider a worst case, in which the coefficients do not depend on the delay and are given by

$$\gamma_{i,n}^{j,k} = \begin{cases} 1, & j = k; \\ \dfrac{2}{K^2 \sin^2\left(\pi/K(k-j)\right)}, & j = 1, \ldots, K,\ j \neq k. \end{cases} \tag{7}$$

The derivation of (6) and (7) involves a few approximations. For example, the ICI coefficients are not user dependent—thus we could drop the subscripts i and n, but we keep the same notation as (3) for consistency—and the cyclic prefix between consecutive blocks is not considered. Also note that the ICI coefficients are symmetric, i.e $\gamma_{i,n}^{k,(k-j)} = \gamma_{i,n}^{k,(k+j)}$. To the best of our knowledge, this is so far the only attempt to calculate the ICI coefficients.

III. ICI Coefficients

In this section we provide a new derivation of the ICI coefficients and compare the obtained ICI coefficients with those of [15]. This section is divided in two parts. First, we obtain the ICI coefficients as a function of the delay η. Second, we obtain the ICI coefficients averaged over η.

In the following, lower-case boldface letters denote vectors, while upper-case boldface is used for matrices. When we refer to DMT symbols, bracketed subscripts refer to time (not to user) and superscripts to tones. Hence $a_{(i)}^k$ should be read as a quantity in the ith block at the kth tone. The vector $\mathbf{a}_{(i)} = \begin{bmatrix} a_{(i)}^1 & \cdots & a_{(i)}^K \end{bmatrix}^{\mathrm{T}}$ is representative for the ith symbol. The DMT block has length $K + L_{\mathrm{cp}}$, where L_{cp} is the length of the cyclic prefix (CP)—we refer to a block as the symbol plus the CP. Other notation includes $\mathrm{E}[\cdot]$ as expectation, $(\cdot)^{\mathrm{H}}$ as conjugate transpose, $\lfloor \cdot \rfloor$ as rounding down and $\mathrm{diag}\{\mathbf{a}\}$ as a matrix

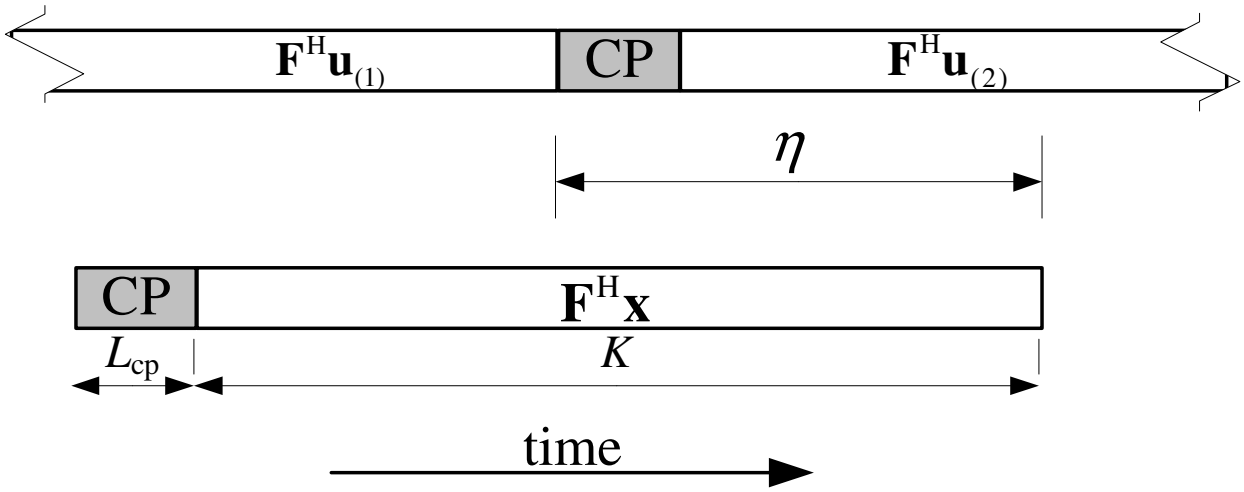

Fig. 1. DMT reception in time for victim user n.

with $\mathbf{a}$ on the main diagonal. Also $\mathbf{0}_{N\times K}$ is the $N \times K$ matrix of zeros and $\mathbf{I}_K$ is the $K \times K$ identity matrix.

A. ICI coefficients as a function of the delay η

Referring to Fig. 1, we consider a victim user n and one interferer i. The victim user has DMT symbol denoted by $\mathbf{x} \in \mathbb{C}^K$, while the interferer is represented by $\mathbf{u} \in \mathbb{C}^K$. Users are not synchronized, and the delay is η, $0 \leq \eta \leq 1$, indicating a fraction of the DMT block length. DMT symbols $\mathbf{u}_{(1)}$ and $\mathbf{u}_{(2)}$ interfere with the reception of the victim user. Mathematically, reception for the victim user is given by

$$\begin{aligned}\mathbf{r} &= \mathbf{F}\widetilde{\mathbf{C}}\mathbf{G}_{n,n}\mathbf{C}\mathbf{F}^{\mathrm{H}}\mathbf{x} + \mathbf{F}\widetilde{\mathbf{C}}\mathbf{G}_{i,n}\mathbf{S}_{(1)}\mathbf{C}\mathbf{F}^{\mathrm{H}}\mathbf{u}_{(1)} \\ &\qquad + \mathbf{F}\widetilde{\mathbf{C}}\mathbf{G}_{i,n}\mathbf{S}_{(2)}\mathbf{C}\mathbf{F}^{\mathrm{H}}\mathbf{u}_{(2)} + \mathbf{z} \\ &= \operatorname{diag}\{\mathbf{h}_{n,n}\}\,\mathbf{x} + \mathbf{F}\widetilde{\mathbf{C}}\mathbf{G}_{i,n}\mathbf{S}_{(1)}\mathbf{C}\mathbf{F}^{\mathrm{H}}\mathbf{u}_{(1)} \\ &\qquad + \mathbf{F}\widetilde{\mathbf{C}}\mathbf{G}_{i,n}\mathbf{S}_{(2)}\mathbf{C}\mathbf{F}^{\mathrm{H}}\mathbf{u}_{(2)} + \mathbf{z}. \end{aligned} \tag{8}$$

Here $\mathbf{F}$ and $\mathbf{F}^{\mathrm{H}} \in \mathbb{C}^{K\times K}$ represent the DFT and IDFT matrices, respectively; $\mathbf{G}_{i,n} \in \mathbb{C}^{(K+L_{\mathrm{cp}})\times(K+L_{\mathrm{cp}})}$ is a Toeplitz matrix with first column $\begin{bmatrix}\mathbf{g}_{i,n}^{\mathrm{T}} & \mathbf{0}_{1\times(K+L_{\mathrm{cp}}-L)}\end{bmatrix}^{\mathrm{T}}$ and first row $\begin{bmatrix}g_{i,n}(1) & \mathbf{0}_{1\times(K+L_{\mathrm{cp}}-1)}\end{bmatrix}$, where $\mathbf{g}_{i,n} \in \mathbb{C}^L$ is the L-tap channel impulse response from transmitter i to receiver n and is considered constant in time; $\mathbf{h}_{i,n} = \begin{bmatrix}h_{i,n}^1 & \cdots & h_{i,n}^K\end{bmatrix}^{\mathrm{T}} \in \mathbb{C}^K$ is the corresponding channel frequency response;

$\mathbf{z} \in \mathbb{C}^K$ is the background Gaussian noise vector; the matrices

$$\widetilde{\mathbf{C}} = \left[\begin{array}{c:c} \mathbf{0}_{K\times L_{\rm cp}} & \mathbf{I}_K \end{array}\right]$$

and

$$\mathbf{C} = \left[\begin{array}{c:c} \mathbf{0}_{L_{\rm cp}\times(K-L_{\rm cp})} & \mathbf{I}_{L_{\rm cp}} \\ \hdashline \multicolumn{2}{c}{\mathbf{I}_K} \end{array}\right],$$

where $\widetilde{\mathbf{C}} \in \mathbb{N}^{K\times(K+L_{\rm cp})}$ and $\mathbf{C} \in \mathbb{N}^{(K+L_{\rm cp})\times K}$, respectively remove and insert the CP. The operation $\widetilde{\mathbf{C}}\mathbf{G}_{n,n}\mathbf{C}$ results in a square circulant matrix, which, if $L_{\rm cp} \geq L$, is then diagonalized by pre- and post-multiplication with the IDFT and DFT matrices. We assume that the CP is longer than both the direct and crosstalk channel impulse responses. The matrices $\mathbf{S}_{(1)}$ and $\mathbf{S}_{(2)}$ capture the effect of the time offset. Define $\omega = \lfloor \eta(K+L_{\rm cp}) \rfloor$ as the number of samples in delay, then these matrices are given by

$$\mathbf{S}_{(1)} = \left[\begin{array}{c:c} \mathbf{0}_{(K+L_{\rm cp}-\omega)\times\omega} & \mathbf{I}_{(K+L_{\rm cp}-\omega)} \\ \hdashline \multicolumn{2}{c}{\mathbf{0}_{\omega\times(K+L_{\rm cp})}} \end{array}\right] \tag{9}$$

and

$$\mathbf{S}_{(2)} = \left[\begin{array}{c:c} \multicolumn{2}{c}{\mathbf{0}_{(K+L_{\rm cp}-\omega)\times(K+L_{\rm cp})}} \\ \hdashline \mathbf{I}_\omega & \mathbf{0}_{\omega\times(K+L_{\rm cp}-\omega)} \end{array}\right]. \tag{10}$$

Here $\mathbf{S}_{(1)},\ \mathbf{S}_{(2)} \in \mathbb{N}^{(K+L_{\rm cp})\times(K+L_{\rm cp})}$. If η is equal to zero or one, then the system is synchronized and $\mathbf{S}_{(1)} = \mathbf{I}_{(K+L_{\rm cp})}$ and $\mathbf{S}_{(2)} = \mathbf{0}_{(K+L_{\rm cp})\times(K+L_{\rm cp})}$ or vice-versa. For $0 < \eta < 1$, the operation $\widetilde{\mathbf{C}}\mathbf{G}_{i,n}\mathbf{S}_{(1)}\mathbf{C}$

(and $\widetilde{\mathbf{C}}\mathbf{G}_{i,n}\mathbf{S}_{(2)}\mathbf{C}$) fails to produce a circulant matrix, and therein lies the effect of the asynchronicity.

Observe that we can write one element of $\mathbf{r}$ in (8) as

$$r^k = h_{n,n}^k x^k + \sum_j \mathbf{A}[k,j]u_{(1)}^j + \sum_j \mathbf{B}[k,j]u_{(2)}^j + z^k, \tag{11}$$

where we define

$$\mathbf{A} = \mathbf{F}\widetilde{\mathbf{C}}\mathbf{G}_{i,n}\mathbf{S}_{(1)}\mathbf{C}\mathbf{F}^{\mathrm{H}}, \tag{12}$$

$$\mathbf{B} = \mathbf{F}\widetilde{\mathbf{C}}\mathbf{G}_{i,n}\mathbf{S}_{(2)}\mathbf{C}\mathbf{F}^{\mathrm{H}}. \tag{13}$$

In (11), the $[k,j]$ elements of $\mathbf{A}$ and $\mathbf{B}$ account for the ICI effect when $j \neq k$.

With (12) and (13) in hands and taking into account that the PSD of the crosstalk symbols is $\mathrm{E}\big[\mathbf{u}_{(1)}\mathbf{u}_{(1)}^{\mathrm{H}}\big] = \mathrm{E}\big[\mathbf{u}_{(2)}\mathbf{u}_{(2)}^{\mathrm{H}}\big] = \mathrm{diag}\left\{\mathbf{p}_i\right\}$, we can write

$$\gamma_{i,v}^{j,k}|h_{i,n}^j|^2 p_i^j = \left(|\mathbf{A}[k,j]|^2 + |\mathbf{B}[k,j]|^2\right) p_i^j. \tag{14}$$

Eq. (14) is easily calculable and it offers an accurate model for the ICI as a function of $\mathbf{g}_{i,v}$ and η.

