A Low Complexity Optimal Spectrum Balancing Algorithm for Digital Subscriber Lines [1]

Paschalis Tsiaflakis [a,⋆⋆,1] Jan Vangorp [a] Marc Moonen [a]
Jan Verlinden [b]

[a] *Katholieke Universiteit Leuven, Department of Electrical Engineering, Kasteelpark Arenberg 10, 3001 Leuven/Heverlee, Belgium*

[b] *Alcatel Bell, DSL Experts Team, Francis Wellesplein 1, 2018 Antwerpen, Belgium*

Abstract

In modern DSL systems, multi-user crosstalk is a major source of performance degradation. Optimal Spectrum Balancing (OSB) is a centralized algorithm that mitigates the effect of crosstalk by allocating optimal transmit spectra to all interfering DSL modems. By the use of Lagrange multipliers the algorithm decouples the spectrum management problem into per-tone optimization problems. The remaining issues are then finding the Lagrange multipliers that enforce the constraints and solving the per-tone optimization problems. Finding the optimal Lagrange multipliers can become complex when more than two users are considered. Starting from the single-user case, this paper presents a number of properties, which are then extended to the multi-user case and lead to an efficient search algorithm for the Lagrange multipliers. Simulations show that the number of Lagrange multiplier evaluations is as small as 20-50, independent of the number of users. Secondly, the complexity of the per-tone optimization problems grows exponentially with the number of lines in the binder. For multiple user scenarios this becomes computationally intractable. This paper presents an efficient branch and bound approach for the per-tone optimization problem. Simulations show enormous complexity reductions, especially for a large number of users.

Key words: Asymmetric digital subscriber line, xDSL, crosstalk, dynamic spectrum management, optimal spectrum balancing, multi-user power loading, multi-user bit loading

1 Introduction

The ever increasing demand for higher data rates forces DSL systems to use higher frequencies, e.g. up to 30 MHz for VDSL2 [1]. At these frequencies, electromagnetic coupling becomes particularly harmful and causes crosstalk between systems operating in the same bundle. This crosstalk, typically 10-15dB larger than the background noise, is a dominant source of performance degradation in DSL systems currently under development [2].

There are two strategies for dealing with this crosstalk: crosstalk cancellation and spectrum management. Crosstalk cancellation can remove the crosstalk completely with minimal noise enhancement [3] [4] [5], but requires signal level cooperation between receivers or transmitters. Unfortunately, in an unbundled scenario this cooperation is not available. In this case crosstalk can be mitigated through the use of spectrum management.

Current DSL systems use a Static Spectrum Management (SSM) approach where fixed spectral masks ensure that crosstalk levels remain within an acceptable range [6] [7]. Because these spectral masks are designed for worst case loop characteristics, this approach can be extremely suboptimal. Dynamic Spectrum Management (DSM) overcomes this problem by designing the transmit spectrum of each modem according to the topology of the network. In this way spectra take into account the current requirements of all users, causing as little disturbance as possible.

One of the first DSM algorithms proposed is Iterative Waterfilling (IW) [8]. In this algorithm, each user waterfills its spectrum against the noise and interference. By repeating this in an iterative fashion over the different users, the procedure converges to a "selfish" optimum. IW is a low complexity distributed algorithm, meaning it does not need any form of centralized control. Although IW significantly outperforms SSM, it is found to be suboptimal. This is especially so in heavily unbalanced scenarios, where some lines cause

[1] This research work was carried out at the ESAT laboratory of the Katholieke Universiteit Leuven, in the frame of the Belgian Programme on Interuniversity Attraction Poles, initiated by the Belgian Federal Science Policy Office, IUAP P5/22 and P5/11, and in the frame of the IWT project: 'SOPHIA 'Stabilization and Optimization of the PHysical layer to Improve Applications' and was partially sponsored by Alcatel-Bell. The scientific responsibility is assumed by its authors.
**Paschalis Tsiaflakis is a research assistant with the F.W.O. Vlaanderen.
* Corresponding author.
Email adresses: paschalis.tsiaflakis@esat.kuleuven.be (Paschalis Tsiaflakis), jan.vangorp@esat.kuleuven.be (Jan Vangorp), marc.moonen@esat.kuleuven.be (Marc Moonen), jan.vj.verlinden@alcatel.be (Jan Verlinden).

much more crosstalk than others (e.g. near-far scenario).

The Optimal Spectrum Balancing (OSB) algorithm [9] [10] provides a computationally tractable way to calculate optimal transmit spectra. By optimizing a weighted rate sum, this algorithm can make every possible trade off between the rates of different users. The damage done to other modems in the network is taken into account explicitly, avoiding the "selfish" optimum and thereby improving on the performance of IW. However, this can only be done when complete information about the channel is available (direct channels as well as crosstalk channels), making OSB only suitable with centralized control in a Spectrum Management Center (SMC).

OSB uses a Lagrange multiplier approach to decouple the spectrum management problem into K per-tone optimization problems, where K is the number of tones. The remaining issues are then finding the Lagrange multipliers that enforce the constraints and solving the non-convex per-tone optimization problems.

Finding the optimal Lagrange multipliers can become complex when more than two users are considered. Starting from the single-user case, this paper presents a number of properties, which are then extended to the multi-user case and lead to an efficient search algorithm for the Lagrange multipliers.

The per-tone optimization problems are non-convex with many local optima. Hence, an exhaustive search was proposed to find the global optimum. Unfortunately the set size of feasible solutions is exponential in the number of users N. This becomes computationally intractable for more than four users. To cope with this complexity, research has been focussing on reducing the complexity without sacrificing performance. In [11] [12] the authors present an iterative approach to solve the per-tone optimization problems. It is claimed that these algorithms are computationally tractable for multiple users and lead to near-optimal performance. Unfortunately, it is not possible to compare their performance with the optimal performance for scenarios with more than four users.

In this paper a branch and bound approach is presented to solve the per-tone optimization problem. Branch and bound [13] [14] is a general method for finding optimal solutions of various optimization problems. Branch and bound operations will be proposed which require a limited amount of computation, keeping the total computational complexity low. The complexity of the proposed branch and bound procedure will be compared with the complexity of the exhaustive search.

The paper is organized as follows. Section 2 describes the system model for the crosstalk environment. In section 3, different viewpoints on the spectrum management problem are discussed. It is shown that the dual problem decou-

ples the primal spectrum management problem into many per-tone problems. Finally, it is discussed how to solve the dual problem. Section 4 starts with a discussion of the single-user spectrum management problem from a dual decomposition viewpoint. Some interesting properties and relations are derived. These relations are then extended to the multi-user case. The major contribution of this section is the derivation of the Lagrange multiplier update strategy based on these relations. This approach gives a certain freedom to define an intelligent step size selection strategy leading to a fast algorithm to find the optimal Lagrange multipliers. Section 5 presents the low complexity branch and bound approach for solving the per-tone optimization problem. Finally, in section 6 simulations results are shown for the proposed OSB solution and the overall complexity reduction is evaluated.

2 System Model

Most current DSL systems use Discrete Multi-Tone (DMT) modulation. The available frequency band is divided in a number of parallel subchannels or tones. We assume that the modems in the network are synchronized. Each tone is capable of transmitting data independently from other tones, and so the transmit power and the number of bits can be assigned individually for each tone. This gives a large flexibility in optimally shaping the transmit spectrum.

Transmission for a binder of N users can be modelled on each tone k by

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{z}_k \qquad k = 1 \ldots K.$$

The vector $\mathbf{x}_k = [x_k^1, x_k^2, \ldots, x_k^N]^T$ contains the transmitted signals on tone k for all N users. $[\mathbf{H}_k]_{n,m} = h_k^{n,m}$ is an $N \times N$ matrix containing the channel transfer functions from transmitter m to receiver n. The diagonal elements are the direct channels, the off-diagonal elements are the crosstalk channels. $\mathbf{z}_k$ is the vector of additive noise on tone k, containing thermal noise, alien crosstalk, RFI,.... The vector $\mathbf{y}_k$ contains the received symbols.

We denote the transmit power as $s_k^n \triangleq \Delta_f E\{|x_k^n|^2\}$, the noise power as $\sigma_k^n \triangleq \Delta_f E\{|z_k^n|^2\}$. The vector containing the transmit power of user n on all tones is $\mathbf{s}^n \triangleq [s_1^n, s_2^n, \ldots, s_K^n]^T$. The DMT symbol rate is denoted as f_s, the tone spacing as Δ_f.

It is assumed that each modem treats interference from other modems as noise. When the number of interfering modems is large, the interference is well approximated by a Gaussian distribution. Under this assumption the achievable bit loading of user n on tone k, given the transmit spectra $\mathbf{s}_k \triangleq$

$[s_k^1, s_k^2, \ldots, s_k^N]^T$ of all modems in the system, is

$$b_k^n \triangleq \log_2\left(1 + \frac{1}{\Gamma}\frac{|h_k^{n,n}|^2 s_k^n}{\sum_{m\neq n}|h_k^{n,m}|^2 s_k^m + \sigma_k^n}\right), \tag{1}$$

where Γ denotes the SNR-gap to capacity, which is a function of the desired BER, the coding gain and noise margin. The data rate for user n is

$$R^n = f_s \sum_k b_k^n.$$

Conversely, when the bit loading $\mathbf{b}_k \triangleq [b_k^1, b_k^2, \ldots, b_k^N]^T$ of all the users is given for a specific tone k, the required transmit power for the modems in the system can be calculated by [10]

$$\mathbf{s}_k = \left(\mathbf{D}_k - \mathbf{\Lambda}_k \mathbf{A}_k\right)^{-1} \mathbf{\Lambda}_k \sigma_k \tag{2}$$

$$\begin{aligned}
\mathbf{D}_k &\triangleq \operatorname{diag}\{|h_k^{1,1}|^2, |h_k^{2,2}|^2, \ldots, |h_k^{N,N}|^2\} \\
\mathbf{\Lambda}_k &\triangleq \operatorname{diag}\{2^{b_k^1} - 1, 2^{b_k^2} - 1, \ldots, 2^{b_k^N} - 1\} \\
[\mathbf{A}_k]_{n,m} &\triangleq a_k^{n,m} \quad \text{with} \quad a_k^{n,m} \triangleq \begin{cases} 0 & n = m \\ \Gamma|h_k^{n,m}|^2 & n \neq m \end{cases} \\
\sigma_k &\triangleq \Gamma[\sigma_k^1, \sigma_k^2, \ldots, \sigma_k^N]^T
\end{aligned}$$

The total power used by user n is then

$$P^n = \sum_k s_k^n.$$

Note that, based on **Formulae (1)** and **(2)**, every $\mathbf{s}_k$ corresponds to a $\mathbf{b}_k$, and vice versa. Finding optimal transmit spectra $\mathbf{s}_k$ (referred to as "power loading") is therefore equivalent to finding optimal bit loads $\mathbf{b}_k$ (referred to as "bit loading"). The Optimal Spectrum Balancing problem (section 3) is mostly presented as a power loading problem. The branch and bound approach in section 5 will be presented as a bit loading problem.

