AN EFFICIENT SEARCH ALGORITHM FOR THE LAGRANGE MULTIPLIERS OF OPTIMAL SPECTRUM BALANCING IN MULTI-USER xDSL SYSTEMS

Paschalis Tsiaflakis, Jan Vangorp, Marc Moonen

Department of Electrical Engineering
Katholieke Universiteit Leuven, Belgium
{Paschalis.Tsiaflakis, Jan.Vangorp}@esat.kuleuven.be

Jan Verlinden, Katleen Van Acker

DSL Experts Team
Alcatel Bell, Belgium
{Jan.VJ.Verlinden}@alcatel.be

ABSTRACT

In modern DSL systems, multi-user crosstalk is a major source of performance degradation. Optimal Spectrum Balancing (OSB) is a centralized algorithm that optimally allocates the available transmit power over frequencies, thereby mitigating the effect of crosstalk. OSB uses Lagrange multipliers to enforce constraints that are coupled over frequencies. However, finding the optimal Lagrange multipliers can become complex when more than two users are considered. This paper presents a number of properties of the Lagrange multipliers which lead to an efficient search algorithm. Simulations show that the required number of Lagrange multiplier evaluations is independent of the number of users and much smaller compared to the number of evaluations of currently known search algorithms.

1. INTRODUCTION

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

To mitigate crosstalk, current DSL systems use a Static Spectrum Management (SSM) approach where fixed spectral masks ensure that crosstalk levels remain within an acceptable range [1]. 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) [2], a low complexity distributed algorithm. Each user iteratively waterfills its spectrum against the noise and interference. Although IW significantly outperforms SSM, it is not optimal. This is especially so in heavily unbalanced scenarios, where some lines cause much more crosstalk than others (e.g. near-far scenario).

The Optimal Spectrum Balancing (OSB) algorithm [3] [4] 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 a 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 Lagrange multipliers to enforce constraints that are coupled over frequencies. However, finding the optimal Lagrange multipliers can become complex when more than two users are considered. In [3] [4] a bisection method is proposed which has an exponential complexity in the number of users. [5] avoids this exponential complexity by using a subgradient search. However, the stepsize has to be small to guarantee convergence, resulting in a large number of evaluations. This paper develops some insights in the Lagrange multipliers which lead to an efficient search algorithm. Simulations show that the required number of Lagrange multiplier evaluations is independent of the number of users and much smaller compared to the number of evaluations of currently known search algorithms.

2. OPTIMAL SPECTRUM BALANCING

2.1. 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. 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. $\mathbf{z}_k = [z_k^1, z_k^2, \ldots, z_k^N]^T$ 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 vector containing the transmit power of all users on tone k is $\mathbf{s}_k \triangleq [s_k^1, s_k^2, \ldots, s_k^N]^T$.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,

---

Paschalis Tsiaflakis is a research assistant with the F.W.O. Vlaanderen.
Jan Vangorp is a research assistant with the ESAT/SISTA laboratory.
This research work was carried out at the ESAT laboratory of the Katholieke Universiteit Leuven, in the frame of IUAP P5/22 and P5/11, FWO project nr.G.0196.02, IWT project 030054 and CELTIC/IWT project 040049: 'BANITS 'Broadband Access Networks Integrated Telecommunications' and was partially sponsored by Alcatel-Bell.

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 of all modems in the system, is

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

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

$$R^n = f_s \sum_k b_k^n \quad \text{and} \quad P^n = \sum_k s_k^n.$$

2.2. 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. Note that when using a spectrum management model where both these constraints are present, one of them can be made inactive by merely setting it such that the other constraint is more restrictive.

A second type of constraint is a *rate constraint* $R^{n,target}$ 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.

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 all available power to load a maximum number of bits on tones. The rate is thus limited by the total power and spectral mask constraints.

$$\begin{array}{lll} \max_{s^1,\ldots,s^N} & \sum_{n=1}^N R^n & \\ \text{s. t.} & P^n \leq P^{n,tot} & n = 1\ldots N \\ & 0 \leq s_k^n \leq s_k^{n,mask} & n = 1\ldots N, k = 1\ldots K \\ & R^n \geq R^{n,target} & n = 1\ldots N \end{array} \qquad (2)$$

For the *power adaptive mode*, a similar formulation can be used [6]. Only now the sum of the total powers is minimized while the same constraints hold.

