**ATIS Committee NIPP-NAI**
**Working Group NAI (DSL Access)** **NIPP-NAI-2007-056R1**
**Minneapolis, MN: May 14, 2007**

**CONTRIBUTION**

TITLE: **DSM Level 2 Distributed Band Preference**

SOURCE:

| | | | |
|---|---|---|---|
| **J. Cioffi** | **Stanford** | +1-650-723-2150 | cioffi@stanford.edu |
| **W. Lee** | | | |
| **Y. Kim** | | | |
| **M. Brady** | **ASSIA** | +1-650-654-3400 | mbrady@assia-inc.com |


PROJECT: **T1E1.4, DSM, Information only**

**ABSTRACT**

**This contribution examines preferential treatment of selected frequency bands in DMT DSL's bit-swapping and loading procedures. The result is an example of a nearly optimally performing DSM Level 2 distributed loading method that is a minor adjustment to water-filling when the margin-cap (or "politeness") indication is on. The gains of this distributed method over Level 1 DSM occur with only the use of an indication of "margin cap" on or off by an SMC, completely consistent with the margin-cap definition in the DSM report. With an additional transfer of infrequently, but centrally, computed band preferences, optimal performance can be obtained in all situations. Such preferences can be easily distributed by re-interpretation of existing non-politeness-pertinent VDSL2 fields when politeness (margin cap) is on.**



- CONTACT: J.M. Cioffi, Cioffi@stanford.edu, Tel: +1-650-723-2150 ; Fax: +1-650-724-3652

# DSM Level 2 Distributed Band Preference (056R1)

**J. Cioffi, W. Lee, Y. Kim**

Dept of EE, Stanford University
363 Packard EE, Stanford, CA 94305-9515
Cioffi@stanford.edu,
P: +1-650-723-2150 ; Fax: +1-650-724-3652

**M. Brady**

ASSIA Inc.
203 Redwood Shores Parkway , Suite 250
Redwood City, CA 94065-1198
gginis@assia-inc.com ,
+1-650-654-3400 , Fax: +1-650-654-3404

**ABSTRACT**

**This contribution examines preferential treatment of selected frequency bands in DMT DSL's bit-swapping and loading procedures. The result is an example of a nearly optimally performing DSM Level 2 distributed loading method that is a minor adjustment to water-filling when the margin-cap (or "politeness") indication is on. The gains of this distributed method over Level 1 DSM occur with only the use of an indication of "margin cap" on or off by an SMC, completely consistent with the margin-cap definition in the DSM report. With an additional transfer of infrequently, but centrally, computed band preferences, optimal performance can be obtained in all situations. Such preferences can be easily distributed by re-interpretation of existing non-politeness-pertinent VDSL2 fields when politeness (margin cap) is on.**



## 1. Introduction

Level 2 DSM's spectrum balancing can effect large data rate gains through an Spectrum Maintenance Center's (SMC's) central imposition of politeness upon users that otherwise transmit too much power in their victim's essential spectrum band. A theoretical bound on Level 2 performance was introduced in [1], but unfortunately called for a highly complex centralized bit-swapping mechanism that could not realistically be considered for implementation. Reference [2] simultaneously introduced that specification of preferred transmission band for some DSLs, but otherwise allowing them to adapt autonomously with water-fill-like loading algorithms, offered a practical alternative called "band preference." Many references (see the appendix) have subsequently attempted to address mechanisms for such band preference, but have commonly still required a level of message passing that render dubious their realization. This contribution provides a distributed loading algorithm that requires an SMC to indicate only which of two closely related loading algorithms should be used autonomously by each user. The distributed band-preference performance is essentially that of [1], thus facilitating practical realization of Level 2 DSM.

Fortunately, after a period of debate, the current released DSM report [3] allows for indication of an exceptionally polite loading algorithm under the name of "margin-cap mode" or "margin-cap indication." This margin-cap indicator is used in the band-preference method of this contribution to alert certain DSL's to be exceptionally polite while others may and be less polite by operating with the margin-cap mode off. Whether using band preference (margin cap on) or not (margin cap off), all spectrum and power limits are always observed.

Normal water-filling-like loading algorithms can be implemented via bit-swapping to select spectra that obeys power-spectral-density masks, power limits, and margin limits (at all frequencies is still good even) when the margin-cap indication is off. However, when the SMC directs a user to observe margin-cap, the loading algorithm bears an additional restriction that it should load more bits to bands that water-filling would otherwise have limited. When the SMC alerts specific users to behave in said polite manner, those users give preference to frequency bands of relatively lower nominal SNR if their specified DSL service data rate can be so achieved. In so doing, those polite users reduce crosstalk into their victims autonomously as best as they can.

Rather than completely detail the loading algorithms in the main body, this contribution provides a complete description in the appendix of the algorithms. Instead, the main body focuses on the basic concept in Section 2, showing the use of the margin-cap mode. Section 2 also suggests an optional re-use of some unused fields in VDSL2 that are better used to pass some information that can provide additional gains but still avoid central swapping. Section 3 then provides performance results.

## 2. Band Preference Basics

Figure 1 illustrates the basic band preference concept. This example has been well studied in DSM and

