Optimal Multi-user Spectrum Management for Digital Subscriber Lines

Raphael Cendrillon, Wei Yu, Marc Moonen, Jan Verlinden, Tom Bostoen

***Abstract*— Crosstalk is a major issue in modern DSL systems such as ADSL and VDSL. Static spectrum management, the traditional way of ensuring spectral compatibility, employs spectral masks which can be overly conservative and lead to poor performance.**

In this paper we present a centralized algorithm for optimal spectrum management (OSM) in DSL. The algorithm uses a dual decomposition to solve the spectrum management problem in an efficient and computationally tractable way. The algorithm shows significant performance gains over existing DSM techniques, e.g. in an upstream VDSL scenario the centralized OSM algorithm can outperform a distributed DSM algorithm such as *iterative waterfilling* by up to 380%

***Index Terms*— Asymmetric digital subscriber line (ADSL), dual decomposition, dynamic spectrum management, interference channel, Lagrange, non-convex optimization, power control, spectral compatibility, power backoff, remote terminal, very-high bit-rate digital subscriber line (VDSL)**

I. INTRODUCTION

CROSSTALK is a major issue in modern DSL systems such as ADSL and VDSL. Typically 10-20 dB larger than the background noise, crosstalk is *the* dominant source of performance degradation.

Whilst it is possible to do *crosstalk cancellation*[1][2], in many scenarios this may not be feasible due to complexity issues or as a result of unbundling. In this case the effects of crosstalk must be mitigated through *spectral management*. With spectral management the transmit spectra of the modems within a network are limited in some way to minimize the negative effects of crosstalk.

Static spectrum management (SSM) is the traditional approach. In SSM spectral masks are employed which are identical for all modems. To ensure widespread deployment, these masks are based on worst case scenarios[3]. As a result they can be overly restrictive and lead to poor performance.

Dynamic spectrum management (DSM), a new paradigm, overcomes this problem by designing the spectra of each modem to match the specific topology of the network[4]. These spectra are adapted based on the direct and crosstalk channels seen by the different modems. They are customized to suit each modem in each particular situation.

A DSM algorithm known as *iterative waterfilling* was recently proposed[5] and demonstrates the spectacular performance gains which are possible. An unanswered question at this point is: How much better can we do?

In this paper we address this question. We focus on the problem of spectrum management where a centralized spectrum management center (SMC) is responsible for setting the spectra of the modems within the network. We present an algorithm for *optimal spectrum management* (OSM) in the DSL interference channel. This algorithm can achieve the best possible trade-off between the rates of the modems within the network, allowing operation at any point on the rate region boundary.

The algorithm is suitable for direct application when a SMC is available. In the absence of a SMC this algorithm is also useful as it provides an upper bound on the performance of all other DSM algorithms, both centralized and distributed. The spectra generated by the algorithm also give insight into the design of distributed DSM algorithms.

One may ask, if centralized control is available (via a SMC) why not do full-blown crosstalk cancellation which leads to greater performance than with DSM alone. The fundamental difference between DSM and crosstalk cancellation is complexity. DSM involves only setting the PSD levels of currently-available modems. This can be done without any change to the modem hardware. Crosstalk cancellation uses signal level co-ordination, requiring an entirely new design of the *DSL access multiplexer* (DSLAM) and *customer premises* (CP) modems. DSM can potentially be applied right now, where-as it may be several years before systems with crosstalk cancellation become economically viable.

Optimal spectrum management has been investigated previously. Unfortunately the resulting optimisation is non-convex which leads to an exponential complexity in the number of tones K in the system. In ADSL $K = 256$ whilst in VDSL $K = 4096$. This results in a computationally intractable problem.

The fundamental problem is that the total power constraints on the modems couple the optimisation across frequency. As such the optimisation must be done jointly across all tones which leads to an exponential complexity in K. We overcome this problem through use of the dual decomposition method. This technique allows us to replace the constrained optimisation problem with an unconstrained maximization of a Lagrangian. The Lagrangian incorporates the constraints implicitly into the cost function, removing the need for the con-

R. Cendrillon and M. Moonen are with the Katholieke Universiteit Leuven, Belgium. Email: {raphael.cendrillon, marc.moonen}@esat.kuleuven.ac.be

Wei Yu is with the University of Toronto, Canada. He is supported by Bell Canada University Laboratories, Communications and Information Technology Ontario (CITO), Natural Science and Engineering Research Council (NSERC) of Canada, and the Canada Research Chairs Program. Email: weiyu@comm.utoronto.ca

Jan Verliden and Tom Bostoen are with the DSL Research and Innovation group, Alcatel Bell, Belgium.

This work was carried out in the frame of IUAP P5/22, *Dynamical Systems and Control: Computation, Identification and Modelling* and P5/11, *Mobile multimedia communication systems and networks*; the Concerted Research Action GOA-MEFISTO-666, *Mathematical Engineering for Information and Communication Systems Technology*; IWT SOLIDT Project, *Solutions for xDSL Interoperability, Deployment and New Technologies*; FWO Project G.0196.02, *Design of efficient communication techniques for wireless time-dispersive multi-user MIMO systems* and was partially sponsored by Alcatel-Bell.

straints to be explicitly enforced. As a result the optimisation can be decoupled across frequency and an optimal solution can be found in a per-tone fashion. This leads to a linear rather than exponential complexity in K and a computationally tractable problem.

In [6] an attempt was made to formulate an optimal spectrum management algorithm based on simulated annealing. Due to the complexity of the problem the PSDs of each modem were forced to be flat within each transmission band. Only the level of the PSD in each band could be varied which led to a sub-optimal solution. Even with this restriction the algorithm had a large complexity $\mathcal{O}(e^K)$. Furthermore since the algorithm is based on simulated annealing it is not possible to guarantee that the global optimum has been obtained.

Sub-optimal DSM algorithms, both distributed [5], [7], [8] and centralized [9] have also been proposed.

II. System Model

Assuming that *discrete multi-tone* (DMT) modulation is employed we can model transmission independently on each tone

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{z}_k \quad (1)$$

The vector $\mathbf{x}_k \triangleq \left[x_k^1, \cdots, x_k^N\right]$ contains transmitted signals on tone k. There are N lines in the binder and x_k^n is the signal transmitted onto line n at tone k. $\mathbf{y}_k$ and $\mathbf{z}_k$ have similar structures. $\mathbf{y}_k$ is the vector of received signals on tone k. $\mathbf{z}_k$ is the vector of additive noise on tone k and contains thermal noise, alien crosstalk, RFI etc. Recall that $1 \leq k \leq K$ where K is the number of tones within the system. We denote the noise PSD on line n as $\sigma_k^n \triangleq \mathcal{E}\left\{|z_k^n|^2\right\}$. $\mathbf{H}_k$ is the $N \times N$ channel transfer matrix on tone k. $h_k^{n,m} \triangleq [\mathbf{H}_k]_{n,m}$ is the channel from TX m to RX n on tone k. The diagonal elements of $\mathbf{H}_k$ contain the direct-channels whilst the off-diagonal elements contain the crosstalk channels. We denote the transmit PSD $s_k^n \triangleq \mathcal{E}\left\{|x_k^n|^2\right\}$. For convenience we denote the vector containing the PSD of user n on all tones as $\mathbf{s}_n \triangleq [s_1^n, \ldots, s_K^n]$.

In (1) we assume that each tone operates independently and free from inter-carrier interference (ICI). If the modems in the network are not synchronized this will not strictly be the case. Sidelobes from the FFT operation in the receiver will give rise to ICI between users. In VDSL however, due to the transmit and receive windowing employed the effect of sidelobes is minimal.