2.3. Dual Decomposition

The rate adaptive optimization problem (2) is a non-convex problem and therefore difficult to solve. Because the constraints are coupled across the tones, all possible transmit spectra have to be searched exhaustively to find the global optimum. 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 in case of discrete or continuous loading respectively.

In [3] [4] it was shown that this complexity can be reduced by using the method of dual decomposition, using Lagrange multipliers to move constraints coupled over tones into the unconstrained part of the optimization problem:

$$\mathbf{s}^{1,opt},\ldots,\mathbf{s}^{N,opt} = \underset{\mathbf{s}^1,\ldots,\mathbf{s}^N}{\text{argmax}} \sum_{n=1}^N \omega_n R^n + \sum_{n=1}^N \lambda_n\big(P^{n,tot} - \sum_{k=1}^K s_k^n\big) \qquad (3)$$

$$\begin{array}{lll} \text{with} & 0 \leq s_k^n \leq s_k^{n,mask} & n = 1\ldots N, k = 1\ldots K \\ & \lambda_n \geq 0, \omega_n \geq 0 & n = 1\ldots N \end{array}$$

This optimization problem is decoupled over the tones, thus the spectrum management problem can now be solved in a per-tone fashion.

Given $\omega_n, \lambda_n, n = 1\ldots N$ this maximization problem can be easily solved by performing an exhaustive search on each tone over all possible bit or power loading combinations for the users. This results in transmit spectra for all users. For random ω_n's and λ_n's, the power and rate constraints are generally not satisfied. By choosing appropriate values for the Lagrange multipliers, these constraints can be enforced.

From (3), it can be observed that the λ_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 λ_n's can be viewed as setting a cost for power. The ω_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.

3. LAGRANGE MULTIPLIER SEARCH ALGORITHM

We first investigate the case where $N = 2$. These results are then extended to the multi-user case.

Given $\boldsymbol{\omega} = (\omega_1, \omega_2)$ and $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$, the dual problem, decoupled over the tones, can be solved easily by performing an exhaustive search for each tone over all possible bit or power loading combinations for the users. The optimal solution is then a bit and power loading corresponding to total powers and data rates

$$\big(P^{1,\boldsymbol{\omega},\boldsymbol{\lambda}}, R^{1,\boldsymbol{\omega},\boldsymbol{\lambda}}, P^{2,\boldsymbol{\omega},\boldsymbol{\lambda}}, R^{2,\boldsymbol{\omega},\boldsymbol{\lambda}}\big)$$

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 (3). 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 1**. Here $\big(P^{1,\boldsymbol{\omega},\boldsymbol{\lambda}}, P^{2,\boldsymbol{\omega},\boldsymbol{\lambda}}\big)$ 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 $\big(P^{1,\boldsymbol{\omega},\boldsymbol{\lambda}} = P^{1,tot}, P^{2,\boldsymbol{\omega},\boldsymbol{\lambda}} = P^{2,tot}\big)$, this solution satisfies the total power constraints and so has a weighted rate sum larger than every other possible loading in the marked rectangle. Thus the primal problem would be solved if also the rate constraints are satisfied. Hence we would like 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 $R^{2,\boldsymbol{\omega},\boldsymbol{\lambda}} \geq R^{2,target}$.

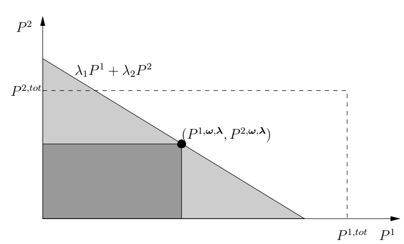

Fig. 1. 2-user power plane

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

$$\begin{aligned} & \omega_{1,A} R^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} + \omega_{2,A} R^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} \\ & \quad -\lambda_{1,A} P^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} - \lambda_{2,A} P^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} \\ \leq \;& \omega_{1,A} R^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A} + \omega_{2,A} R^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A} \\ & \quad -\lambda_{1,A} P^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A} - \lambda_{2,A} P^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A} \end{aligned} \tag{4}$$