3 Optimal Spectrum Balancing

3.1 The Spectrum Management Problem

The spectrum management problem amounts to finding optimal transmit spectra for a bundle of interfering DSL lines, following a certain criterion and subject to a number of constraints.

First of all, there is a *total power constraint* $P^{n,tot}$ for each user. This constraint ensures the user's total power does not exceed the maximum allowed total transmit power. On top of this constraint there can be a *spectral mask constraint* $s_k^{n,mask}$ for each tone to guarantee electromagnetic compatibility with other systems.

A second type of constraint is a *rate constraint* for each user. Typically service providers offer a number of profiles and guarantee a certain Quality of Service. The rate constraint then indicates a minimum data rate required by the user.

The spectrum management problem can be viewed from different angles, each time leading to a different criterion that is then to be optimized by the spectrum management algorithm. Either rate, margin or power of the users can be optimized. An optimal solution has to be found within the domain set out by the various constraints.

- In *rate adaptive mode*, the spectrum management problem is to maximize the sum of the data rates of the users. This will be done by using the available power to load a maximum number of bits on tones. The rate is thus limited by the total power and spectral mask constraints. It is possible that there are many solutions to this problem. Several trade-offs between the rates of individual users may exist, all resulting in the same total sum rate. Rate constraints can be used to select a solution with reasonable rates for all users.

$$\begin{aligned} &\max_{\mathbf{s}^1,\ldots,\mathbf{s}^N} \sum_{n=1}^{N} R^n \\ &\text{subject to } P^n \leq P^{n,tot} \qquad n = 1 \ldots N \\ &\qquad\qquad 0 \leq s_k^n \leq s_k^{n,mask} \quad n = 1 \ldots N, k = 1 \ldots K \\ &\qquad\qquad R^n \geq R^{n,target} \qquad n = 1 \ldots N \end{aligned} \tag{3}$$

  In this optimization problem, the first set of constraints indicate total power constraints per user. The second set of constraints are spectral mask constraints and the third set are rate constraints per user.
- In *power adaptive mode*, the spectrum management problem is to minimize

the total power needed by all users and still meet the rate constraints. This has to be done without violating the spectral mask constraints and total power constraints.

$$\begin{aligned} &\min_{\mathbf{s}^1,\ldots,\mathbf{s}^N} \sum_{n=1}^{N} P^n \\ &\text{subject to } P^n \leq P^{n,tot} \qquad n = 1 \ldots N \\ &\qquad\qquad 0 \leq s_k^n \leq s_k^{n,mask} \quad n = 1 \ldots N, k = 1 \ldots K \\ &\qquad\qquad R^n \geq R^{n,target} \qquad n = 1 \ldots N \end{aligned} \tag{4}$$

Total power constraints are represented in the first set of constraints, the second set of constraints are spectral mask constraints and the third set are rate constraints.

- In *margin adaptive mode*, the spectrum management algorithm will use the available power to tightly satisfy the rate constraints. In this mode, the noise margin is maximized, thus minimizing the bit error rate for the requested data rate. In single-user mode, this can be achieved by first using the spectrum management algorithm in power adaptive mode and then assigning all remaining power over the used tones by scaling with a constant factor, while not violating spectral mask constraints. In the multi-user case however, this strategy should be revised because adding power to a user causes more crosstalk to other users.

3.2 Dual Decomposition

We will now focus on the *rate adaptive* formulation of the spectrum management problem. The power adaptive form can be treated in a similar way.

The rate adaptive optimization problem (3) is a non-convex problem and therefore difficult to solve. To find the global optimum one must exhaustively search through all possible transmit spectra. This leads to an exponential complexity in both the number of users and tones, namely $\mathcal{O}(B^{NK})$ where B is the number of possibilities for the bit or power loading for each tone and each user.

With $K = 256$ in ADSL and $K = 4096$ in VDSL, exhaustively searching all possible transmit spectra is seen to be computationally intractable. The reason behind this exponential complexity in the number of tones is that the total power constraints and rate constraints are coupled across tones. Therefore transmit spectra have to be searched jointly across tones.

In [9] [10] it was shown that this complexity can be reduced by focussing on the dual problem to decouple the optimization problem across tones. By

using Lagrange multipliers to move constraints coupled over tones into the unconstrained part of the optimization problem, the spectrum management problem can be solved in a per-tone fashion. By choosing appropriate values for the Lagrange multipliers, the constraints can still be enforced.

Following the approach of [15] we first formulate the rate adaptive spectrum management problem (3) in a slightly different way. Instead of optimizing the sum rate of the users in a general fashion, while trying to satisfy individual rate constraints, we now enforce a fixed ratio between the rates for all users. Solutions to this problem are more restricted than solutions to (3) because now rates in excess of the individual rate constraints cannot be assigned randomly but are divided over the users in proportion to their rate constraints. The resulting spectrum management problem is the maximization of the so-called base rate R while satisfying total power and spectral mask constraints. This set of solutions is further restricted by a rate constraint for the users expressed as a fixed proportion β^n of the base rate. The parameters β^n are given fixed numbers and depend on the importance that we wish to put on the users with respect to each other.

$$\begin{aligned} &\max_{\mathbf{s}^1,\ldots,\mathbf{s}^N,R} R \\ &\text{subject to} \quad P^n \leq P^{n,tot} \qquad n = 1 \ldots N \\ &\qquad\qquad 0 \leq s_k^n \leq s_k^{n,mask} \quad n = 1 \ldots N, k = 1 \ldots K \\ &\qquad\qquad R^n \geq \beta^n R \qquad n = 1 \ldots N \end{aligned} \tag{5}$$

Within the framework of [16], optimization problem (5) can be referred to as the primal problem. The dual problem corresponding with this primal problem can be formulated as follows

$$\begin{aligned} &\min_{\boldsymbol{\omega},\boldsymbol{\lambda}} \quad g(\boldsymbol{\omega},\boldsymbol{\lambda}) \\ &\text{subject to } \lambda_n \geq 0, \omega_n \geq 0, \quad n = 1 \ldots N \end{aligned} \tag{6}$$

with

$$\begin{aligned} &g(\boldsymbol{\omega},\boldsymbol{\lambda}) = \max_{\mathbf{s}^1,\ldots,\mathbf{s}^N,R} J(\mathbf{s}^1,\ldots,\mathbf{s}^N,R,\boldsymbol{\omega},\boldsymbol{\lambda}) \\ &\text{subject to} \quad 0 \leq s_k^n \leq s_k^{n,mask} \quad n = 1 \ldots N, k = 1 \ldots K. \end{aligned} \tag{7}$$

where J is called the Lagrangian and can be formulated as follows

$$\begin{aligned} J &= R + \textstyle\sum_{n=1}^N \omega_n \Big(R^n - \beta^n R\Big) + \sum_{n=1}^N \lambda_n \Big(P^{n,tot} - \sum_{k=1}^K s_k^n\Big) \\ &= \Big(1 - \textstyle\sum_{n=1}^N \omega_n \beta^n\Big) R + \sum_{n=1}^N \omega_n R^n + \sum_{n=1}^N \lambda_n \Big(P^{n,tot} - \sum_{k=1}^K s_k^n\Big) \end{aligned} \tag{8}$$

If $\left(1 - \sum_n \omega_n \beta^n\right) > 0$ the maximization over R results in $R = \infty$. If $\left(1 - \sum_n \omega_n \beta^n\right) < 0$ the maximization over R results in $R = 0$. There only exists a non-trivial solution if

$$\left(1 - \sum_n \omega_n \beta^n\right) = 0. \tag{9}$$

This results in the following Lagrangian:

$$J = \sum_{n=1}^{N} \omega_n R^n + \sum_{n=1}^{N} \lambda_n \Big(P^{n,tot} - \sum_{k=1}^{K} s_k^n\Big). \tag{10}$$

This Lagrangian is decoupled across the tones:

$$\begin{aligned} J &= \textstyle\sum_{k=1}^{K} \Bigg(\underbrace{\sum_{n=1}^{N} \omega_n f_s b_k^n - \sum_{n=1}^{N} \lambda_n s_k^n}_{J_k} \Bigg) + \underbrace{\sum_{n=1}^{N} \lambda_n P^{n,tot}}_{\text{constant}} \\ &= \textstyle\sum_{k=1}^{K} J_k + \text{constant}. \end{aligned}$$

The constant has no influence on the maximization and can be discarded. Then (for a particular choice of $\lambda_n, \omega_n, n = 1 \ldots N$) the problem is reduced to a maximization of a sum across tones, which is equal to the sum of independent maximizations. This is referred to as the dual decomposition into independent per-tone optimization problems. The original complexity of $\mathcal{O}(B^{NK})$, exponential in K, is now reduced to a linear complexity in K: $\mathcal{O}(KB^N)$.

In a three-user VDSL system with 14 possible bit loadings per tone, the complexity is reduced from one maximization over $14^{3\times 4096}$ possibilities to 4096 maximizations over 14^3 possibilities. This is a spectacular reduction in complexity.

Solving the dual problem (see also section 3.3) does not necessarily correspond to solving the original constrained problem. Optimization theory [16] states that the solution to the dual problem provides only an upper bound to the solution of the original problem. The difference between the upper bound and the optimum of the primal problem is called the duality gap. In [15] it is shown that for multi-carrier systems like DMT that satisfy certain conditions, the duality gap is zero. This means the solution to the dual problem is also the solution to the original problem. In practice, it is found that, while the necessary conditions are not strictly satisfied, the dual problem formulation leads to adequate solutions, and so is currently the method of choice.

3.3 Solving the Dual Problem

In the previous section the spectrum management problem defined by the primal problem (5) was transformed into following dual problem:

$$\begin{aligned} &\min_{\boldsymbol{\omega},\boldsymbol{\lambda}} \quad g(\boldsymbol{\omega},\boldsymbol{\lambda}) \\ &\text{subject to } \lambda_n \geq 0, \omega_n \geq 0, \quad n = 1 \dots N \\ &\qquad\qquad \left(1 - \textstyle\sum_n \omega_n \beta^n\right) = 0 \end{aligned} \tag{11}$$

with

$$g(\boldsymbol{\omega},\boldsymbol{\lambda}) = \sum_{k=1}^{K} \underbrace{\max_{\mathbf{s}_k^1,\dots,\mathbf{s}_k^N} \sum_{n=1}^{N} \omega_n f_s b_k^n - \sum_{n=1}^{N} \lambda_n s_k^n}_{\text{per-tone optimization problem}} \tag{12}$$

$$\text{subject to} \quad 0 \leq s_k^n \leq s_k^{n,mask} \; n = 1 \dots N, k = 1 \dots K.$$

Basically the dual problem consists of two optimization problems. First, the optimal Lagrange multipliers ω_n and $\lambda_n (n = 1 \dots N)$ need to be find such that the total power constraints and the data rate constraints of (5) are satisfied. Secondly, for given $\omega_n, \lambda_n (n = 1 \dots N)$ K independent nonconvex per-tone optimization problems need to be solved to obtain the optimal transmit spectra for all users on all tones.