We assume that each modem can only support a maximum bitloading of $b_{\max}$. $b_{\max}$ lies in the range 8-15 in current standards[10][11][12]. Under this assumption the achievable bitloading of user n on tone k is

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

Γ is the SNR-gap to capacity and is a function of the desired BER, coding gain and noise margin[13]. The data-rate on line n is thus

$$R_n = \sum_k b_k^n$$

III. Spectrum Management

A. The Spectrum Management Problem

We restrict our attention to the two user case for ease of explanation. Extensions to more than two users are straightforward. The spectrum management problem for the two user case is defined as

$$\max_{\mathbf{s}_1,\mathbf{s}_2} R_2 \quad \text{s.t.} \quad R_1 \geq R_1^{\text{target}} \quad (3)$$

B. Constraints

The optimisation (3) is typically subject to a *total power constraint* on each modem

$$\sum_k s_k^n \leq P_n, \; n = 1, 2 \quad (4)$$

This arises from limitations on each modem's analog front-end. *Spectral mask constraints* may also apply

$$s_k^n \leq s_k^{n,\text{mask}}, \; \forall k, n \quad (5)$$

C. Continuous Bitloading

We first consider the case where the modems can support any possible bitloading. This is of more theoretical interest since it determines the theoretical capacity of the channel. We assume that the modems can control their transmit PSDs with an accuracy of Δ_s dB.

The total power (4) and spectral mask constraints (5) allow us to upper bound the transmit power on any tone

$$s_k^n \leq s_k^{\text{n,max}}$$

where $s_k^{n,\max} = \min(P_n,\; s_k^{n,\text{mask}})$. In the absence of spectral masks $s_k^{n,\max} = P_n$. Hence we can limit the range of s_k^n

$$s_k^n \in \{0, \Delta_s, \ldots, s_k^{n,\max}\}$$

So on tone k there are $q_k = \prod_n \left(s_k^{n,\max}\Delta_s^{-1} + 1\right)$ possible PSD pairs.

D. Discrete Bitloading

In practice DSL modems can only support a fixed set of discrete bitloadings. This allows us to reduce our search space to the PSDs corresponding to these exact bitloadings, reducing complexity considerably without affecting optimality. To find the PSDs corresponding to a particular bitloading we proceed as follows. Define

$$\mathbf{A} \triangleq \begin{bmatrix} 0 & \alpha_k^{1,2} \\ \alpha_k^{2,1} & 0 \end{bmatrix}$$

where $\alpha_k^{n,m} \triangleq \Gamma |h_k^{n,m}|^2$. Also define $\boldsymbol{\sigma}_k \triangleq \Gamma[\sigma_k^1,\; \sigma_k^2]^T$, $\mathbf{D}_k \triangleq \text{diag}\{|h_k^{1,1}|^2, \ldots, |h_k^{N,N}|^2\}$ and $\Lambda_k \triangleq \text{diag}\{2^{b_k^1} - 1,\; 2^{b_k^2} - 1\}$. The PSD pair required to support a particular bitloading pair b_k^1, b_k^2 is then

$$\begin{bmatrix} s_k^1 \\ s_k^2 \end{bmatrix} = (\mathbf{D}_k - \Lambda_k \mathbf{A}_k)^{-1} \Lambda_k \boldsymbol{\sigma}_k \quad (6)$$

In the following we use $s_k^n(b_k^1, b_k^2)$ to denote the PSD of user n corresponding to the bitloadings b_k^1, b_k^2 as calculated by (6). Hence we can limit the range of PSD pairs (s_k^1, s_k^2)

$$(s_k^1, s_k^2) \in \left\{ \left(s_k^1(b_k^1, b_k^2), s_k^2(b_k^1, b_k^2) \right) \mid b_k^n \in \{0, \dots b_{\max}\} \, \forall n \right\}$$

So on tone k there are $q_k = (b_{\max} + 1)^2$ possible PSD pairs.

E. Exhaustive Search

At this point we could propose a simplistic algorithm to find the optimal PSDs based on an exhaustive search. On tone k there are p_k possible PSD pairs. Taking all possible PSD levels across all tones results in $\prod_k q_k$ possible PSD pairs. We determine the feasibility of each PSD pair based on any power constraints as described in Sec. III-B, and on the target rate constraint for user 1. We then choose the PSD pair which maximizes the data-rate of user 2.

Unfortunately whilst this algorithm is simple to implement, its complexity in the discrete bitloading case is $\mathcal{O}((b_{\max} + 1)^{2K})$. With $K = 256$ in ADSL and $K = 4096$ in VDSL, this results in a computationally intractable problem. In the continuous bitloading case the complexity can be even higher.

F. Dual Decomposition

As we saw in the previous section, an exhaustive search for the optimal PSDs leads to a computationally intractable problem. The reason behind this is as follows. The total power constraint on each line causes the power allocation problem to become coupled across frequency. As such we must jointly search the PSDs across all tones. This results in an exponential complexity in K and an intractable problem.

To overcome this we replace the power constrained optimisation (3), with an unconstrained optimisation of a Lagrangian (9). In the Lagrangian the total power constraints are enforced through the use of the Lagrangian multipliers λ_1 and λ_2 which form part of the cost function. When λ_1 and λ_2 are chosen correctly, maximizing the Lagrangian will implicitly enforce the power constraints. The power constraints need not be explicitly enforced and the problem can be decoupled across frequency.

When the problem is decoupled we can solve the optimisation by maximizing the Lagrangian independently on each tone. This leads to a complexity which is linear rather than exponential in K and the problem becomes computationally tractable. This is the main innovation in this paper.

We begin in Sec. III-G by replacing the original optimisation problem (3), with a weighted rate-sum maximization (7). With a correctly chosen weight w, maximizing (7) implicitly enforces the target rate constraint on user 1. The weight w is in itself a form of Lagrangian multiplier.

In Sec. III-H we append the Lagrangian multipliers to the weighted rate-sum to form the Lagrangian. We will see that maximizing this Lagrangian is equivalent to solving the original optimisation problem (3). We will also see that this Lagrangian can be decoupled and maximized independently on each tone.

Using a Lagrangian to solve a constrained optimisation in an unconstrained way is a commonly used approach in convex optimisation theory and is known as the dual decomposition method. The dual decomposition has been applied in other communication problems with convex cost functions such as joint routing and resource allocation[14] and power allocation in the vector multiple access channel[15]. In this work we show that the dual decomposition method can also be applied to non-convex optimizations.

G. Rate Regions

The rate region is a plot of all possible operating points (rate pairs) that can be supported by a multi-user channel. Operating points on the boundary of the region are said to be optimal. It is our goal to find the PSDs corresponding to these points.

We start by considering the following optimization problem where our goal is to maximize the weighted rate-sum

$$\max_{\mathbf{s}_1, \mathbf{s}_2} wR_1 + (1 - w)R_2 \tag{7}$$

We will show that solving this problem is equivalent to solving the original spectrum management problem (3). We start with the following Lemma.

Lemma 1: For any particular w, there exists some R_1^{target} for which the weighted rate-sum optimization (7) is equivalent to the original spectrum management problem (3).

Proof: The proof will be made by illustration. Examining the rate region in Fig. 1 we see that there is only one point $A \triangleq (R_1^a, R_2^a)$ which maximizes the weighted rate-sum for a given w. Assume that there exists some other point in the rate region $B \triangleq (R_1^b, R_2^b)$ such that $R_2^b \geq R_1^a$ and $R_2^b > R_2^a$. This would imply that point B leads to a larger weighted rate-sum than point A, but this is contradicted by the optimality of A in the weighted rate-sum (7). Hence R_2^a is the highest rate for line 2 which will allow the target rate for line 1 to be achieved. This implies that point A is optimal in terms of the original spectrum management problem (3) for the target rate $R_1^{\text{target}} = R_1^a$. ■

Theorem 1: For any rate region define $\mathbb{X}$ as the set of rate pairs within the rate region, and $\mathbb{Y}$ as the boundary of the convex hull of $\mathbb{X}$. Consider an operating point $C \triangleq (R_1^c, R_2^c)$ which is achievable $C \in \mathbb{X}$ and on the boundary of the convex hull of the rate region $C \in \mathbb{Y}$. So $C \in \mathbb{X} \cap \mathbb{Y}$. This is depicted in Fig. 2. For any such C there exists some w such that the PSDs at point C can be found through a weighted rate-sum maximization. Furthermore C is optimal in terms of the original spectrum management problem (3) for $R_1^{\text{target}} = R_1^c$.