However, for comparing the ICI coefficients to those of [15], we want the ICI PSD to be captured by a multiplication of the type $\mathbf{M}_{i,n}\mathrm{diag}\big\{|\mathbf{h}_{i,n}|^2\big\}\mathrm{diag}\left\{\mathbf{p}_i\right\}$, where $\mathbf{M}_{i,n} \in \mathbb{R}^{K\times K}$ is the ICI coefficients matrix and $|\mathbf{h}_{i,n}|^2 = \begin{bmatrix}|h_{i,n}^1|^2 & \cdots & |h_{i,n}^K|^2\end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^K$. If we follow the notation of [15], each row of $\mathbf{M}_{i,n}$ would contain the ICI coefficients for one victim tone, i.e. $\mathbf{M}_{i,n} = \begin{bmatrix}\boldsymbol{\gamma}_{i,n}^1 & \boldsymbol{\gamma}_{i,n}^2 & \cdots & \boldsymbol{\gamma}_{i,n}^K\end{bmatrix}^{\mathrm{T}}$, where $\boldsymbol{\gamma}_{i,n}^k = \begin{bmatrix}\gamma_{i,n}^{1,k} & \cdots & \gamma_{i,n}^{K,k}\end{bmatrix}^{\mathrm{T}}$. Calculating the PSD of the interference term in (8), we obtain

$$\mathbf{M}_{i,n}\mathrm{diag}\big\{|\mathbf{h}_{i,n}|^2\big\}\mathrm{diag}\left\{\mathbf{p}_i\right\} = \left(|\mathbf{F}\widetilde{\mathbf{C}}\mathbf{G}_{i,n}\mathbf{S}_{(1)}\mathbf{C}\mathbf{F}^{\mathrm{H}}|^2 + |\mathbf{F}\widetilde{\mathbf{C}}\mathbf{G}_{i,n}\mathbf{S}_{(2)}\mathbf{C}\mathbf{F}^{\mathrm{H}}|^2\right)\mathrm{diag}\left\{\mathbf{p}_i\right\},$$

and hence

$$\mathbf{M}_{i,n} = \left(|\mathbf{A}|^2 + |\mathbf{B}|^2\right)\mathrm{diag}\big\{|\mathbf{h}_{i,n}|^2\big\}^{-1}, \tag{15}$$

where $\mathbf{A}$ and $\mathbf{B}$ are defined in (12) and (13) and where the $[k,j]$th element of $|\mathbf{A}|$ is $|\mathbf{A}[k,j]|$.

With $\mathbf{M}_{i,n}$ calculated as in (15) we can calculate $\alpha_{i,n}^{j,k}$ with (4) and crosstalk with (3) and (2). Notice in both (14) and (15) that we need the crosstalk channel impulse response, $\mathbf{g}_{i,n}$, to compute (14) or (15).[2] As a consequence, the *ICI coefficients are channel dependent*, i.e. different crosstalk channels have different ICI coefficients. They are also frequency dependent: The columns of $\mathbf{M}_{i,n}$ are similar,

[2]Notice that if the channel frequency response is for some tone close to being zero, i.e. for some k, $h_{i,v}^k \approx 0$, calculating (15) might be inaccurate. However, even in this case (14) offers an accurate model for the ICI.

but they are not delayed replicas of one another, e.g. $\gamma_{i,n}^{j,k}$ is usually slightly different than $\gamma_{i,n}^{j+1,k+1}$. It can be shown that the only exception to these two facts is the case of frequency flat channels, i.e. when $\mathbf{g}_{i,n} = \begin{bmatrix} \nu & \mathbf{0}_{K+L_{\text{cp}}-1\times 1} \end{bmatrix}^{\text{T}}$ for a given complex number ν. Notice that in this case $\mathbf{G}_{i,n} = \nu \mathbf{I}_K$. As a consequence, the ICI coefficients are the same for every complex ν. For the frequency flat case, they are also not frequency dependent, i.e. $\gamma_{i,n}^{j,k} = \gamma_{i,n}^{j+1,k+1}$.

In Fig. 2, we plot the ICI coefficients for tone 112 of the 224 tones of an ADSL downstream system with AWG 24 cable for a delay of $\eta = 0.5$. The crosstalk channel for this example is 1 km long and was calculated according to [20]. We use a CP of 32 samples [21]. The plot shows ICI coefficients calculated with (6) and (7), following the model of [15]; and (15) in this paper. Observe that the coefficients of (6) for $\eta = 0.5$ are usually optimistic and the coefficients of (7) for the worst case are usually pessimistic. For instance, for the coefficients of (6) for $\eta = 0.5$ every second tone has coefficient equal to zero. For the worst case, in certain tones the difference between the coefficients in (7) and (15) is more than 25 dB. Also note that, although there are many similarities, the ICI coefficients calculated with (15) are not perfectly symmetric.

In Fig. 3, we illustrate the change in the coefficients when we vary η for tone 112 of the 224 tones for the same ADSL system. For this plot, we assume a frequency flat crosstalk channel. We also only show the ICI coefficients of the 12 closest tones. As mentioned, the coefficients are now symmetric and not frequency dependent. In the figure, we can better see how the ICI coefficients spread power in frequency as η increases. The case with $\eta = 0.5$ is where the coefficients are the most spread. For this case, the direct coefficients are approximately -3 dB and the neighboring coefficients are about -7.1 dB. For $\eta = 1/4$, these values are, respectively, -2.3 and -9.1 dB. For $\eta = 1/8$, we obtain -1.2 and -14.1 dB. In this same figure, we again show the worst case model of (7).

B. ICI coefficients as the expected value of a function of η

On the previous section η was considered a fixed variable. In this section, we consider it to be a random variable, and we calculate the crosstalk as the expected value of a function of η. Let $\mathbf{M}_{i,v}(\eta)$ be a function of the random variable η. It is defined similarly to (15), i.e.

$$\mathbf{M}_{i,v}(\eta) = \left(|\mathbf{A}|^2 + |\mathbf{B}^2|\right) \text{diag}\{|\mathbf{h}_{i,n}|^2\}^{-1}.$$

We remind that the dependence on the delay η is through the definition of (9) and (10). Also, let $f_\eta(H)$ be a given probability distribution function. The expected value of a function of a random variable is given by the inner product of the function and the probability density function of the random variable

(see [22], Eq. 5-55), i.e.

$$\mathrm{E}\left[\mathbf{M}_{i,v}(\eta)\right] = \int_{-\infty}^{+\infty} \mathbf{M}_{i,v}(H) f_{\eta}(H) dH. \tag{16}$$

We can rewrite (16) in a more convenient form by noticing that the matrices $\mathbf{S}_{(1)}$ and $\mathbf{S}_{(2)}$ in (9) and (10) depend on $\left\lfloor \eta(K + L_{\mathrm{cp}}) \right\rfloor$. Hence, we define a discrete random variable $\omega = \left\lfloor \eta(K + L_{\mathrm{cp}}) \right\rfloor$. We consider that η is uniformly distributed between 0 and 1, which leads us to conclude that ω is also uniformly distributed. Mathematically, we have $\Pr(\omega = \Omega) = {}^{1}/_{K+L_{\mathrm{cp}}},\ \Omega = \{0, 1, \ldots, K + L_{\mathrm{cp}} - 1\}$. In this way, we can rewrite (16) as a simple average, i.e.

$$\widetilde{\mathbf{M}}_{i,v} \triangleq \mathrm{E}\left[\mathbf{M}_{i,v}(\omega)\right] = \sum_{\Omega=0}^{K+L_{\mathrm{cp}}-1} \mathbf{M}_{i,v}(\Omega) \frac{1}{K + L_{\mathrm{cp}}}. \tag{17}$$

With $\widetilde{\mathbf{M}}_{i,v}$ in hands, we can calculate crosstalk with (4), (3) and (2).

In Fig. 2, we plot the ICI coefficients for tone 112 of the same 1 km crosstalk channel mentioned on Section III-A.

Eq. (17) is useful because it is independent of the specific delay between two users. In practice calculating the ICI coefficients with (17) may be more interesting, since it is very likely that the delay between the transmission of two users changes over time and is not known accurately. Here, we take an approach of assuming we know nothing about the delay, and thus we assume the probability distribution of η to be uniform. Other options are possible.

Because the delay is a source of uncertainty, using the ICI coefficients in (17) adds some robustness to the subsequent PSD design. Recently, other such sources of uncertainty in the parameters of the problem were considered. For example, references [23]–[25] deal with the impact of errors in the direct and crosstalk channel estimation. Uncertainty in the delay η should be considered alongside with uncertainty in channel estimation for claims of robust solutions. In this paper, we consider all channel transfer functions, both direct and crosstalk, to be known perfectly.

IV. Novel Algorithm

This section contains the derivation of our proposed algorithm for the solution of the asynchronous DMT DSM problem. The proposed algorithm is coined *multiple virtual lines-DSB*, and we motivate this choice shortly. In Section IV-A, we explain the algorithm for the two user case. At this point, we want to emphasize some important design aspects and to talk about a more complicated scenario would just get in the way. In Section IV-B, we convey the intuition of partially decoupling the problem by transforming the asynchronous victim in a batch of synchronous victims. In Section IV-C, we extend the solution to the N user case (which is straightforward), and in Section IV-D we talk about algorithm design.

We need some definitions before proceeding. We refer to the set of user as $\mathcal{N} = \{1, \ldots, N\}$ and the set of tones as $\mathcal{K} = \{1, \ldots, K\}$. The set $\mathcal{K}_n$ is defined as the set of active tones of user n, i.e. $\mathcal{K}_n = \{k \mid p_n^k \neq 0\}$. Of critical importance to the following sections is the crosstalk damage ratio (CDR). The CDR first appeared in [26] and is defined by

$$\mathrm{CDR}_n^k = 1 - \frac{b_n^k}{b_n^k(\sigma)}. \tag{18}$$

Here b_n^k is given by (1) and $b_n^k(\sigma)$ is similar, but it only takes into account the background noise—no crosstalk is considered, i.e. $b_n^k(\sigma) = \log(1 + p_n^k(\sigma_n^k)^{-1})$. The CDR is a measure of the impact of crosstalk damage on bit loading. It ranges in a continuum from zero to 1: The closer it is to zero, the smaller the damage inflicted by crosstalk; and the closer it is to 1, the larger the damage. We also need to introduce one slight variant of (18),

$$\mathrm{CDR}_{i,n}^{j,k} = 1 - \frac{b_{i,n}^{j,k}}{b_n^k(\sigma)} \tag{19}$$

Here, $b_n^k(\sigma)$ is the same as before and $b_{i,n}^{j,k}$ is the bit loading when considering the specific damage from interferer i on tone j to the victim n on tone k. To write $b_{i,n}^{j,k}$ in the same way as (1), we only need to substitute XT_n^k by $\mathrm{xt}_{i,n}^{j,k}$, i.e $b_{i,n}^{j,k} = \log(1 + p_n^k(\sigma_n^k + \mathrm{xt}_{i,n}^{j,k})^{-1})$.

A. Two user case

In this section, for the sake of appropriately emphasizing the important concepts of our proposal, we focus on a two user scenario.

For our proposal, we apply an iterative approach, like the one suggested in [8]–[11]. Instead of solving (5) directly, we optimize the transmit spectrum of one user when the other user has its transmit spectrum fixed. For the two user case, we compute $\mathbf{p}_1$ while keeping $\mathbf{p}_2$ fixed, then $\mathbf{p}_2$ while keeping $\mathbf{p}_1$ fixed and so on. The problem to be solved by, say, user 2 is

$$\begin{aligned} \hat{\mathbf{p}}_2 = \arg\max_{\mathbf{p}_2} & \sum_k w_2 b_2^k + \sum_k w_1 b_1^k \\ \text{subject to} \quad & \sum_k p_2^k \leq P_2^{\max} \end{aligned} \tag{20}$$

Consider the set

$$\mathcal{P} = \{\mathbf{p}_2 | \sum_k p_2^k \leq P_2^{\max}\}.$$

For notational simplicity, in the following discussion we avoid the explicit mention of the constraints in (20). Instead, we write that the optimization problem is only valid on the set $\mathcal{P}$. For the same reason,

we also omit the weights. We recover both the weights and the explicit mention of the constraints when suitable.