Given $\omega_n, \lambda_n (n = 1 \dots N)$ the nonconvex per-tone optimization problem can be solved by performing an exhaustive search over all possible bit or power loading combinations for the users. This results in transmit spectra for all users. For random $\omega_n{}'$s and $\lambda_n{}'$s, the power and rate constraints are generally not satisfied. By choosing appropriate values for the Lagrange multipliers, these constraints can be enforced. Looking at (12), it can be seen that the $\lambda_n{}'$s influence the resulting spectra. A larger λ_n for user n results in a larger penalty in the cost function when power is allocated to s_k^n. Therefore the $\lambda_n{}'$s can be viewed as setting a cost for power. The $\omega_n{}'$s have a similar intuitive interpretation. A larger ω_n for user n results in an increased importance attached to its rate. The larger ω_n, the higher the rate allocated to user n compared to other users.

In [10] it is explained that in order to solve the primal problem the Lagrange multipliers $\lambda_n (n = 1 \dots N)$ need to be found such that the total power constraints are tight ($P^n = P^{n,tot}$) or λ_n is equal to zero (in the case of a non-active power constraint $P^n \leq P^{n,tot}$) for each n. The $\omega_n{}'$s will be treated more loosely. The last constraint of (11) is related with the last constraint of (5) which enforces a fixed proportion β^n to the different users n. This constraint

is not so important in our application and we will mainly use the $\omega_n{}'$s to put more importance on some lines. The last constraint of (11) is thus removed keeping in mind that the $\omega_n{}'$s can be chosen to satisfy data rate constraints like in (3). Thus, the $\omega_n{}'$s can be seen as weights given to the different users.

Starting from the single-user case, section 4 derives a number of properties of the Lagrange multipliers based on simple mathematics. These properties are then extended to the multi-user case and lead to a very efficient and intuitive search algorithm for the Lagrange multipliers. Section 5 focuses on the complexity of the discrete per-tone optimization problem. As the dimensionality of the solution set is exponential in the number of users, an exhaustive search becomes computationally intractable for more than four users. A branch and bound approach will be presented to reduce this complexity significantly.

4 An Efficient Lagrange Multiplier Search Algorithm

4.1 Single-user Spectrum Management

In this section, the single-user spectrum management problem will be discussed within the framework provided by the dual decomposition method. For the single-user case, the spectrum management problem can be formulated as follows:

$$\begin{aligned} &\max_{\mathbf{s}} \quad R \\ &\text{subject to } \textstyle\sum_k s_k \le P^{tot} \\ &\qquad\qquad 0 \le s_k \le s_k^{mask} \;\; k = 1 \ldots K \end{aligned} \tag{13}$$

where now superscript '1' (for user 1) has been omitted. Using a Lagrange multiplier λ to incorporate the total power constraint, the dual problem is formulated as

$$\begin{aligned} &\max_{\mathbf{s}} \quad R + \lambda(P^{tot} - \textstyle\sum_k s_k) \\ &\text{subject to } 0 \le s_k \le s_k^{mask} \qquad k = 1 \ldots K \end{aligned} \tag{14}$$

Recalling that λ represents a cost for power, larger λ's will result in less power being used, e.g. $\lambda = \infty$ leads to $P = \sum_k s_k = 0$ and thus $R = 0$. For any other λ, the optimal spectrum can be found through a per-tone exhaustive search over all possible bit or power loadings.

This relation can be represented on the power axis shown in **Figure 1**. This axis shows the total power obtained when optimizing the dual problem for a particular λ. If for a particular λ a total power and rate (P^λ, R^λ) is obtained,

then optimality of this solution implies that a total power P less than P^λ, i.e. $P < P^\lambda$ must correspond to a rate R smaller than R^λ, i.e. $R < R^\lambda$. This then ensures that if a λ is found that makes the power constraint tight, the primal problem is solved.



Fig. 1. Power axis

Secondly, it is easily proven that if λ is decreased, then both the achieved bit rate and the consumed power for the optimal solution do not decrease, i.e

$$P^{\lambda_B} \geq P^{\lambda_A} \text{ if } \quad \lambda_B < \lambda_A$$
$$R^{\lambda_B} \geq R^{\lambda_A} \text{ if } \quad \lambda_B < \lambda_A$$

where $P^{\lambda_A}, P^{\lambda_B}$ and $R^{\lambda_A}, R^{\lambda_B}$ are the optimal total power and rate when $\lambda = \lambda_A$ and $\lambda = \lambda_B$ respectively. This can be proven as follows. Optimality of $(P^{\lambda_A}, R^{\lambda_A})$ for λ_A implies that

$$R^{\lambda_B} - \lambda_A P^{\lambda_B} \leq R^{\lambda_A} - \lambda_A P^{\lambda_A}, \tag{15}$$

because the right-hand side indeed uses the spectra that maximize the Lagrangian for λ_A. Any other spectra, e.g. those corresponding to $(P^{\lambda_B}, R^{\lambda_B})$, result in a smaller value for the Lagrangian. A similar statement can be made on the optimality of $(P^{\lambda_B}, R^{\lambda_B})$ for λ_B:

$$R^{\lambda_A} - \lambda_B P^{\lambda_A} \leq R^{\lambda_B} - \lambda_B P^{\lambda_B}. \tag{16}$$

Taking the sum of (15) and (16) results in

$$\underbrace{(\lambda_B - \lambda_A)}_{\Delta\lambda}\underbrace{(P^{\lambda_B} - P^{\lambda_A})}_{\Delta P} \leq 0. \tag{17}$$

Assuming $\lambda_B < \lambda_A$, (17) implies

$$P^{\lambda_B} \geq P^{\lambda_A}. \tag{18}$$

Then by using this in (16) we obtain

$$R^{\lambda_B} \geq R^{\lambda_A}. \tag{19}$$

Relation (17) can be used to construct a simple procedure to find the λ that makes the total power constraint tight. By starting with a random λ, e.g.

$\lambda = 0$, we maximize the Lagrangian to obtain a bit and power loading. If the total power then exceeds the total power constraint, λ has to increase ($\Delta\lambda > 0$ so that $\Delta \mathrm{P} \leq 0$). When the total power is below the total power constraint, λ has to decrease ($\Delta\lambda < 0$ so that $\Delta \mathrm{P} \geq 0$). Note that λ has to remain positive. If by decreasing λ we end up with $\lambda = 0$, the total power constraint is effectively made inactive by some other per-tone power constraint, e.g. spectral mask constraints. An update formula for λ is then as follows

$$\Delta\lambda = \mu\Big(\sum_k s_k - P^{tot}\Big) \quad \Rightarrow \quad \lambda^{t+1} = \left[\lambda^t + \mu\Big(\sum_k s_k - P^{tot}\Big)\right]^+, \tag{20}$$

where $[x]^+$ means $\max(0, x)$. By varying the step size μ according to **Algorithm 1** the λ can be found which makes the power constraint tight. This procedure is illustrated in **Figure 2**. Starting from some initial λ (e.g. $\lambda = 0$), μ is always doubled, creating a trajectory of points on the power axis towards the target power. If for a new point the distance to the target is larger then the distance of the currently best known point, a new trajectory is started by reinitializing (20) with the λ and $\sum_k s_k$ of the best known point.

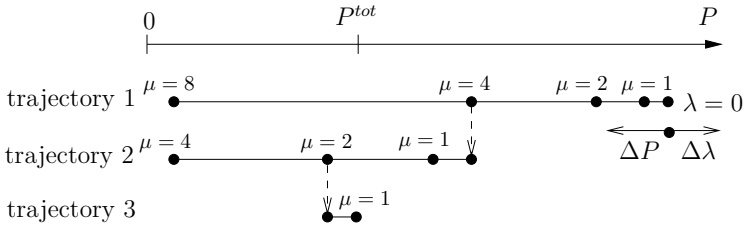

Fig. 2. Single-user search procedure

Algorithm 1 Single-user λ search algorithm

```
while distance > tolerance do
  λ = best λ so far
  μ = 1
  while distance ≤ previousDistance do
    previousDistance = distance
    μ = μ × 2
    Δλ = −μ(P^tot − P^λ)
    [P^{λ+Δλ}, s_k] = calculateLoading([λ + Δλ]^+)   (per-tone exhaustive search)
    distance = |P^tot − P^{λ+Δλ}|
  end while
end while
```

Formula (20) also corresponds to the subgradient approach of [15], be it that it was derived here based on a simple problem-specific property (17). This in particular allows for alternative stepsize selection strategies, such as **Algorithm 1**. Using the theory provided by [15], a formula similar to (17) can also be derived. One can plot the maximum achievable rate as a function of the total power as in **Figure 3**. This plot can be obtained by performing the maximization of the Lagrangian $\max_{\mathbf{s}} R - \lambda P$ for all possible λ's. For each λ this results in an optimal total power usage $P = P^{opt}$ and a corresponding optimal rate $R = R^{opt}$. It is argued in [15] that this function is concave, based on the assumption that the number of tones is *infinite*. Solving $\max_{\mathbf{s}} R - \lambda P$ for some given λ corresponds to finding the point on the curve where the difference between R^{λ} and λP^{λ} is largest. From **Figure 3** it can be seen that this corresponds to finding the point on the curve where the tangent is equal to λ. Hence for the optimal solution

$$\frac{\partial R(P = P^{\lambda})}{\partial P} = \lambda. \tag{21}$$

Mathematically, if the maximum achievable rate as a function of the optimal power budget is twice differentiable, the concavity of this curve can be expressed as

$$\frac{\partial^2 R}{\partial^2 P} \leq 0. \tag{22}$$

Based on (21) and (22), for a *small* ΔP the same relation as (17) is then derived as follows:

$$\frac{\partial^2 R}{\partial^2 P} \leq 0 \Rightarrow \frac{\partial}{\partial P}\Big(\underbrace{\frac{\partial R}{\partial P}}_{\lambda}\Big) \leq 0 \Rightarrow \frac{\partial \lambda}{\partial P} \leq 0 \overset{\Delta P \Delta P \geq 0}{\Rightarrow} \Big(\underbrace{\frac{\partial \lambda}{\partial P}\Delta P}_{\Delta\lambda}\Big)\Delta P \leq 0.$$

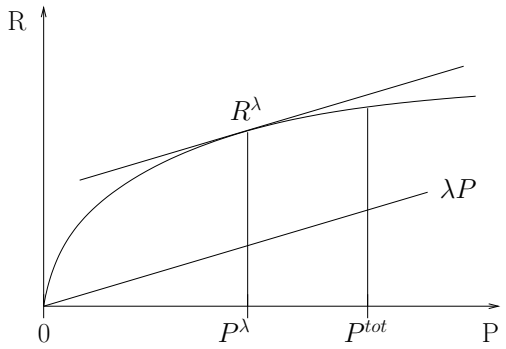

Fig. 3. Optimal rate vs power

Finally, note that the λ of the dual problem formulation (14) without the spectral mask constraint is related to the well-known waterfilling solution [17] in the continuous bit loading case. The waterfilling solution is obtained by

differentiating $R + \lambda(P^{tot} - \sum_k s_k)$ with respect to s_k. This leads to

$$s_k + \frac{\Gamma \sigma_k}{|h_k|^2} = \frac{1}{\lambda \ln(2)}.$$

The constant at the right-hand side is the waterfilling level, which is seen to be inversely related to the Lagrange multiplier. A higher waterfilling level corresponds to more power allocated, which can be obtained by setting a small λ.