Proof: C is on the boundary of $\mathbb{Y}$ and $\mathbb{Y}$ is convex. So there exists no point $D \triangleq (R_1^d, R_2^d) \in \mathbb{Y}$ such that $R_1^d \geq R_1^c$ and $R_2^d > R_2^c$. So for some w

$$wR_2^c + (1 - w)R_2^c \geq wR_2^d + (1 - w)R_2^d, \; \forall (R_1^d, R_2^d) \in \mathbb{Y}$$

Now since all points in $\mathbb{X}$ are on the interior of $\mathbb{Y}$

$$wR_2^c + (1 - w)R_2^c \geq wR_2^d + (1 - w)R_2^d, \; \forall (R_1^d, R_2^d) \in \mathbb{X}$$

So the point C gives the maximum weighted rate-sum of all achievable points within the rate region $\mathbb{X}$ for some particular weight w. Hence the point C is optimal in the weighted rate-sum (7) for that w.

Now Lemma 1 implies that any weighted rate-sum optimization has an equivalent spectrum management problem (3). So C is optimal in terms of (3), in this case for $R_1^{\text{target}} = R_1^c$. ■

Corollary 1: For any convex rate-region, all optimal operating points on the boundary of the rate region can be found through a weighted rate-sum optimization.

Proof: In a convex rate region, the boundary of the convex hull $\mathbb{X}$, contains the entire boundary of the rate region and $\mathbb{X} \cap \mathbb{Y} = \mathbb{Y}$. All optimal operating points in terms of the original spectrum management problem (3) lie on the boundary of the rate region. Hence Theorem 1 implies that all optimal operating points can be found through a weighted rate-sum optimization. ■

Theorem 1 implies that any achievable operating point on the boundary of the convex hull of the rate region can be found through a weighted rate-sum optimization. If the rate region is close to being convex, then the majority of the optimal operating points can be found. Thankfully this is the case in DSL channels as we now explain.

In the wireline medium there is some correlation between the channels on neighbouring tones. If we sample the channel finely enough then neighbouring tones will see almost the same channels (both direct and crosstalk).

Imagine that the tone spacing is fine enough such that $h_k^{n,m} \simeq h_{k+l}^{n,m}$, $0 \leq l \leq L-1$. Consider two points in the rate region, $A = (R_1^a, R_2^a)$ and $B = (R_1^b, R_2^b)$ and their corresponding PSDs $(s_k^{1,a}, s_k^{2,a})$ and $(s_k^{1,b}, s_k^{2,b})$. It is possible to operate at a point $E = (\frac{l}{L}R_1^a + \frac{L-l}{L}R_1^b, \frac{l}{L}R_2^a + \frac{L-l}{L}R_2^b)$ for any $0 \leq l \leq L-1$ as depicted in Fig. 2. This is done by setting the PSDs to $(s_k^{1,a}, s_k^{2,a})$ on tones $k \in \{pL+1, \ldots pL+l\}$ for all integer values of p, and to $(s_k^{1,b}, s_k^{2,b})$ on all other tones.

For example, to operate at a point 2/3 between A and B (on the side closer to A), we require $l = 2$ and $L = 3$. Thus we set the PSDs to $(s_k^{1,a}, s_k^{2,a})$ on tones $k \in \{1, 2, 4, 5, 7, 8, \ldots, K\}$ and to $(s_k^{1,b}, s_k^{2,b})$ on tones $k \in \{3, 6, 9, \ldots, K\}$. For this to work the tone spacing must be small enough such that the channel is approximately flat over $L = 3$ neighbouring tones. That is, we must have $h_k^{n,m} \simeq h_{k+1}^{n,m} \simeq h_{k+2}^{n,m}$, $\forall k \in \{1, 4, \ldots, K\}$.

For large L (small tone spacing), practically any operating point between A and B can be achieved. Thus for any two points in the rate region, any point between them is also within the rate region. This is the definition of a convex set. As such the rate region is convex in DMT systems with small tone spacings.

In ADSL and VDSL the tone spacing is 4.3125 kHz. In both measured and empirical wireline channels we have found this tone spacing to be small enough such that the rate regions are convex or very nearly-convex.

Note that one should not confuse convexity of the rate-region with convexity of the cost function (7). In practice the rate regions are seen to be very nearly-convex, however the optimisation problem is highly non-convex, exhibiting many local maxima. For this reason conventional convex optimisation techniques cannot be applied and an exhaustive search is required on each tone.

H. The Lagrangian

So far we have shown that the spectrum management problem (3) can be solved through a weighted rate-sum optimization (7). We also saw that in DSL the rate region is approximately convex, allowing almost all optimal operating points to be found. We now focus on solving the weighted rate-sum optimization in a computationally tractable way.

We can incorporate the total power constraints (4) into the optimization problem by defining the Lagrangian

$$L \triangleq wR_1 + (1-w)R_2 + \lambda_1(P_1 - \sum_k s_k^1) + \lambda_2(P_2 - \sum_k s_k^2) \quad (8)$$

λ_n is the Lagrangian multiplier for user n and is chosen such that either the power constraint on user n is tight $\sum_k s_k^n = P_n$ or $\lambda_n = 0$. The constrained optimization (7) can now be solved via the unconstrained optimization

$$\max_{\mathbf{s}_1, \mathbf{s}_2} L(w, \lambda_1, \lambda_2, s_k^1, s_k^2) \quad (9)$$

Define the Lagrangian on tone k

$$L_k \triangleq wb_k^1 + (1-w)b_k^2 - \lambda_1 s_k^1(b_k^1, b_k^2) - \lambda_2 s_k^2(b_k^1, b_k^2)$$

Clearly the Lagrangian (8) can be decomposed into a sum across tones of L_k and a term which is independent of s_k^1 and s_k^2

$$L = \sum_k L_k + \lambda_1 P_1 + \lambda_2 P_2$$

As a result we can split the optimization into K per-tone optimizations which are coupled only through w, λ_1 and λ_2. This will lead to a linear, rather than exponential complexity in K and a computationally tractable search.

IV. Optimal Spectrum management

The *optimal spectrum management* (OSM) algorithm is listed as Alg. 1. Spectral mask constraints can be incorporated into the optimisation by setting L_k to $-\infty$ if $s_k^1 > s_k^{1,\max}$ or $s_k^2 > s_k^{2,\max}$. If discrete bitloading is employed then the maximization in the function *optimize_s* is limited to the PSD pairs corresponding to valid bitloading pairs.

The algorithm operates as follows. We need to search through both λ_1 and λ_2 to find values which place sufficient importance on the total power constraint terms within the Lagrangian (8). Varying w allows us to map out all optimal, achievable points on the convex hull of the rate region.

The algorithm contains three loops, an outer loop which varies w, an intermediate loop which searches for λ_1 and an inner loop which searches for λ_2. Bisection is used in each search.

When searching for λ_n, we first find a value of λ_n which ensures that the power constraint of user n is satisfied. This value is stored in $\lambda_n^{\max}$. Note that a larger λ_n places more emphasis on the power constraint of user n in the Lagrangian. As a result, using a larger λ_n will result in a lower total power for user n.