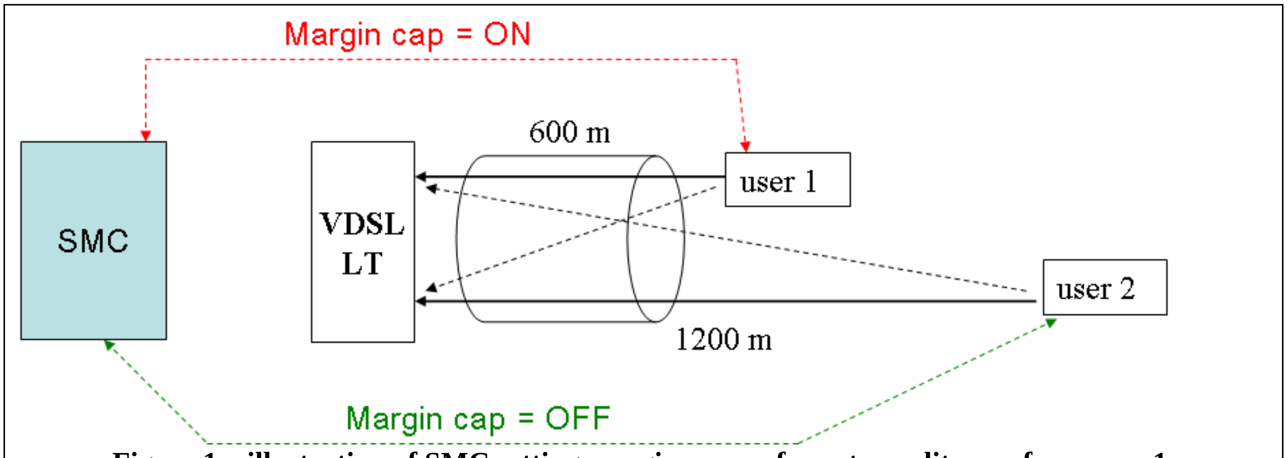



**Figure 1 – illustration of SMC setting margin cap on for extra politeness from user 1.**

corresponds to a situation where Level 2 DSM can provide an enlarged set of possible data rates for both users (see Annex C of the DSM Report [3]). The SMC recognizes this "near/far" like situation for upstream DSL via data collected on the DSM-D interface and subsequently requests that User 1 be polite by indicating margin-cap mode on the DSM-C interface. (The DSM-C and DSM-D interfaces are not explicitly shown, but correspond to the dashed red and green lines in Figure 1.). User 2 need not be extra polite and does the best it can within the existing VDSL2 spectrum and power limits. User 1 can be polite by avoiding frequency bands that victimize user 2, which would be lower frequencies in this case. User 1 then attempts its data rate at higher frequencies than it would nominally use in a crosstalk-free environment. It uses higher frequencies first and achieves its data rate by expanding into lower frequencies only insofar as is necessary to achieve its SMC-specified data rate.

There are a variety of loading algorithms that can be used with band preference. The basic water-filling algorithm is:

$$Power(n) = K - \frac{Noise(n)}{\Gamma \cdot |Channel(n)|^2} \quad , \qquad (1)$$

where the power is for all tones n such that it is either zero or the difference between a frequency-independent constant K and the scaled (by the gap $\Gamma$) channel noise-to-signal ratio. This method is best for single-user DSL and well-known throughout the DSL industry. There are various practical approximations to it, but the basic spectrum looks as in (1) when margin-cap is OFF. The constant *K* is the "water level."

This contribution investigates a scaled water-filling that differs in that the constant *K* is different for each of a small set of frequency bands.

$$Power(n) = \frac{K}{\alpha_m} - \frac{Noise(n)}{\Gamma \cdot |Channel(n)|^2} \quad , \qquad (2)$$

where $\alpha_m$ is a positive scale factor that may be different in selected frequency bands $m=1,\ldots,M$. The scalar $\alpha_m$ may be correctly interpreted as a "band preference" factor with smaller values $\alpha_m < \alpha_k$ corresponding to a statement of increased preference to load in band *m* over band *k*. Normal water filling can maximize a sum of data rates over the DMT tones of a DSL modem. Scaled water-filling maximizes instead a weighted sum of data rates over the DMT tones, where the weighting factors in the sum for each band are the scale factors $\frac{1}{\alpha_m}$. Clearly if all the scale factors are the same value, then scaled water-filling is water-filling. In effect, the water-level is different in different bands for scaled water-filling according to the preference specified by the scale factors $\alpha_m$. Those using various practical approximation to water-filling, typically with "incremental energy" tables will find that those tables are simply scaled by the $\alpha_m$ in different bands and their algorithms proceed just as simply and efficiently as before.

This scaling begs a question: "Where does the loading algorithm get the $\alpha_m$?" This question is answered in two ways in this contribution:

(1) from an SMC (which then requires additional parameters)
(2) computes them itself when margin cap is on

The latter self-computation method is called Algorithm 2 in the appendix and analyzed there. Essentially, though, the algorithm will initially run normal water filling and then "de-load" the best frequencies by shifting (swapping) the bits on the best tones to bands (upper frequencies in the example of User 1 in Figure 1). Once the maximum number of such swaps has occurred without violation of any power or power-spectral-density constraints, the bands and the scale factors are determined autonomously by any modem with margin-cap on. This method's peformance is very close to having an SMC determine the $\alpha_m$ and distribute them, but not quite optimum. If it is desirable to get the last performance gains, then an SMC can centrally and infrequently determine $\alpha_m$ for each user and distribute them. Such distribution could be done for instance in VDSL2 by re-interpreting other parameters not used with a margin-cap observing modem – for instance, virtual noise would not be used by an SMC that is using politeness instead with margin cap, so that field is easily sufficient to instead pass scale factors to the modems if the last small performance improvement is desired.

An interesting point of scaled water-filling is that the modems continue to operate and swap. If conditions change, the autonomous method above (Algorithm 2 and method 2) simply re-determines a new set of $\alpha_m$ that represent its best effort at politeness for the changed condition. In no situation would the modems be disabled from swapping, and in all situations near maximum politeness is maintained. Intermittent noises and impulse noises should be handled with FEC (and not by raising phantom or virtual noise levels, which is impolite behavior for any multi-user environment)>

## 3. Some Performance Results

First, one should recall that when most lines in a binder have about the same length that there is little advantage to optimum spectrum balancing (OSB) over simple single-water-level water-filling loading algorithms, each implemented with a maximum margin constraint. Thus, in such situations the SMC would either disable margin-cap mode or enable it possibly if scale factors are distributed (in which case they'd all likely be 1). Such methods are here (and elsewhere) best known as "iterative water-filling" or IWF. However, when lines have greatly different lengths, IWF does not provide the full improvement over impolite present-day DSLs that often do not observe any margin-capped politeness. Figure 1 is a classic example of such a situation for upstream VDSL. For clarity, DBPSM is obtained by having each modem compute its own weights (as defined in Section 2), while BPSM is when the SMC computes the weights.

Thus, Figure 2 provides the upstream rate regions for this situation of Figure 1. As is clear, the IWF with 6 dB margin limits works as shown (which is much better than no margin limits, a common situation today even in VDSL2 where the longer user essentially gets no data rate. However, this contribution concentrates on the additional improvement for Level 2. The DBPSM (Distributed Band Preference Spectrum Management) obtains nearly the same rate region. By distributing the scale factors by an SMC, an additional 5% rate gain can be obtained for the shorter user.