4.2 Multi-user Spectrum Management

In the previous section, interesting relations (17), (18) and (19) were derived between λ, the total power and the data rate. These relations gave rise to a simple procedure for searching the λ and leading to a total power tightly satisfying the total power constraint. In this section these relations will be investigated for the multi-user case. We will focus on the two-user case first, and then extend the results to the general N-user case.

For the N-user case, the spectrum management problem can be formulated as in **Formula (3)**. **Formula (11)** and (12) specified the per-tone dual formulation of the generic spectrum management problem. For a two-user case the Lagrange dual function can be formulated as

$$\begin{aligned} &\text{argmax}_{\mathbf{s}^1,\mathbf{s}^2}\, \omega_1 R_1 + \omega_2 R_2 + \lambda_1\Big(P^{1,tot} - \textstyle\sum_k s_k^1\Big) + \lambda_2\Big(P^{2,tot} - \textstyle\sum_k s_k^2\Big) \\ &\text{subject to} \quad 0 \le s_k^n \le s_k^{mask} \qquad n = 1,2; k = 1 \ldots K \end{aligned} \tag{23}$$

Given $\boldsymbol{\omega} = (\omega_1, \omega_2)$ and $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$, this Lagrange dual function, decoupled over the tones, can be solved by performing an exhaustive search for each tone over all possible bit or power loading combinations for the users.

$$\text{for } k = 1 \ldots K, \qquad s_k^{1,opt}, s_k^{2,opt} = \underset{s_k^1, s_k^2}{\text{argmax}} \sum_{n=1}^{2} \omega_n f_s b_k^n - \sum_{n=1}^{2} \lambda_n s_k^n \tag{24}$$

$$\text{subject to } 0 \le s_k^n \le s_k^{n,mask} \qquad n = 1,2$$

The optimal solution is then a bit and power loading corresponding to total powers and data rates

$$\Big(P^{1,\boldsymbol{\omega},\boldsymbol{\lambda}}, R^{1,\boldsymbol{\omega},\boldsymbol{\lambda}}, P^{2,\boldsymbol{\omega},\boldsymbol{\lambda}}, R^{2,\boldsymbol{\omega},\boldsymbol{\lambda}}\Big) \tag{25}$$

The optimality of this solution implies that for this $\boldsymbol{\lambda}$ and $\boldsymbol{\omega}$ there exists no other bit or power loading giving a larger value to the Lagrangian. This then implies that for a weighted total power budget $\lambda_1 P^1 + \lambda_2 P^2$ smaller than $P^{\boldsymbol{\omega},\boldsymbol{\lambda}} \triangleq \lambda_1 P^{1,\boldsymbol{\omega},\boldsymbol{\lambda}} + \lambda_2 P^{2,\boldsymbol{\omega},\boldsymbol{\lambda}}$, it is impossible to achieve a weighted rate sum (with weights ω_1, ω_2) that is larger than $R^{\boldsymbol{\omega},\boldsymbol{\lambda}} \triangleq \omega_1 R^{1,\boldsymbol{\omega},\boldsymbol{\lambda}} + \omega_2 R^{2,\boldsymbol{\omega},\boldsymbol{\lambda}}$.

This is shown graphically in the power plane of **Figure 4** (which is the two-user version of **Figure 1**). $\left(P^{1,\boldsymbol{\omega},\boldsymbol{\lambda}}, P^{2,\boldsymbol{\omega},\boldsymbol{\lambda}}\right)$ is the optimal power allocation for a given $\boldsymbol{\lambda}$ and $\boldsymbol{\omega}$. Every loading corresponding to total powers in the marked triangle then has a smaller weighted rate sum.

If a $\boldsymbol{\lambda}$ is found such that $\left(P^{1,\boldsymbol{\omega},\boldsymbol{\lambda}} = P^{1,tot}, P^{2,\boldsymbol{\omega},\boldsymbol{\lambda}} = P^{2,tot}\right)$, this solution satisfies the total power constraints and so has a weighted rate sum larger than every other possible loading in the marked rectangle (i.e. a subset of the marked triangle). Thus the primal problem is solved once the rate constraints are satisfied. If one of the λ_n is zero and the corresponding total power is smaller than the total power constraint, this means that the total power constraint is not active for user n and we should not further try to make this total power tight. This case is handled later in **Algorithm 2**

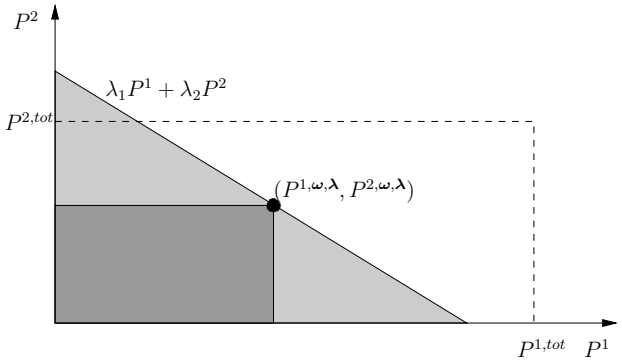

Fig. 4. Two-user power plane

So our aim is to tune $\boldsymbol{\lambda}$ and $\boldsymbol{\omega}$ such that the total power constraints and rate constraints are satisfied: $P^{1,\boldsymbol{\omega},\boldsymbol{\lambda}} = P^{1,tot}$, $P^{2,\boldsymbol{\omega},\boldsymbol{\lambda}} = P^{2,tot}$, $R^{1,\boldsymbol{\omega},\boldsymbol{\lambda}} \geq R^{1,target}$ and $R^{2,\boldsymbol{\omega},\boldsymbol{\lambda}} \geq R^{2,target}$. The same strategy as in the single-user case can now be followed to first derive a generalization of **Formula (17)**, and then an update strategy for $\boldsymbol{\lambda}$ and $\boldsymbol{\omega}$.

Starting from two optimal solutions $(R^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A}, P^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A}, R^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A}, P^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A})$ and $(R^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B}, P^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B}, R^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B}, P^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B})$ corresponding to $(\boldsymbol{\omega}_A, \boldsymbol{\lambda}_A)$ and $(\boldsymbol{\omega}_B, \boldsymbol{\lambda}_B)$ respectively, optimality for $(\boldsymbol{\omega}_A, \boldsymbol{\lambda}_A)$ implies

$$\begin{aligned} &\omega_{1,A} R^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} + \omega_{2,A} R^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} - \lambda_{1,A} P^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} - \lambda_{2,A} P^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} \\ &\leq \omega_{1,A} R^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A} + \omega_{2,A} R^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A} - \lambda_{1,A} P^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A} - \lambda_{2,A} P^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A} \end{aligned} \tag{26}$$

Optimality for $(\boldsymbol{\omega}_B, \boldsymbol{\lambda}_B)$ implies

$$\begin{aligned} &\omega_{1,B} R^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A} + \omega_{2,B} R^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A} - \lambda_{1,B} P^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A} - \lambda_{2,B} P^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A} \\ &\leq \omega_{1,B} R^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} + \omega_{2,B} R^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} - \lambda_{1,B} P^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} - \lambda_{2,B} P^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} \end{aligned} \tag{27}$$

Taking the sum of (26) and (27) results in

$$\begin{aligned} &-\underbrace{\left(\omega_{1,B} - \omega_{1,A}\right)}_{\Delta\omega_1} \underbrace{\left(R^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} - R^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A}\right)}_{\Delta R^1} - \underbrace{\left(\omega_{2,B} - \omega_{2,A}\right)}_{\Delta\omega_2} \underbrace{\left(R^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} - R^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A}\right)}_{\Delta R^2} \\ &+\underbrace{\left(\lambda_{1,B} - \lambda_{1,A}\right)}_{\Delta\lambda_1} \underbrace{\left(P^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} - P^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A}\right)}_{\Delta P^1} + \underbrace{\left(\lambda_{2,B} - \lambda_{2,A}\right)}_{\Delta\lambda_2} \underbrace{\left(P^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} - P^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A}\right)}_{\Delta P^2} \leq 0 \end{aligned} \tag{28}$$

Relation (28) for two users can be extended straightforward to the N-user case:

$$\begin{bmatrix} -(\Delta\boldsymbol{\omega})^T & (\Delta\boldsymbol{\lambda})^T \end{bmatrix} \begin{bmatrix} \Delta\mathbf{R} \\ \Delta\mathbf{P} \end{bmatrix} \leq 0. \tag{29}$$

where $\boldsymbol{\Delta\lambda} = [\Delta\lambda_1, \ldots, \Delta\lambda_N]^T$ and $\boldsymbol{\Delta\omega} = [\Delta\omega_1, \ldots, \Delta\omega_N]^T$ are vectors containing the $\Delta\lambda$'s and $\Delta\omega$'s for the N users, $\boldsymbol{\Delta P} = [\Delta P^1, \ldots, \Delta P^N]^T$ and $\boldsymbol{\Delta R} = [\Delta R^1, \ldots, \Delta R^N]^T$ are vectors with the corresponding total powers and data rates, Δx^n is equal to $x^{n,B} - x^{n,A}$.

Two special cases can be derived from **Formula (29)**.

$$\text{fixed } \boldsymbol{\omega} \quad (\Delta\boldsymbol{\omega} = 0) \Rightarrow (\Delta\boldsymbol{\lambda})^T \Delta\mathbf{P} \leq 0 \tag{30}$$

$$\text{fixed } \boldsymbol{\lambda} \quad (\Delta\boldsymbol{\lambda} = 0) \Rightarrow (\Delta\boldsymbol{\omega})^T \Delta\mathbf{R} \geq 0 \tag{31}$$

Note that (17) is the single-user version of (30).