Once $\lambda_n^{\max}$ is found the algorithm proceeds to bisection. Note that after the algorithm has completed, for each user either $\sum_k s_k^n = P_n$ or the corresponding Lagrangian multiplier

Algorithm 1 Optimal Spectrum Management

```
Main Function
  for w = 0...1
    s_1, s_2 = optimize_λ_1(w)
  end

Function s_1, s_2 = optimize_λ_1(w)
  λ_1^max = 1, λ_1^min = 0
  while Σ_k s_k^1 > P_1
    λ_1^max = 2λ_1^max
    s_1, s_2 = optimize_λ_2(w, λ_1^max)
  end
  repeat
    λ_1 = (λ_1^max + λ_1^min)/2
    s_1, s_2 = optimize_λ_2(w, λ_1)
    if Σ_k s_k^1 > P_1, then λ_1^min = λ_1, else λ_1^max = λ_1
  until convergence

Function s_1, s_2 = optimize_λ_2(w,λ_1)
  λ_2^max = 1, λ_2^min = 0
  while Σ_k s_k^2 > P_2
    λ_2^max = 2λ_2^max
    s_1, s_2 = optimize_s(w, λ_1, λ_2^max)
  end
  repeat
    λ_2 = (λ_2^max + λ_2^min)/2
    s_1, s_2 = optimize_s(w, λ_1, λ_2)
    if Σ_k s_k^2 > P_2, then λ_2^min = λ_2, else λ_2^max = λ_2
  until convergence

Function s_1, s_2= optimize_s(w, λ_1, λ_2)
  for k = 1...K
    s_k^1, s_k^2 = arg max_{s_k^1, s_k^2} L_k(s_k^1, s_k^2, w, λ_1, λ_2)
  end
```

is driven to zero ($\lambda_n = 0$). Thus the Lagrangian and the original objective become equivalent. More rigorously,

Theorem 2: For each w Alg. 1 returns a PSD pair which is optimal for the spectrum management problem (3). That is, for some R_1^{target}

$$\begin{aligned} \mathbf{s}_1, \mathbf{s}_2 \quad &= \quad \arg\max_{\mathbf{s}_1, \mathbf{s}_2} R_2 \\ \text{s.t.} \quad & R_1 \geq R_1^{\text{target}} \\ & \sum_k s_k^n \leq P_n, \ \forall n \end{aligned} \tag{10}$$

Varying w from 0 to 1 allows all optimal operating points which lie on the convex hull of the rate region to be found. If the rate region is convex then all optimal operating points can be found.

Proof: See Appendix ■

If the function *optimize_s* finds multiple PSD pairs which yield the same value for the Lagrangian L_k, then all PSD pairs are stored. This ensures that if the rate-region contains a flat section, then all operating points on this section are discovered.

Note that by solving the optimization independently on each tone we require only $\mathcal{O}(K(b_{\max}+1)^2)$ evaluations of L_k each time the function *optimize_s* is called in the discrete bitloading case. So the complexity becomes linear rather than exponential in K. In contrast solving the problem jointly across all tones would have required $(b_{\max}+1)^{2K}$ evaluations of $s_k^1(b_k^1, b_k^2)$ and $s_k^2(b_k^1, b_k^2)$ which is computationally intractable. A similar comparison can be made in the continuous bitloading case.

In this paper we have only presented the algorithm and optimality proof for 2 user channels. Extensions to more than 2 users are straight-forward and follow naturally from the algorithm and proof presented here. In the general case of N users we have a target rate constraint on the first $N-1$ users and the goal is to maximize the rate of the Nth user. This can be shown to be equivalent to maximizing a weighted rate-sum $\sum_n w_n R_n$. Here the weight for the Nth user is arbitrarily defined as $w_N \triangleq 1 - \sum_{n=1}^{N-1} w_n$. To enforce the total power constraints on all users N Lagrangian multipliers are required $\lambda_1, \ldots, \lambda_N$. The algorithm has a similar form to Alg. 1 however we sweep through all values of $w_1, \ldots, w_N$ in *Main Function*. In Alg. 1 bi-section is only done on λ_1 and λ_2. With N users bi-section must be done on $\lambda_1, \ldots, \lambda_N$.

In section III-G we showed that not all points in a non-convex rate region can be found through a weighted rate-sum optimization. However it should be made clear that this does not affect the optimality of Alg. 1. As shown in Theorem 2, at convergence any PSDs generated by Alg. 1 are optimal and lead to a rate pair on the boundary of the convex hull of the rate region. A non-convex rate region only implies that not all optimal operating points can be found through Alg. 1. In practice this is not a problem since, as we discussed in section III-G the rate regions in DSL are very close to being convex.

V. PERFORMANCE

We now examine the performance of OSM when compared with other spectrum management techniques. For all simulations we use 0.5 mm (24-Gauge) lines. The target symbol error probability is 10^{-7} or less. The coding gain and noise margin are set to 3 dB and 6 dB respectively. We use continuous bitloading and set Δ_s to 0.1 dBm/Hz. We do not constrain the maximum bitloading.

A. Near-far Problem in Upstream VDSL

We simulate upstream transmission in VDSL with 4×600 m. lines and 4×1200 m. lines as depicted in Fig. 3. A maximum transmit power of 11.5 dBm is available to each modem. The usual PSD constraint is not applied in the OSM or iterative waterfilling algorithms, except for below 1.1 MHz for the protection of ADSL and other services. A spectral mask is applied to the flat PBO, reference noise and reference PSD method and is set at -60 dBm/Hz[11]. Alien crosstalk is incorporated into the background noise using ETSI model A. FDD bandplan 998 is used with the frequency bands corresponding to amateur radio frequencies notched off. For more details see [11].

Fig. 4 shows the rate regions corresponding to various spectrum management algorithms. Included are iterative waterfilling[5], and the reference PSD method which is currently adopted in VDSL standards[10][11]. For comparison we also include flat PBO where each modem transmits the minimum possible flat PSD required to support their target rate.

The reference noise method is also included. In the reference noise method each modem sets its transmit PSD such that the crosstalk it induces on the victim modem is equal to the background noise seen by that modem, the so-called *reference noise*[8].

Note that iterative waterfilling, the reference PSD method and flat PBO are all distributed algorithms and require no centralized control. In contrast OSM is a centralized algorithm requiring knowledge of the direct and crosstalk channel attenuations within the network. OSM is suitable for direct application when a spectrum management center (SMC) is available. In the absence of a SMC, OSM is still a useful tool in the design of distributed dynamic spectrum management algorithms, providing both an upper bound on performance and insight into spectrum management.

The PSDs corresponding to a 5 Mbps service on the 1200 m. lines are depicted in Fig. 5 and Fig. 6. Under the 998 FDD bandplan there are two separate upstream bands: 3.75 - 5.2 MHz and 8.5 - 12 MHz. CP modems may not use frequencies between these bands since they are reserved for downstream transmission by the central office (CO) modems.

The OSM PSD of the 600 m. lines in the first upstream band is quite flat with some power back-off (PBO) applied. In the second upstream band PBO is not required since the 1200 m. lines are not active. There the OSM PSD on the 600 m. lines increases with frequency as the crosstalk coupling between the four different 600 m. lines rises.

The iterative waterfilling algorithm yields almost equal PSDs in both bands. It does not exploit the fact that the 1200 m. lines are inactive in the high frequencies. This leads to poorer performance.

As shown in Tab. I using OSM instead of iterative waterfilling allows the data-rate on the 600 m. lines to be increased from 7.7 Mbps to 15 Mbps whilst still maintaining a 5 Mbps service on the 1200 m. lines. This is a gain of over 190%

B. Remote Terminal Distributed Downstream ADSL

We simulate downstream transmission in ADSL with a 5 km CO distributed line and a 3 km RT distributed line. The RT is located 4 km from the CO as depicted in Fig. 7.