The spectra of the two users for both OSB and DBPSM are shown in Figures 3(a) and 3(b) where it is clear that the small difference is caused by the approximately flattening of the lower used frequency band in DBPSM. That said, with distribution of scale factors, BPSM is the same as OSB.

A natural question arises to multiple users, which are easily accommodated in DBPSM (unlike OSB, which has severe complexity limitations that grow exponentially with the number of users and can be intractable to compute). However, it is difficult to plot multiple-user rate regions in this case, so Table 1 below provides some sample results:

Table 1(a) - Data rate for 1200 m lines (one 600 m line, two 1200 lines)

| 600 m line | OSB | BPSM | DBPSM | IWF |
|---|---|---|---|---|
| 35 Mbps | 4.48 Mbps | 4.31 Mbps | 3.89 Mbps | 1.61 Mbps |
| 33 Mbps | 4.90 Mbps | 4.76 Mbps | 4.52 Mbps | 1.75 Mbps |

Table 1(b) - Data rate for 1200 m line (two 600 m lines, one 1200 line)

| 600 m line | OSB | BPSM | DBPSM | IWF |
|---|---|---|---|---|
| 24.5 Mbps | 4.27 Mbps | 4.17 Mbps | 4.01 Mbps | 1.55 Mbps |
| 22.8 Mbps | 5.08 Mbps | 4.77 Mbps | 4.70 Mbps | 1.80 Mbps |

## 4. Discussion and Conclusions

Band-Preference loading algorithms when the margin-cap mode is set by an SMC can be a very effective way to gain the full benefit of Level 2 DSM. More complicated situations can be investigated and two users only were used for simplicity of presentation. The appendix discusses partitioning of more users into two groups, strong and weak. Strong users then would have margin caps set and band preference algorithms applied while weak users would not be asked to use band preference and thus margin cap would be off.

Level 2 DSM extends Level 1 DSM's basic line management, diagnostics, and simple iterative water-filling spectrum management to effect large data rate gains in situations of mixed-length DSL binders. DBPSM provides a realizable and simple method to obtain such best Level 2 performance.

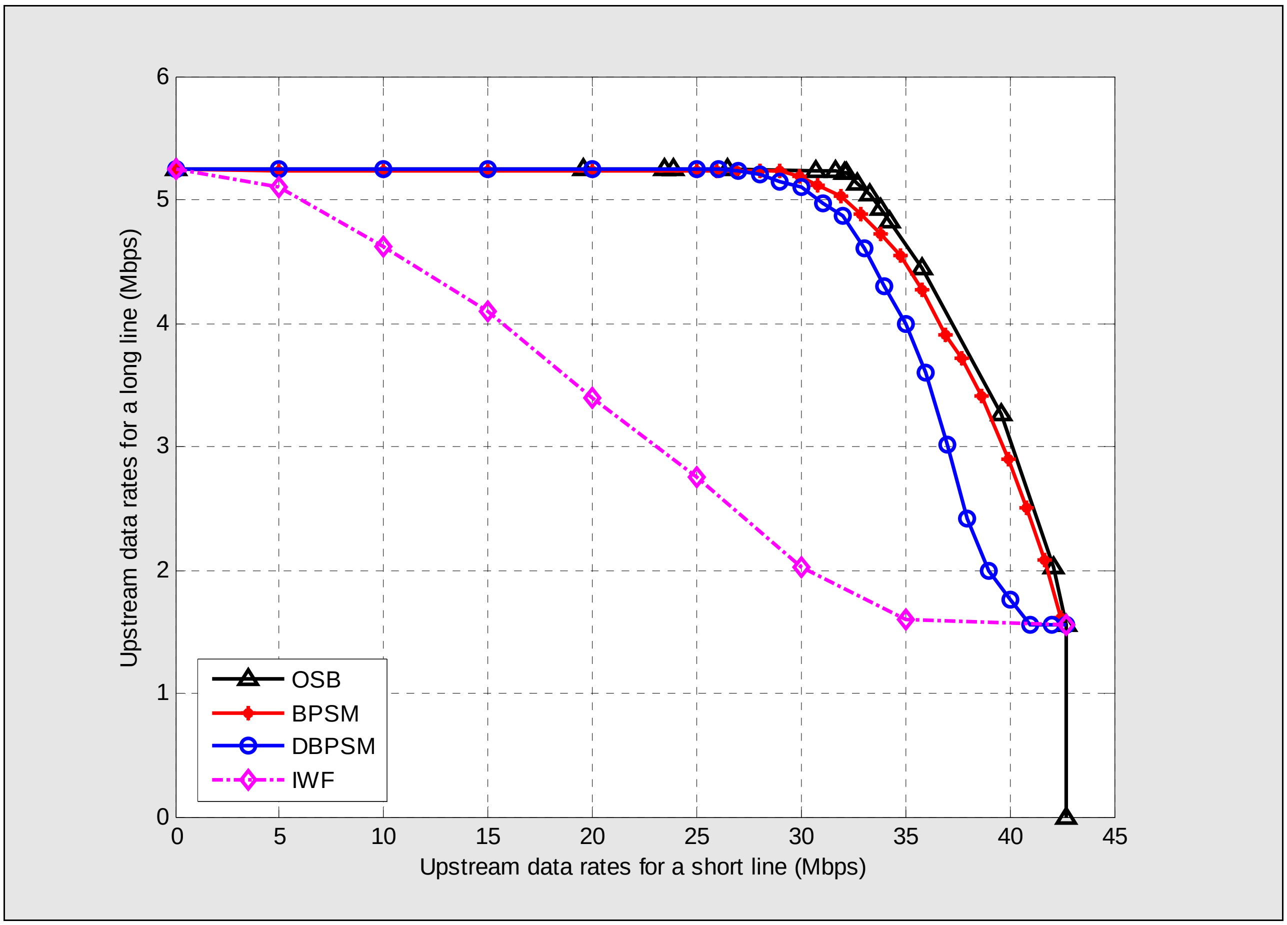


Figure 2 – Performance of various methods for the situation of Figure 1.