Relation (29) can be used to construct a simple procedure to find the $\boldsymbol{\omega}$ and $\boldsymbol{\lambda}$ that enforce the constraints. To simplify the graphical illustration of the procedure, we limit ourselves to a procedure that only updates the $\boldsymbol{\lambda}$ Lagrange multipliers. However, this procedure can be straightforwardly extended to also include the update of the $\boldsymbol{\omega}$ Lagrange multipliers by extending the vectors as in **Formula (29)**.

In the two-user power plane, **Formula (30)** can be represented graphically as two vectors with a non-positive inner product, as in **Figure 5**. When searching for the Lagrange multipliers that make the total power constraint tight, we need to make changes to the $\boldsymbol{\lambda}$ such that in the power plane we end up in the point where every user is at maximum power. Because of the non-positive

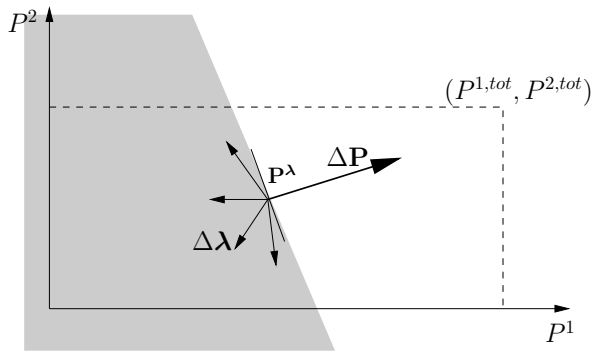

Fig. 5. Two-user power plane

inner product of $\Delta\boldsymbol{\lambda}$ and $\Delta\mathbf{P}$, $\Delta\boldsymbol{\lambda}$ for the desired $\Delta\mathbf{P}$ must be somewhere in the gray half plane opposite to the $\Delta\mathbf{P}$ vector.

This relation between $\Delta\boldsymbol{\lambda}$ and $\Delta\mathbf{P}$ can be used to steer the used power towards the total power constraint. By changing the current $\boldsymbol{\lambda}$ vector by a $\Delta\boldsymbol{\lambda}$ in the opposite direction of the desired $\Delta\mathbf{P}$, **Formula (30)** guarantees that the step taken with $\Delta\mathbf{P}$ will get the used power closer to the power constraint $(P^{1,tot}, P^{2,tot})$, as long as $\Delta\boldsymbol{\lambda}$ is not too large. This is shown in **Figure 6**, where the $\Delta\boldsymbol{\lambda}$ brings $\mathbf{P^\lambda}$ to the next point inside the shaded circle.

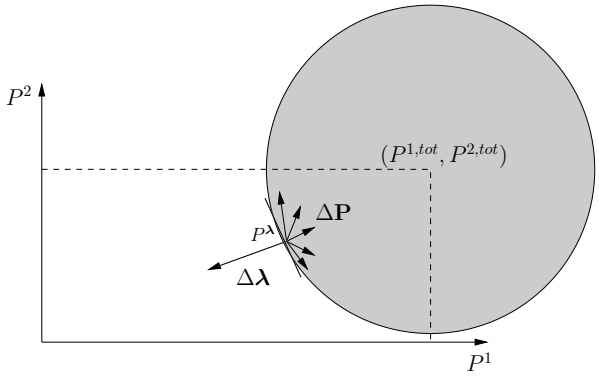

Fig. 6. Two-user power plane

Mathematically this procedure can be captured in the following update formula:

$$\Delta\boldsymbol{\lambda} = -\mu\Big(\mathbf{P}^{tot} - \mathbf{P}^{\boldsymbol{\lambda}}\Big) \quad \Rightarrow \quad \boldsymbol{\lambda}^{t+1} = \left[\boldsymbol{\lambda}^t - \mu\Big(\mathbf{P}^{tot} - \sum_k \mathbf{s}_k\Big)\right]^+ \tag{32}$$

which also applies to the N-user case.

By starting with a small μ, e.g. $\mu = 1$, the used power makes a small step closer to the desired total power. As long as the used power keeps getting closer to the desired total power, μ can be increased, e.g. doubled. A trajectory of points is then followed, each point with a total power closer to the power constraint.

At some point, a $\Delta\boldsymbol{\lambda}$ will be selected taking the used power further from the desired total power than the previous point found along the trajectory. Then this last step has to be discarded and a new trajectory is started using a new direction $\left(\mathbf{P}^{tot} - \mathbf{P}^{\boldsymbol{\lambda}}\right)$. This procedure is formally represented in **Algorithm 2**. The outer loop of this algorithm iterates over the trajectories while the inner loop follows one of the trajectories. A possible evolution of the total power using this strategy is shown in **Figure 7** for the two-user case.

Remark that if some of the used powers are smaller than the total power constraints ($P^n < P^{n,tot}$) and the corresponding Lagrange multipliers λ_n are zero, the $[\lambda_n + \Delta\lambda_n]^+$ operation keeps these Lagrange multipliers λ_n to zero. The algorithm will so converge to the case that either λ_n becomes zero or $P^n = P^{n,tot}$ for each n. In this case, the Lagrangian objective is identical to the original primal objective, thus solving the original primal problem [10].

Formula (32) again corresponds to the subgradient approach of [15], be it that it was derived here based on a simple problem-specific property (30). This in particular allows for alternative stepsize selection strategies, such as **Algorithm 2**.

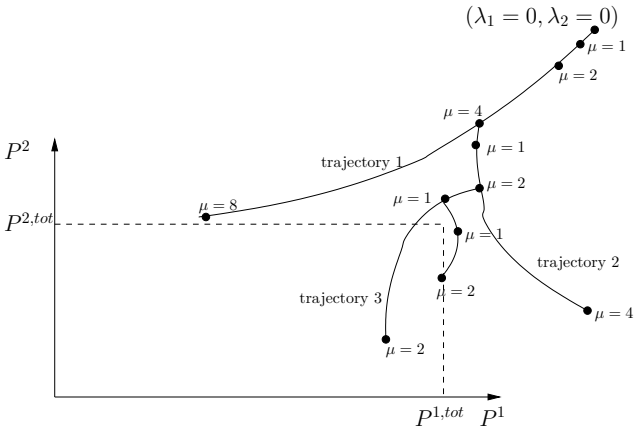

Fig. 7. Trajectories of total power in the power plane

In the framework of [15], a relation similar to (30) results from the concavity of the function that gives the optimal weighted rate $R = \omega_1 R^1 + \omega_2 R^2$ (two-user case) in function of the used total power $(\mathrm{P}^1, \mathrm{P}^2)$. Because of this concavity, for *any* x and y the following relation must hold:

$$\begin{bmatrix} x & y \end{bmatrix} \underbrace{\begin{bmatrix} \frac{\partial^2 R}{\partial^2 P^1} & \frac{\partial^2 R}{\partial P^1 \partial P^2} \\ \frac{\partial^2 R}{\partial P^2 \partial P^1} & \frac{\partial^2 R}{\partial^2 P^2} \end{bmatrix}}_{\begin{bmatrix} \frac{\partial \lambda_1}{\partial P^1} & \frac{\partial \lambda_2}{\partial P^1} \\ \frac{\partial \lambda_1}{\partial P^2} & \frac{\partial \lambda_2}{\partial P^2} \end{bmatrix}} \begin{bmatrix} x \\ y \end{bmatrix} \leq 0$$

Algorithm 2 Multi-user $\boldsymbol{\lambda}$ search algorithm

```
while distance > tolerance do
  λ = best λ so far
  μ = 1
  while distance ≤ previousDistance do
    previousDistance = distance
    μ = μ × 2
    Δλ = −μ(P^tot − P^λ)
    [P^{λ+Δλ}, s^n] = calculateLoading([λ + Δλ]^+)   (per-tone exhaustive search)
    distance = ‖P^tot − P^{λ+Δλ}‖
  end while
end while
```

By choosing $x = \Delta P^1$ and $y = \Delta P^2$, for small ΔP^1, ΔP^2 relation (30) is obtained.

5 A Low Complexity Branch and Bound Approach to the Per-Tone Optimization Problem

The per-tone optimization problem shown in (12) is non-convex with many local optima. In section 3 it was mentioned that an exhaustive search was proposed [10] to find the global optimum. Unfortunately the set size of feasible solutions is exponential in the number of users N. This becomes intractable for more than four users. In this section a branch and bound procedure is presented to solve the per-tone optimization problem. Note that the focus will be on the bit loading instead of the power loading. Every bit loading corresponds to a particular power loading, as can be seen from **Formula (2)**. The variables of the optimization problems of the previous sections can be changed from powers into bits leading to the following equivalent per-tone optimization problem:

$$\mathbf{b}_k^{opt} = \underset{\mathbf{b}_k}{\operatorname{argmax}} \; \underbrace{\sum_{n=1}^{N} \omega_n f_s b_k^n}_{1} - \underbrace{\sum_{n=1}^{N} \lambda_n s_k^n}_{2} \tag{33}$$

$$\text{subject to } 0 \leq s_k^n \leq s_k^{n,mask} \; n = 1 \ldots N.$$

Branch and bound [13] [14] is a general method for finding optimal solutions for various optimization problems. A branch and bound procedure has two ingredients. The first ingredient is a smart way of covering the feasible region by several smaller subregions. This leads to a branching operation. The second

ingredient is bounding, which is a way of finding upper and lower bounds for the optimal solution within a feasible subregion. The core of the approach is the simple observation that (for a maximization task) if the upper bound of a subregion A is smaller than the lower bound for some other subregion B, then subregion A may be safely discarded from the search. The efficiency of the method depends critically on the effectiveness of the used branching and bounding operations. There is no universal bounding operation that works for all problems. In the remainder of this section a branch and bound procedure will be presented particularly for the per-tone optimization problem (33).

The procedure starts from an **initial solution set** containing all feasible solutions. A possible initial solution set is based on a maximum bit loading of every user as follows

$$B = \Big\{ \mathbf{b}_k \in \mathcal{R}^N, \forall n : 0 \leq b_k^n \leq b_k^{n,MAX} \Big\}, \tag{34}$$

where $b_k^{n,MAX}$ is the maximum bit loading for user n on tone k. Based on the fact that the power loading increases with increasing bit loading (see Appendix) a tighter initial solution set can be defined by (35)

$$B = \Big\{ \mathbf{b}_k \in \mathcal{R}^N, \forall n : 0 \leq b_k^n \leq \min(\log_2\Big(1 + \frac{1}{\Gamma}\frac{|h_k^{n,n}|^2 s_k^{n,mask}}{\sigma_k^n}\Big), b_k^{n,MAX}) \Big\}. \tag{35}$$

The **branching operation** splits every region in 2^N rectangular subregions. This is done by splitting each dimension (bit loading range for 1 particular user) of the considered region in two. One has several options for splitting a dimension: two equal parts, two parts delimited by integer bit loadings, two parts delimited by fractional bit loadings,.... In the case of integer bit loading the branching can be stopped when all the regions have dimensions smaller than or equal to 1 bit. **Figure 8** depicts the branching operation starting from an initial solution set for a three-user case. It can be remarked that this "rectangular" branching admits a region (e.g. region I) to be fully characterized by its two extreme points $\mathbf{b}_k^{min}$ and $\mathbf{b}_k^{max}$ as shown in **Figure 8** (with $b_k^{n,min} \leq b_k^n \leq b_k^{n,max}$ for all n, in the considered region).