In ADSL sidelobes have a more significant impact however OSM still yields significant gains over the other spectrum management algorithms.

A maximum transmit power of 20.4 dBm is applied to each modem[12]. The usual PSD constraint is not applied in the OSM or iterative waterfilling algorithms. A spectral mask is applied to the flat PBO and reference noise method and is set at -40 dBm/Hz[12]. Background noise includes crosstalk from 10 ISDN, 4 HDSL, and 10 SSM (legacy) ADSL disturbers.

Fig. 8 shows the rate regions corresponding to various spectrum management algorithms. For comparison the rate regions with iterative waterfilling, the reference noise method and flat PBO are shown. No PBO method for RT distributed ADSL modems has been defined in standardization and this is still an open issue[3]. A method for reducing the downstream transmit PSD known as the power cutback method is currently implemented in ADSL modems to prevent the receiver ADCs from being overloaded[12]. However on the 3 km RT distributed line this technique does not cause any reduction in the transmit PSD.

The PSDs corresponding to a 1 Mbps service on the CO distributed line are depicted in Fig. 9 and Fig. 10. The OSM PSD on the RT line decreases with frequency to reflect the increase in crosstalk coupling. This continues until 440 kHz where the CO line becomes inactive due to its low SNR above this frequency. Once the CO line is inactive we see a sudden increase in the OSM PSD on the RT line.

Examining the flat PBO PSD for the RT line we see that a large amount of PBO is required to protect the CO line. This is the case since the degree of PBO in the flat PBO algorithm cannot be varied with frequency as it was with OSM.

The iterative waterfilling algorithm gives similar results. Slightly less PBO is required since the CO line PSD has been boosted on the active tones as shown in Fig. 9. However the amount of PBO required is still much larger than with OSM. The iterative waterfilling algorithm does not exploit the fact that crosstalk coupling is low at low frequencies. It also does not exploit the fact that the CO line is inactive above 440 kHz. Both of these facts could have been exploited to increase the transmit PSD on the RT line at low and high frequencies, leading to increased performance without a large degradation in the performance of the CO line. Due to this the iterative waterfilling algorithm gives poor performance.

It has been shown that the reference noise method is near-optimal when the *signal to interference-plus-noise ratio* (SINR) is high[8]. This is the case in low frequencies. For this reason we see that the reference noise PSD matches the OSM PSD quite closely at frequencies below 440 kHz. This could be exploited to create a low-complexity, near-optimal spectrum management algorithm.

As shown in Tab. II using OSM instead of iterative waterfilling allows us to increase the data-rate on the RT distributed line from 3.6 Mbps to 7.4 Mbps whilst still maintaining a 1 Mbps service on the CO distributed line. This corresponds to a gain of over 200%

Note that the optimal rate regions for both the ADSL and VDSL scenarios are convex as we discussed in Sec. III-G.

C. Discrete Bitloading

We also ran the same simulations with discrete bitloading, with each modem forced to adopt an integer bitloading value. We set the maximum bitloading $b_{\max}$ to 14. All other simulation parameters were the same. In the iterative waterfilling algorithm we use the Levin-Campello algorithm to ensure integer bitloadings on each tone[13].

In the VDSL scenario, using OSM instead of iterative waterfilling allows the data-rate on the 600 m. lines to be increased from 3.4 Mbps to 13 Mbps whilst still maintaining a 5 Mbps service on the 1200 m. lines. This corresponds to a gain of over 380%

In the ADSL scenario, using OSM instead of iterative waterfilling allows us to increase the data-rate on the RT distributed line from 3.1 Mbps to 7.3 Mbps whilst still maintaining a 1 Mbps service on the CO distributed line.

VI. Conclusions

In this paper we presented an algorithm for optimal spectrum management (OSM) in DSL. This algorithm calculates the spectra required for the modems within a network to achieve maximal performance, thereby operating on the rate region boundary. The algorithm can operate under a combination of total power and/or spectral mask constraints and can use either continuous or discrete bitloading.

Through the use of a dual decomposition the algorithm solves the spectrum management problem independently on each tone. The result is a computationally tractable and efficient algorithm.

Simulations show that the algorithm yields significant gains over existing spectrum management techniques, e.g. in an upstream VDSL scenario the OSM algorithm can outperform another DSM algorithm *iterative waterfilling* by up to 380%

OSM is a centralized algorithm requiring a spectrum management center (SMC) for direct implementation. In future work it will be interesting to develop distributed DSM algorithms based on the insight gained from the OSM algorithm. The goal is to find a simple, distributed algorithm which yields near-optimal performance in a broad range of scenarios.

Whilst this paper has focused on the problem of spectrum management in DSL, the algorithm is also applicable to any communication system where inter-user interference is a problem. OSM could also be applied to broadband cable networks, high-speed Ethernets or fixed wireless links.

Appendix
Proof of Theorem 2

To prove the optimality of Alg. 1 as stated in Theorem 2, we first show that the algorithm converges. We then show that at convergence maximizing the Lagrangian is equivalent to maximizing the weighted rate-sum (7). This implies the optimality of the PSDs generated by the algorithm.

To prove the convergence of Alg. 1, we first examine the convergence of a related routine. This routine finds the optimal value for λ_n, thereby ensuring that the total power constraint on user n (4) is satisfied. At this value of λ_n the routine finds the optimal PSD for user n. As we show in Corollary 2 and 3, the algorithms *optimize_*λ_1 and *optimize_*λ_2 can be seen as special cases of this routine for specific values of the optimisation function $f(\mathbf{s}_n)$. In Lemma 3 we prove that this routine converges. This in turn implies the convergence of *optimize_*λ_1 and *optimize_*λ_2.

We first define the cost function

$$G(\mathbf{s}_n, \lambda_n) \triangleq f(\mathbf{s}_n) + \lambda_n (P_n - \sum_k s_k^n) \qquad (11)$$

Denote the optimal power allocation for a given λ_n

$$\mathbf{s}_n(\lambda_n) \triangleq \arg\max_{\mathbf{s}_n} G(\mathbf{s}_n, \lambda_n)$$

with $s_k^n(\lambda_n) \triangleq [\mathbf{s}_n(\lambda_n)]_k$. The routine for user n is then

Routine for user n

$\lambda_n^{\max} = 1$, $\lambda_n^{\min} = 0$
while $\sum_k s_k^n > P_n$
  $\lambda_n^{\max} = 2\lambda_n^{\max}$
  $\mathbf{s}_n = \arg\max_{\mathbf{s}_n} f(\mathbf{s}_n) + \lambda_n^{\max}(P_n - \sum_k s_k^n)$
end
repeat
  $\lambda_n = (\lambda_n^{\max} + \lambda_n^{\min})/2$
  $\mathbf{s}_n = \arg\max_{\mathbf{s}_n} f(\mathbf{s}_n) + \lambda_n(P_n - \sum_k s_k^n)$
  if $\sum_k s_k^n > P_n$, then $\lambda_n^{\min} = \lambda_n$, else $\lambda_n^{\max} = \lambda_n$
until convergence

To prove the convergence of this routine we make use of the following Lemma.

Lemma 2: $\sum_k s_k^n(\lambda_n)$ is monotonic decreasing in λ_n.

