Distributed spectrum coordination for DSL broadband access networks ⋆

Paschalis Tsiaflakis, Marc Moonen

Electrical Engineering (ESAT-SCD), Katholieke Universiteit Leuven, Belgium (e-mail: paschalis.tsiaflakis@esat.kuleuven.be, marc.moonen@esat.kuleuven.be)

Abstract: Digital subscriber line (DSL) technology is the dominating broadband internet access technology with a market share of over 66% of all broadband subscribers worldwide. The main performance bottleneck of modern DSL networks is the crosstalk interference generated among different lines (i.e. twisted pair cables) inside the same cable bundle. Multi-user spectrum coordination is recognized as a key technology to tackle this crosstalk problem. It consists of optimizing the users' transmit spectra so as to mitigate the destructive impact of crosstalk, leading to spectacular data rate performance gains as well as power savings. The corresponding optimization problems are however large-scale nonconvex problems. In this paper we provide a survey of recent developments on spectrum coordination techniques where the main focus is on distributed spectrum coordination. Here, individual users optimize their transmit spectra based on local measures while they are steered by a centralized controller (spectrum management center) so as to obtain a good global network performance. It is shown how state-of-the-art techniques from mathematical programming are used to obtain distributed algorithms with very low complexity and yet good performance. Finally simulation results are shown that demonstrate the impact of spectrum coordination for DSL broadband access.

Keywords: dynamic spectrum management, digital subscriber lines, broadband access networks, distributed optimization, dual decomposition, Lagrange relaxation, proximal center method, iterative convex approximation, distributed spectrum balancing

1. INTRODUCTION

The internet plays a prominent role in our current information-driven society. One crucial component here is the broadband access network, which connects the subscribers to their service provider, which is then connected to the internet core network. One popular way to provide these broadband access connections makes use of the existing telephone network infrastructure. Here multiple twisted pairs are bundled into large cable bundles that typically consist of 20-100 twisted pairs each, and where many cable bundles are deployed from the telephone central office (CO) or a remote terminal (RT) towards the customers.

Digital subscriber line (DSL) technology is a technology that can turn each of these twisted pair connections into a true broadband connection, which is capable of delivering data rates of several Mb/s over a distance of multiple kms. DSL is the dominating broadband access technology with a market share of 66% of all broadband subscribers worldwide. The number of DSL subscribers is expected to grow to 330 million in 2012.

⋆ 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, GOA AMBioRICS, FWO project G.0235.07('Design and evaluation of DSL systems with common mode signal exploitation'), FWO project G.0226.06, and the Belgian Federal Science Policy Office IUAP DYSCO. The scientific responsibility is assumed by its authors.

To accommodate the many new broadband services (IPTV, video-phone, youtube.com, . . .), it is necessary that DSL technologies (ADSL, ADSL2(+), VDSL(2)) are designed that can also guarantee the corresponding quality of service (QoS) requirements (data rates, stability, power usage, delay, . . .). However, the major obstacle for further performance improvement of DSL access networks is the electromagnetic coupling amongst different lines in the same cable bundle, also referred to as crosstalk. Different twisted pair lines indeed act as antennas when used at the high DSL frequencies, creating a very challenging multi-user interference environment.

Multi-user coordination techniques, also referred to as dynamic spectrum management (DSM) techniques [Song et al. (2002); Starr et al. (2003)], are indispensable here to tackle the crosstalk problem by coordinating the different interfering users (i.e. lines, modems). Two main types of DSM techniques can be distinguished: (i) multi-user signal coordination [Ginis and Cioffi (2002); Cendrillon et al. (2006a)], where the data signals are coordinated over the different users so as to *remove* crosstalk, and (2) multi-user spectrum coordination, which consists of coordinating the user's transmit spectra, i.e. the user's transmit powers over different frequency bands, so as to *prevent* the destructive impact of crosstalk. Both techniques can lead to spectacular data rate performance gains as well as power savings. In this text the focus will be on multi-user spectrum coordination also referred to as spectrum balancing, spectrum management, or multi-carrier power control.

Optimal multi-user spectrum coordination basically corresponds to solving an optimization problem, also referred to as the spectrum management problem [Cendrillon et al. (2006b)], where the variables are the transmit spectra of the users. One crucial component in DSM is the spectrum management center (SMC). An SMC can be seen as a set of computer workstations that collects data from the network environment, i.e., channel characteristics (line attenuations as well as crosstalk levels), noise characteristics, etc. Based on the obtained information, the SMC can then solve the corresponding spectrum management problem to obtain the optimal combination of transmit spectra. This SMC-centric approach can be seen as *centralized spectrum coordination*. However this is not always practical because of different reasons, (i) different service providers can have different SMCs that each coordinate only part of the lines in a cable bundle, (ii) the centralized setting may not react fast enough to changes in the network because of the large communication overhead and the typically slow network monitoring [Cioffi et al. (2004)]. Therefore it makes sense to move to a more distributed setting, where users monitor their local environment using local measures and react to rapid changes in the network environment (noises turning on or off) by configuring their transmit spectra. In addition and on a slower pace, they can be steered by a centralized controller in the SMC towards a better global network performance. We will refer to this approach as *distributed spectrum coordination.*

The goal of this paper is to provide a survey of recent developments in distributed multi-user spectrum coordination. In Section 2, the system model is presented for multi-user spectrum coordination over a DSL cable bundle. In Section 3, spectrum management problem formulations are discussed where the most common formulation, i.e. constrainted weighted rate sum maximization, is presented. Furthermore it is shown how typical spectrum coordination solutions tackle this problem. In Section 4, the focus is on distributed spectrum coordination, where the approach of iterative convex approximation is discussed and elaborated for one particular distributed spectrum coordination algorithm, i.e. distributed spectrum balancing (DSB) [Tsiaflakis et al. (2008a)]. Finally in Section 5 the impact of applying spectrum coordination is demonstrated using a realistic DSL scenario.

2. SYSTEM MODEL

We consider multi-user spectrum coordination over a cable bundle consisting of $\mathcal{N} = \{1, \ldots, N\}$ interfering DSL users. The available bandwidth is divided into $\mathcal{K} = \{1, \ldots, K\}$ independent tones (i.e., frequency bands or carriers), where the exact number of tones K depends on the considered DSL technology (e.g. VDSL can use up to 4096 tones). Assuming standard synchronous discrete multitone (DMT) modulation, the cable bundle transmission can be modeled independently on each tone k by

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{z}_k.$$

The vector $\mathbf{x}_k = [x_k^1, \ldots, x_k^N]^T$ contains the transmitted signals on tone k, where x_k^n refers to the signal transmitted by user n on tone k. Vectors $\mathbf{z}_k$ and $\mathbf{y}_k$ have similar structures; $\mathbf{z}_k$ refers to the additive noise on tone k, containing thermal noise, alien crosstalk, radio frequency interference (RFI), etc, and $\mathbf{y}_k$ refers to the received signals on tone k. $\mathbf{H}_k$ is an $N \times N$ channel matrix with $[\mathbf{H}_k]_{n,m} = h_k^{n,m}$ referring to the channel gain from transmitter m to receiver n on tone k. The diagonal elements are the direct channels and the off-diagonal elements are the crosstalk channels.