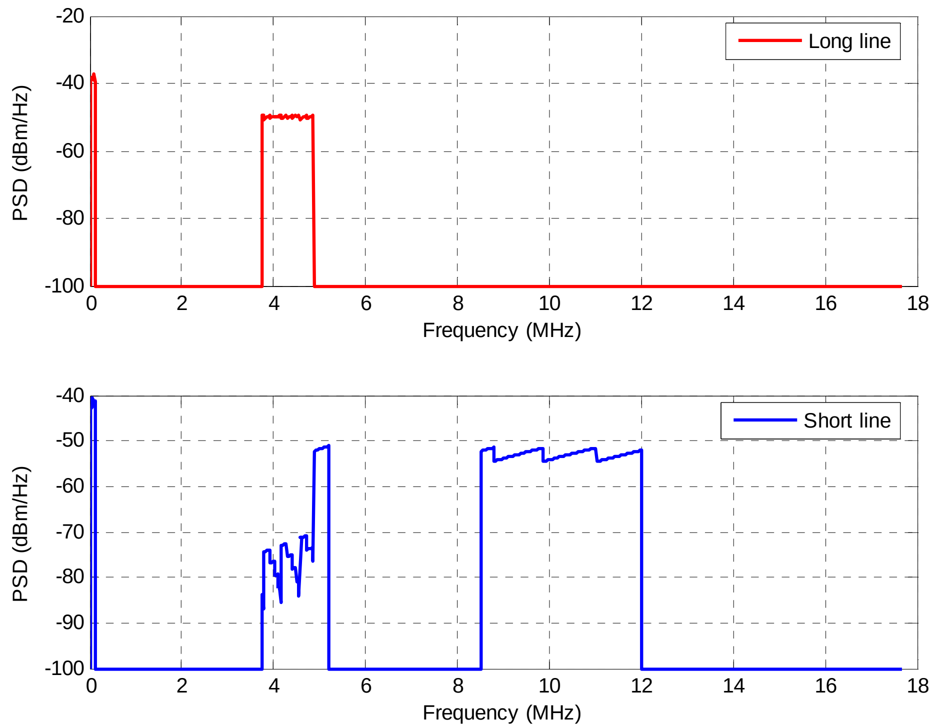



**Figure 3(a) – OSB spectra for Figure 1.**

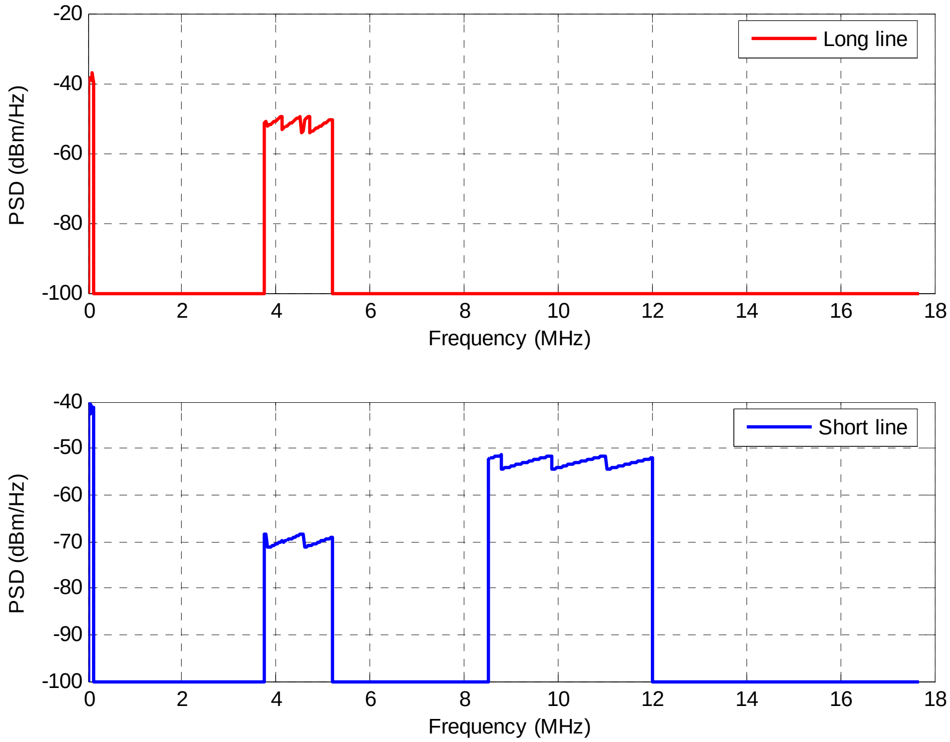



**Figure 3(b) – DBPSM spectra for Figure 1**

## 5. References


[1] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, T. Bostoen, "Optimal Spectrum Management," *ANSI T1E1.4 Working Group (DSL Access) Meeting*, contrib. 2004-459, North Carolina, May 2004. (Now part of the DSM Report T1E1.4/2003-018RE).

[2] J. Cioffi and M. Mohseni, ``Preference in Water-filling with DSM,'' ANSI contribution T1E1.4-2003-321R1, San Diego, CA, December 8, 2003.

[3] DSM (Draft) Technical Report, *ATIS NIPP-NAI Contribution 2006-028R2*, June 2006, Savannah, GA.

# Distributed Band-Preference Dynamic Spectrum Management in a DSL Environment

Wooyul Lee, Youngjae Kim, Mark H. Brady and John M. Cioffi

Department of Electrical Engineering
Stanford University, Stanford, CA 94305
Email: {wylee, youngjae, mhbrady, cioffi}@stanford.edu

***Abstract*— This paper introduces an algorithm for spectrum management for digital subscriber line (DSL) systems based on band preference (BPSM). The proposed method influences the usage of spectrum through band-preference factors that subtly modify the loading algorithm of DSL modems. Ad-hoc algorithms for computing such band-preference factors are discussed. When the band-preference factors are not centrally computed, a fully distributed band preference algorithm (DBPSM) is proposed to compute those factors with only local information. Simulation results in a practical very high speed DSL (VDSL) environment show that the performance of BPSM is better than that of Iterative Water-filling (IWF) [1] and is close to that of Optimal Spectrum Balancing (OSB) [2], even with a small number of control parameters. The results also show that the performance of DBPSM is also close to that of BPSM.**

## I. Introduction

While DSL systems are widespread in today's data access networks, there still exist several barriers to achieving higher data rates. Chief among these barriers is Far-End Crosstalk (FEXT), which is the electro-magnetic interference from other same-direction users in the binder. In order to mitigate FEXT, current ADSL systems rely on a Static Spectrum Management (SSM) scheme to set power spectral density masks (PSDMASKs) for all the modems [3]. PSDMASKs limit each modem's transmitted power so that its FEXT into other users can be guaranteed to be lower than an acceptable level. However, this form of static spectrum management must be designed conservatively, and thus its overall performance is much lower than what can be achieved by Dynamic Spectrum Management (DSM).