Proof: Consider two Lagrangian multipliers λ_n^a and λ_n^b and their corresponding optimal PSDs $\mathbf{s}_n^a \triangleq \mathbf{s}_n(\lambda_n^a)$ and $\mathbf{s}_n^b \triangleq \mathbf{s}_n(\lambda_n^b)$. Denote the elements of these PSDs as $s_k^{n,a}$ and $s_k^{n,b}$ respectively. Let

$$\lambda_n^b \geq \lambda_n^a \qquad (12)$$

Define

$$\begin{aligned} A &\triangleq f(\mathbf{s}_n^b) + \lambda_n^b (P_n - \sum_k s_k^{n,b}) \\ B &\triangleq f(\mathbf{s}_n^a) + \lambda_n^b (P_n - \sum_k s_k^{n,a}) \\ C &\triangleq f(\mathbf{s}_n^a) + \lambda_n^a (P_n - \sum_k s_k^{n,a}) \\ D &\triangleq f(\mathbf{s}_n^b) + \lambda_n^a (P_n - \sum_k s_k^{n,b}) \end{aligned}$$

Now $G(\mathbf{s}_n^b, \lambda_n^b) \geq G(\mathbf{s}_n^a, \lambda_n^b)$ by the optimality of $\mathbf{s}_n^b$ in $G(\mathbf{s}_n, \lambda_n^b)$. Hence $A \geq B$. Similarly the optimality of $\mathbf{s}_n^a$ in $G(\mathbf{s}_n, \lambda_n^a)$ implies $C \geq D$. Consider 3 cases:

In the first case let $P_n - \sum_k s_k^{n,a} \geq 0$. Combining this with (12) implies $B \geq C$. Now $A \geq B \geq C \geq D$ implies $A - D \geq B - C$. Hence

$$(\lambda_n^b - \lambda_n^a)(P_n - \sum_k s_k^{n,b}) \quad \geq \quad (\lambda_n^b - \lambda_n^a)(P_n - \sum_k s_k^{n,a})$$

which implies

$$\sum_k s_k^{n,a} \geq \sum_k s_k^{n,b} \qquad (13)$$

In the second case let $P_n - \sum_k s_k^{n,b} \leq 0$. Combining this with (12) implies $D \geq A$. Now $C \geq D \geq A \geq B$ implies $C - B \geq D - A$. Hence

$$(\lambda_n^a - \lambda_n^b)(P_n - \sum_k s_k^{n,a}) \geq (\lambda_n^a - \lambda_n^b)(P_n - \sum_k s_k^{n,b})$$

which again implies (13).

In the third case let $P_n - \sum_k s_k^{n,a} < 0$ and $P_n - \sum_k s_k^{n,b} > 0$. This implies $P_n - \sum_k s_k^{n,b} > P_n - \sum_k s_k^{n,a}$ and again leads to (13).

So in all cases a larger λ_n leads to a smaller $\sum_k s_k^n$. This implies that $\sum_k s_k^n$ is monotonic decreasing in λ_n. ∎

Lemma 3: Routine for user n converges. At convergence

$$\mathbf{s}_n = \arg\max_{\mathbf{s}_n} f(\mathbf{s}_n) \text{ s.t. } \sum_k s_k^n \le P_n$$

Proof: The routine consists of two stages: a preamble that determines $\lambda_n^{\max}$ and the actual routine itself.

The preamble clearly converges since $\sum_k s_k^n(\lambda_n) \to 0$ as $\lambda_n \to \infty$.

The convergence of the main part of *routine for user n* can be shown as follows: $\lambda_n^{\max} - \lambda_n^{\min}$ decreases by half in each iteration. Thus, λ_n converges to a fixed value. Let's now consider two cases, depending on whether $\sum_k s_k^n(\lambda_n^{\min}) > P_n$ or not.

Suppose that $\sum_k s_k^n(\lambda_n^{\min}) > P_n$ at $\lambda_n^{\min} = 0$, then since the preamble ensures that $\sum_k s_k^n(\lambda_n^{\max}) \le P_n$, throughout the algorithm it is always the case that $\sum_k s_k^n(\lambda_n^{\min}) > P_n$ and $\sum_k s_k^n(\lambda_n^{\max}) \le P_n$. Since $\lambda_n^{\max} \ge \lambda_n \ge \lambda_n^{\min}$, $\lambda_n^{\min}$ and $\lambda_n^{\max}$ converge to a fixed value, and since $\sum_k s_k^n(\lambda_n)$ is monotonic in λ_n, this implies that $\sum_k s_k^n(\lambda_n)$ must converge to P_n. On the other hand, suppose that $\sum_k s_k^n(\lambda_n^{\min}) \le P_n$ at $\lambda_n^{\min} = 0$. Then, λ_n will converge to 0.

Hence the algorithm will converge and at convergence either $\lambda_n = 0$ or $\sum_k s_k^n(\lambda_n) = P_n$. So at convergence

$$G(\mathbf{s}_n, \lambda_n) = f(\mathbf{s}_n)$$

In the routine

$$\begin{aligned} \mathbf{s}_n &= \arg\max_{\mathbf{s}_n} G(\mathbf{s}_n, \lambda_n) \\ &= \arg\max_{\mathbf{s}_n} f(\mathbf{s}_n) \text{ s.t. } \sum_k s_k^n \le P_n \end{aligned}$$

To see this clearly $\mathbf{s}_n$ satisfies the constraint. Further, if there is some other feasible $\mathbf{s}'_n$ that does better than $\mathbf{s}_n$ for the objective function $f(\mathbf{s}_n)$ then $\mathbf{s}'_n$ should do better than $\mathbf{s}_n$ for the objective $G(\mathbf{s}_n, \lambda_n)$ also. This is contradicted by the optimality of $\mathbf{s}_n$ in $G(\mathbf{s}_n, \lambda_n)$. Hence $\mathbf{s}_n$ must be optimal in $f(\mathbf{s}_n)$. ■

Lemma 4: The function *optimize_s* yields PSDs $\mathbf{s}_1$ and $\mathbf{s}_2$ which maximize the Lagrangian.

Proof: From function *optimize_s*

$$s_k^1, s_k^2 = \arg\max_{s_k^1, s_k^2} L_k(w, \lambda_1, \lambda_2, s_k^1, s_k^2)$$

Since $L = \sum_k L_k + \lambda_1 P_1 + \lambda_2 P_2$, and since optimisation of the Lagrangian is unconstrained (recall that the constraints are incorporated into the cost function and need not be explicitly enforced) this implies

$$\mathbf{s}_1, \mathbf{s}_2 = \arg\max_{\mathbf{s}_1, \mathbf{s}_2} L(w, \lambda_1, \lambda_2, \mathbf{s}_1, \mathbf{s}_2)$$

■

Define the rates of user 1 and user 2 with the PSDs $\mathbf{s}_1$ and $\mathbf{s}_2$ as $R_1(\mathbf{s}_1, \mathbf{s}_2)$ and $R_2(\mathbf{s}_1, \mathbf{s}_2)$.

Corollary 2: The function *optimize_*λ_2 converges. At convergence

$$\begin{aligned} \mathbf{s}_2 = \; & \arg\max_{\mathbf{s}_2} \max_{\mathbf{s}_1} wR_1(\mathbf{s}_1, \mathbf{s}_2) + (1-w)R_2(\mathbf{s}_1, \mathbf{s}_2) \\ & + \lambda_1(P_1 - \sum s_k^1) \\ \text{s.t. } & \sum_k s_k^2 \le P_2 \end{aligned} \tag{14}$$

Proof: Let $n = 2$ and $f(\mathbf{s}_2) \triangleq \max_{\mathbf{s}_1} wR_1(\mathbf{s}_1, \mathbf{s}_2) + (1-w)R_2(\mathbf{s}_1, \mathbf{s}_2) + \lambda_1(P_1 - \sum_k s_k^1)$. Lemma 4 implies that *optimize_*λ_1 and the routine are equivalent. Hence Lemma 3 implies *optimize_*λ_2 converges, and that at convergence (14) is satisfied. ■

Corollary 3: The function *optimize_*λ_1 converges. At convergence

$$\begin{aligned} \mathbf{s}_1 = \; & \arg\max_{\mathbf{s}_1} \max_{\mathbf{s}_2} wR_1(\mathbf{s}_1, \mathbf{s}_2) + (1-w)R_2(\mathbf{s}_1, \mathbf{s}_2) \\ \text{s.t. } & \sum_k s_k^1 \le P_1, \; \sum_k s_k^2 \le P_2 \end{aligned} \tag{15}$$

Proof: Let $n = 1$ and $f(\mathbf{s}_1) = \max_{\mathbf{s}_2} wR_1(\mathbf{s}_1, \mathbf{s}_2) + (1-w)R_2(\mathbf{s}_1, \mathbf{s}_2)$ s.t. $\sum_k s_k^2 \le P_2$. Then Lemma 3 and Corollary 2 imply *optimize_*λ_1 converges, and that at convergence (15) is satisfied. ■

From Lemma 1, for any particular w, there exists some R_1^{target} for which the weighted rate-sum optimization (7) is equivalent to the original spectrum management problem (10). Hence for any particular w the weighted rate-sum optimization leads to an optimal operating point.