Observe in (20) that, because of the ICI effect, the problem is highly coupled in frequency. Hence, we cannot use the dual decomposition approach of [4] to decouple the problem and solve for every tone separately.

Using (18), we can write (20) as

$$\hat{\mathbf{p}}_2 = \arg\max_{\mathbf{p_2}\in\mathcal{P}} \sum_k b_2^k - \sum_{k\in\mathcal{K}_1} \mathrm{CDR}_1^k b_1^k(\sigma) + \sum_{k\in\mathcal{K}_1} b_1^k(\sigma). \tag{21}$$

Here we use the set $\mathcal{K}_1$ for summations of quantities for user 1. Since the last term of (21) does not depend on $\mathbf{p}_2$, it can be removed.

Notice that

$$\mathrm{CDR}_{2,1}^{j,k} = \frac{1}{b_1^k(\sigma)} \left[\log\left(\frac{\sigma_1^k + \mathrm{xt}_{2,1}^{j,k}}{\sigma_1^k} \right) - \log\left(\frac{p_1^k + \sigma_1^k + \mathrm{xt}_{2,1}^{j,k}}{p_1^k + \sigma_1^k} \right) \right]. \tag{22}$$

Now consider the sum in j of all $\mathrm{CDR}_1^{j,k}$. We can write this sum as

$$\begin{aligned}
&\sum_j \mathrm{CDR}_{2,1}^{j,k} = \mathrm{CDR}_{2,1}^{1,k} + \cdots + \mathrm{CDR}_{2,1}^{K,k} \\
&= \frac{1}{b_1^k(\sigma)} \left[\log\left(\frac{\sigma_1^k + \mathrm{xt}_{2,1}^{1,k}}{\sigma_1^k} \right) + \cdots + \log\left(\frac{\sigma_1^k + \mathrm{xt}_{2,1}^{K,k}}{\sigma_1^k} \right) \right. \\
&\left. - \log\left(\frac{p_1^k + \sigma_1^k + \mathrm{xt}_{2,1}^{1,k}}{p_1^k + \sigma_1^k} \right) + \cdots - \log\left(\frac{p_1^k + \sigma_1^k + \mathrm{xt}_{2,1}^{K,k}}{p_1^k + \sigma_1^k} \right) \right] \\
&= \frac{1}{b_1^k(\sigma)} \left[\log\left(\frac{\prod_j (\mathrm{xt}_{2,1}^{j,k} + \sigma_1^k)}{(\sigma_1^k)^K} \right) \right. \\
&\qquad \left. - \log\left(\frac{\prod_j (p_1^k + \sigma_1^k + \mathrm{xt}_{2,1}^{j,k})}{(p_1^k + \sigma_1^k)^K} \right) \right] \\
&= \frac{1}{b_1^k(\sigma)} \left[\log\left(\frac{(\sigma_1^k)^K + (\sigma_1^k)^{K-1} \sum_j \mathrm{xt}_{2,1}^{j,k}}{(\sigma_1^k)^K} + \epsilon_1^k \right) \right. \\
&\qquad \left. - \log\left(\frac{(p_1^k + \sigma_1^k)^K + (p_1^k + \sigma_1^k)^{K-1} \sum_j \mathrm{xt}_{2,1}^{j,k}}{(p_1^k + \sigma_1^k)^K} \right) + \tau_n^k \right] \\
&= \frac{1}{b_1^k(\sigma)} \left[\log\left(\frac{\sigma_1^k + \mathrm{XT}_1^k}{\sigma_1^k} + \epsilon_1^k \right) \right. \\
&\qquad \left. - \log\left(\frac{p_1^k + \sigma_1^k + \mathrm{XT}_1^k}{p_1^k + \sigma_1^k} + \tau_1^k \right) \right], \ k \in \mathcal{K}_1
\end{aligned} \tag{23}$$

Note that (23) is only valid for the tones on the set $\mathcal{K}_1$. If $k \notin \mathcal{K}_1$, $\mathrm{CDR}_1^{j,k}$ should be zero for all j, because, since no power is allocated in k, no crosstalk damage can be caused. In (23),

$$\epsilon_1^k = (\sigma_1^k)^{-2}\mathrm{XXT}_1^k + (\sigma_1^k)^{-3}\mathrm{XXXT}_1^k + \cdots + (\sigma_1^k)^{-K} \underbrace{\mathrm{X} \cdots \mathrm{X}}_{K \text{times}} \mathrm{T}_1^k \tag{24}$$

$$\tau_1^k = (p_1^k + \sigma_1^k)^{-2}\mathrm{XXT}_1^k + (p_1^k + \sigma_1^k)^{-3}\mathrm{XXXT}_1^k + \cdots + (p_1^k + \sigma_1^k)^{-K} \underbrace{\mathrm{X} \cdots \mathrm{X}}_{K \text{times}} \mathrm{T}_1^k \tag{25}$$

where

$$\mathrm{XXT}_1^k = \sum_{j=1}^{K} \sum_{q=j+1}^{K} \mathrm{xt}_{2,1}^{j,k} \mathrm{xt}_{2,1}^{q,k}, \tag{26}$$

$$\mathrm{XXXT}_1^k = \sum_{j=1}^{K} \sum_{q=j+1}^{K} \sum_{m=q+1}^{K} \mathrm{xt}_{2,1}^{j,k} \mathrm{xt}_{2,1}^{q,k} \mathrm{xt}_{2,1}^{m,k}. \tag{27}$$

Eq. (26) represents the sum of the second order products, without repetition, of the crosstalk originating from user 2 ro user 1 for all tones. Likewise, (27) represents the sum of the third order products. In the definition of ϵ_1^k in (24) and τ_1^k in (25) we have products of orders 2 to K. For the ith order, $i = 2, \ldots, K$, there are $K!/(i!(K-i)!)$ terms in the sum.

Using the relations $\log(a+b) = \log(a) + \log(1 + {}^{b}\!/_{a})$, $-\log(a+b) = \log({}^{1}\!/_{a}) + \log({}^{a}\!/_{a+b})$ and (22), we can write

$$\sum_j \mathrm{CDR}_{2,1}^{j,k} = \mathrm{CDR}_1^k + \frac{1}{b_1^k(\sigma)} \log \underbrace{\left(\frac{\left(\sigma_1^k(\epsilon_1^k+1) + \mathrm{XT}_1^k\right)\left(p_1^K + \sigma_1^k + \mathrm{XT}_1^k\right)}{\left(\sigma_1^k + \mathrm{XT}_1^k\right)\left((p_1^k + \sigma_1^k)(\tau_n^k+1) + \mathrm{XT}_1^k\right)} \right)}_{\triangleq \chi_1^k}, \quad (28)$$

$k \in \mathcal{K}_1$. The interesting fact that (28) conveys is that, for tones in the set $\mathcal{K}_1$ (i.e, active tones for user 1), CDR is "summable" up to an error term, the error being smaller when the crosstalk is smaller. This fact was first mentioned in [12], but here it is spelled out more accurately. In (28), we define χ_1^k, which we call the coupling term. Notice that for a situation when crosstalk is small, or, more precisely, when $\epsilon_1^k \ll (\sigma_1^k)^{-1}(\sigma_1^k + \mathrm{XT}_1^k)$ and $\tau_1^k \ll (p_1^k + \sigma_1^k)^{-1}(p_1^k + \sigma_1^k + \mathrm{XT}_1^k)$, then $\chi_1^k \approx 1$ and we can write $\sum_j \mathrm{CDR}_{2,1}^{j,k} \approx \mathrm{CDR}_1^k$.

Using (28) in (21), we obtain

$$\hat{\mathbf{p}}_2 = \arg\max_{\mathbf{p}_2 \in \mathcal{P}} \sum_k b_2^k - \sum_{k \in \mathcal{K}_1} \sum_j \mathrm{CDR}_{2,1}^{j,k} b_1^k(\sigma) + \sum_{k \in \mathcal{K}_1} \log\left(\chi_1^k\right),$$

and, by inverting the order of the summations for the term with $\mathrm{CDR}_1^{j,k}$, we get

$$\hat{\mathbf{p}}_2 = \arg\max_{\mathbf{p}_2 \in \mathcal{P}} \sum_k b_2^k - \sum_j \sum_{k \in \mathcal{K}_1} \mathrm{CDR}_{2,1}^{j,k} b_1^k(\sigma) + \sum_{k \in \mathcal{K}_1} \log\left(\chi_1^k\right). \quad (29)$$

We now use (19) to rearrange (29) as

$$\hat{\mathbf{p}}_2 = \arg\max_{\mathbf{p}_2 \in \mathcal{P}} \sum_j \left(b_2^j + \sum_{k \in \mathcal{K}_1} b_{2,1}^{j,k} \right) + \sum_{k \in \mathcal{K}_1} \log\left(\chi_1^k\right). \quad (30)$$

Notice that here we have separated the highly coupled problem of (21) into two parts, a decoupled (synchronous) and a coupled (asynchronous) one. The term $b_2^j + \sum_{k \in \mathcal{K}_1} b_{2,1}^{j,k}$ in (30) depends only on the

power allocated on tone j of user 2, and thus corresponds to the decoupled part. The term $\sum_{k\in\mathcal{K}_1}\log(\chi_1^k)$ corresponds to the coupled part: Since it involves the elements XT_1^k, ϵ_1^k and τ_1^k, it depends on the whole vector $\mathbf{p}_2$.

Stepping back a moment for a brief digression, we remark that we have departed from (20) and reached (30). Eq. (20) contains the term b_1^k, which depended on $\sum_j \mathrm{xt}_{2,1}^{j,k}$—denote this by $b_1^k(\sum_j \mathrm{xt}_{2,1}^{j,k})$. Eq. (30) contains the term $\sum_j b_{2,1}^{j,k}+\mathrm{error}$, in which each term of the sum depends on $\mathrm{xt}_{2,1}^{j,k}$—denote this by $\sum_j b_{2,1}^{j,k}(\mathrm{xt}_{2,1}^{j,k})+\mathrm{error}$. So basically what we have accomplished is that *the summation term in j was moved outside of the bit loading calculation*. This allows us to partially decouple the problem. What allowed us to do this was the summability property of CDR in (28). The manipulations in between entail in an error, which is represented by the term $\sum_{k\in\mathcal{K}_1}\log(\chi_1^k)$ in (30).

Notice that, if the conditions $\epsilon_1^k \ll (\sigma_1^k)^{-1}(\sigma_1^k+\mathrm{XT}_1^k)$ and $\tau_1^k \ll (p_1^k+\sigma_1^k)^{-1}(p_1^k+\sigma_1^k+\mathrm{XT}_1^k)$ are respected for every active tone of user 1, then the term $\sum_{k\in\mathcal{K}_1}\log(\chi_1^k)$ in the optimization in (30) is very small. For such a case, it can be said that (30) almost fully decouples the problem. For all other cases, i.e when the term with $\sum_{k\in\mathcal{K}_1}\log(\chi_1^k)$ is considerable, then the optimization has considerable coupling between power allocated on different tones of user 2.

However, for the typical DSL scenario, in at least a couple of tones the decoupling is close to complete. The coupling term χ_1^k as defined in (28) is larger as the crosstalk to tone k is larger. In DSL, crosstalk is highly frequency selective, and is most damaging only in a limited range of tones. It is only in these most damaging tones that the coupling term plays an important role. For the remaining tones, in which crosstalk is small or moderate, CDR is approximately summable and crosstalk is almost fully decoupled. For the typical DSL scenario, it is very common for a good percentage of tones to be close to full decoupling. We demonstrate this in Section V.

We again remark that (30) is not fully decoupled, since the last term of it still depends on the vector $\mathbf{p}_2$. However, it is certainly *more decoupled* than (20). There is an error due to the not full decoupling in (30), which is represented by the term $\sum_{k\in\mathcal{K}_1}\log(\chi_1^k)$ in (30).