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

$$\begin{aligned} & \omega_{1,B} R^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A} + \omega_{2,B} R^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A} \\ & \quad -\lambda_{1,B} P^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A} - \lambda_{2,B} P^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A} \\ \leq \;& \omega_{1,B} R^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} + \omega_{2,B} R^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} \\ & \quad -\lambda_{1,B} P^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} - \lambda_{2,B} P^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} \end{aligned} \tag{5}$$

Taking the sum of (4) and (5) results in

$$\begin{aligned} & -\underbrace{(\omega_{1,B} - \omega_{1,A})}_{\Delta\omega_1} \underbrace{(R^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} - R^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A})}_{\Delta R^1} \\ & -\underbrace{(\omega_{2,B} - \omega_{2,A})}_{\Delta\omega_2} \underbrace{(R^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} - R^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A})}_{\Delta R^2} \\ & +\underbrace{(\lambda_{1,B} - \lambda_{1,A})}_{\Delta\lambda_1} \underbrace{(P^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} - P^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A})}_{\Delta P^1} \\ & +\underbrace{(\lambda_{2,B} - \lambda_{2,A})}_{\Delta\lambda_2} \underbrace{(P^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} - P^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A})}_{\Delta P^2} \leq 0 \end{aligned} \tag{6}$$

Relation (6) for two users can be extended straightforwardly to the multi-user case:

$$\begin{bmatrix} -(\Delta\boldsymbol{\omega})^T & (\Delta\boldsymbol{\lambda})^T \end{bmatrix} \begin{bmatrix} \Delta\mathbf{R} \\ \Delta\mathbf{P} \end{bmatrix} \leq 0. \tag{7}$$

$\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_N]^T$ and $\boldsymbol{\omega} = [\omega_1, \ldots, \omega_N]^T$ are vectors containing the λ's and ω's for the N users, $\mathbf{P} = [P^1, \ldots, P^N]^T$ and $\mathbf{R} = [R^1, \ldots, R^N]^T$ are vectors with the corresponding total powers and data rates.

Two special cases can be derived from formula (7).

$$\text{fixed } \boldsymbol{\omega} \;\; (\Delta\boldsymbol{\omega} = 0) \quad \Rightarrow \quad (\Delta\boldsymbol{\lambda})^T \Delta\mathbf{P} \leq 0 \tag{8}$$

$$\text{fixed } \boldsymbol{\lambda} \;\; (\Delta\boldsymbol{\lambda} = 0) \quad \Rightarrow \quad (\Delta\boldsymbol{\omega})^T \Delta\mathbf{R} \geq 0 \tag{9}$$

Relation (7) can be used to construct a simple procedure to find the $\boldsymbol{\omega}$ and $\boldsymbol{\lambda}$ that make the rate and total power constraints tight. 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 (7).

In the power plane, formula (8) can be represented graphically as two vectors with a non-positive inner product, as in **figure 2(a)**. 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 inner product of $\Delta\boldsymbol{\lambda}$ and $\Delta\mathbf{P}$, the $\Delta\boldsymbol{\lambda}$ for a 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 (8) guarantees that the step taken with $\Delta\mathbf{P}$ will get the used power

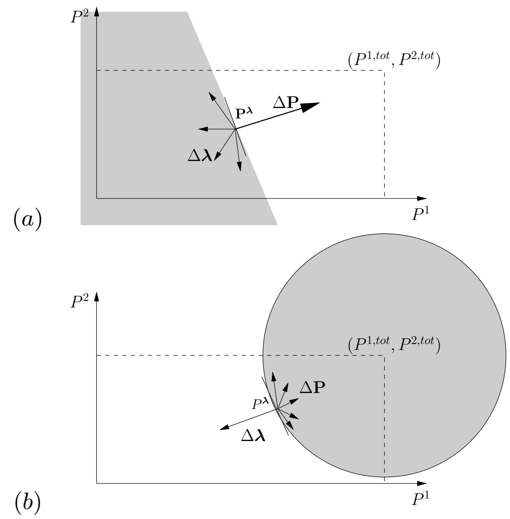

Fig. 2. 2-user power plane

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 2(b)**, where the $\Delta\boldsymbol{\lambda}$ brings $\mathbf{P}^{\boldsymbol{\lambda}}$ to the next point inside the shaded circle.