Corollary 3 implies that for each value of w in Alg. 1, $\mathbf{s}_1$ and $\mathbf{s}_2$ maximize a weighted-rate sum. Hence they are also an optimal solution to (10). Furthermore, Theorem 1 states that by varying w from 0 to 1 we can map out all achievable operating points on the boundary of the convex hull of the rate region.

REFERENCES

[1] G. Ginis and J. Cioffi, "Vectored Transmission for Digital Subscriber Line Systems," *IEEE J. Select. Areas Commun.*, vol. 20, no. 5, pp. 1085–1104, June 2002.

[2] R. Cendrillon, M. Moonen, G. Ginis, K. Van Acker, T. Bostoen, and P. Vandaele, "Partial Crosstalk Cancellation Exploiting Line and Tone Selection in Upstream DMT-VDSL," *accepted for publication in EURASIP Journal on Applied Signal Processing*.

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

[4] J. Cioffi, *DSL Advances*. Prentice Hall, 2002, ch. 11 - Dynamic Spectrum Management.

[5] W. Yu, G. Ginis, and J. Cioffi, "Distributed Multiuser Power Control for Digital Subscriber Lines," *IEEE J. Select. Areas Commun.*, vol. 20, no. 5, pp. 1105–1115, June 2002.

[6] G. Cherubini, "Optimum upstream power back-off and multiuser detection for VDSL," in *Proc. IEEE Globecom*, 2001, pp. 375–380.

[7] K. Jacobsen, "Methods of upstream power backoff on very high-speed digital subscriber lines," *IEEE Commun. Mag.*, pp. 210–216, Mar. 2001.

[8] F. Sjoberg, S. Wilson, and M. Isaksson, "Power back-off in the upstream of VDSL," in *Proc. Radiovetenskaplig Koferens (RVK'99)*, vol. 1, Karlskrona, Sweden, June 1999, pp. 436–440.

[9] J. Lee, V. Sonalkar, and J. Cioffi, "A multi-user rate and power control algorithm for VDSL," in *Proc. IEEE Globecom*, 2003, pp. 1264–1268.

[10] *Very-high bit-rate Digital Subscriber Lines (VDSL) Metallic Interface*, ANSI Std. T1E1.4/2003-210R1, 2003.

[11] *Transmission and Multiplexing (TM); Access transmission systems on metallic access cables; VDSL; Functional Requirements*, ETSI Std. TS 101 270-1, Rev. V.1.3.1, 2003.

[12] *Asymmetrical digital subscriber line (ADSL) transceivers*, ITU Std. G.992.1, 1999.

[13] T. Starr, J. Cioffi, and P. Silverman, *Understanding Digital Subscriber Line Technology*. Prentice Hall, 1999.

[14] L. Xiao, M. Johansson, and S. Boyd, "Simultaneous routing and resource allocation via dual decomposition," July 2003, submitted to IEEE Transactions on Communications.

[15] W. Yu, "A Dual Decomposition Approach to the Sum Power Gaussian Vector Multiple Access Channel Sum Capacity Problem," in *Proc. 37th Annual Conf. on Information Sciences and Systems (CISS)*, March 2003.

Scheme	1200m. Rate	600m. Rate
Ref. PSD	2.7 Mbps	13.1 Mbps
Flat PBO	3.9 Mbps	0.0 Mbps
Ref. Noise	3.9 Mbps	0.0 Mbps
It. W.f.	5.0 Mbps	7.7 Mbps
OSM	5.0 Mbps	15.0 Mbps

TABLE I
ACHIEVABLE RATES IN UPSTREAM VDSL

Scheme	CO Rate	RT Rate
Flat PBO	1.0 Mbps	1.7 Mbps
Ref. Noise	1.0 Mbps	2.3 Mbps
It. W.f.	1.0 Mbps	3.6 Mbps
OSM	1.0 Mbps	7.4 Mbps

TABLE II
ACHIEVABLE RATES IN DOWNSTREAM ADSL

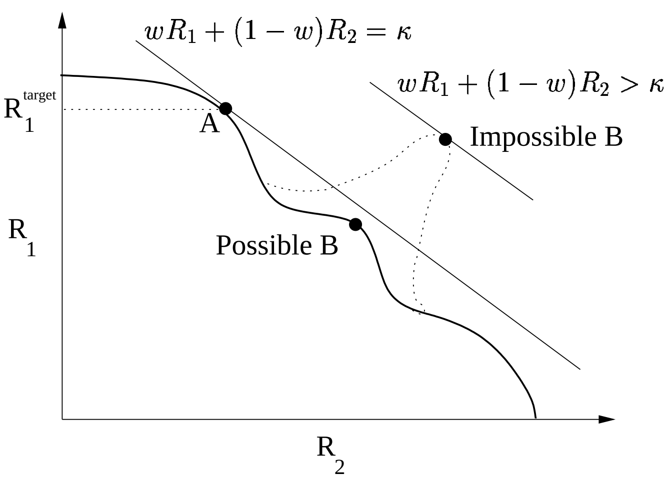

Fig. 1. Optimality of A in the weighted rate-sum (7) implies optimality in the spectrum management problem (3)

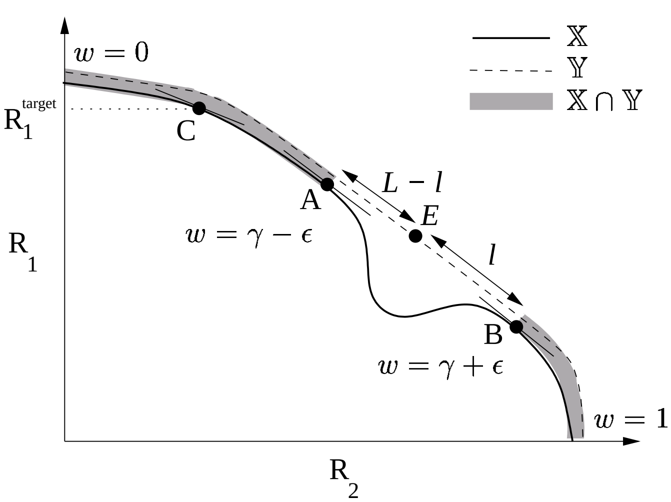

Fig. 2. Operating points in $\mathbb{X} \cap \mathbb{Y}$ can be found through a weighted rate-sum optimization