The transmit power of user n on tone k is denoted as $s_k^n \triangleq \Delta_f E\{|x_k^n|^2\}$, where Δ_f refers to the tone spacing. The vector $\mathbf{s}_k \triangleq \{s_k^n, n \in \mathcal{N}\}$ denotes the transmit powers of all users on tone k. The vector $\mathbf{s}^n \triangleq \{s_k^n, k \in \mathcal{K}\}$ denotes the transmit spectrum of user n, i.e. the transmit powers of user n over all tones, which is also referred to as the transmit power spectral density of user n. The received noise power by user n on tone k is denoted as $\sigma_k^n \triangleq \Delta_f E\{|z_k^n|^2\}$.

Note that we assume no signal coordination at the transmitters and at the receivers, and that the interference is treated as additive white Gaussian noise. Under this standard assumption, the number of bits that user n can transmit on tone k, given the transmit spectra $\mathbf{s}_k$ of all users on tone k, is given by

$$b_k^n \triangleq b_k^n(\mathbf{s}_k) \triangleq \log_2\left(1 + \frac{1}{\Gamma}\frac{|h_k^{n,n}|^2 s_k^n}{\sum\limits_{m \neq n} |h_k^{n,m}|^2 s_k^m + \sigma_k^n}\right) \text{ bits/Hz,} \tag{1}$$

where Γ denotes the so-called signal-to-noise ratio (SNR) gap to capacity, which is a function of the desired bit error rate (BER), the coding gain and noise margin [Starr et al. (1999)]. The DMT symbol rate is denoted as f_s. The achievable total data rate for user n and the total power used by user n are equal to, respectively:

$$R^n \triangleq f_s \sum_{k \in \mathcal{K}} b_k^n, \text{ and } P^n \triangleq \sum_{k \in \mathcal{K}} s_k^n. \tag{2}$$

3. SPECTRUM COORDINATION

3.1 Spectrum management problems

Multi-user spectrum coordination consists of carefully allocating transmit spectra $\mathbf{s}^n$ over the different users $n, \in \mathcal{N}$, so as to pursue certain objectives subject to constraints in resources and QoS requirements. However the presence of crosstalk among the users significantly complicates this coordination. More particularly, the data rate R^n for user n depends on the transmit spectra of all other users in a very nonconvex way.

The particular objectives and constraints should be chosen by the service providers depending on the needs of their customers. The objectives are generally functions of the data rates R^n and the transmit powers P^n, where one typically strives for large data rates and small transmit powers. Also the constraints are generally defined for the data rates and transmit powers where minimum data rate requirements and maximum power constraints are common. A number of constraints are enforced by DSL standardization and should always be satisfied, namely the per-user total power constraints (3) and the so-called spectral mask constraints (4), i.e.

$$P^n \leq P^{n,\text{tot}}, \qquad n \in \mathcal{N} \tag{3}$$

$$0 \leq s_k^n \leq s_k^{n,\text{mask}}, \qquad n \in \mathcal{N}, k \in \mathcal{K}, \tag{4}$$

where $P^{n,\text{tot}}$ denotes the available total power budget for user n and $s_k^{n,\text{mask}}$ denotes the spectral mask for user n on tone k. The set of all possible data rate combinations that satisfy the constraints (3)(4) can be characterized by the achievable rate region $\mathcal{R}$, a concept that has its origin in information theory, i.e.

$$\mathcal{R} = \Big\{(R^n : n \in \mathcal{N}) | R^n = f_s \sum_{k \in \mathcal{K}} b_k^n(\mathbf{s}_k), \text{ s.t. (3) and (4)}\Big\} \tag{5}$$

The most common multi-user spectrum coordination problem formulation [Cendrillon et al. (2006b); Yu and Lui (2006); Lui and Yu (2005)], which will be referred to as the constrained weighted rate sum maximization (cWRS) formulation, is given as follows, where w_n is the weight given to user n:

$$\begin{aligned} \max_{\{\mathbf{s}^n, n \in \mathcal{N}\}} \quad & \sum_{n \in \mathcal{N}} w_n R^n \quad (= f_0) \\ \text{s.t.} \quad & P^n \leq P^{n,\text{tot}}, \quad n \in \mathcal{N} \qquad \text{(cWRS)} \\ & 0 \leq s_k^n \leq s_k^{n,\text{mask}}, \quad n \in \mathcal{N}, k \in \mathcal{K} \end{aligned} \tag{6}$$

However, recently many other spectrum management problem formulations have been proposed leading to a much larger modelling freedom for service providers to configure their network: e.g. extension to power-aware spectrum coordination (Green DSL in [Tsiaflakis et al. (2009d,c)], [Wolkerstorfer et al. (2008)]), extension to fairness in data rates [Luo and Zhang (2008)], extension to fairness in power usage [Tsiaflakis et al. (2009c)], and even extension to cross-layer spectrum coordination taking queue scheduling into account [Tsiaflakis et al. (2008b)]. It is shown that the solutions to these different problem formulations have a similar structure. Here, due to space limitations, we restrict ourselves to a discussion for cWRS. However extension to the other problem formulations is straightforward, and for further details we refer to [Tsiaflakis et al. (2009d,c, 2008b)].

3.2 Spectrum coordination solutions

From an optimization point of view, cWRS can be classified as an NP-hard nonconvex optimization problem [Luo and Zhang (2008)]. In fact, depending on the particular channel and noise characteristics, the objective function can exhibit many locally optimal solutions with a large difference in objective function value. The number of variables, i.e. transmit powers over all tones k and for all users n, is equal to NK, where the number of users N ranges between 2-100 and the number of tones K can go up to 4096.