With (30) in hands, we can now move towards a solution. The next step is to write the Lagrangean dual of (30), to take the derivative of (30) in relation to p_2^q, set it to zero and then solve for p_2^q.

B. Transformation of the asynchronous victim to a batch of synchronous victims

However, before proceeding we turn our attention to one useful interpretation of (30) that uses the concept of virtual lines (VL). VL's (or reference lines) were first introduced in [7] for the synchronous case as a means to de-centralize processing in the network. For the proposal of [7], each user n should

optimize its transmit spectrum with a fictitious VL. If the VL of user n faithfully represents the damage inflicted from n to other users in the network, then a good solution is found. For the synchronous two user scenario, a good VL for user 2 accurately represents the crosstalk damage user 2 causes to user 1. Further work on VL's includes [12], [27].

Define $K_1 = |\mathcal{K}_1|$. Consider a batch of K_1 VL's. VL j, $j \in \mathcal{K}_1$, should represent the damage caused to tone j of user 1. As in (1), define $\bar{\alpha}_{2,j}^k$, $\bar{\sigma}_j^k$ and $\bar{p}_j^k$ respectively as the crosstalk channel, the background noise and the PSD of VL j, tone k—every time we refer to VL's we use the $\bar{(\cdot)}$ symbol. The parameters of the jth VL should match the appropriate parameters of user 1, so let

$$\bar{\alpha}_{2,j}^k \triangleq \alpha_{2,1}^{k,j},\ j,k \in \mathcal{K} \tag{31}$$

$$\bar{\sigma}_j \triangleq \bar{\sigma}_j^k = \sigma_1^j,\ j,k \in \mathcal{K} \tag{32}$$

$$\bar{p}_j \triangleq \bar{p}_j^k = p_1^n,\ j,k \in \mathcal{K}. \tag{33}$$

In (32) and (33), since the background noise and the PSD of the jth VL are frequency flat, we simplify notation by suppressing the superscripts. As a consequence of (31), we can write $\overline{\mathrm{xt}}_{2,j}^k = \mathrm{xt}_{2,1}^{k,j}$, $\overline{\mathrm{XT}}_j = \mathrm{XT}_1^j$, etc. With (31), (32) and (33) in hands, we can re-write (30) as

$$\hat{\mathbf{p}}_2 = \arg\max_{\mathbf{p}\in\mathcal{P}} \sum_j b_2^j + \sum_{j\in\mathcal{K}_1}\sum_k \bar{b}_j^k \tag{34}$$

where

$$\bar{b}_j^k = \log\left[\left(1 + \frac{\bar{p}_j}{\bar{\sigma}_j + \bar{\alpha}_{2,j}^k p_2^k}\right)\bar{\chi}_j\right], \tag{35}$$

is the bit loading for the jth VL. In (35), we have $\bar{\chi}_j \triangleq \chi_1^j$. More precisely,

$$\bar{\chi}_j = \frac{\left((\bar{\sigma}_j+1)\bar{\epsilon}_j^k + \overline{\mathrm{XT}}_j\right)(\overline{p}_j + \bar{\sigma}_j + \overline{\mathrm{XT}}_j)}{(\bar{\sigma}_j + \overline{\mathrm{XT}}_j)\left((\overline{p}_j+\bar{\sigma}_j)(\overline{\tau}_j+1) + \overline{\mathrm{XT}}_j\right)}.$$

Here $\bar{\epsilon}_j$ and $\bar{\tau}_j$ are defined similarly to (24) and (25), respectively. Also,

$$\overline{\mathrm{XXT}}_j = \sum_{m=1}^{K}\sum_{k=j+1}^{K} \overline{\mathrm{xt}}_{2,j}^m \overline{\mathrm{xt}}_{2,j}^k$$

and likewise for higher orders.

Consider the case when $\bar{\chi}_j \approx 1\ \forall j$. Then, as already mentioned, we can consider that the problem is almost fully decoupled. Consider for a moment that this is the case. Then we can interpret (34) as transforming an asynchronous victim (user 1) in a batch of K_1 (almost) synchronous VL's. Because of (31), (32) and (33), the optimization for user 2 results in the same $\hat{\mathbf{p}}_2$ when optimizing with the asynchronous user 1 or with the batch of VL's. Fig. (4) illustrates this transformation for a simple case

with three tones. As the figure shows, we can interpret the situation as the asynchronous crosstalk from user 2 on tone q to user 1 on tone j being transformed to synchronous crosstalk from user 2 on tone q to the jth VL on tone q.

For the general case, when the coupling term may be considerable for some active tones, then the transformation is a bit different. We can interpret the transformation as substituting the asynchronous victim by a batch of VL's that suffer both from synchronous and asynchronous crosstalk. The larger the coupling term $\bar{\chi}_1^k$, the more the VL's are affected by asynchronous crosstalk. The smaller $\bar{\chi}_1^k$, the more the VL's are affected by synchronous crosstalk only.

C. N user case

We now continue with the N user case. The equivalent of (28) is

$$\sum_{i\neq n}\sum_{j}\mathrm{CDR}_{i,n}^{j,k}=\mathrm{CDR}_{n}^{k}+\frac{1}{b_n^k(\sigma)}\log\left(\chi_n^k\right), \tag{36}$$

$k\in\mathcal{K}_n$. In (36), χ_n^k is still given by (28) (just substitute n for '1'), and ϵ_n^k and τ_n^k are still given by (24) and (25). However, in this case XXT_n^k is, similarly to (26), given by all possible second order multiplications (without repetition) of crosstalk originating from different users. Higher order terms are similarly defined. For the ith order, $i=2,\ldots,K$, there are now $^{((N-1)K)!}/_{i!((N-1)K)!}$ terms in the sum.

We skip some steps of the argument due to space limitations. Also, the way to proceed is very much analogous to the two user case. We write the equivalent of (30) as (here we finally re-introduce the weights and the explicit mention of the constraints)

$$\begin{aligned}\hat{\mathbf{p}}_i=\arg\max_{\mathbf{p}_i}&\sum_j w_i b_i^j+\sum_j\sum_{n\neq i}\sum_{k\in\mathcal{K}_n} w_n b_{i,n}^{j,k}\\&+\sum_{n\neq i}\sum_{k\in\mathcal{K}_n} w_n\log\left(\chi_n^k\right).\end{aligned} \tag{37}$$

$$\text{subject to}\quad \sum_k p_i^k\leq P_i^{\max}$$

We now continue by considering the Lagrangean dual of (37).

$$\begin{aligned}L(\mathbf{p}_i,\lambda_i)=&\sum_j\Big(w_i b_i^j+\sum_{n\neq i}\sum_{k\in\mathcal{K}_n} w_n b_{i,n}^{j,k}\Big)\\&+\sum_{n\neq i}\sum_{k\in\mathcal{K}_n} w_n\log\left(\chi_n^k\right)-\lambda_i\big(\sum_k p_i^k-P_i^{\max}\big).\end{aligned} \tag{38}$$

If we were dealing with a synchronous problem, the next step would be to use a dual decomposition [4] and then solve the decomposed Lagrangean for every tone separately. The situation here is however

more difficult. We cannot perfectly decompose (38) because of the term with χ_n^k. As already mentioned, this term depends on the whole vector $\mathbf{p}_i$.

By taking the derivative of the Lagrangean in (38) with respect to p_i^q, setting the result to zero and solving for p_i^q, we find the Karush-Kuhn-Tucker (KKT) stationarity condition for the optimization. The outcome is a power allocation formula that is similar to the waterfilling formula, i.e.

$$p_i^q = \frac{w_i}{\lambda_i + t_i^q} - (\sigma_i^q + \mathrm{XT}_i^q) \tag{39}$$

where

$$t_i^q = \sum_{n \neq i} w_n \sum_{k \in \mathcal{K}_n} \left(\frac{\alpha_{i,n}^{q,k} (\mathrm{SNIR}_{i,n}^{q,k})^2}{p_n^k (\mathrm{SNIR}_{i,n}^{q,k} + 1)} - \frac{\dot{\chi}_{i,n}^{q,k}}{\chi_n^k} \right). \tag{40}$$

Here

$$\mathrm{SNIR}_{i,n}^{q,k} = \frac{p_n^k}{\sigma_n^k + \mathrm{xt}_{i,n}^{q,k}}$$

$$\dot{\chi}_{i,n}^{q,k} = \frac{\partial \chi_n^k}{\partial p_i^q}$$

The variable t_i^q is a per-tone penalty that distorts the waterfilling formula so that damage to other users is considered. This variable should be large for tone q if there is potential for large crosstalk damage to other users by allocating power on tone q—accordingly, [7] calls a similar power allocation strategy frequency selective waterfilling and [6] calls t_i^q a taxation term. Eq. (39) should jointly hold for all users and tones in the network. In the next section, we explain one way to do such power updates iteratively so that the KKT system is solved and a local optimum is reached.

We detail efficient methods to calculate χ_n^k and $\dot{\chi}_{i,n}^{q,k}$ in the Appendix. In particular, in (39) we are interested in the ratio $\dot{\chi}_{i,n}^{q,k}/\chi_n^k$, which is given by

$$\frac{\dot{\chi}_{i,n}^{q,k}}{\chi_n^k} = \frac{p_n^k \alpha_{i,n}^{q,k} (\mathrm{XT}_n^k - \mathrm{xt}_{i,n}^{q,k})(p_n^k + 2\sigma_n^k + \mathrm{XT}_n^k + \mathrm{xt}_{i,n}^{q,k})}{(\sigma_n^k + \mathrm{XT}_n^k)(p_n^k + \sigma_n^k + \mathrm{XT}_n^k)} \times \frac{(p_n^k + \sigma_n^k + \mathrm{xt}_{i,n}^{q,k})}{(\sigma_n^k + \mathrm{xt}_{i,n}^{q,k})}. \tag{41}$$

The equation above is also derived in the Appendix.

Clearly, the interpretation of transforming the asynchronous victims into batches of VL's is also possible for the N user case. For such a case, however, the setting of the VL's is a bit more complicated. Now we must not assume any kind of structure in the network, so we should set the VL's in the most general

form possible. The equivalent of (31), (32) and (33) for the N user case are

$$\bar{\alpha}^{k}_{i,(n,j)} = \alpha^{k,j}_{i,n},\ \ j,k \in \mathcal{K}; i,n \in \mathcal{N}, i \neq n \tag{42}$$

$$\bar{\sigma}_{(n,j)} \triangleq \bar{\sigma}^{k}_{(n,j)} = \sigma^{j}_{n},\ \ j,k \in \mathcal{K}; n \in \mathcal{N} \tag{43}$$

$$\bar{p}_{(n,j)} \triangleq \bar{p}^{k}_{(n,j)} = p^{j}_{n},\ \ j,k \in \mathcal{K}; n \in \mathcal{N}. \tag{44}$$

Here the notation is a bit more complicated than for the two user case. In (42), we are referring to the crosstalk channel from user i to VL j, where j is in the batch of VL's that corresponds to the damage caused to user n. Now every user i has $N-1$ batches of VL's, each one of them corresponding to a victim user. For example, consider a case with 4 users. For, say, user 3 ($i = 3$), there are three batches of VL's, each batch mimicking the damage caused to each victim ($n = 1$, $n = 2$ and $n = 4$). Since each batch has $K_n = |\mathcal{K}_n|$ VL's, the optimization for user 3 on a given tone involves $K_1 + K_2 + K_4$ VL's. If we refer to the channel gain from $i = 3$ to the VL 10 on the batch of VL's corresponding to victim $n = 2$ on tone 100, we write $\bar{\alpha}^{100}_{3,(2,10)}$.

With (42), (43) and (44), we can substitute the appropriate values in (39) and (40) to reach an equivalent power allocation formula.

D. Algorithm

We can now finally specify a full algorithm. Since our power loading formula is very similar to that of DSB [10] and the algorithm can be interpreted as transforming the asynchronous users into multiple VL's, we name our algorithm MVL-DSB (multiple virtual lines DSB). MVL-DSB is detailed in Algorithm 1.