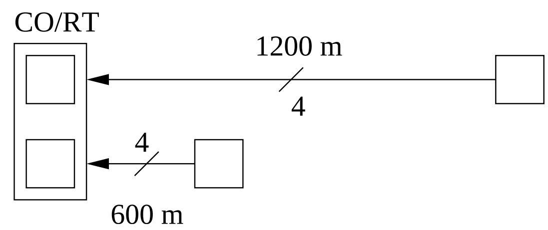

Fig. 3. Upstream VDSL Scenario

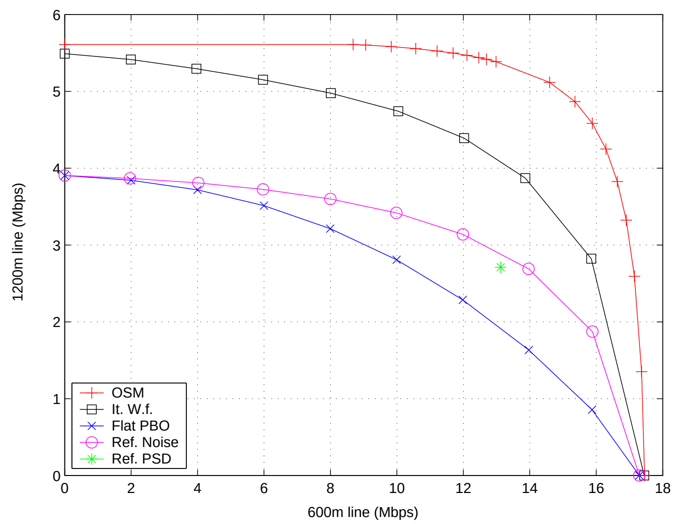

Fig. 4. Rate Regions in Upstream VDSL

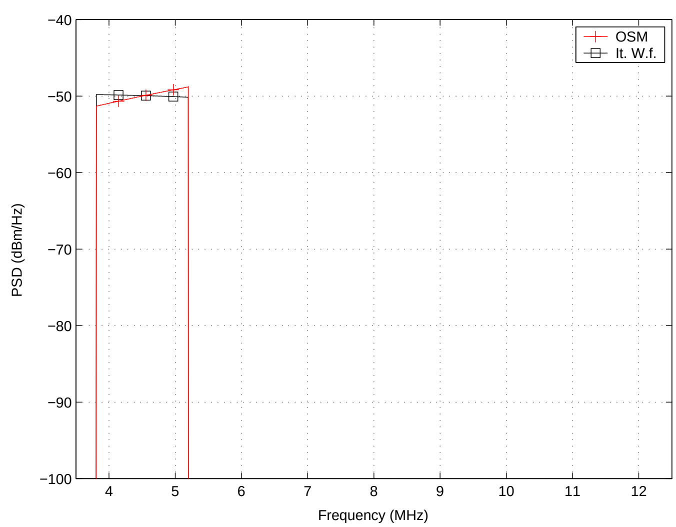

Fig. 5. PSDs on 1200 m. lines in Upstream VDSL (1200m. Line @ 5 Mbps)

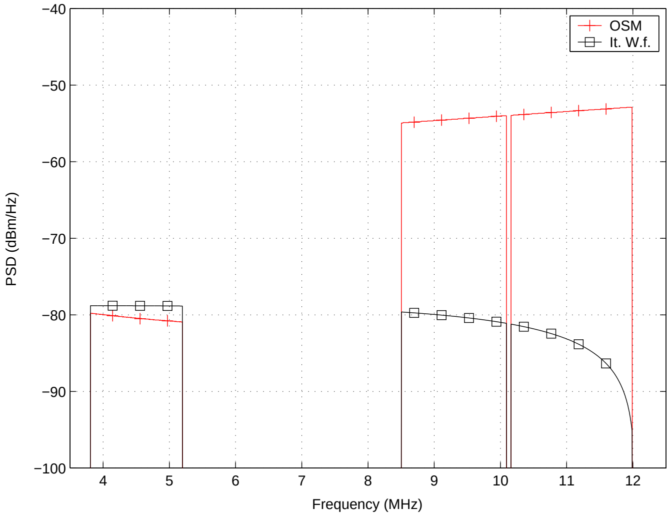

Fig. 6. PSDs on 600 m. lines in Upstream VDSL (1200m. Line @ 5 Mbps)

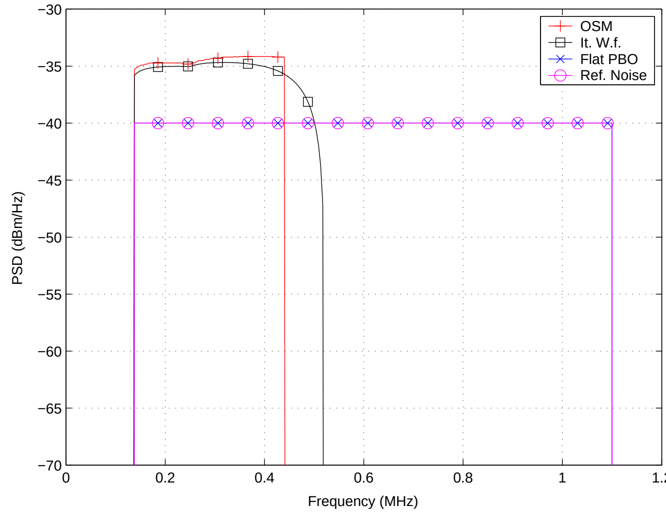

Fig. 9. PSDs on CO line in Downstream ADSL (CO Line @ 1 Mbps)

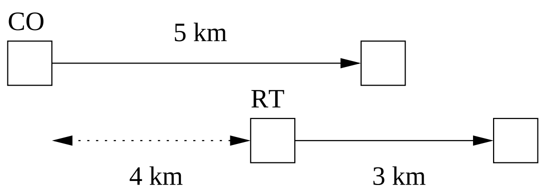

Fig. 7. Downstream ADSL Scenario

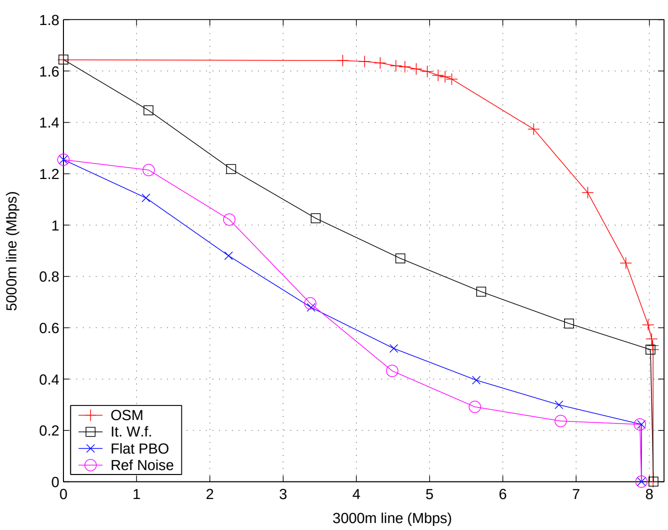

Fig. 8. Rate Regions in Downstream ADSL

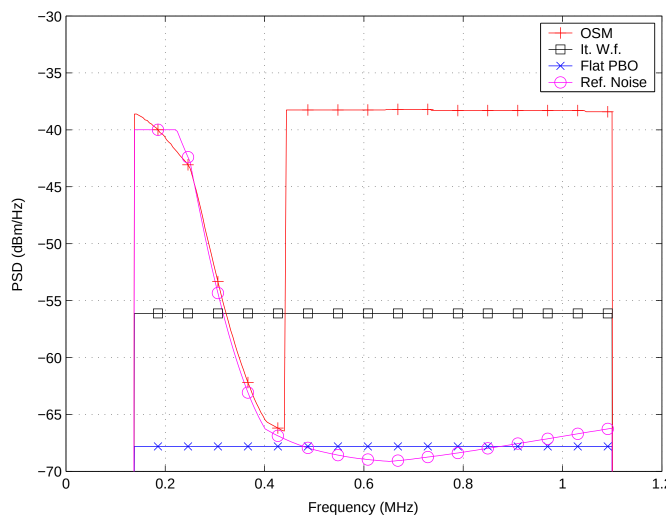

Fig. 10. PSDs on RT line in Downstream ADSL (CO Line @ 1 Mbps)