The main approach to solve cWRS is via its dual formulation, sometimes also called the (Lagrange) dual problem or Lagrange relaxation, which is given as

$$\begin{aligned} \min_{\boldsymbol{\lambda}} \quad & g(\boldsymbol{\lambda}) \\ \text{s.t.} \quad & \boldsymbol{\lambda} \geq 0 \end{aligned} \tag{7}$$

$$\begin{aligned} \text{with } g(\boldsymbol{\lambda}) &= \max_{\{\mathbf{s}_k \in \mathcal{S}_k, k \in \mathcal{K}\}} \mathcal{L}(\boldsymbol{\lambda}, \mathbf{s}_k, k \in \mathcal{K}) \\ &= \max_{\{\mathbf{s}_k \in \mathcal{S}_k, k \in \mathcal{K}\}} \sum_{n \in \mathcal{N}} w_n R^n - \sum_{n \in \mathcal{N}} \lambda_n (P^n - P^{n,\text{tot}}) \end{aligned} \tag{8}$$

where $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_N]^T$ are the Lagrange multipliers, $g(\boldsymbol{\lambda})$ is the dual objective function in $\boldsymbol{\lambda}$ and corresponds to an optimization problem for fixed $\boldsymbol{\lambda}$, $\mathcal{L}(\boldsymbol{\lambda}, \mathbf{s}_k, k \in \mathcal{K})$ is the Lagrangian, and $\mathcal{S}_k = \{\mathbf{s}_k \in \mathbb{R}^N : 0 \leq s_k^n \leq s_k^{n,\text{mask}}, n \in \mathcal{N}\}$. The dual problem consists of a master problem (7), where the unknowns are the Lagrange multipliers $\boldsymbol{\lambda} \in \mathbb{R}_+^N$, and a slave problem (8), where the Lagrange multipliers are fixed and the unknowns are the transmit powers $s_k^n, n \in \mathcal{N}, k \in \mathcal{K}$.

In [Luo and Zhang (2008); Yu and Lui (2006)] the authors proved that asymptotic zero duality holds, i.e. when K goes to infinity the solution of cWRS is equal to the solution of its dual formulation (7) - (8). In practice K is quite large and so zero duality gap is a valid assumption.

The master problem (7) is a convex problem. However it is non-differentiable because the slave problem can have multiple globally optimal solutions for given Lagrange multipliers $\boldsymbol{\lambda}$. In [Yu and Lui (2006); Cendrillon and Moonen (2005)] an iterative subgradient approach is proposed for the master problem as follows

$$\boldsymbol{\lambda} = \Big[\boldsymbol{\lambda} + \delta(\sum_{k \in \mathcal{K}} \mathbf{s}_k(\boldsymbol{\lambda}) - \mathbf{P}^{\text{tot}})\Big]^+, \tag{9}$$

where $[\mathbf{x}]^+$ denotes the projection of $\mathbf{x} \in \mathbb{R}^N$ onto $\mathbb{R}_+^N$, $\mathbf{P}^{\text{tot}} = [P^{1,\text{tot}}, \ldots, P^{N,\text{tot}}]^T$, $\mathbf{s}_k(\boldsymbol{\lambda})$ denotes the optimal solution of the slave problem (8) for given $\boldsymbol{\lambda}$, and δ is the stepsize that can be adapted using different strategies [Yu and Lui (2006); Tsiaflakis et al. (2007)] so as to converge to the optimal Lagrange multipliers, that satisfy the complementarity conditions $\boldsymbol{\lambda}^{\boldsymbol{T}}(\sum_k \mathbf{s}_k(\boldsymbol{\lambda}) - \mathbf{P}^{\text{tot}}) = 0$.

Because of the separability of the objective function of cWRS and the use of the Lagrange relaxation approach, the slave problem (8), i.e. $g(\boldsymbol{\lambda})$ for fixed $\boldsymbol{\lambda}$, can be decomposed into independent per-tone problems as follows:

$$\begin{aligned} & g(\boldsymbol{\lambda}) = \sum_{k \in K} g_k(\boldsymbol{\lambda}) = \sum_{k \in K} \max_{\mathbf{s}_k \in \mathcal{S}_k} \mathcal{L}_k(\boldsymbol{\lambda}, \mathbf{s}_k) \\ & \text{with } \mathcal{L}_k(\boldsymbol{\lambda}, \mathbf{s}_k) = f_s b_k(\mathbf{s}_k) - \sum_{n \in \mathcal{N}} \lambda_n s_k^n + \sum_{n \in \mathcal{N}} \lambda_n P^{n,\text{tot}}/K \\ & \text{and } b_k(\mathbf{s}_k) = \sum_{n \in \mathcal{N}} w_n b_k^n(\mathbf{s}_k) \end{aligned} \tag{10}$$

For each tone k, this corresponds to a per-tone nonconvex problem $g_k(\boldsymbol{\lambda})$, for fixed $\boldsymbol{\lambda}$, in a much smaller dimension N. This decomposition over tones is also known as dual decomposition and reduces the computational complexity of cWRS from exponential in K to only linear in K, with the addition of N unknown Lagrange multipliers.

The dual decomposition approach for finding the solution of cWRS is summarized in Algorithm 1. The different spectrum coordination algorithms proposed in literature mainly differ in how the nonconvex per-tone problems $g_k(\boldsymbol{\lambda})$, i.e. line 3 of Algorithm 1 are tackled. For instance, optimal spectrum balancing (OSB) [Cendrillon et al. (2006b)] uses an exhaustive discrete search, modified prismatic branch-and-bound (PBnB) [Xu et al. (2008)] uses a prismatic branch and bound algorithm, branch-and-bound optimal spectrum balancing (BB-OSB) [Tsiaflakis et al. (2007)] uses a branch and bound algorithm, iterative spectrum balancing (ISB) [Lui and Yu (2005); Cendrillon and Moonen (2005)] uses a coordinate descent discrete search, etc.

Algorithm 1 Dual decomposition approach for cWRS

1: Initialize $\boldsymbol{\lambda}, \{\mathbf{s}_k(\boldsymbol{\lambda}), k \in \mathcal{K}\}$
2: **while** $\boldsymbol{\lambda}^T(\sum_k \mathbf{s}_k(\boldsymbol{\lambda}) - \mathbf{P}^{\text{tot}}) \neq 0$ **do**
3: $\quad \forall k \in \mathcal{K} : \mathbf{s}_k(\boldsymbol{\lambda}) = \underset{\mathbf{s}_k \in \mathcal{S}_k}{\operatorname{argmax}}\ \mathcal{L}_k(\boldsymbol{\lambda}, \mathbf{s}_k) \quad (= g_k(\boldsymbol{\lambda}))$
4: $\quad \boldsymbol{\lambda} = \left[\boldsymbol{\lambda} + \delta(\sum_{k\in\mathcal{K}} \mathbf{s}_k(\boldsymbol{\lambda}) - \mathbf{P}^{\text{tot}})\right]^+$
5: **end while**

Generally, spectrum coordination algorithms can be divided into three categories: (1) centralized algorithms (e.g. OSB, ISB, PBnB, BB-OSB), which are executed by a centralized controller (SMC) that has full information on the network environment, (2) distributed algorithms (e.g. successive convex approximation for low-complexity (SCALE) [Papandriopoulos and Evans (2006)], (convex approximation) distributed spectrum balancing ((CA-)DSB) [Tsiaflakis et al. (2008a)], multiple starting point distributed spectrum balancing (MS-DSB) [Tsiaflakis et al. (2008a)], modified iterative water-filling (MIW) [Yu (2007)]), that consist of local procedures run by the users and that are steered by a centralized controller through message passing, and (3) autonomous algorithms (e.g. iterative water-filling (IW) [Yu et al. (2002)], autonomous spectrum balancing (ASB) [Cendrillon et al. (2007)]), which can be seen as fully distributed algorithms where each user optimizes its transmit spectrum based on local information and some a priori-information, and without any message passing. In the next section we will focus on distributed spectrum coordination algorithms and discuss the derivation of one particular algorithm, i.e. distributed spectrum balancing (DSB), which is also presented in [Tsiaflakis et al. (2008a)].