After the branching operation, a number of the generated subregions can be discarded from the search because they consist only of power loading combinations that do not satisfy the spectral mask constraints of (33). In the Appendix it is proven that the point $\mathbf{b}_k^{min}$ of **Figure 8** corresponds to a power loading that is smaller than every other power loading in the considered subregion I. If this point $\mathbf{b}_k^{min}$ does not satisfy the spectral mask constraints than no point of the considered subregion satisfies the spectral mask constraints and so the subregion can be discarded from the search. **Formula (2)** can be used to check the spectra corresponding to $\mathbf{b}_k^{min}$.

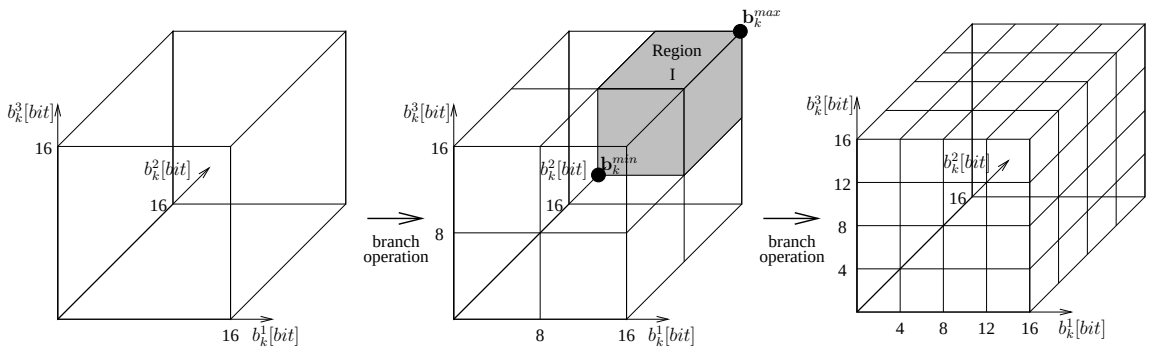

Fig. 8. Branching operation for a three-user case starting from an initial solution set based on the maximum bit loading with $b_k^{n,MAX} = 16$. Each axis corresponds to the bit loading of a particular user for tone k. The middle figure emphasizes the fact that a region can be fully characterized by its two points $\mathbf{b}_k^{min}$ and $\mathbf{b}_k^{max}$.

The **first bound operation** is the calculation of a **lower bound** for the maximum of a subregion. In fact, the cost function value of any bit loading combination of the considered subregion can be used as a lower bound. As the spectra corresponding to $\mathbf{b}_k^{min}$ have already been calculated to check the spectral mask constraints, the cost function value of this point can be used as a lower bound. The calculation of the lower bound (36) then corresponds to a simple sum of products.

$$\text{lower bound}_{region} = \sum_{n=1}^{N} \omega_n f_s b_k^{n,min} - \sum_{n=1}^{N} \lambda_n s_k^{n,min} \tag{36}$$

The **second bound operation** is the calculation of an **upper bound** for the maximum of a subregion. **Formula (33)** consists of two terms. Term 1 increases with increasing bit loading. Term 2 decreases with increasing bit loading because the power loading increases with increasing bit loading (see Appendix). An exact upper bound of term 1 is $\sum_{n=1}^{N} \omega_n f_s b_k^{n,max}$. An exact upper bound of term two is $-\sum_{n=1}^{N} \lambda_n s_k^{n,min}$. The sum of the upper bounds of term 1 and term 2 is an upper bound for the sum of the two terms.

$$\text{upper bound}_{region} = \sum_{n=1}^{N} \omega_n f_s b_k^{n,max} - \sum_{n=1}^{N} \lambda_n s_k^{n,min}. \tag{37}$$

Note that the spectra $s^{n,min}, \forall n$ have already been calculated. The calculation of the upper bound (37) then again corresponds to a simple sum of products. As can be seen, the bound operations represent a small computational complexity. The effectiveness of these bounds in discarding regions is checked by simulations. This is presented in the simulations section.

The complete branch and bound procedure for one tone of optimization prob-

lem (33) is presented in **Algorithm 3**.

Algorithm 3 Branch and bound procedure for the per-tone optimization problem

1: Start from an initial solution set containing all feasible solutions as defined by (34) or (35)
2: Branching: split the region(s) into subregions as shown in **Figure 8**
3: Discard the subregions which do not satisfy the spectral mask constraints
4: Bounding: determine lower and upper bounds for the maximum of each subregion as defined by (36) and (37)
5: Determine maximum lower bound MAX_low of all subregions.
6: Remove all subregions with an upper bound smaller than MAX_low.
7: **if** desired accuracy achieved **then** go to step 9
8: **else** go to step 2
9: Exhaustive search on all remaining subregions

The general spectrum management problem takes total power constraints into account. Suffice it to say that in some cases, the spectral mask constraints are so constraining that the total power constraints are always met, namely when $\sum_k s_k^{n,mask} \leq P^{n,tot}, \forall n$. In this case (33) reduces to term 1 and the calculation of the lower bound and upper bound is simplified to:

$$\text{lower bound}_{region} = \sum_{n=1}^{N} \omega_n f_s b_k^{n,min} \tag{38}$$

$$\text{upper bound}_{region} = \sum_{n=1}^{N} \omega_n f_s b_k^{n,max}. \tag{39}$$

Computational Complexity Analysis

The overall computational complexity of the proposed branch and bound procedure is determined by the number of subregions to be considered and the computational complexity required for each considered subregion.
For general branch and bound procedures the bound operations for a particular subregion require to solve an optimization problem (e.g. linear optimization problem, convex relaxation optimization problem, ...). In the proposed branch and bound procedure the computational complexity required for each subregion is reduced to a minimum. The lower bound computation (36) requires one function evaluation of (33). The main part of this computation is calculating the spectra corresponding with $b_k^{n,min}$ (2). The multiplications and summations can be neglected with respect to this computation. The upper bound computation (37) can reuse this part of the computation of the lower bound (33) and requires a few extra multiplications and summations which can be neglected in terms of computational complexity. So the computational complexity for each subregion is approximately one function evaluation of (33).

Because of this reason the proposed branch and bound scheme is called a low complexity branch and bound scheme.
The number of subregions to be considered is determined by the branching operation and the number of subregions that are discarded. In the worst-case none of the subregions will be discarded, which is very unlikely. In this case the branching operations will generate B^N subregions. This means that in the worst-case when no subregions are discarded the computational complexity consists of calculating B^N function evaluations of (33) which is the same as the computational complexity of the exhaustive search. Typically a lot of subregions get discarded based on the bounds and the spectral mask constraints leading to a much more efficient scheme than the exhaustive search.

Speed Up by Use of Correlation

The effectiveness of the presented branch and bound procedure depends on how fast regions are discarded. As already explained a region is discarded from the search when this region has an upper bound smaller than the lower bound of another region. The presence of a region with a large lower bound thus speeds up the convergence. Typically there is a lot of correlation between neighbouring tones , i.e. the optimal bit or power loading of tone k does not change a lot from the optimal bit or power loading of tone $k+1$. When the optimal bit or power loading of tone k is known, the subregion containing this optimal solution for tone k can be used in the branch and bound procedure for tone $k+1$ as a good approximation of the optimum solution for tone $k+1$, having a large lower bound. Because of the presence of this subregion with a large lower bound more subregions will be discarded in a early stage leading to a much smaller set of subregions to consider. This use of tone correlation speeds up the algorithm significantly.
A similar observation is that the optimal bit or power loading of tone k does not change much whenever the Lagrange multipliers λ_n are updated (towards the values that enforce the total power constraints). A similar use of correlation over Lagrange multiplier updates for tone k can then be used to speed up the algorithm.

6 Simulation Results

In section 5 we presented a low complexity branch and bound procedure for solving the non-convex per-tone optimization problem. Together with the Lagrange search algorithm presented in section 4, this composes a low complexity solution for Optimal Spectrum Balancing. In this section we will evaluate the complexity reduction of the proposed solution with respect to the OSB algorithm using an exhaustive search [10].
The ADSL downstream scenarios use a line diameter of 0.5 mm (24 AWG).

The maximum transmit power is 20.4 dBm [18]. The SNR gap Γ is 12.9 dB, corresponding to a coding gain of 3 dB, a noise margin of 6 dB and a target symbol error probability of 10^{-7}. The maximum bit loading is 16 bits. The tone spacing Δ_f is 4.3125 kHz and the DMT symbol rate f_s is 4 kHz. The simulations are performed in Matlab on a dual Opteron 250 with 4 GB RAM and a 2.4 GHz processor.

6.1 Symmetric Scenario

The first scenario is a symmetric scenario and is shown in **Figure 9**. A slightly modified version of **Algorithm 2** is used for finding the optimal Lagrange multipliers that enforce the constraints. The intuitive interpretation of **Algorithm 2** allows for the number of $\boldsymbol{\lambda}$-evaluations to be further reduced. By starting a trajectory with step size $\mu = 1$, a number of $\boldsymbol{\lambda}$-evaluations is wasted to increase the μ to a magnitude for which the total power starts converging towards the total power constraint. Instead, one could start a trajectory with a μ inspired by the best μ of the previous trajectory, avoiding unnecessary $\boldsymbol{\lambda}$-evaluations.
Table 1 compares the computational complexity of our branch and bound OSB solution, starting with initial set (35), with the computational complexity of the exhaustive search OSB solution. The simulations are done for a two- up to a seven-user scenario. The spectral mask constraints are set to -30dBm/Hz.