We consider updates to be parallel, i.e. all users change their PSDs at the same time. We follow the same design of the algorithms in [3], [6], [10], i.e. users update their powers according to (39). After all users have updated, the process is repeated taking into account the new powers. We repeat the process until the whole KKT system is solved.

Referring to the algorithm, in line 1, we initialize the matrix $\mathbf{P} = \mathbf{0}_{K \times N}$. It is not known how to optimally initialize the power matrix, but this choice gives good results. Other options are of course possible.

The MVL-DSB comprises two main loops. The outermost loop contains lines 2-13. The loop is basically the DSB algorithm with a different calculation for t^{q}_{i}. Before power allocation itself, we calculate the per-tone penalties t^{q}_{n} for all users (line 4) and the crosstalk plus background noise, called int^{q}_{n}, for all users (line 3). The 'for' loop in lines 5-12 contains the power allocation formula (line 7). Observe that, since t^{q}_{i} and int^{q}_{i} are calculated outside the 'for' loop, the power updates are done in parallel and take

Algorithm 1: MVL-DSB

Input: $P_n^{\max}, w_n,\ n \in \mathcal{N}$

$\sigma_n^k,\ n \in \mathcal{N},\ k \in \mathcal{K}$

$\alpha_{i,n}^{j,k}\ j,k \in \mathcal{K},\ i,n \in \mathcal{N}, i \neq n$

Output: $\mathbf{P}$

1 Initialize $\mathbf{P} = \mathbf{0}_{K \times N}$;

2 **repeat**

3 $\quad \text{int}_n^q = \sigma_n^q + \text{XT}_n^q, n \in \mathcal{N}, q \in \mathcal{K}$;

4 $\quad$ Calculate t_n^q (40), $n \in \mathcal{N}, q \in \mathcal{K}$;

5 $\quad$ **for** $i = 1, \ldots, N$ **do**

6 $\quad\quad$ **repeat**

7 $\quad\quad\quad p_i^q = \frac{w_i}{\lambda_i + t_i^q} - \text{int}_i^q,\ q \in \mathcal{K}$;

8 $\quad\quad\quad$ **if** $\sum_k p_i^k > P_n^{\max}$ **then**

9 $\quad\quad\quad\quad$ increase λ_n;

10 $\quad\quad\quad$ **else**

11 $\quad\quad\quad\quad$ decrease λ_n;

12 $\quad\quad$ **until** $|\sum_k p_i^k - P_n^{\max}| / P_n^{\max} < c_1$ *or* $\lambda_n < c_2$.

13 **until** *PSD convergence.*

into account the PSDs of the previous iteration. The inner loop comprises lines 6 to 12. In this loop we adjust the Lagrange multiplier λ_i. This variable is adjusted until the maximum power is met or it reaches zero. In line 12 the constants c_1 and c_2 are two small non-negative numbers.

Tough we don't have a formal proof of convergence, the algorithm has been experimented in extensive simulations and has always converged.

E. Complexity

The fact that the power loading formula of the MVL-DSB is a type of waterfilling formula means that it can be implemented very efficiently. The most computationally expensive step of MVL-DSB is the calculation of the per-tone penalties t_n^k in (39). For a specific user and tone, the calculation of this variable must cycle through the K tones of $N-1$ users, which implies a number of operations of the order of $(N-1)K$. This should be repeated for all users and tones, which implies a number of operations

of the order $N(N-1)K^2$. Hence, the complexity of MVL-DSB is $O(N^2K^2)$.

The MIW algorithm, which serves as benchmark in the experiment section in this paper, also has complexity of $O(N^2K^2)$. In fact, the two algorithms are very similar in terms of computational cost. The MIW also needs to calculate a similar t_n^k. The only difference is that the MVL-DSB needs to compute the additional term $\dot{\chi}_{i,n}^{q,k}/\chi_n^k$ in (41). This incurs an additional but marginal increase in complexity in comparison to MIW.

F. Discrete bit loading

Practical systems are limited to having discrete bit loadings. MVL-DSB can easily take that into account. Consider the set

$$\widetilde{\mathcal{P}}_i^q = \{p_n^q | b_i^q = 0 \text{ or } b_i^q = 1 \ \ldots \text{ or } b_i^q = b_{\max}\}. \tag{45}$$

This set contains the power levels for the discrete bit loadings for user i on tone q—here $b_{\max}$ is a sufficiently big positive integer. It contains $b_{\max}+1$ elements. Considering $\widetilde{\mathcal{P}}_i^k$ to be the new feasible set for each tone, we can re-write the problem as

$$\hat{\mathbf{p}}_i = \arg \max_{\substack{\mathbf{p}_i \\ p_i^q \in \widetilde{\mathcal{P}}_i^q}} L(\mathbf{p}_i, \lambda_i), \tag{46}$$

where $L(\mathbf{p}_i, \lambda_i)$ is given by (38). Solving (46) tone by tone, we get

$$\hat{p}_i^q = \arg \max_{p_i^q \in \widetilde{\mathcal{P}}_n^q} L^q(p_i^q, \lambda_i) + \sum_{n \neq i} \sum_{k \in \mathcal{K}_n} w_n \log\left(\chi_n^k\right) + \lambda_i P_i^{\max}, \tag{47}$$

$q \in \mathcal{K}$, where

$$L^q(p_i^q, \lambda_i) = w_i b_i^q + \sum n \neq i \sum_{k \in \mathcal{K}_n} w_n b_{i,n}^{q,k} - \lambda_i p_i^q.$$

We can solve (47) by exhaustive search on $\widetilde{\mathcal{P}}_i^q$. By substituting lines 3 and 4 in Algorithm 1 by (45) and line 7 by (47), we have a discrete bit loading version of the algorithm.

V. EXPERIMENTS

In this section we compare the performance of MVL-DSB to that of the MIW.

All simulations in this section consider downstream ADSL. Cables of 0.5 mm (AWG 24) are used and an SNR gap of 12.8 dB is considered. The values for Δ_f and f_s are set to 4.3125 kHz and 4 kHz, respectively. Modems have at their disposal a maximum power of 20.4 dBm. For each line, noise model ANSI A is adopted.

We first consider the near-far downstream ADSL scenario depicted in Fig. 5. This scenario is the main testing ground for any algorithm and it encompasses a good deal of the complexity of the DMT DSM problem. One near user, connected to a central office (CO), shares a binder with a far user connected to a remote terminal (RT). Crosstalk in the RT-CO direction can be prohibitively high, to the point where transmission on the CO may be seriously compromised. A good solution is one in which the RT protects the CO from excessive crosstalk so that both users socially share the system's resources. For this experiment, we set $\eta = 0.5$ and calculate the ICI coefficients with (15). We remark that both the MVL-DSB and the MIW use the same ICI model. Referring to Fig. 5, we set $l_1 = 5$ km and $l_2 = 3$ km. We have simulated three different values for d_2, $d_2 = 3.5$, 4 and 4.5 km.

We compare the performance of MVL-DSB to the MIW algorithm [6]. The MIW departs from (5) and directly writes the corresponding Lagrangean. By taking the derivative of this Lagrangean, [6] also reaches a frequency selective waterfilling formula, like (39). The difference between the two algorithms is that for the MVL-DSB, the power allocation formula (39) departed from the Lagrangean in (38), which was more decoupled across frequency than the Lagrangean originating from (5). This decoupling is a consequence of the CDR summability property derived in Section IV. Hence the power allocation formula of the MVL-DSB is more precise.

Fig 6 shows the three resulting rate regions (RR) for the MIW and for the MVL-DSB algorithms. As expected, the smaller the d_2, the larger the RR. We notice from the plot that the MVL-DSB consistently outperforms the MIW, the difference being larger as d_2 increases. For example, consider the case when $d_2 = 4.5$ km. While providing 1.2 Mbps to the CO, we can achieve 4.1 Mbps on the RT with the MIW and 5.3 Mbps with the MVL-DSB. For the case $d_2 = 4$ km, these values are, respectively, 6.2 and 6.7 Mbps. For $d_2 = 3.5$ km, these values are, respectively, 7.4 and 7.6 Mbps

The MVL-DSB not only compares favorably with the MIW in rate performance, it also uses a smaller amount of power to achieve a given point in the RR. Because of the ICI coefficients, the DMT DSM problem involves some degree of power minimization too. For the RR's in Fig. 6, the RT does not allocate full power for all points. Accordingly, Fig. 7 depicts how much total power was used for the RT to achieve each point in the RR with $d_2 = 4.5$ in Fig 6. Total power is defined as $p_n^{\text{tot}} = \sum_k p_n^k$. The difference is not that significant, but we notice that the MVL-DSB uses less power than the MIW in a broad range of scenarios.

We now simulate a 4 user scenario. Define the vectors $\mathbf{l} = \begin{bmatrix} l_1 & \cdots & l_N \end{bmatrix}^{\mathrm{T}}$, $\mathbf{d} = \begin{bmatrix} d_1 & \cdots & d_N \end{bmatrix}^{\mathrm{T}}$, with $d_1 = 0$, by definition. We thus set the scenario as $\mathbf{l} = \begin{bmatrix} 5 & 4 & 3.5 & 3 \end{bmatrix}^{\mathrm{T}}$ km and $\mathbf{d} = \begin{bmatrix} 0 & 2 & 3 & 4 \end{bmatrix}^{\mathrm{T}}$

km. Here we consider the averaged ICI coefficients of (17) for all user pairs. Again, the same ICI model is used for both the MVL-DSB and the MIW. We plot the RR for the two algorithms in Fig. 8. For all points, we set $R_2 = R_3 = 2$ Mbps. We again notice the better performance of the MVL-DSB throughout the whole RR.

For $R_4 = 5$ Mbps, we plot the resulting power allocation in Fig. 9. User 1, the one with the longer line, has the lower frequencies. As is typical for DSM solutions, all other users protect user 1 on these lower frequencies. A particularly interesting observation for the asynchronous case is the gently increasing transmit powers for users 2, 3 and 4 after user 1 stops transmission. For a synchronous scenario, the transmit power changes abruptly, i.e. as soon as user 1 stops transmission, other users can transmit at full power. The ripples on the transmit spectra in Fig. 9 are due to the ICI coefficients.

Finally, we illustrate how the CDR summability property decouples the problem, which is related to how close χ_n^k is to 1. In Fig. 10, we plot χ_n^k for all users. As argued in Section IV-B, the dimension of this variable also determines how well the transformation from the asynchronous victim to the batch of synchronous VL's is. We notice that, even in this case where there are 4 users, in a good part of the spectrum and for most users χ_n^k is very close to 1. The most inaccurate transformation is for user 1. That is to be expected, since it is this user who suffers the most from crosstalk. We notice that, even in this case where there are 3 considerable interferers to user 1, in a good part of the spectrum χ_1^k is still very close to 1. Hence in this part of the spectrum, which ranges approximately from 0.13 to 0.35 Mhz, the transformation is almost perfect.

VI. Conclusion

Ten years of research on the synchronous DMT DSM problem have brought remarkable results. There is an optimal algorithm and several others with good performance and simpler logistics requirements. Some algorithms, like the iterative waterfilling (IWF) [3] and the OSB have been very influential into areas of research that far surpass the original DSM problem. The research community has successfully played its part, and now commercial products originated from this research are already having an impact in high speed data communications.

Not only are there good algorithms, but a lot is also known about the theory behind the synchronous DMT DSM problem. In our opinion, this accumulated knowledge on the synchronous DMT DSM problem is the perfect platform for deeper investigations of the more complex DMT DSM asynchronous problem. The DMT DSM asynchronous problem has been understandably relegated to a second priority in the recent past, but now the accumulated knowledge on the synchronous problem provides firmer ground to

go forward. That is what we have attempted to do on this paper. Our proposed algorithm, the MVL-DSB, relies heavily on virtual lines and on the frequency selective power allocation formula that were both first proposed to the DMT DSM synchronous problem. Our algorithm is based on the summability property of the crosstalk damage ratio (CDR). This property can be readily applied to any problem in which there is similar coupling in the optimization. The MVL-DSB is seen to consistently outperform the state-of-the-art. In this paper, we have also derived accurate ICI coefficients.