4. DISTRIBUTED SPECTRUM COORDINATION

Distributed spectrum coordination algorithms aim at near-optimal network performance, i.e. the transmit power allocation should be close to the optimal transmit power allocation of cWRS. To achieve this, the starting point in the design of distributed spectrum coordination algorithms is an optimization procedure that can efficiently tackle the cWRS problem. In a second step, the challenge is to decompose this optimization procedure into a distributed implementation that can be executed by the DSL infrastructure. In Section 4.1 we review a general optimization procedure to tackle nonconvex problems, namely the *iterative convex approximation approach.* In Sections 4.2, 4.3 and 4.4, it is shown how a particular convex approximation of cWRS can be decomposed into a distributed implementation that maps very well on the existing DSL infrastructure, resulting in a powerful distributed spectrum coordination algorithm.

4.1 Iterative convex approximation

To tackle nonconvex spectrum management problems an iterative convex approximation approach can be used. The basic idea here is to generate and solve a sequence of convex approximations so as to converge to a solution of the original nonconvex problem. This approach is formally summarized in Algorithm 2 where $\mathbf{F}$ refers to the original nonconvex spectrum management problem, e.g. cWRS, $\tilde{\mathbf{F}}$ refers to a convex approximation of $\mathbf{F}$, $\{\mathbf{s}_{k,\text{cvx}}(i), k \in \mathcal{K}\}$ are the transmit spectra over all users and all tones computed in iteration i, and $\mathbf{F}_{\text{cvx}}(\{\mathbf{s}_{k,\text{cvx}}(i), k \in \mathcal{K}\})$ refers to the convex approximation for a given point $\{\mathbf{s}_{k,\text{cvx}}(i), k \in \mathcal{K}\}$.

Algorithm 2 Generic iterative convex approximation for spectrum coordination

1: Initialize $i := 0, \{\mathbf{s}_{k,\text{cvx}}(i), k \in \mathcal{K}\}$
2: Approximate $\mathbf{F}$ by a convex approximation $\tilde{\mathbf{F}} = \mathbf{F}_{\text{cvx}}(\{\mathbf{s}_{k,\text{cvx}}(i), k \in \mathcal{K}\})$
3: **repeat**
4: $\quad i := i + 1$
5: $\quad \{\mathbf{s}_{k,\text{cvx}}(i), k \in \mathcal{K}\} :=$ Solve $\tilde{\mathbf{F}}$
6: $\quad$ Tighten convex approximation:
$\qquad \tilde{\mathbf{F}} = \mathbf{F}_{\text{cvx}}(\{\mathbf{s}_{k,\text{cvx}}(i), k \in \mathcal{K}\})$
7: **until** convergence

Under certain conditions on the chosen convex approximations $\mathbf{F}_{\text{cvx}}(\{\mathbf{s}_{k,\text{cvx}}(i), k \in \mathcal{K}\})$, which are described in [Chiang et al. (2007)], Algorithm 2 converges to a solution, that satisfies the Karush-Kuhn-Tucker (KKT) conditions of the original spectrum coordination problem $\mathbf{F}$. Different convex approximations have been proposed in literature that satisfy these conditions. For instance in [Chiang et al. (2007)], three types of approximations were proposed based on geometric programming: logarithmic approximation, which is also used by the SCALE algorithm, single condensation, and double condensation. Here, we focus on a convex approximation that was proposed in [Tsiaflakis et al. (2008a)], i.e. distributed spectrum balancing (DSB). Note that DSB, CA-DSB, and MIW are very similar algorithms that have been derived based on slightly different viewpoints but eventually come down to the same basic approach. In the next sections, we explain how DSB chooses its particular iterative convex approximation procedure and how a distributed implementation is obtained for this. A similar derivation can also be done for the other schemes based on geometric programming. We refer to [Chiang et al. (2007); Papandriopoulos and Evans (2006)] for further details.

4.2 DSB: Convex approximation

The convex approximation for DSB can be best explained by reformulating the nonconvex objective of cWRS as a difference of concave functions as follows:

$$f_0 = \sum_{n\in\mathcal{N}} w_n f_s \sum_{k\in\mathcal{K}} \log_2(|h_k^{n,n}|^2 s_k^n + \sum_{m\neq n} \Gamma|h_k^{n,m}|^2 s_k^m + \Gamma\sigma_k^n) \underbrace{- \sum_{n\in\mathcal{N}} w_n f_s \sum_{k\in\mathcal{K}} \log_2(\sum_{m\neq n} \Gamma|h_k^{n,m}|^2 s_k^m + \Gamma\sigma_k^n)}_{\text{A =non}-\text{concave part}} \tag{11}$$

Note that term A causes the non-concavity of the objective f_0. In DSB this convex term is approximated by a lower affine hyperplane which is tangent in point $\{\mathbf{s}_{k,\text{cvx}}(i - 1), k \in \mathcal{K}\}$ in iteration i. This approximation leads to the following convex approximation $\mathbf{F}_{\text{cvx}}(\{\mathbf{s}_{k,\text{cvx}}(i - 1), k \in \mathcal{K}\})$ of cWRS:

$$
\begin{aligned}
\max_{\{\mathbf{s}_k \in \mathcal{S}_k, k \in \mathcal{K}\}} \quad & \sum_{k \in \mathcal{K}} b_{k,cvx}(\mathbf{s}_k) \qquad (= f_{\text{cvx}}) \\
\text{s.t.} \quad & \sum_{k \in \mathcal{K}} s_k^n \leq P^{n,\text{tot}}, n \in \mathcal{N}
\end{aligned}
\tag{12}
$$

with

$$
\begin{aligned}
&b_{k,\text{cvx}}(\mathbf{s}_k) = \\
&\sum_{n \in \mathcal{N}} w_n f_s \log_2(|h_k^{n,n}|^2 s_k^n + \sum_{m \neq n} \Gamma |h_k^{n,m}|^2 s_k^m + \Gamma \sigma_k^n) \\
&\qquad\qquad - \sum_{n \in \mathcal{N}} w_n f_s (\sum_{m \neq n} a_k^{m,n} s_k^m + c_k^n)
\end{aligned}
$$

and

$$
a_k^{m,n} = \frac{\Gamma |h_k^{n,m}|^2 / \log(2)}{\sum_{p \neq n} \Gamma |h_k^{n,p}|^2 s_{k,\text{cvx}}^p (i-1) + \Gamma \sigma_k^n} \tag{13}
$$

The parameters $a_k^{m,n}$ and c_k^n are the approximation parameters where in fact c_k^n can be disregarded as it does not have any influence on the obtained transmit spectra. The approximation parameters $a_k^{m,n}$ that are used in iteration i, are computed based on the optimal solution $\{\mathbf{s}_{k,\text{cvx}}(i-1), k \in \mathcal{K}\}$ obtained in the previous iteration $i-1$ so as to obtain successive convex approximations, cfr. line 6 of Algorithm 2.