The computational complexity of both the branch and bound procedure and the exhaustive search procedure consists of calculating expression (33) for given Lagrange multipliers λ_n, ω_n and for a particular bit loading. The main part of this computation corresponds to calculating the spectra for a particular bit loading, as defined by **Formula (2)**. The 'Users' column indicates the number of users of the scenario. The 'Nb of Traj' and 'Nb of $\boldsymbol{\lambda}$-ev' columns indicate the number of trajectories and the number of $\boldsymbol{\lambda}$-evaluations respectively that the Lagrange Search algorithm has to perform to find the values that enforce the constraints. The 'Nb ev of (2)' indicates the total number of evaluations of (2) needed to converge and is a good measure of the complexity of the algorithm. For a two-user case, the total number of evaluations for the exhaustive search is calculated as follows. The maximum number of allocated bits is 16. This means that every user can be assigned 17 possible bit loadings. For two-users there are $289 (= 17^2)$ possible bit loadings. ADSL downstream uses 224 tones. There are 43 λ-evaluations needed to converge. This makes a total of 43 x 224 x $17^2 (= 2783648)$ possibilities that have to be calculated to obtain the optimal bit loading. The ' % ev of (2)' column indicates the percentage of evaluations of expression (2) with respect to the number of evaluations of expression (2) by the exhaustive search. The 'C.r.f.'

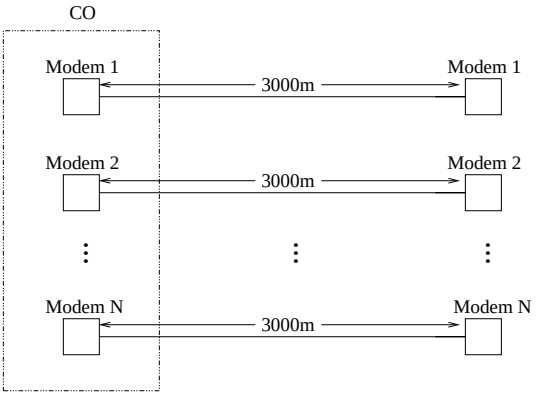

Fig. 9. Symmetric N-user ADSL scenario

column gives the complexity reduction factor with respect to the exhaustive search. This is the inverse of the previous column and is a good indication of how much less computation is needed to converge to the optimal solution.

As both procedures use the same Lagrange search, the number of $\boldsymbol{\lambda}$-evaluations and trajectories is equal. It can be seen that this number is smaller than 50, independent of the number of users. This is much faster than the bisection method proposed in [10] where the number of $\boldsymbol{\lambda}$-evaluations is exponential in the number of users. Furthermore it can be seen that the complexity reduction factor increases as the number of users increases. For seven users there is a complexity reduction of 368. Using the exhaustive search, it is impossible to simulate scenarios with more than four users on the standard platform used here. With the proposed branch and bound procedure one can simulate up to *seven-user* scenarios on the same platform within a reasonable time. Typical simulation times are 3 hours for a five-user scenario, 1 day for a six-user scenario and 1 week for a seven-user scenario. The optimal power spectra or bit loads that are obtained can provide insights needed for the design of more practical suboptimal loading schemes.

Table 2 shows the results for the spectrum management problem without total power constraints and with spectral mask constraints set to -40dBm/Hz. For the branch and bound procedure, **Algorithm 3** changes in step 4 using **Formula (38)** and **(39)** instead of **Formula (36)** and **(37)**. The computational complexity of the exhaustive search procedure also reduces as the second term of (11) is removed. This can be viewed as the evaluation of the (36) and (37) with $\boldsymbol{\lambda}$ equal to zero. It can be seen that the complexity reduction factors are now even larger than those of **Table 1**. For an *eight-user* scenario there is a complexity reduction of 18518. One very interesting remark is that for all the solutions of **Table 2** the resulting total powers for every user satisfy the total power constraints. In other words, the spectral mask constraints are so constraining that the total power constraints are not active.

Table 1
Computational complexity comparison of branch and bound procedure and exhaustive search procedure for symmetric ADSL Downstream scenario with power constraints and spectral mask constraints

			Branch and Bound procedure			Exhaustive Search procedure		
Users	Nb of Traj	Nb of λ-ev	Nb ev of (2)	% ev of (2)	C.r.f.	Nb ev of (2)	% ev of (2)	C.r.f.
2	11	43	367769	13.2	7.6	2783648	100	1
3	6	30	1840265	5.57	17.9	33015360	100	1
4	9	36	16620093	2.47	40.48	673513344	100	1
5	9	36	120037408	1.053	94.97	11.4e9	100	1
6	14	47	1.49e9	0.586	171	2.54e11	100	1
7	13	44	1.0993e+10	0.272	368	4.044e12	100	1

Table 2
Computational complexity comparison of branch and bound procedure and exhaustive search procedure for symmetric ADSL Downstream scenario without power constraints and with spectral mask constraints

			Branch and Bound procedure			Exhaustive Search procedure		
Users	Nb of Traj	Nb of λ-ev	Nb ev of (2)	% ev of (2)	C.r.f.	Nb ev of (2)	% ev of (2)	C.r.f.
2	1	1	2638	4.07	25	64736	100	1
3	1	1	8409	0.764	131	1100512	100	1
4	1	1	44929	0.236	423	19e6	100	1
5	1	1	258480	0.081	1231	318e6	100	1
6	1	1	1673325	0.031	3226	5.4e9	100	1
7	1	1	11989628	0.013	7673	92e9	100	1
8	1	1	84312076	0.0054	18518	1.56e12	100	1

6.2 Asymmetric Scenario

The second scenario is an asymmetric scenario and is show in **Figure 10**. The simulations are performed for a two-user case (N = 2) up to an eight-user case (N = 8). The four-user scenario, for example, consists of modem 1 up to modem 4. **Table 3** shows the results for the spectrum management problem with total power constraints and spectral mask constraints set to -30dBm/Hz. The complexity reduction factors increase with increasing number of users. For a seven-user case there is a complexity reduction with a factor 538. **Table 4** shows the results for the spectrum management problem without total power constraints and with spectral mask constraints set to -40dBm/Hz. For an eight-user scenario there is a complexity reduction with a factor 12345.

Figures 11 and **12** show the power loading and bit loading respectively for the seven-user asymmetric scenario with total power constraints and spectral mask constraints set to -30dBm/Hz.

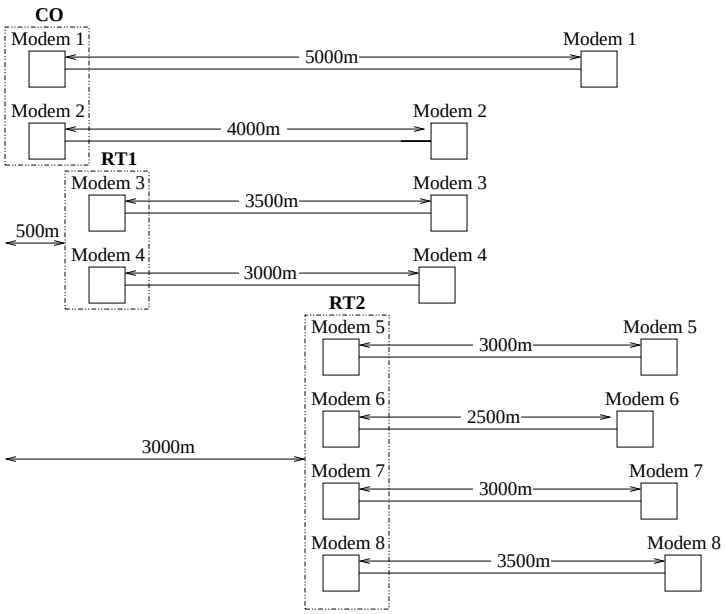

Fig. 10. Asymmetric N-user ADSL scenario

Table 3
Computational complexity comparison of branch and bound procedure and exhaustive search procedure for asymmetric ADSL Downstream scenario with power constraints and spectral mask constraints

			Branch and Bound procedure			Exhaustive Search procedure		
Users	Nb of Traj	Nb of λ-ev	Nb ev of (2)	% ev of (2)	C.r.f.	Nb ev of (2)	% ev of (2)	C.r.f.
2	7	33	149818	7.01	14.3	2136288	100	1
3	3	25	665511	2.42	41.3	27512800	100	1
4	5	29	8161797	1.5	66.7	542552416	100	1
5	11	43	125016093	0.9125	109.6	13.7e9	100	1
6	5	28	681798051	0.45	222	151e9	100	1
7	4	25	4.2708e9	0.186	538	2.297e12	100	1

Table 4
Computational complexity comparison of branch and bound procedure and exhaustive search procedure for asymmetric ADSL Downstream scenario without power constraints and with spectral mask constraints

			Branch and Bound procedure			Exhaustive Search procedure		
Users	Nb of Traj	Nb of λ-ev	Nb ev of (2)	% ev of (2)	C.r.f.	Nb ev of (2)	% ev of (2)	C.r.f.
2	1	1	1676	2.6	38.5	64736	100	1
3	1	1	5924	0.54	185	1100512	100	1
4	1	1	38619	0.2	500	19e6	100	1
5	1	1	367166	0.11	909	318e6	100	1
6	1	1	3254672	0.06	1667	5.4e9	100	1
7	1	1	22638706	0.0246	4065	92e9	100	1
8	1	1	129249600	0.0081	12345	1.56e12	100	1

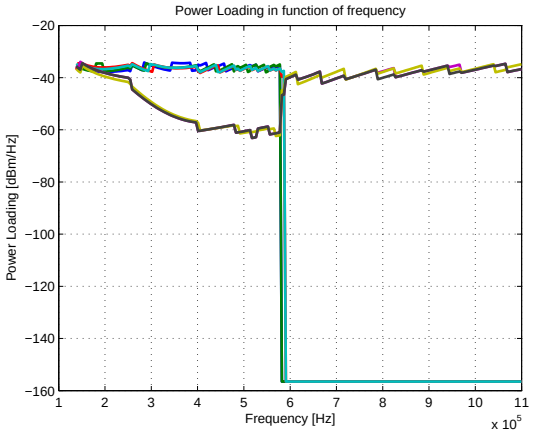

Fig. 11. Asymmetric seven-user ADSL downstream optimal power loading

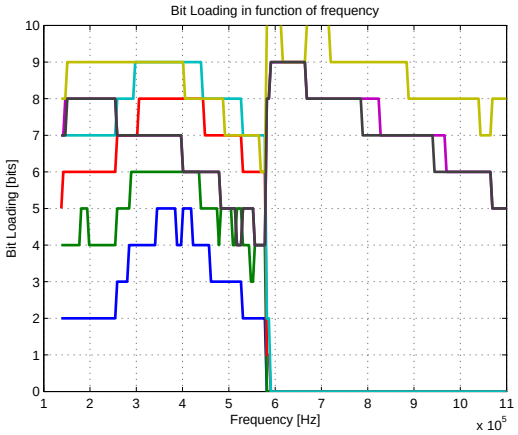

Fig. 12. Asymmetric seven-user ADSL downstream optimal bit loading

7 Conclusion

In modern DSL systems, multi-user crosstalk is a major source of performance degradation. Optimal Spectrum Balancing (OSB) is a centralized algorithm that mitigates the effect of crosstalk by allocating optimal transmit spectra to all interfering DSL modems. By the use of a Lagrange multiplier approach the spectrum management problem is decoupled into non-convex per-tone optimization problems. The remaining issues are finding the Lagrange multipliers that enforce the constraints and solving the per-tone optimization problems. Starting from the single user case, this paper presented a number of properties, which were then extended to the multi-user case and lead to an efficient search algorithm for the Lagrange multipliers. Simulations showed that this algorithm succeeds in finding the optimal Lagrange multipliers in less than 50 iterations, independent of the number of users. Secondly, a low complexity branch and bound procedure is presented to solve the per-tone optimization problems. Simulations showed enormous complexity reductions with respect

to the exhaustive search. The combination of the two proposed algorithms creates an OSB solution with a significantly reduced complexity. This reduction makes it possible to simulate realistic *eight-user* scenarios within a reasonable time on a standard platform whereas before it was only possible to simulate up to four-user scenarios on the same platform. The obtained optimal power spectra or bit loads can provide insights needed for the design of practical suboptimal loading schemes.