Techniques for DSM may be stratified into three levels of coordination [4]. In Level 1 DSM, such as Iterative Water-filling (IWF), each user views other users' signals as noise and seeks to maximize its data rate in a fully distributed manner. Because each user runs a single-user water-filling process in IWF, it does not require any central controller and has a much lower complexity compared to other DSM algorithms. However, Level 1 DSM does not perform well particularly in a near-far situation, which is commonly encountered in remote-terminal-deployed ADSL systems and upstream VDSL systems. In these situations, the users close to the central office have strong crosstalk channel gains as well as strong direct channel gains. Therefore, the strong crosstalk signals can significantly interfere with other users' signals unless the spectrum is appropriately managed. However, because each modem maximizes its own data rate without any information about the other modems, that management is not feasible in Level 1 DSM. To address this problem, a Spectrum Management Center (SMC) coordinates the spectra of all modems centrally in Level 2 DSM, which is mainly studied in this paper. In Level 3 DSM, complete coordination, or 'vectoring' occurs if all modems terminate at the same multiplexor, resulting in a vectored or "MIMO" channel [5]. Therefore, FEXT can be cancelled at the receiver (transmitter) for upstream (downstream) transmission via the QR decomposition method. The achievable rate of this FEXT-cancellation method is known to be close to the information theoretic bounds of the MIMO channel.

This paper considers Level 2 DSM, for which much work has been undertaken. At this level, the "optimal spectrum balancing" (OSB) algorithm attempts to maximize the weighted sum rate of all users [2]. Because the signal is not coordinated in Level 2 DSM, FEXT is considered as a noise without being cancelled. Under this assumption, the problem to maximize the weighted sum rate is non-convex, and it is difficult to find optimal solutions. An exhaustive search method is infeasible for this problem because of the large number of tones and users. However, using Lagrange dual-decomposition method, the problem can be decomposed into per-tone optimization problems in the OSB, which has a linear complexity in terms of tones. Although the solution of the dual problem is generally different from the solution of the original non-convex problem, the gap between these solutions is shown to diminish as the number of tones increases [6].

The exponential complexity in the number of users, however, prevents the application of the OSB for a practical number of users. Several methods for reducing the OSB's exponential complexity have been reported. An iterative approach is considered in [7], [8], and a message-passing algorithm is proposed where the complexity is reduced through successive convex relaxations in [9]. In this level of DSM, several discrete bit-loading algorithms have been also proposed. The problem of how to minimize the total power was studied in [10], [11] and an efficient discrete bit-loading algorithm that initializes with a continuous algorithm such as the "SCALE" was proposed in [12].

However, these methods in Level 2 DSM require central controllers to compute and update PSD or messages, which results in significant control overheads that may be limiting when system parameters change rapidly. Band Preference Spectrum Management (BPSM) avoids these problems by instead relying on the inherent adaptive capability of each of the DSL modems. A central controller infrequently (e.g. on a daily or weekly basis) communicates to each modem which frequency bands are preferable (and conversely undesirable) for loading. Cognizant of these "band preferences", each DSL modem then autonomously adapts to any subsequent channel variations. Thus, BPSM significantly reduces control overhead while allowing a largely distributed implementation. [13] discusses a different form of BPSM based on setting PSDMASKs.

Other techniques for mitigating the control and overhead problem have been studied in [14]. In that work, instead of solving a global optimization problem for all the users, each user solves a local problem that maximizes the rate of a reference line while achieving its own rate target. The reference line particularly represents weak users in the network, and the information about the reference line including the channel gains and background noise is infrequently distributed. Therefore, each user can autonomously determine its PSD without any centralized control, although the overall performance depends on a proper selection of the reference line.

The novel approach of the BPSM algorithm proposed here is to employ power-scaling factors instead of a PSDMASK [13] or the reference line [14]. These scaling factors are, heuristically speaking, penalties that are given to tones. During the bit-loading process, a modem usually finds the tone that requires minimum energy to load a new bit. Under the proposed algorithm, the modem instead finds the tone that requires minimum *penalized* energy. If a central controller determines that it is desirable for some tones to load a smaller number of bits (*e.g.* to protect other users from FEXT), large scaling factors may be given to those tones. In this way, the spectrum can be managed without direct control of each modem. To further reduce the controlling overhead, adjacent tones are grouped into one band and those tones in the band share one scaling factor. Because adjacent tones tend to have similar direct and crosstalk channel gain as well as receive similar interference from alien systems such as Integrated Services Digital Network (ISDN), High data rate DSL (HDSL), and other ADSL lines, grouping those tones degrades the performance of BPSM at the minimum. Therefore, the SMC can manage the DSL network with a minimum distribution of control messages while each modem still keeps its ability to adapt to any frequent changes.

To decrease further the amount of control messages, Distributed BPSM (DBPSM) is also proposed in this paper. In the previously mentioned methods, the necessary information for spectrum management is computed and distributed by the SMC, and each user determines its PSD based on the received information. However, that information may not be available when control channels can not accommodate those messages. In this case, the one possible option is to let each user run the water-filling algorithm, which does not require any information although the performance is significantly degraded in near-far situations. Recently, IWF with adaptive band was proposed to improve the performance of these situations by adjusting available tones of strong users [15]. In the DBPSM, each user autonomously determines its own scaling factors assuming that each user knows whether it has to be polite to others users or not. Because the strong users interfere with weak users and degrade their performance, they should be polite to weak users by avoiding strong interference to the tones that the weak users prefer or mainly use. However, the strong users are usually unaware of the existence of the weak users, and can not decide by themselves whether they should be polite or not. Therefore, the SMC helps the decision of each user by sending a bit that indicates the situation. Once the users know that they should be polite, they load bits in a polite way to other users. Meanwhile, the users who are not requested to be polite run the normal water-filling algorithm. In this way, the SMC only needs to send a bit to each user, and each user autonomously manages its spectrum based on its own information, which may be enough because of the similarity of DSL channel statistics.