The particular choice of approximation parameters (13) leads to a convex approximation (12) whose objective function f_{cvx} has the following relation with respect to the objective f_0 of cWRS:

(1) $f_{\text{cvx}}(\mathbf{s}_k, k \in \mathcal{K}) \leq f_0(\mathbf{s}_k, k \in \mathcal{K}) \qquad \forall k, \mathbf{s}_k \in \mathcal{S}_k$

(2) $f_{\text{cvx}}(\mathbf{s}_{k,\text{cvx}}(i), k \in \mathcal{K}) = f_0(\mathbf{s}_{k,\text{cvx}}(i), k \in \mathcal{K})$, for all iterations i

(3) $\nabla f_{\text{cvx}}(\mathbf{s}_{k,\text{cvx}}(i), k \in \mathcal{K}) = \nabla f_0(\mathbf{s}_{k,\text{cvx}}(i), k \in \mathcal{K})$, for all iterations i

These three properties correspond to the conditions that were given in [Chiang et al. (2007)] to guarantee convergence of the iterative convex approximation approach to a solution that satisfies the KKT conditions of cWRS.

4.3 DSB: Solving the convex approximation

The convex approximation $\mathbf{F}_{\text{cvx}}$ (12) for each iteration i has to be solved in an efficient way, cfr. line 5 of Algorithm 2. Despite its large dimension, i.e. NK, it can be observed that the corresponding objective function is separable over the tones whereas the per-user total power constraints are coupling the problem over the tones. This is actually the key motivation for using a Lagrange relaxation approach, namely to move the coupled constraints to the objective with the introduction of Lagrange multipliers, and to obtain a so-called dual decomposition. The resulting dual problem is given in (14) and (15). It consists of a master convex problem in the Lagrange multipliers (14) and K independent convex slave problems in the transmit powers $\mathbf{s}_k$ of much smaller dimension N (15), i.e. $g_{k,\text{cvx}}(\boldsymbol{\lambda})$ for fixed $\boldsymbol{\lambda}$.

$$
\min_{\boldsymbol{\lambda} \geq 0} \; g_{\text{cvx}}(\boldsymbol{\lambda}) \qquad (\mathbf{F}_{\text{cvx,dual}}) \tag{14}
$$

$$
\begin{aligned}
\text{with} \quad & g_{\text{cvx}}(\boldsymbol{\lambda}) = \sum_{k \in \mathcal{K}} g_{k,\text{cvx}}(\boldsymbol{\lambda}) = \sum_{k \in \mathcal{K}} \max_{\mathbf{s}_k \in \mathcal{S}_k} \mathcal{L}_{k,\text{cvx}}(\mathbf{s}_k, \boldsymbol{\lambda}) \\
& \mathcal{L}_{k,\text{cvx}}(\mathbf{s}_k, \boldsymbol{\lambda}) = b_{k,\text{cvx}}(\mathbf{s}_k) - \sum_{n \in \mathcal{N}} \lambda_n s_k^n + \sum_{n \in \mathcal{N}} \lambda_n P^{n,\text{tot}} / K
\end{aligned}
\tag{15}
$$

In the next two sections, two approaches are proposed to solve the dual problem. As the duality gap is zero, the solution of the dual problem corresponds to the optimal solution of the convex approximation $\mathbf{F}_{\text{cvx}}$.

Subgradient based dual decomposition

The first approach is to use a standard subgradient-based procedure. This approach is summarized in Algorithm 3, where line 4 is the subgradient update of the Lagrange multipliers with the stepsize δ. Note that this algorithm is similar to Algorithm 1 proposed for cWRS, except for line 3 that now consists of solving K independent convex problems $g_{k,\text{cvx}}$ instead of nonconvex problems g_k.

Algorithm 3 Dual decomposition approach for $\mathbf{F}_{\text{cvx}}$

1: Initialize $\boldsymbol{\lambda}$
2: **while** $\boldsymbol{\lambda}^{\boldsymbol{T}}(\sum_k \mathbf{s}_k(\boldsymbol{\lambda}) - \mathbf{P}^{\text{tot}}) \neq 0$ **do**
3: $\quad \forall k \in \mathcal{K} : \mathbf{s}_k(\boldsymbol{\lambda}) = \underset{\mathbf{s}_k}{\text{argmax}}\, \mathcal{L}_{k,\text{cvx}}(\mathbf{s}_k, \boldsymbol{\lambda}) (= g_{k,\text{cvx}}(\boldsymbol{\lambda}))$
4: $\quad \boldsymbol{\lambda} = \left[\boldsymbol{\lambda} + \delta(\sum_{k \in \mathcal{K}} \mathbf{s}_k(\boldsymbol{\lambda}) - \mathbf{P}^{\text{tot}})\right]^+$
5: **end while**

The per-tone problems $g_{k,\text{cvx}}$ can be solved using state-of-the-art iterative methods (e.g. Newton's method). Alternatively, DSB uses an iterative fixed point update strategy. Based on the KKT conditions of $\mathbf{F}_{\text{cvx}}$ one can indeed derive the fixed point update formula [Tsiaflakis et al. (2008a)]:

$$
s_k^n = \left[\frac{w_n f_s / \log(2)}{\lambda_n + \sum_{m \neq n} \left(\omega_m f_s a_k^{n,m} - \frac{w_m f_s |h_k^{m,n}|^2 / \log(2)}{\frac{|h_k^{m,m}|^2 s_k^m}{\Gamma} + \sum_{p \neq m} |h_k^{m,p}|^2 s_k^p + \sigma_k^m} \right)} - \frac{\sum_{m \neq n} \Gamma |h_k^{n,m}|^2 s_k^m + \Gamma \sigma_k^n}{|h_k^{n,n}|^2} \right]_{s_k^{n,\text{mask}}}^{+} . \tag{16}
$$

where $[x]_a^+$ refers to $\min(\max(x,0),a)$. Note that the spectral mask constraints are taken into account by simple projection. Per-tone problem $g_{k,\text{cvx}}$ can now be solved by successively updating the transmit powers, using (16), for all users n over multiple iterations (outer iterations). It is experimentally shown that this iterative fixed point update works very well for solving the per-tone problems, in the sense that it converges in very few outer iterations (3-5) within a satisfying accuracy. Proving the convergence of this type of fixed point updates turns out to be very difficult. In [Cendrillon et al. (2007); Yu et al. (2002); Tsiaflakis et al. (2008a)], only few convergence conditions are provided, but it was also shown that convergence is always observed in simulations with realistic DSL scenarios.