It would be interesting for new investigations of the asynchronous DMT DSM problem to include analysis of the convexity of the rate region and an estimation of the duality gap. These analysis can draw inspiration from recent studies of the synchronous problem [28] [29].

APPENDIX

For calculating (40), we need to have the values of χ_n^k and of $\dot{\chi}_{i,n}^{q,k}$. However, calculating χ_n^k by its definition in (36) is too computationally expensive. That is so because calculating the ith order crosstalk in ϵ_n^k and τ_n^k in (24) and (25) involves adding ${}^{((N-1)K!)}/_{((N-1)K-i)!i!}$ terms. Instead, we can re-write (36) as

$$\sum_{i\neq n}\sum_{j}\mathrm{CDR}_{i,n}^{j,k}b_n^k(\sigma)=\mathrm{CDR}_n^k+\frac{1}{b_n^k(\sigma)}\log\left(\chi_n^k\right).$$

Then,

$$\begin{aligned}\log(\chi_n^k)&=\sum_{i\neq n}\sum_{j}(b_n^k(\sigma)-b_{i,n}^{j,k})-(b_n^k(\sigma)-b_n^k)\\&=\sum_{i\neq n}\sum_{j}\log\left(\frac{p_n^k+\sigma_n^k}{\sigma_n^k}\frac{\sigma_n^k+\mathrm{xt}_{i,n}^{j,k}}{p_n^k+\sigma_n^k+\mathrm{xt}_{i,n}^{j,k}}\right)\\&\qquad-\log\left(\frac{p_n^k+\sigma_n^k}{\sigma_n^k}\frac{\sigma_n^k+\mathrm{XT}_n^k}{p_n^k+\sigma_n^k+\mathrm{XT}_n^k}\right).\end{aligned}$$

Finally,

$$\begin{aligned}\chi_n^k=\left(1+\frac{p_n^k}{\sigma_n^k}\right)^{K(N-1)-1}&\left(1+\frac{p_n^k}{\sigma_n^k+\mathrm{XT}_n^k}\right)\\&\times\prod_{i\neq n}\prod_{j}\frac{\sigma_n^k+\mathrm{xt}_{i,n}^{j,k}}{p_n^k+\sigma_n^k+\mathrm{xt}_{i,n}^{j,k}}.\end{aligned}\tag{48}$$

We obtain $\dot{\chi}_{i,n}^{q,k}$, $i \neq n$, by the taking the derivative of χ_n^k in p_i^q, i.e.

$$\dot{\chi}_{i,n}^{q,k} = \left(1+\frac{p_n^k}{\sigma_n^k}\right)^{K(N-1)-1} \prod_{\substack{u\neq n\\ u\neq i}} \prod_j \left(\frac{\sigma_n^k + \mathrm{xt}_{u,n}^{j,k}}{p_n^k + \sigma_n^k + \mathrm{xt}_{u,n}^{j,k}}\right) \times \prod_{j\neq q}\left(\frac{\sigma_n^k + \mathrm{xt}_{i,n}^{j,k}}{p_n^k + \sigma_n^k + \mathrm{xt}_{i,n}^{j,k}}\right) \times \frac{\partial}{\partial p_i^q}\left[\left(\frac{\sigma_n^k + \mathrm{xt}_{i,n}^{q,k}}{p_n^k + \sigma_n^k + \mathrm{xt}_{i,n}^{q,k}}\right)\left(1+\frac{p_n^k}{\sigma_n^k + \mathrm{XT}_n^k}\right)\right]$$

Solving, we obtain

$$\dot{\chi}_{i,n}^{q,k} = \left(1+\frac{p_n^k}{\sigma_n^k}\right)^{K(N-1)-1} \prod_{\substack{u\neq n\\ u\neq i}} \prod_j \left(\frac{\sigma_n^k + \mathrm{xt}_{u,n}^{j,k}}{p_n^k + \sigma_n^k + \mathrm{xt}_{u,n}^{j,k}}\right) \times \prod_{j\neq q}\left(\frac{\sigma_n^k + \mathrm{xt}_{i,n}^{j,k}}{p_n^k + \sigma_n^k + \mathrm{xt}_{i,n}^{j,k}}\right) \frac{p_n^k \alpha_{i,n}^{q,k}(\mathrm{XT}_n^k - \mathrm{xt}_{i,n}^{q,k})}{(p_n^k + \sigma_n^k + \mathrm{xt}_{i,n}^{q,k})^2} \times \frac{(p_n^k + 2\sigma_n^k + \mathrm{XT}_n^k + \mathrm{xt}_{i,n}^{q,k})}{(\sigma_n^k + \mathrm{XT}_n^k)^2}. \quad (49)$$

Calculating (48) and (49) is easy and computationally cheap. By dividing (49) by (48), we arrive at (41).

REFERENCES

[1] R. B. Moraes, P. Tsiaflakis, and M. Moonen, "Dynamic spectrum management in DSL with asynchrnous crosstalk," in *IEEE Int. Conf. Acoust., Speech, Signal Process.*, Prague, Czech Republic, 2011.

[2] 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.

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

[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, 2006.

[5] W. Yu, R. Lui, and R. Cendrillon, "Dual optimization methods for multiuser orthogonal frequency division multiplex systems," in *IEEE Global Telecomm. Conf.*, Dallas, USA, 2004.

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

[7] 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.

[8] R. Cendrillon and M. Moonen, "Iterative spectrum management for digital subscriber lines," in *IEEE Int. Conf. Commun.*, Seoul, Korea, 2005.

[9] R. Lui and W. Yu, "Low-complexity near-optimal spectrum balancing for digital subscriber lines," in *IEEE Int. Conf. Commun.*, Seoul, Korea, 2005.

[10] 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.

[11] 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.

[12] 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.

[13] Y. Noam and A. Leshem, "Iterative power pricing for distributed spectrum coordination in DSL," *IEEE Trans. Commun.*, vol. 57, no. 4, pp. 948–953, 2009.

[14] F. Sjöberg, M. Isaksson, R. Nilsson, P. Ödling, S. K. Wilson, , and P. O. Börjesson, "Zipper: A duplex method for VDSL based on DMT," *IEEE Trans. Commun.*, vol. 47, no. 8, pp. 1245–1252, 1999.

[15] V. M. K. Chan and W. Yu, "Multiuser spectrum optimization for discrete multitone systems with asynchronous crosstalk," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5425–5435, 2007.

[16] G. Scutari, D. P. Palomar, and S. Barbarossa, "Distributed totally asynchronous iterative waterfilling for wideband interference channel with time/frequency offset," in *IEEE Int. Conf. Acoust., Speech, Signal Process.*, Honolulu, USA, 2007.

[17] T. Starr, M. Sorbara, J. M. Cioffi, and P. Silverman, *DSL Advances*. Prentice Hall, 1999.

[18] W. Henkel, G. Tauböck, P. Ödling, P. O. Börjesson, and N. Petersson, "The cyclic prefix of OFDM/DMT—an analysis," in *Int. Zurich Seminar on Broadband Commun.*, Zurich, Switzerland, 2002.

[19] ANSI 77E7.4/99.438R2, "Noise models for VDSL performance verification)," 1999.

[20] ETSI Std. TS 101 270-1, "Transmission and multiplexing (TM); acess transmission systems on matellic acess cables; very-high bit-rate digital subscriber line transceivers (VDSL); part 1: Functional requirements," 2003.

[21] ITU std. G.992.2, "Asymmetrical digital subscriber line transceivers 2 (ADSL2)," 2002.

[22] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes, Fourth edition*. McGraw-Hill Inc., New York, 2001.

[23] M. Wolkerstorfer, D. Statovci, and T. Nordström, "Robust spectrum management for DMT-based systems," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3238–3250, 2010.

[24] N. Lindqvist, F. Lindqvist, B. Dortschy, E. Pelaes, and A. Klautau, "Impact of crosstalk estimation on the dynamic spectrum management performance," in *IEEE Global Telecommun. Conf.*, New Orleans, USA, 2008.

[25] M. Ding, B. L. Evans, and I. Wong, "Effect of channel estimation error on bit rate performance of time domain equalizers," in *Asilomar Conf. on Signals, Syst., and Comput.*, Pacific Grove, USA, 2004.

[26] R. Moraes, B. Dortschy, A. Klautau, and J. R. i Riu, "Semi-blind power allocation for digital subscriber lines," in *IEEE Int. Conf. Commun.*, Beijing, China, 2008.

[27] C. Leung, S. Huberman, and T. Le-Ngoc, "Autonomous spectrum balancing using multiple reference lines for digital subscriber lines," in *IEEE Global Telecomm. Conf.*, Miami, USA, 2010.

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

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

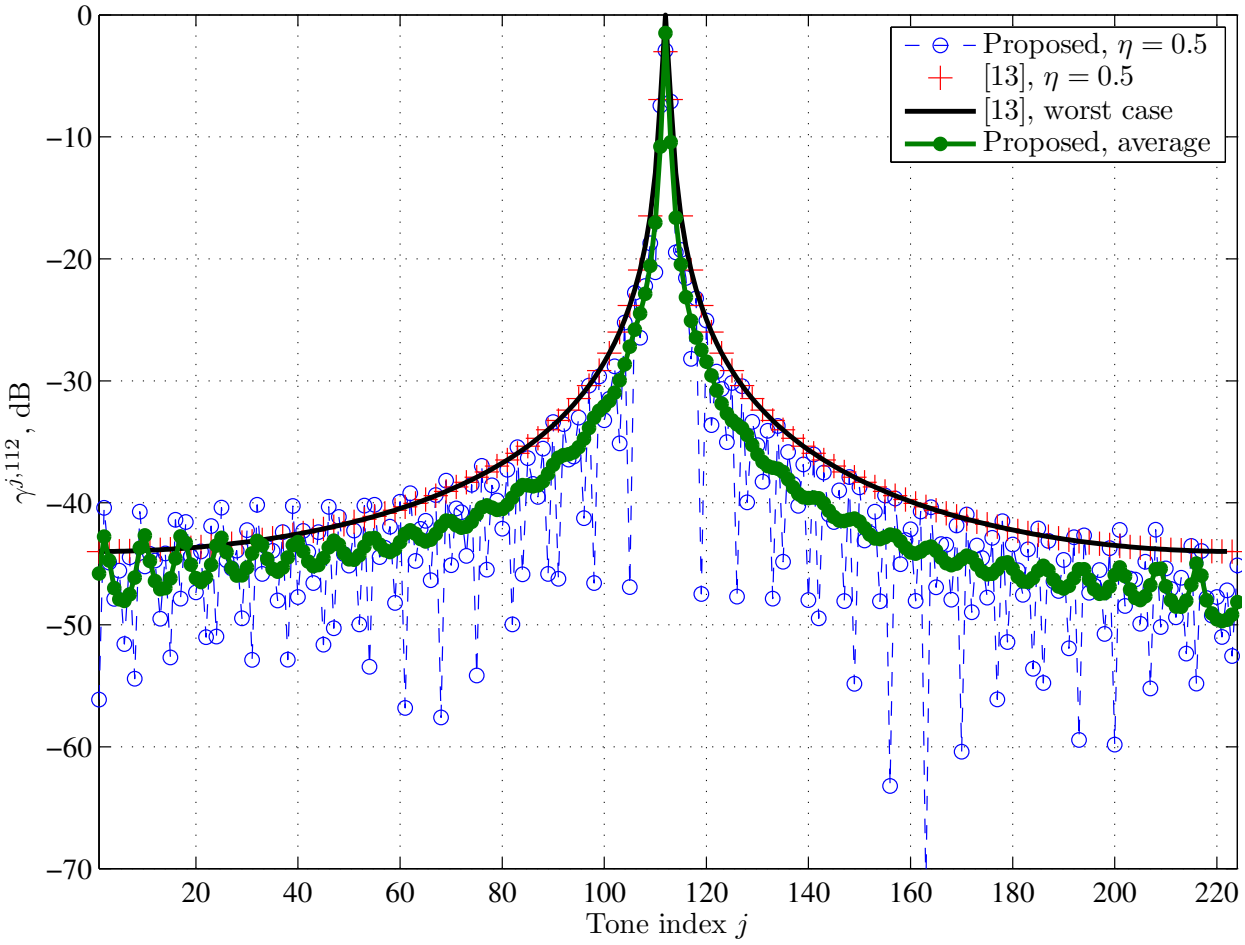

Fig. 2. ICI coefficients from (15), (6) and from (7). For the fisrt two plots, $\eta = 0.5$. The crosstalk channel is obtained with [19]