The remainder of the paper is organized as follows: Section II introduces the system model of multi-user DSL systems and formulates the problem. Section III details the proposed band-preference algorithm and Section IV shows an ad-hoc algorithm to compute the scaling factors. Section V introduces a fully distributed BPSM. Section VI presents simulation results and Section VII concludes the paper.

## II. System Model And Problem Definition

This paper considers a multi-user Discrete Multi-Tone based (DMT) DSL system of $L$ users, which models a copper-wire binder group. For each tone, the channel can be expressed as a linear system as follows:

$$y_n^i = \sum_{i=1}^{L} H_n^{i,j} x_n^i + n_n^i \quad (i = 1, \cdots L,\ n = 1, \cdots, N), \quad (1)$$

where $H_n^{i,j}$ is the $(i,j)^{\text{th}}$ entry of the channel matrix that represents crosstalk from the transmitter $j$ to the receiver $i$, $y_n^i$ is the output of user $i$, $x_n^j$ is the input of user $j$, $n_n^i$ is the noise of user $i$ at tone $n$, and $N$ is the total number of used tones.

In this model, no signal coordination is assumed between lines and the signals from other users are treated as noise; such a multi-user channel is often called an "interference channel". Under this assumption, the rate of user $i$ is proportional to:

$$\begin{aligned} b_n^i &= \log_2 \left( 1 + \frac{1}{\Gamma} \frac{h_n^{i,i} p_n^i}{\sigma_n^2 + \sum_{j \neq u} h_n^{i,j} p_n^j} \right) \\ &= \log_2 \left( 1 + \frac{1}{\Gamma} \cdot g_n^i \cdot p_n^i \right) \quad \text{(bits/dim)} \\ R^i &= \sum_{n=1}^{N} b_n^i, \end{aligned} \quad (2)$$

where $h_n^{i,j} = |H_n^{i,j}|^2$, $g_n^i = h_n^{i,i} / \left(\sigma_n^2 + \sum_{j\neq i} h_n^{i,j} p_n^j\right)$ is the normalized channel gain, $p_n^i$ is the transmit power user $i$ at tone n, and $\Gamma$ is the implementation gap.

## III. Band-Preference Algorithm

### A. Power Scaling Factors, $\alpha_n$

The case of a single DSL modem is first considered. A good DMT modem, in the absence of PSD masks, loads bits to approximate the following "water-filling" condition in a tone set $E = \{1, \ldots, N\}$

$$p_n^{wf} = \left(K_1 - \frac{\Gamma}{g_n}\right)^+, \quad n \in E, \qquad \sum_{n\in E} p_n^{wf} = P,\ K_1 \geq 0 \tag{3}$$

where $P$ is a total power constraint, and $(x)^+ \triangleq \max(x, 0)$. We will say that water-filling is the process by which $p_n^{WF}$ and $K_1$ satisfying (3) are found (for given $\Gamma,\ g_n$ and $P$). Scaling factors that modify (3) are introduced as follows:

*Definition 1:* The *scaled water-filling condition* is said to hold when the following conditions are satisfied

$$p_n^{swf} = \left(\frac{K_2}{\alpha_n} - \frac{\Gamma}{g_n}\right)^+, \quad n \in E, \qquad \sum_{n\in E} p_n^{swf} = P,\ K_2 \geq 0 \tag{4}$$

Accordingly, the process of finding $p_n^{wf}$ and $K_2$ that satisfy (4) (for given $\Gamma, \mathbf{g}, \boldsymbol{\alpha}$ and $P$) is termed scaled water-filling. For $\boldsymbol{\alpha} = \mathbf{1}$, scaled water-filling is equivalent to water-filling.

In (4), the factors $\boldsymbol{\alpha}$ may be interpreted as a tone-dependant penalty that is useful for controlling the modem's power on that tone. For all $n$ such that $\alpha_n = \infty$, observe that one must have $p_n = 0$ in order that the scaled-water-filling condition holds. An intuitive interpretation is that setting the penalty on tone $n$ (namely $\alpha_n$) to $\infty$ has the effect of disabling tone $n$.

The water-filling and scaled water-filling conditions may be generalized to the setting where the modem has PSD masks $C_n$. This is a strict generalization of (3) and (4) because PSD masks are redundant if larger than the total power constraint ($\mathbf{C} \succ \mathbf{1} \cdot P$). This generalized setting will be considered in the remainder of the paper. The water-filling condition (3) generalizes to

$$\begin{aligned}
&K_2 \geq 0, \nu \succeq 0,\\
&0 \geq \left(K_1 - \frac{\Gamma}{g_n}\right), \nu_n = 0, \quad \text{if } p_n = 0\\
&p_n = \left(K_1 - \frac{\Gamma}{g_n}\right), \nu_n = 0, \quad \text{if } 0 < p_n < C_n,\\
&p_n = \left(\frac{1}{1/K_1 + \nu_n} - \frac{\Gamma}{g_n}\right), \nu_n \geq 0, \quad \text{if } p_n = C_n,
\end{aligned} \tag{5}$$

for each $n$. The (generalized) scaled water-filling condition (4) is therefore defined[1] as

*Definition 2:* For fixed scaling factors $\boldsymbol{\alpha} \in \overline{\mathbf{R}}^N$, $\boldsymbol{\alpha} \succeq \mathbf{1}$, the (generalized) *scaled water-filling condition* is defined as

$$\begin{aligned}
&K_2 \geq 0,\ \nu \succeq 0,\\
&0 \geq \left(\frac{K_2}{\alpha_n} - \frac{\Gamma}{g_n}\right), \nu_n = 0, \quad \text{if } p_n = 0\\
&p_n = \left(\frac{K_2}{\alpha_n} - \frac{\Gamma}{g_n}\right), \nu_n = 0, \quad \text{if } 0 < p_n < C_n,\\
&p_n = \left(\frac{1}{1/K_2 + \nu_n} \cdot \frac{1}{\alpha_n} - \frac{\Gamma}{g_n}\right), \nu_n \geq 0, \quad \text{if } p_n = C_n,
\end{aligned} \tag{6}$$

for each $n$. Note again that under the scaled water-filling condition (6), $\alpha_n = \infty$ implies that $p_n = 0$, and that (6) with $\boldsymbol{\alpha} = \mathbf{1}$ is equivalent to (5).