Improved dual decomposition based optimization

Subgradient-based dual decomposition approaches are however known to exhibit slow convergence, i.e. convergence order of $\mathcal{O}(\frac{1}{\epsilon^2})$ with ϵ referring to the required accuracy of the approximation of the optimum of $\mathrm{F_{cvx}}$ [Nesterov (2004)]. Furthermore the stepsize parameter δ of the subgradient update (line 4 of Algorithm 3) is very difficult to tune so as to guarantee fast convergence.

An alternative improved dual decomposition approach for spectrum coordination was proposed in [Tsiaflakis et al. (2009a,b)], which is based on recent developments in mathematical programming, namely the proximal center method [Necoara and Suykens (2008)], advanced smoothing schemes and optimal gradient-based schemes [Nesterov (2004)]. More specifically, the improved approach combines a per-tone smoothing technique for obtaining a Lipschitz continuous dual objective, with an optimal gradient-based scheme to update the Lagrange multipliers. This results in the improved dual decomposition algorithm, as shown in Algorithm 4, where the smoothness parameter $c = \epsilon/(\sum_k \sum_n (s_k^{n,\mathrm{mask}})^2/2)$, the Lipschitz constant $L_c = K/c$, ϵ denotes the required accuracy, and $i_{\max}$ denotes the number of iterations. For further details, we refer to [Tsiaflakis et al. (2009a,b)].

Algorithm 4 Improved dual decomposition scheme for $\mathrm{F_{cvx}}$ (12)

1: $i := 0$, $\mathrm{tmp} := 0$
2: initialize $i_{\max}$, $\boldsymbol{\lambda}^i$
3: **for** $i = 0 \dots i_{\max}$ **do**
4: $\quad \forall k : \mathbf{s}_k^{i+1} = \underset{\mathbf{s}_k \in \mathcal{S}_k}{\mathrm{argmax}} b_{k,\mathrm{cvx}}(\mathbf{s}_k) - \sum_{n \in \mathcal{N}} \left(\lambda_n^i s_k^n - c\frac{(s_k^n)^2}{2}\right)$
5: $\quad d\bar{g}_c^{i+1} = \sum_{k \in \mathcal{K}} \mathbf{s}_k^{i+1} - \mathbf{P}^{\mathrm{tot}}$
6: $\quad \mathbf{u}^{i+1} = [\frac{d\bar{g}_c^{i+1}}{L_c} + \boldsymbol{\lambda}^i]^+$
7: $\quad \mathrm{tmp} := \mathrm{tmp} + \frac{i+1}{2} d\bar{g}_c^{i+1}$
8: $\quad \mathbf{v}^{i+1} = [\frac{\mathrm{tmp}}{L_c}]^+$
9: $\quad \boldsymbol{\lambda}^{i+1} = \frac{i+1}{i+3}\mathbf{u}^{i+1} + \frac{2}{i+3}\mathbf{v}^{i+1}$
10: **end for**
11: Build $\hat{\boldsymbol{\lambda}} = \boldsymbol{\lambda}^{i_{\max}+1}$
and $\hat{\mathbf{s}}_k = \sum_{i=0}^{i_{\max}} \frac{2(i+1)}{(i_{\max}+1)(i_{\max}+2)} \mathbf{s}_k^{i+1}$

Note that this approach is fully automatic, i.e. it automatically and optimally tunes its stepsize in contrast to the difficult stepsize tuning required for the subgradient-based approach. Furthermore it is proven in [Tsiaflakis et al. (2009a,b)], that Algorithm 4 converges one order of magnitude faster than the standard subgradient approach (Algorithm 3), i.e. $\mathcal{O}(\frac{1}{\epsilon})$ instead of $\mathcal{O}(\frac{1}{\epsilon^2})$, resulting in a much faster convergence of DSB. This has also been confirmed experimentally.

Finally note that the per-tone problem, i.e. line 4 of Algorithm 4, consists of an extra term. The fixed point update formula for the transmit powers (16) can be simply modified so as to take this extra term into account. For further details, we refer to [Tsiaflakis et al. (2009a,b)].

4.4 DSB: Distributed implementation

The derivation of the distributed implementation of DSB is based on the procedures given in Sections 4.2 and 4.3 that follow an iterative convex approximation approach to solve cWRS, namely the approximation step (13), the update of the Lagrange multipliers (e.g. the subgradient approach of Algorithm 3), and the iterative updates of the transmit powers (16). By combining (13) and (16), we obtain the following expression for the transmit power updates

$$s_k^n = \left[\frac{w_n f_s / \log(2)}{\lambda_n + \sum\limits_{m \neq n} \frac{w_m f_s \Gamma |h_k^{m,n}|^2}{\log(2)} \left(\frac{1}{\mathrm{int}_k^m} - \frac{1}{\mathrm{rec}_k^m} \right)} - \frac{\mathrm{int}_k^n}{|h_k^{n,n}|^2} \right]_{s_k^{n,\mathrm{mask}}}^+ \tag{17}$$

where $\mathrm{int}_k^n = \sum_{m \neq n} \Gamma |h_k^{n,m}|^2 s_k^m + \Gamma \sigma_k^n$ refers to the interference received by user n on tone k, and $\mathrm{rec}_k^n = |h_k^{n,n}|^2 s_k^n + \sum_{m \neq n} \Gamma |h_k^{n,m}|^2 s_k^m + \Gamma \sigma_k^n$ refers to the received signal by user n on tone k. Formula (17) can then be reformulated as follows

$$\begin{aligned} & s_k^n = \left[\frac{w_n f_s / \log(2)}{\lambda_n + W_k^n} - \frac{\mathrm{int}_k^n}{|h_k^{n,n}|^2} \right]_{s_k^{n,\mathrm{mask}}}^+ \\ \text{with} \quad & W_k^n = \sum_{m \neq n} \frac{w_m f_s \Gamma |h_k^{m,n}|^2}{\log(2)} V_k^m \\ \text{and} \quad & V_k^m = \left(\frac{1}{\mathrm{int}_k^m} - \frac{1}{\mathrm{rec}_k^m} \right) \end{aligned} \tag{18}$$

We can now differentiate between local information and non-local information. Local information is all information that is accessible by the users locally, i.e. (1) constants w_n, f_s, (2) local variables $\lambda_n, \{s_k^n, k \in \mathcal{K}\}$, local state information $\{V_k^n, k \in \mathcal{K}\}$, and local measures $\{\mathrm{int}_k^n, \mathrm{rec}_k^n, h_k^{n,n}, k \in \mathcal{K}\}$. Note that these local measures are already available in current DSL modems. Non-local information is information that is based on local information of the other users, such as $W_k^n, k \in \mathcal{K}, n \in \mathcal{N}$. This information is not available to the users. However in the distributed setting this information can be gathered by the centralized controller (SMC) through message passing. Each of the users then regularly transmits its local state information $\{V_k^n, k \in \mathcal{K}\}$ to the SMC. Based on this, together with the knowledge of the channel and the user constants w_n, f_s, the SMC calculates the $\{W_k^n, k \in \mathcal{K}\}$ for all users n and regularly updates the users with this information. In the meantime all users monitor their local environment, and optimize their transmit powers. The combination of the local algorithms with the continuous updating through message passing basically implements the iterative convex approximation approach and guarantees the convergence to a locally optimal solution of cWRS and so also a good global network performance. The resulting local algorithm run by the individual users is shown in Algorithm 5, whereas the centralized communication steering loop is shown in Algorithm 6.