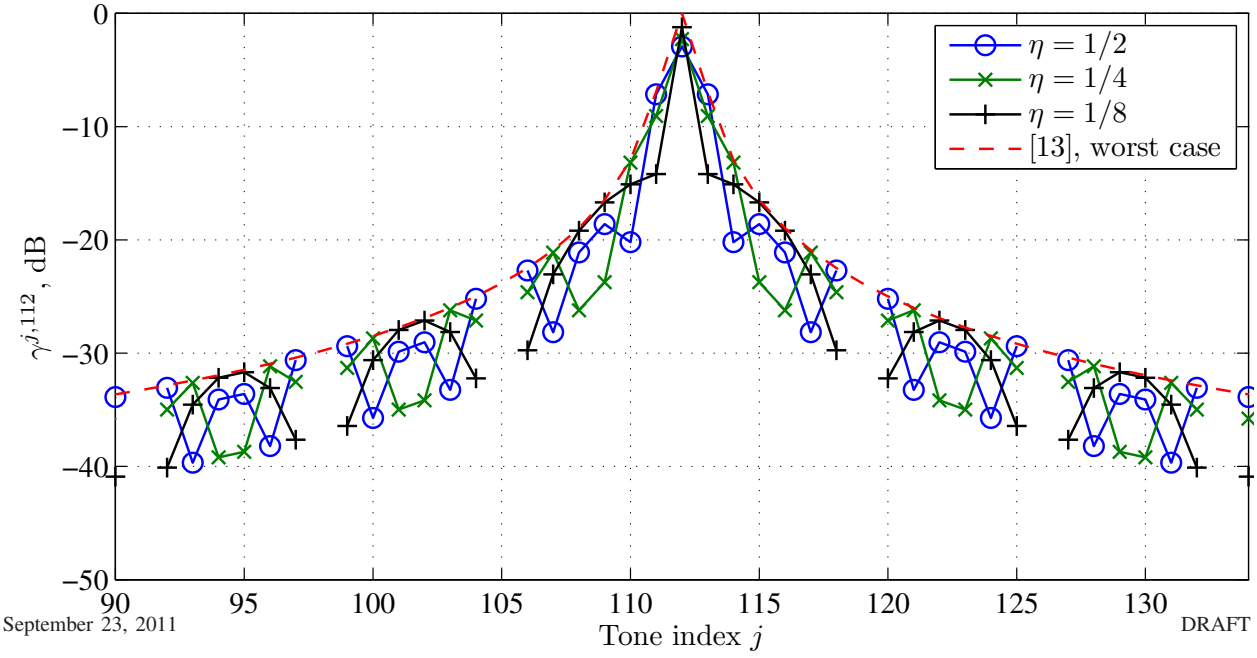

Fig. 3. ICI coefficients for different values of the delay η. In this plot we consider a frequency flat channel.

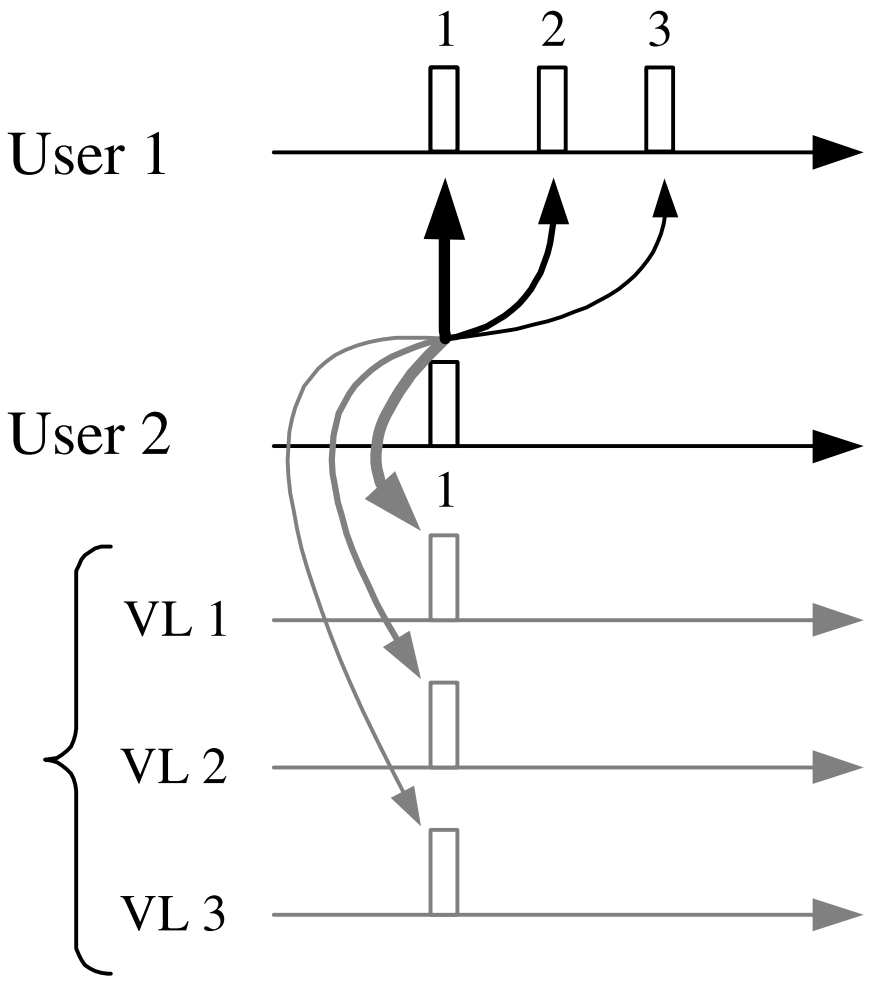

Fig. 4. Here we illustrate the idea of transforming an asynchronous victim to a batch of synchronous victim for the two user case. We only consider three tones. This transformation is only valid when the condition $\chi_n^k \approx 1$ holds true for all tones. We show the equivalent optimization for tone 1 of user 1. Optimizing with the asynchronous user 1 is equivalent to optimizing with 3 synchronous VL's.

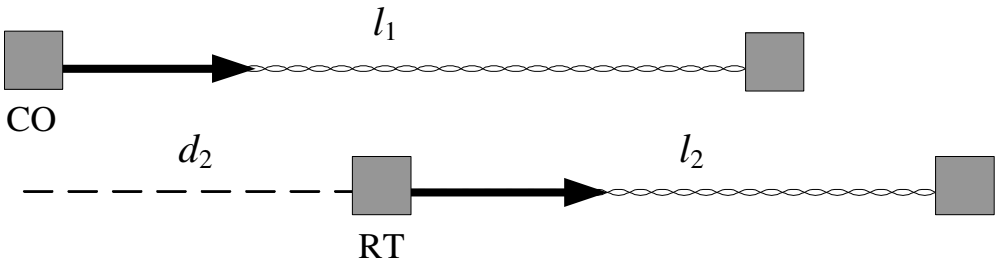

Fig. 5. Near-far downstream ADSL scenario.

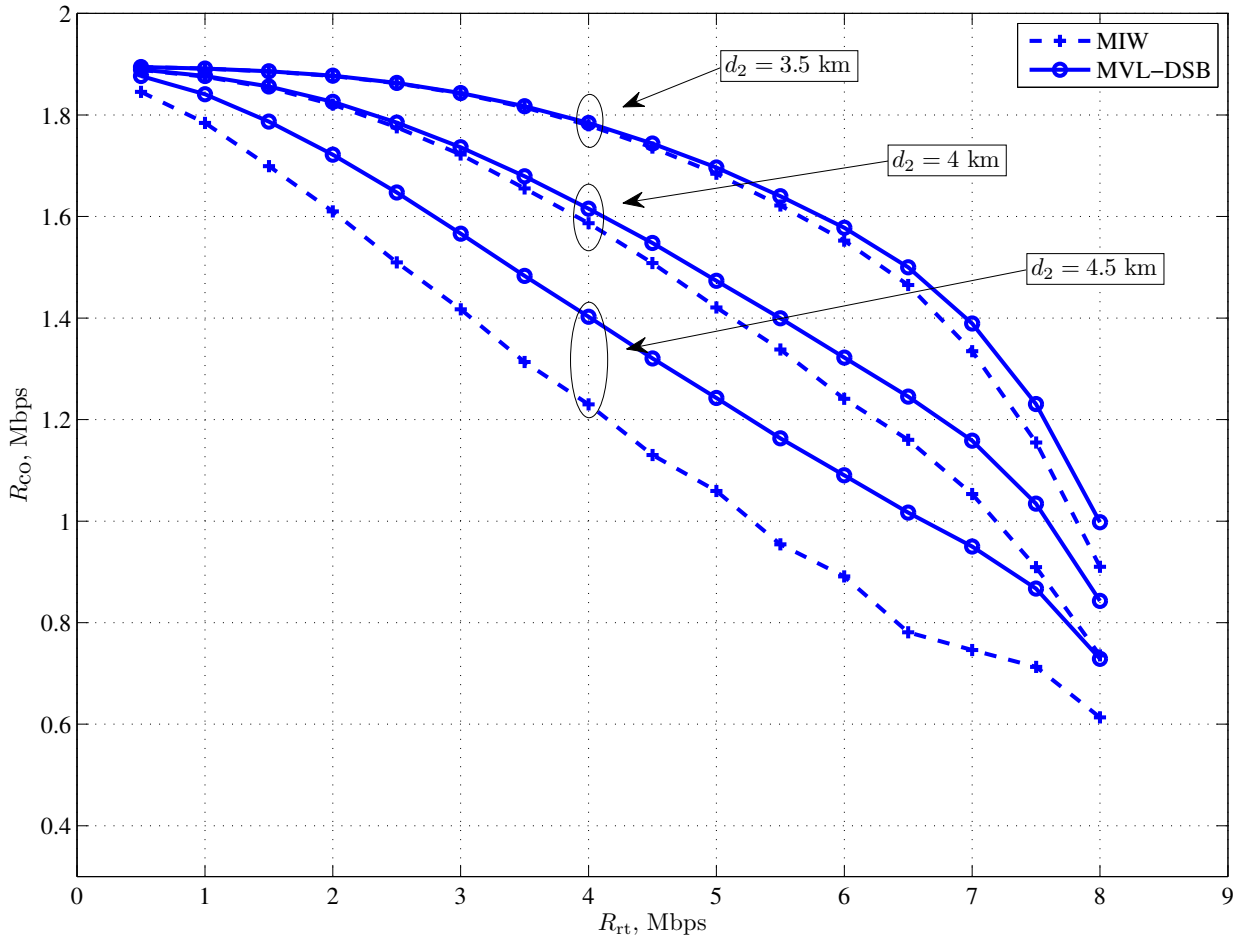

Fig. 6. RR's for asynchronous two user scenario for the MIW and the MVL-DSB algorithms. Referring to Fig. 5, we set $l_1 = 5$ km and $l_2 = 3$ km. The delay is set to $\eta = 0.5$. We change the value of d_2, $d_2 = 3.5$, 4 and 4.5 km.