A Relation Between Bit Loadings and Power Loadings

In this appendix it will be proven that for a multi-user scenario a nondecreasing bit loading corresponds to a nondecreasing power loading. Mathematically this can be formulated as

$$\begin{aligned}
&\text{Given } \mathbf{s}_k \geq 0,\ \tilde{\mathbf{s}}_k \geq 0,\ b_k^n = f(\mathbf{s}_k),\ \tilde{b}_k^n = f(\tilde{\mathbf{s}}_k)\\
&\text{with} \quad \text{f defined by (1)}\\
&\text{if} \quad\ \tilde{\mathbf{b}}_k \geq \mathbf{b}_k \quad \text{(a)}\\
&\text{then} \quad \tilde{\mathbf{s}}_k \geq \mathbf{s}_k
\end{aligned}$$

The '$\geq$'-operator applied to vectors means inequalities hold for all elements in the vector.
The proof will be given for a three-user case but it can be extended trivially to N users. It will be proven that condition (a) implies that $\tilde{s}_k^1 \geq s_k^1$. The proofs for $\tilde{s}_k^2 \geq s_k^2$ and $\tilde{s}_k^3 \geq s_k^3$ are similar.

The proof is a proof by contradiction.
Let us assume that $s_k^1 > \tilde{s}_k^1 \geq 0$. This can be reformulated as

$$\tilde{s}_k^1 = \alpha s_k^1 \quad \text{with} \quad 0 < \alpha < 1. \tag{A.1}$$

Condition (a) can be written as

$$\frac{|h_k^{1,1}|^2 s_k^1}{|h_k^{1,2}|^2 s_k^2 + |h_k^{1,3}|^2 s_k^3 + \sigma_k^1} \leq \frac{|h_k^{1,1}|^2 \tilde{s}_k^1}{|h_k^{1,2}|^2 \tilde{s}_k^2 + |h_k^{1,3}|^2 \tilde{s}_k^3 + \sigma_k^1} \tag{A.2}$$

$$\frac{|h_k^{2,2}|^2 s_k^2}{|h_k^{2,1}|^2 s_k^1 + |h_k^{2,3}|^2 s_k^3 + \sigma_k^2} \leq \frac{|h_k^{2,2}|^2 \tilde{s}_k^2}{|h_k^{2,1}|^2 \tilde{s}_k^1 + |h_k^{2,3}|^2 \tilde{s}_k^3 + \sigma_k^2} \tag{A.3}$$

$$\frac{|h_k^{3,3}|^2 s_k^3}{|h_k^{3,1}|^2 s_k^1 + |h_k^{3,2}|^2 s_k^2 + \sigma_k^3} \leq \frac{|h_k^{3,3}|^2 \tilde{s}_k^3}{|h_k^{3,1}|^2 \tilde{s}_k^1 + |h_k^{3,2}|^2 \tilde{s}_k^2 + \sigma_k^3}. \tag{A.4}$$

Using (A.1) in (A.2) gives

$$|h_k^{1,2}|^2\tilde{s}_k^2 + |h_k^{1,3}|^2\tilde{s}_k^3 + \sigma_k^1 \leq \alpha(|h_k^{1,2}|^2 s_k^2 + |h_k^{1,3}|^2 s_k^3 + \sigma_k^1). \tag{A.5}$$

We can assume that the noise power σ_k^1 is not equal to zero, which gives:

$$|h_k^{1,2}|^2\tilde{s}_k^2 + |h_k^{1,3}|^2\tilde{s}_k^3 < \alpha(|h_k^{1,2}|^2 s_k^2 + |h_k^{1,3}|^2 s_k^3). \tag{A.6}$$

This relation suggests that

$$\tilde{s}_k^2 < \alpha s_k^2 \quad \text{and/or} \quad \tilde{s}_k^3 < \alpha s_k^3.$$

Without loss of generality, let us suppose that $\tilde{s}_k^2 < \alpha s_k^2$. This can be reformulated as

$$\tilde{s}_k^2 = \beta s_k^2 \quad \text{with} \quad 0 < \beta < \alpha < 1. \tag{A.7}$$

Using (A.7) in (A.3) gives

$$|h_k^{2,1}|^2\tilde{s}_k^1 + |h_k^{2,3}|^2\tilde{s}_k^3 + \sigma_k^2 \leq \beta(|h_k^{2,1}|^2 s_k^1 + |h_k^{2,3}|^2 s_k^3 + \sigma_k^2). \tag{A.8}$$

We can assume that the noise power σ_k^2 is not equal to zero, which gives:

$$|h_k^{2,1}|^2\tilde{s}_k^1 + |h_k^{2,3}|^2\tilde{s}_k^3 < \beta(|h_k^{2,1}|^2 s_k^1 + |h_k^{2,3}|^2 s_k^3). \tag{A.9}$$

Note that $\beta < \alpha$ which results in $\tilde{s}_k^3 < \beta s_k^3$. This can be reformulated as

$$\tilde{s}_k^3 = \gamma s_k^3 \quad \text{with} \quad 0 < \gamma < \beta < \alpha < 1. \tag{A.10}$$

Using (A.10) in (A.4) gives

$$|h_k^{3,1}|^2\tilde{s}_k^1 + |h_k^{3,2}|^2\tilde{s}_k^2 + \sigma_k^3 \leq \gamma(|h_k^{3,1}|^2 s_k^1 + |h_k^{3,2}|^2 s_k^2 + \sigma_k^3). \tag{A.11}$$

Based on (A.1), (A.7) and (A.10) we can see that (A.11) is not possible. The initial assumption (A.1) does not hold which proves that $\tilde{s}_k^1 \geq s_k^1 \geq 0$. In a similar way, the proofs can be given for $\tilde{s}_k^2 \geq s_k^2 \geq 0$ and $\tilde{s}_k^3 \geq s_k^3 \geq 0$.

References

[1] Transmission and Multiplexing (TM); Access transmission systems on metallic access cables; Very high speed Digital Subscriber Line (VDSL); Functional Requirements (2003).

[2] Thomas Starr, John M. Cioffi, Peter J. Silverman, Understanding Digital Subscriber Lines, Prentice Hall, 1999.

[3] Raphael Cendrillon, Marc Moonen, Etienne Van den Bogaert and George Ginis, The Linear Zero-Forcing Crosstalk Canceller is Near-optimal in DSL Channels, in: IEEE Global Communications Conference (Globecom), Vol. 4, 2004, pp. 2334–2338.

[4] Raphael Cendrillon, Marc Moonen, George Ginis, Katleen Van Acker, Tom Bostoen, Piet Vandaele, Partial Crosstalk Cancellation for Upstream VDSL, EURASIP Journal on Applied Signal Processing 2004 (10) (2004) 1520–1535.

[5] George Ginis and John M. Cioffi, Vectored Transmission for Digital Subscriber Line Systems, IEEE Journal on Selected Areas in Communications 20 (5) (2002) 1085–1104.

[6] Spectrum Management for Loop Transmission Systems (2003).

[7] Thomas Starr, Massimo Sorbara, John M. Cioffi, Peter J. Silverman, DSL Advances, Prentice Hall, 2003.

[8] Wei Yu, George Ginis and John Cioffi, Distributed Multiuser Power Control for Digital Subscriber Lines, IEEE Journal on Selected Areas in Communications 20 (5) (2002) 1105–1115.

[9] Raphael Cendrillon, Marc Moonen, Jan Verlinden, Tom Bostoen and Wei Yu, Optimal Multi-user Spectrum Management for Digital Subscriber Lines, in: IEEE International Conference on Communications (ICC), Vol. 1, 2004, pp. 1–5.

[10] Raphael Cendrillon, Wei Yu, Marc Moonen, Jan Verlinden and Tom Bostoen, Optimal Multi-user Spectrum Management for Digital Subscriber Lines, IEEE Transactions on Communications 54 (5) (2006) 922–933.

[11] Raphael Cendrillon, Marc Moonen, Iterative Spectrum Balancing for Digital Subscriber Lines, in: International Communications Conference (ICC), 2005.

[12] Raymond Lui and Wei Yu, Low-Complexity Near-Optimal Spectrum Balancing for Digital Subscriber Lines, in: International Communications Conference (ICC), 2005.

[13] Definition: Branch and bound, From Wikipedia, the Free Encyclopedia.

[14] P.M. Pardalos, H.E. Romeijn, Handbook of Global Optimization Volume 2 (Nonconvex Optimization and Its Applications), 1 Edition, Springer, 2002.

[15] Wei Yu, Raymond Lui and Raphael Cendrillon, Dual Optimization Methods for Multiuser Orthogonal Frequency-Division Multiplex Systems, in: IEEE Global Telecommunications Conference (Globecom), Vol. 1, 2004, pp. 225–229.

[16] Stephen Boyd, Lieven Vandenberghe, Convex Optimization, Cambridge University Press, 2004.

[17] Brian S. Krongold, Kannan Ramchandran and Douglas L. Jones, Computationally Efficient Optimal Power Allocation Algorithms for Multicarrier Communication Systems, IEEE Transactions on Communications 48 (1) (2000) 23–27.

[18] Asymmetrical digital subscriber line (ADSL) transceivers (1999).

[19] Paschalis Tsiaflakis, Jan Vangorp, Marc Moonen, Jan Verlinden and Katleen Van Acker, An Efficient Search Algorithm for the Lagrange Multipliers of Optimal Spectrum Balancing in Multi-user xDSL Systems, in: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Vol. 4, 2006, pp. IV–101– IV–104.

[20] Paschalis Tsiaflakis, Jan Vangorp, Marc Moonen and Jan Verlinden, A Low Complexity Branch and Bound Approach to Optimal Spectrum Balancing for Digital Subscriber Lines, in: Proceedings of the IEEE Global Telecommunications Conference (Globecom), 2006.