Note that in the case there is no message passing, this corresponds to the case where all users have no non-local information, i.e. $\{W_k^n = 0, k \in \mathcal{K}\}$ for all users n. It is shown in [Tsiaflakis et al. (2008a)] that this corresponds to the case where all users update their transmit spectra in a selfish way without minding the interference caused to the other users, similarly to what is done by the iterative waterfilling (IW) algorithm. This is an interesting property as it implies that by increasing the centralized control, the network performance improves from a selfish behaviour to

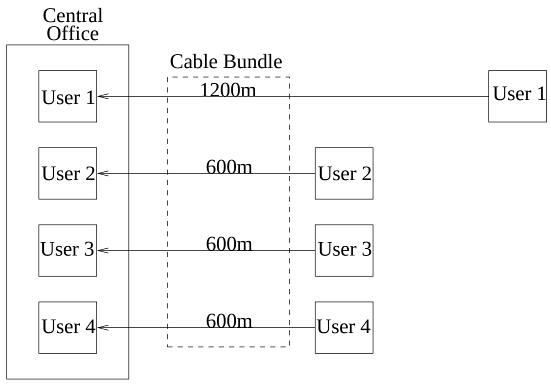

Fig. 1. 4-user upstream VDSL scenario

a more social behaviour that corresponds to a better global network behaviour.

Algorithm 5 Local algorithm for each user n

1: Execute at regular intervals:
2: Measure $\text{int}_k^n, \text{rec}_k^n, k \in \mathcal{K}$
3: Compute $V_k^n, k \in \mathcal{K}$ using (18)
4: **while** $\lambda_n(\sum_{k\in\mathcal{K}} s_k^n - P^{n,\text{tot}}) \neq 0$ **do**
5: Update λ_n (e.g. subgradient update)
6: Update $s_k^n, k \in \mathcal{K}$ using (17)
7: **end while**

Algorithm 6 SMC centralized steering loop:

1: Execute at regular intervals:
2: Monitor network environment (channels, noise)
3: Receive $V_k^n, k \in \mathcal{K}$ of all users n
4: Compute $W_k^n, n \in \mathcal{N}, k \in \mathcal{K}$ using (18)
5: Transmit $W_k^n, k \in \mathcal{K}$ to user n, and that for each user.

5. SIMULATION RESULTS

In this section we will demonstrate the data rate performance gains that result from applying distributed spectrum coordination. The considered scenario is shown in Figure 1, and consists of a four-user upstream VDSL scenario. The twisted pair lines have a diameter of 0.5 mm (24 AWG). The maximum user's total transmit power is 11.5 dBm. The SNR gap Γ is 12.9 dB. The tone spacing Δ_f is 4.3125 kHz. The DMT symbol rate f_s is 4 kHz. The weights w_n are equal to 1 for all users n. The spectral mask constraints are $s_k^{n,\text{mask}} = -60\text{dBm/Hz}, n \in \mathcal{N}, k \in \mathcal{K}$.

This scenario is known as a near-far scenario, where the three near users, i.e. with line length of 600m, can generate a huge amount of interference into the far user, i.e. with line length 1200m. When no spectrum coordination is applied, all users transmit at their spectral mask, i.e. $s_k^n = s_k^{n,\text{mask}}, n \in \mathcal{N}, k \in \mathcal{K}$. In this case the data rates obtained by the users correspond to $R^1 = 0.043$Mb/s, $R^2 = 17.6$Mb/s, $R^3 = 17.6$Mb/s, $R^4 = 17.6$Mb/s.

However when applying spectrum coordination through the DSB algorithm, as discussed in the previous section, the data rates increase up to $R^1 = 1.140$Mb/s, $R^2 = 19.2$Mb/s, $R^3 = 19.2$Mb/s, $R^4 = 19.2$Mb/s. One can observe a huge increase in the data rate of user 1. The corresponding transmit spectra for the four users are shown in Figure 2. One can see that user 1 mainly transmits at the low frequency range, i.e. from available tone 1 up to tone 112. The three near users have the same transmit spectra where in the low frequency range the level is decreased so as to prevent too much crosstalk degradation to user 1. However, in the frequency range above tone 112, the near users transmit at higher power levels as user 1 does not use these frequencies. So one can observe that application of spectrum coordination leads to intelligent spectrum shaping which leads to spectacular data rate performance gains.

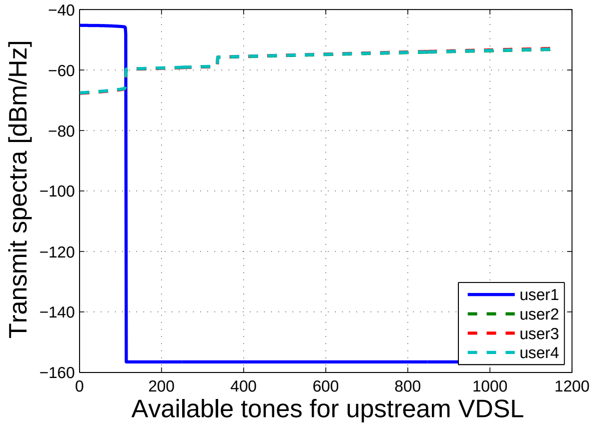

Fig. 2. Transmit spectra for the four users obtained by the DSB algorithm for the DSL scenario of Figure 1

Finally in Figure 3, we plotted the evolution of the data rate of user 1, while the data rates of the other users are fixed at 19.2Mb/s, in function of the number of iterations of message passing, where one iteration corresponds to a system wide update of the $W_k^n, n \in \mathcal{N}, k \in \mathcal{K}$. One can observe that when the number of communication messages increases, the data rate performance increases from 0.2Mb/s up to 1.1Mb/s. After 30 iterations, the distributed algorithm converges to the globally optimal solution of cWRS. This is verified by comparing with a globally optimal spectrum coordination algorithm (BB-OSB). In practice one could make a trade-off between the number of messages (communication overhead) and the data rate performance.