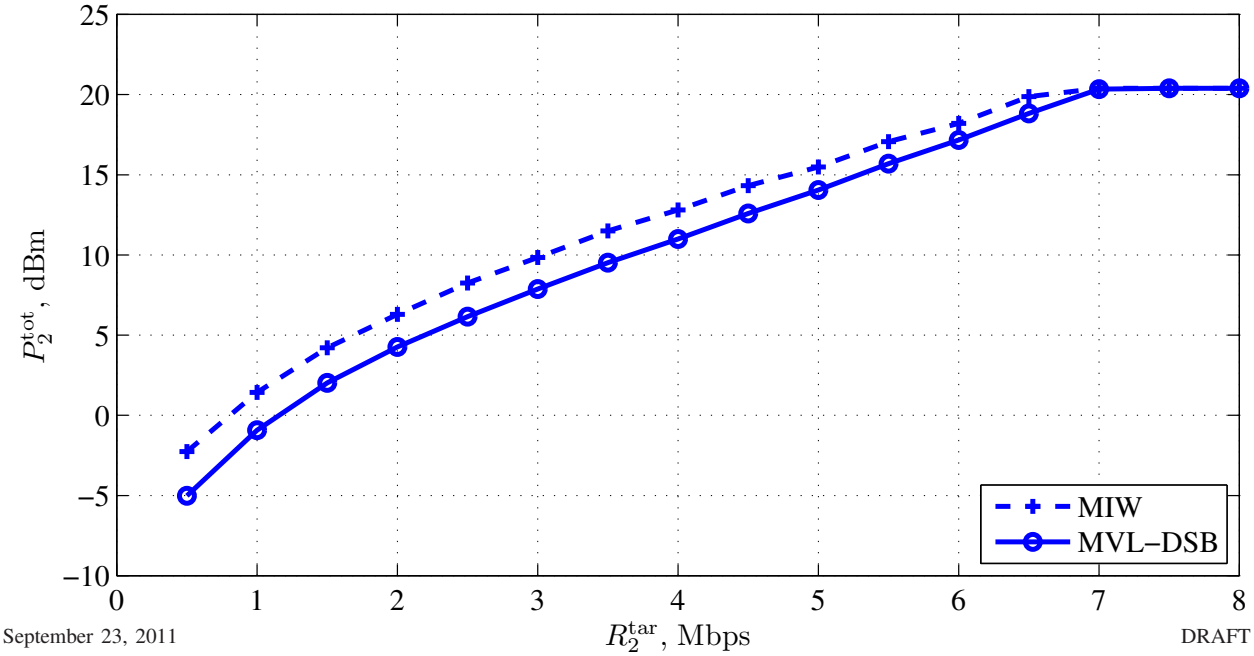

Fig. 7. Rate vs. power plot for the RT. We show the results for the MIW and the MVL-DSB algorithms. Referring to Fig. 5, we set $l_1 = 5$ km, $l_2 = 3$ km and $d_2 = 4.5$ km.

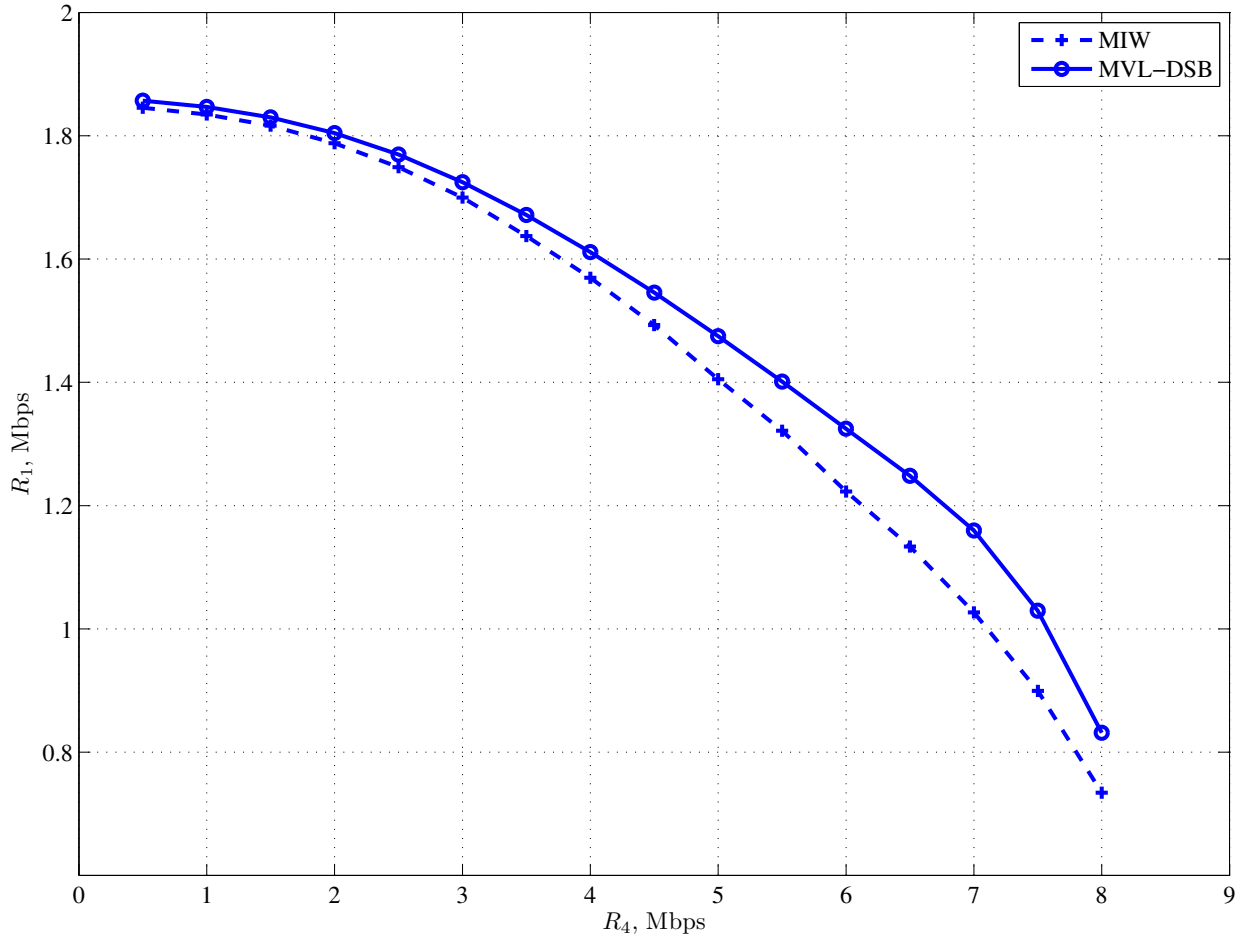

Fig. 8. Rate region for a scenario with $\mathbf{l} = [5 \;\; 4 \;\; 3.5 \;\; 3]^{\mathrm{T}}$ km and $\mathbf{d} = [0 \;\; 2 \;\; 3 \;\; 4]^{\mathrm{T}}$ km. Here we use the averaged ICI coefficients of (17) for all user pairs. For all point in the RR we set $R_2 = R_3 = 2$ Mbps.

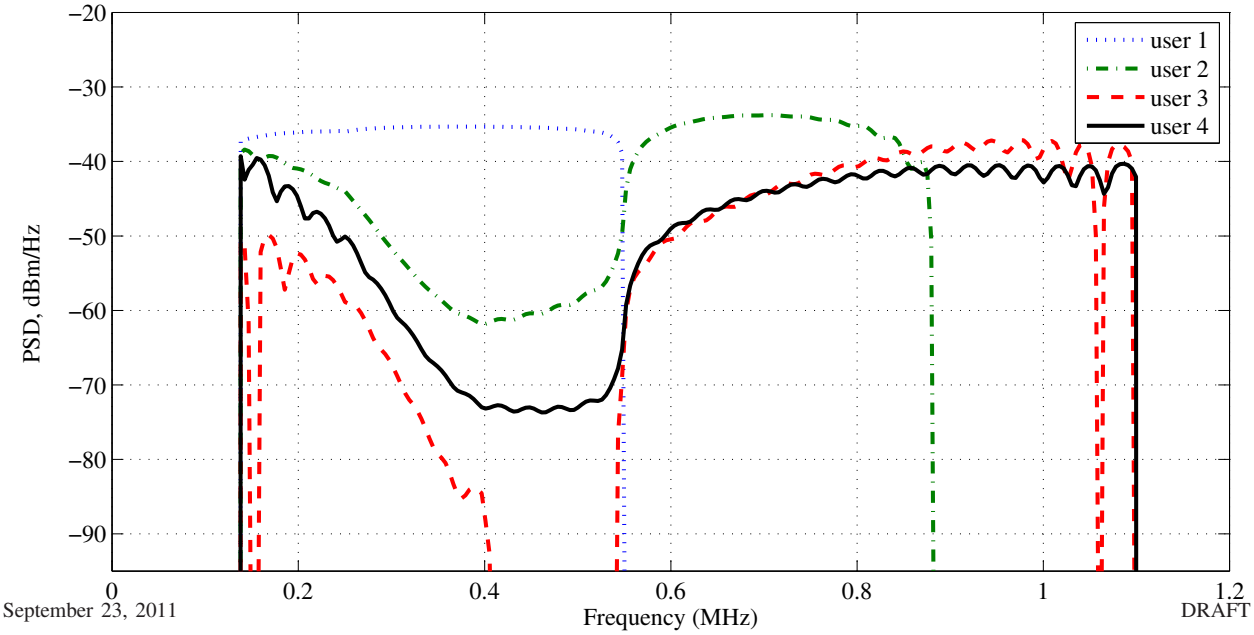

Fig. 9. PSDs corresponding to the point $R_4 = 5$ Mbps for the RR in Fig. 8

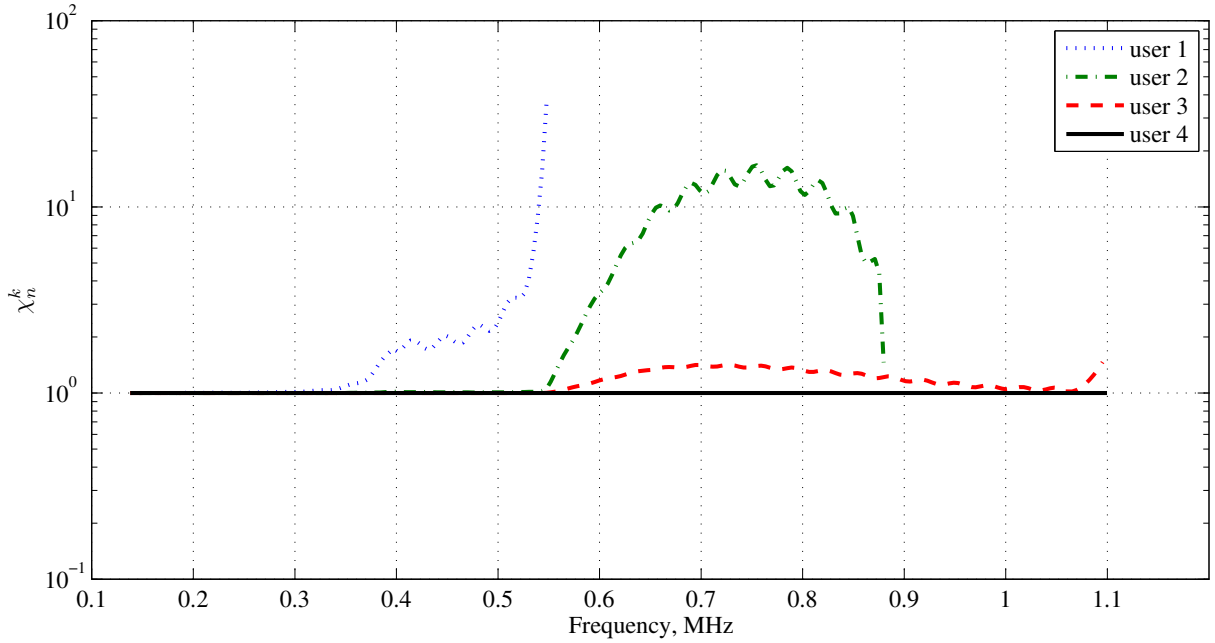

Fig. 10. Quantities corresponding to the point $R_4 = 5$ Mbps for the RR in Fig. 8