The following two theorems show how spectral allocation may be controlled using the scaling factors. The first theorem shows that for any channel and any full-power PSD, there exist scaling constants such that the scaled water-filling condition holds.

*Theorem 1:* For any given PSD $\mathbf{p} \in \mathbb{R}_+^N$ where $\sum_n p_n = P$, and channel gains $\mathbf{g} \in \mathbb{R}_{++}^N$, there exist $\boldsymbol{\alpha} \in \overline{\mathbb{R}}_+^N$, $\boldsymbol{\alpha} \succeq \mathbf{1}$ and $K_2 \in \mathbb{R}_+$ such that the scaled water-filling condition (6) holds.

*Proof:* A constructive proof is given. Choose $K_2 = \max_{n\in E}(p_n + \Gamma/g_n)$, which satisfies $0 \leq K_2 < \infty$. This choice of $K_2$ implies that $0 \leq p_n \leq K_2 - \Gamma/g_n$ for all $n \in E$. For each $n \in E$, either $p_n = 0$, or $p_n > 0$. For $n \in E$ such that $p_n = 0$, choose $\alpha_n = \infty$ and $\nu_n = 0$ to satisfy (6). For $n \in E$ such that $0 < p_n < C_n$, choose $\alpha_n = K_2/(p_n + \Gamma/g_n) < \infty$ and $\nu_n = 0$. It may be verified by substitution that this choice of $\alpha_n$ satisfies (6). Because $p_n \leq K_2 - \Gamma/g_n$, it also holds $\boldsymbol{\alpha} \succeq 1$. ■

In Theorem 1, the dual variables $\nu_n$ associated with the PSD masks may be chosen to always be $0$. This mathematical property may be interpreted as showing that the proper selection of $\alpha_n$ acts as a "virtual PSD mask" and makes PSD mask constraint in Theorem 2 redundant (in this single-user setting).

The following second theorem shows that for every set of channel parameters and fixed scaling constants, there exists exactly one PSD satisfying the scaled water-filling condition.

*Theorem 2:* For any fixed $\boldsymbol{\alpha} \in \overline{\mathbb{R}}_+^N$ where $\boldsymbol{\alpha} \succeq \mathbf{1}$, and channel gains $\mathbf{g} \in \mathbb{R}_{++}^N$, there exists a unique $\mathbf{p} \in \mathbb{R}_+^N$ satisfying (6). Furthermore, $\sum_{n\in E} p_n = P$ unless $\boldsymbol{\alpha} = \mathbf{1} \cdot \infty$.

*Proof:* For $n$ such that $\alpha_n = \infty$, observe that $p_n = 0$ by (6). Define $F$ to be the remaining tone indices, that is, $F \triangleq E - \{n : \alpha_n = \infty\}$. Consider the following convex optimization problem

$$\begin{aligned}
\text{maximize} \quad & \sum_{n\in F} \frac{1}{\alpha_n} \log\left(1 + \frac{g_n p_n}{\Gamma}\right)\\
\text{subject to} \quad & \mathbf{p} \succeq 0,\ n \in F\\
& \sum_{n\in F} p_n \leq P.
\end{aligned} \tag{7}$$

[1] It may be verified that the condition (6) reduces to (4) when $\mathbf{C} \succ \mathbf{1} \cdot P$.

Observe that the objective of (7) is strictly concave in $\mathbf{p}$ because $\log(1+x)$ is strictly concave on $x \in \mathbb{R}_+$. Furthermore, there exists a feasible point to the optimization (7), namely $\mathbf{p} = \mathbf{0}$, and the feasible set is closed and bounded. The optimization problem therefore has a unique optimal value, call it $\mathbf{p}^\star$. Because the objective is strictly increasing in $p_n$ for $n \in F$, it follows that $\sum_{n\in E} p_n = \sum_{n \in F} p_n = P$ (unless $F = \emptyset$).

It is known that the Karhn-Kush-Tucker (KKT) conditions are necessary and sufficient for optimality of a convex optimization problem satisfying these properties [18]. It can be shown by direct computation that the KKT conditions of the optimization (7) are precisely (6). Therefore, because $\mathbf{p}^\star$ is unique optimal solution to (7), $\mathbf{p}^\star$ is also the unique value satisfying (6). ■

### B. Scaled Bit-Loading

Current discrete bit-loading algorithms [19] require only slight modification to include scaling factors for the proposed BPSM scheme. The algorithm is described in Algorithm 1. $b_n$ is the number of bits loaded on tone $n$, $p_n(b_n) \triangleq \left(2^{b_n} - 1\right)\Gamma/g_n$, is the power to load $b_n$ bits on tone $n$, $b_{max}$ is the maximum bits per tone, and $P$ is the maximum power per user. As seen above, the only modification is to scale the incremental energy tables in a DMT modem.

**Algorithm 1** Scaled discrete bit-loading

1: Initialize :
2: $\Delta p'_n(1) \Leftarrow \alpha_n(p_n(1) - p_n(0)),\ b_n = 0,\ \forall n \in E$
3: Iteration:
4: **while** $\min_n(\Delta p'_n(b_n+1)) < \infty$ **do**
5: $\quad m \Leftarrow \operatorname{argmin}_n \Delta p'_n(b_n+1)$
6: $\quad$ **if** $b_m + 1 \leq b_{max}$ , $p_m(b_m+1) \leq C_m$, and $\sum p_n(b_n) + \Delta p_m(b_m+1) < P$ **then**
7: $\quad\quad b_m \Leftarrow b_m + 1$
8: $\quad\quad \Delta p'_m(b_m+1) \Leftarrow \alpha_m(p_m(b_m+1) - p_m(b_m))$
9: $\quad$ **else**
10: $\quad\quad \Delta p'_m(b_m+1) = \infty$
11: $\quad$ **end if**
12: **end while**

The previous bit-loading algorithm [19] was shown to be the optimal discrete bit loading process that maximizes the data rate under the total power constraint. The following theorem shows the efficiency and near-optimality of the proposed scaled bit-loading algorithm.