6. CONCLUSION

In order to cope with the QoS requirements of the many emerging broadband services, current DSL technologies have to be improved. The main bottleneck for data rate performance improvement of current DSL broadband access is crosstalk interference. Spectrum coordination is identified as a key technology to tackle this crosstalk problem and consists of intelligently coordinating the transmit spectra of the interfering users. Here, distributed spectrum coordination is an interesting option as it combines local procedures that can react fast on the channel environment with a centralized control that steers towards a better global network behavior. It is shown how recent advances from mathematical programming can be used to design very efficient distributed spectrum coordination algorithms, that scale with the amount of message passing from a selfish local behavior to a social network behavior.

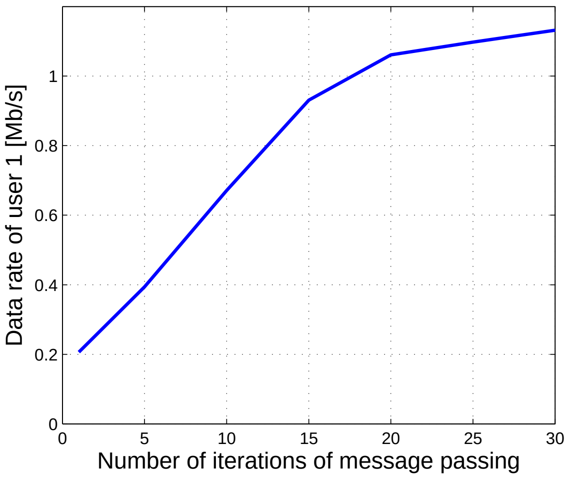

Fig. 3. Evolution of the data rate performance of user 1, for fixed data rates of the other users, for an increasing number of iterations of message passing, i.e. number of network wide updates of W_k^n for all users n and over all tones k

REFERENCES

Cendrillon, R., Ginis, G., Bogaert, E.V.D., and Moonen, M. (2006a). A near-optimal linear crosstalk canceler for upstream vdsl. *IEEE Transactions on Signal Processing*, 54(8), 3136–3146.

Cendrillon, R., Huang, J., Chiang, M., and Moonen, M. (2007). Autonomous spectrum balancing for digital subscriber lines. *IEEE Transactions on Signal Processing*, 55(8), 4241–4257.

Cendrillon, R. and Moonen, M. (2005). Iterative spectrum balancing for digital subscriber lines. In *IEEE Int. Conf. on Communications*, volume 3, 1937–1941.

Cendrillon, R., Yu, W., Moonen, M., Verlinden, J., and Bostoen, T. (2006b). Optimal multiuser spectrum balancing for digital subscriber lines. *IEEE Transactions on Communications*, 54(5), 922–933.

Chiang, M., Tan, C., Palomar, D., O'Neill, D., and Julian, D. (2007). Power control by geometric programming. *IEEE Transactions on Wireless Communications*, 6(7), 2640–2651.

Cioffi, J., Rhee, W., Mohseni, M., and Brady, M. (2004). Band preference in dynamic spectrum management. In *12th European Signal Processing Conference*.

Ginis, G. and Cioffi, J. (2002). Vectored transmission for digital subscriber line systems. *IEEE Journal on Selected Areas in Communications*, 20(5), 1085–1104.

Lui, R. and Yu, W. (2005). Low-complexity near-optimal spectrum balancing for digital subscriber lines. In *IEEE International Conference on Communications*, volume 3, 1947–1951.

Luo, Z.Q. and Zhang, S. (2008). Dynamic spectrum management: Complexity and duality. *IEEE Journal of Selected Topics in Signal Processing*, 2(1), 57–73.

Necoara, I. and Suykens, J. (2008). Application of a smoothing technique to decomposition in convex optimization. *IEEE Transactions on Automatic Control*, 11(53), 2674–2679.

Nesterov, Y. (2004). *Introductory Lectures on Convex Optimization: A Basic Course*. Boston.

Papandriopoulos, J. and Evans, J. (2006). Low-complexity distributed algorithms for spectrum balancing in multi-user DSL networks. In *IEEE International Conference on Communications*, volume 7, 3270–3275.

Song, K., Chung, S., Ginis, G., and Cioffi, J. (2002). Dynamic spectrum management for next-generation DSL systems. *IEEE Communications Magazine*, 40(10), 101–109.

Starr, T., Cioffi, J., and Silverman, P. (1999). *Understanding digital subscriber lines*. Prentice Hall.

Starr, T., Sorbara, M., Cioffi, J., and Silverman, P. (2003). *DSL Advances*, volume Chapter 11. Prentice-Hall.

Tsiaflakis, P., Diehl, M., and Moonen, M. (2008a). Distributed spectrum management algorithms for multiuser DSL networks. *IEEE Transactions on Signal Processing*, 56(10), 4825–4843.

Tsiaflakis, P., Necoara, I., Suykens, J., and Moonen, M. (2009a). An improved dual decomposition approach to DSL dynamic spectrum management. In *17th European Signal Processing Conference (EUSIPCO 2009)*.

Tsiaflakis, P., Necoara, I., Suykens, J., and Moonen, M. (2009b). Improved dual decomposition based optimization for DSL dynamic spectrum management. *submitted to IEEE Transactions on Signal Processing*.

Tsiaflakis, P., Vangorp, J., Moonen, M., and Verlinden, J. (2007). A low complexity optimal spectrum balancing algorithm for digital subscriber lines. *Signal Processing*, 87(7), 1735–1753.

Tsiaflakis, P., Yi, Y., Chiang, M., and Moonen, M. (2008b). Throughput and delay of DSL dynamic spectrum management with dynamic arrivals. In *IEEE Global Telecommunications Conference*, 1–5.

Tsiaflakis, P., Yi, Y., Chiang, M., and Moonen, M. (2009c). Fair greening for DSL broadband access. In *GreenMetrics Workshop in conjuction with ACM sigmetrics (GreenMetrics 2009)*.

Tsiaflakis, P., Yi, Y., Chiang, M., and Moonen, M. (2009d). Green DSL: Energy-efficient DSM. In *IEEE International Conference on Communications*.

Wolkerstorfer, M., Statovci, D., and Nordstrom, T. (2008). Dynamic spectrum management for energy-efficient transmission in DSL. In *11th IEEE Singapore International Conference on Communication Systems*, 1015–1020.

Xu, Y., Le-Ngoc, T., and S.Panigrahi (2008). Global concave minimization for optimal spectrum balancing in multi-user DSL networks. *IEEE Transactions on Signal Processing*, 56(7), pp. 2875–2885.

Yu, W. (2007). Multiuser water-filling in the presence of crosstalk. In *Information Theory and Applications (ITA)*.

Yu, W., Ginis, G., and Cioffi, J. (2002). Distributed multiuser power control for digital subscriber lines. *IEEE Journal on Selected Areas in Communications*, 20(5), 1105–1115.

Yu, W. and Lui, R. (2006). Dual methods for nonconvex spectrum optimization of multicarrier systems. *IEEE Transactions on Communications*, 54(7).