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

$$\Delta\boldsymbol{\lambda} = -\mu\big(\mathbf{P}^{tot} - \mathbf{P}^{\boldsymbol{\lambda}}\big) \;\Rightarrow\; \boldsymbol{\lambda}^{t+1} = \left[\boldsymbol{\lambda}^t - \mu\Big(\mathbf{P}^{tot} - \sum_k \mathbf{s}_k\Big)\right]^+. \tag{10}$$

By starting with a small μ, e.g. $\mu = 1$ appears to work well in practice, 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 $\big(\mathbf{P}^{tot} - \mathbf{P}^{\boldsymbol{\lambda}}\big)$. This procedure is formally represented in **algorithm 1**. 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 3**.

Algorithm 1 Multi-user $\boldsymbol{\lambda}$ search algorithm

while distance $>$ tolerance **do**
  $\boldsymbol{\lambda} =$ best $\boldsymbol{\lambda}$ so far
  $\mu = 1$
  while distance $\leq$ previousDistance **do**
    previousDistance = distance
    $\mu = \mu \times 2$
    $\Delta\boldsymbol{\lambda} = -\mu\big(\mathbf{P}^{tot} - \mathbf{P}^{\boldsymbol{\lambda}}\big)$
    $[\mathbf{P}^{\boldsymbol{\lambda}+\Delta\boldsymbol{\lambda}}, \mathbf{s}^n]$ = calculateLoading($\boldsymbol{\lambda} + \Delta\boldsymbol{\lambda}$)
      (per-tone exhaustive search)
    distance $= \|\mathbf{P}^{tot} - \mathbf{P}^{\boldsymbol{\lambda}+\Delta\boldsymbol{\lambda}}\|$
  end while
end while

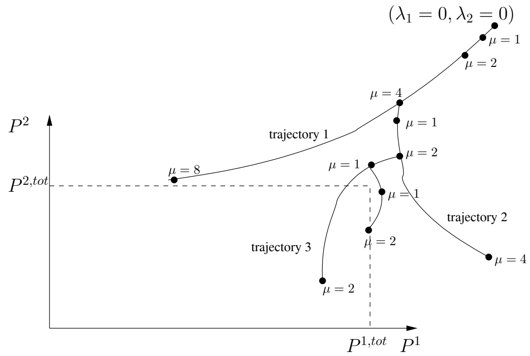

Fig. 3. Trajectories of total power in the power plane

4. SIMULATION RESULTS

The scenarios in this section use a line diameter of 0.5 mm (24 AWG), the maximum transmit power is 20.4 dBm. The SNR gap Γ is set to 12.9 dB, corresponding to a target symbol error probability of 10^{-7}, coding gain of 3 dB and a noise margin of 6 dB. The tone spacing $\Delta_f = 4.3125$ kHz and the DMT symbol rate $f_s = 4$ kHz [7]. All simulations start from initial λ Lagrange multipliers set to zero, while the ω's are fixed. It is seen that in all scenarios, the number of $\boldsymbol{\lambda}$-evaluations for algorithm 1 is roughly independent of the number of users.

4.1. 2-user scenario

The performance of algorithm 1 is first compared to the bisection [3] [4] and subgradient [5] search methods. The scenario for this simulation is a 2-user downstream ADSL system with mixed CO and RT deployment. The CO deployed line has a length of 5 km, the RT deployed line is 3 km. The distance between the CO and RT is 4 km as shown in **figure 4(a)**.

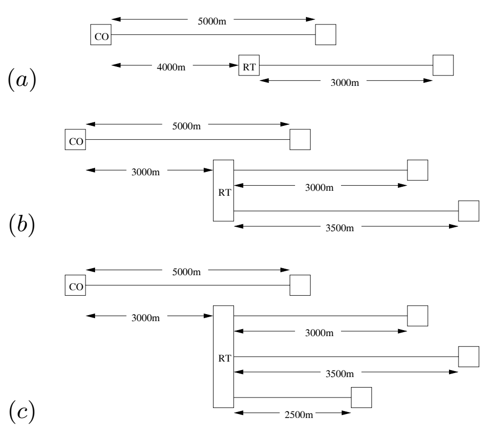

Fig. 4. Downstream ADSL scenarios

Depending on the point in the rate region, about 100 to 150 $\boldsymbol{\lambda}$-evaluations are performed with algorithm 1. Because each $\boldsymbol{\lambda}$-evaluation requires a per-tone exhaustive search to determine the loading on all tones, it is important to keep this number as small as possible. For the bisection method of [3] [4] 400 to 600 $\boldsymbol{\lambda}$-evaluations are needed for this 2-user case. Because of its exponential complexity in the number of users, this will get worse as the number of users increases. For the subgradient method with $\varepsilon = 1$ as suggested in [5], more than 20000 $\boldsymbol{\lambda}$-evaluations are required. However, when the number of users increases, the number of evaluations does not increase exponentially for the subgradient method.

The intuitive interpretation of the algorithm 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. In this way the algorithm converges in less than 40 $\boldsymbol{\lambda}$-evaluations.

4.2. 3-user and 4-user scenarios

The performance of algorithm 1 was also tested on the 3 and 4-user scenarios shown in **figure 4(b) and 4(c)** respectively. The 3-user scenario required 115 $\boldsymbol{\lambda}$-evaluations in 7 trajectories to converge, the 4-user scenario 160 $\boldsymbol{\lambda}$-evaluations in 10 trajectories. For the bisection method, we estimate convergence would take on the order of 8000 and 160000 $\boldsymbol{\lambda}$-evaluations respectively. The subgradient method of [5] with step size $\varepsilon = 1$ is expected to have similar performance as in the 2-user case (more than 20000 $\boldsymbol{\lambda}$-evaluations).

5. CONCLUSION

In this paper, the problem of finding the Lagrange multipliers for OSB is analyzed. Insights into the relation between Lagrange multipliers and constraints lead to a search procedure to find the Lagrange multipliers that enforce the constraints. Here *all* Lagrange multipliers can be updated in parallel, leading to a complexity which is found to be roughly independent of the number of users.

Simulations show that for 2-, 3- and 4-user scenarios, 100-150 evaluations of the Lagrange multipliers are sufficient to enforce the constraints. Moreover, the intuitive interpretation of the search algorithm allows for more efficient updates of the step size. This results in an even faster convergence, typically under 40 evaluations of the Lagrange multipliers. Because of the efficient search procedure, the remaining complexity of OSB is situated at the level of the per-tone exhaustive search. This is an area of current research.

6. REFERENCES

[1] *Spectrum Management for Loop Transmission Systems*, ANSI Std. T1.417, Issue 2, 2003.

[2] W. Yu, G. Ginis and J. Cioffi, "Distributed Multiuser Power Control for Digital Subscriber Lines," *IEEE J. Sel. Area. Comm.*, vol. 20, no. 5, pp. 1105–1115, Jun. 2002.

[3] R. Cendrillon, M. Moonen, J. Verlinden, T. Bostoen and W. Yu, "Optimal Multi-user Spectrum Management for Digital Subscriber Lines," in *IEEE ICC*, vol. 1, Jun. 2004, pp. 1–5.

[4] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden and T. Bostoen, "Optimal Multi-user Spectrum Management for Digital Subscriber Lines," *accepted for IEEE Trans. Comm.*

[5] W. Yu, R. Lui and R. Cendrillon, "Dual Optimization Methods for Multiuser Orthogonal Frequency-Division Multiplex Systems," in *IEEE Globecom*, vol. 1, Dec. 2004, pp. 225–229.

[6] P. Tsiaflakis, J. Vangorp, M. Moonen, J. Verlinden, K. Van Acker and R. Cendrillon, "An efficient Lagrange Multiplier search algorithm for Optimal Spectrum Balancing in crosstalk dominated xDSL systems," *submitted to IEEE J. Sel. Area. Comm.*, 2005.

[7] *Asymmetric digital subscriber line transceivers2 (ADSL2)*, ITU-T Std. G.992.3, 2002.