*Definition 3:* A bit allocation $\mathbf{b}$ is called an *undominated* or *efficient* solution to (7) if the following conditions hold for all $\mathbf{b}'$.

$$\sum_n b'_n/\alpha_n > \sum_n b_n/\alpha_n \Rightarrow \sum_n p_n(b'_n) > \sum_n p_n(b_n),$$
$$\sum_n b'_n/\alpha_n = \sum_n b_n/\alpha_n \Rightarrow \sum_n p_n(b'_n) \geq \sum_n p_n(b_n). \quad (8)$$

Therefore, any bit allocation that achieves a larger objective value in (7) than the efficient bit allocation requires more total power than the efficient bit allocation. In addition, the efficient bit allocation requires the least amount of total power among any bit allocations that have the same objective value in (7).

*Theorem 3:* The allocation generated by the scaled discrete bit-loading algorithm is an *undominated* or *efficient* solution to (7). Furthermore, the terminating value of $\sum_n b_n/\alpha_n$ found by the algorithm is within 1 of the optimal value of (7).

*Proof:* Because $b_n/\alpha_n$ is concave and strictly increasing in $b_n$, and $p_n(b_n) = (2^{b_n} - 1)\Gamma/g_n$ is convex and strictly increasing in $b_n$, the following incremental bit allocation always generates an undominated allocation [20, Thm. 2],[20, §8].

1) Start with the allocation $\mathbf{b} = \mathbf{0}$.
2) $b_m = b_m + 1$, where $m = \operatorname{argmax}_n((b_n+1)/\alpha_n - b_n/\alpha_n)/(p_n(b_n+1) - p_n(b_n))$.
3) If $\sum_n p_n(b_n) > P$, terminate; otherwise go to step 2.

The above condition in step 2 is exactly same with step 5 in the scaled bit-loading algorithm because

$$\begin{aligned}&\alpha_n(p_n(b_n+1) - p_n(b_n))\\ &= (p_n(b_n+1) - p_n(b_n))/((b_n+1)/\alpha_n - b_n/\alpha_n). \quad (9)\end{aligned}$$

Therefore, the scaled bit-loading process generates an undominated allocation. Suboptimality of less than 1 can be shown as a consequence of [20, Thm. 3]. ■

In step 5 of Algorithm 1, the tone that has the minimum incremental energy is found, and a bit is loaded on that tone. The incremental energy can be equivalently expressed as follows.

$$\begin{aligned}\Delta p'_n(b_n+1) &= \alpha_n(p_n(b_n+1) - p_n(b_n))\\ &= \alpha_n\Gamma\left((2^{b_n+1} - 1) - (2^{b_n} - 1)\right)/g_n\\ &= \alpha_n\Gamma\ 2^{b_n}/g_n\\ &= \alpha_n(p_n(b_n) + \Gamma/g_n). \quad (10)\end{aligned}$$

By interpreting $\alpha_n(p_n(b_n) + \Gamma/g_n)$ as a scaled water-level on tone $n$ at the moment when $b_n$ bits are loaded, the proposed bit-loading process loads a bit on the tone that has the minimum scaled water-level, and it is coincident with scaled water-filling condition where the resultant scaled water-level is flat.

### C. Multiuser Use of Band Preference: ISWF

Band preference is designed for deployment in multi-user networks. In this setting, the users' gains $\mathbf{g}$ depend on the power allocations chosen by *other* users. In particular the SMC, with knowledge of the channel and noise, may compute band preference coefficients $\alpha_n^{(i)}$ for each user $i$ and distribute them infrequently to the modems over control channels. Each modem may then implement the proposed scaled bit-loading algorithm to derive its intended PSD. Because the bit-loading process depends on the observed noise, including the interference from other modems, any change in the modem's PSD would affect all other modems and trigger updates in other modems' PSDs. Since this update is performed in an iterative way as in IWF, it is called Iterative Scaled Water-Filling (ISWF). Assuming that ISWF procedure converges, if the channel and noise do not change from the instant when the SMC computes the scaling factors, the PSD computed by each modem will be as expected at the SMC. If however, there is a change, the modem will adapt and a minor difference will be reflected in the power distribution. As noted previously, a scaling factor provides penalty information for a band, and this information may be nearly optimal even when channel changes moderately. BPSM with scaling factors therefore enables DSL modems' rapid adjustment of their power distribution (*i.e.* bit-swapping) to moderate changes without recomputation of all PSDs at the SMC.

### D. Convergence Properties of ISWF

Once the scaling factors are distributed to the modems, the power allocations of ISWF are determined in a distributed and iterative way like those of IWF. Because the resultant power allocations of ISWF are not controlled by any central entities, the convergence properties are of central importance for practical applications of ISWF. Based on the previous work on the convergence of IWF and ASB [1], [16], [17] and [14], the convergence of ISWF is studied here.

For the multi-user case, the vector of user $i$'s scaling factors is $\boldsymbol{\alpha}^i \in \overline{\mathbb{R}}_+^N$, and the concatenated vector of all users' scaling factors is $\boldsymbol{\alpha} \triangleq (\boldsymbol{\alpha}^i)_{i=1}^M \in \overline{\mathbb{R}}_+^{MN}$. Define a $M \times M$ matrix $B = [b_{ij}]$ such that

$$b_{ij} = \max_{n \in F} \left\{ \frac{\Gamma h_n^{i,j} \alpha_n^i}{h_n^{i,i}} \right\}, \quad \text{if } i \neq j \tag{11}$$

$$b_{ii} = \max_{n \in F} \left\{ \alpha_n^i \right\}, \quad \text{if } i = j. \tag{12}$$

Then, consider a splitting of the matrix $B$ as follows:

$$B = B_1 + B_2 + B_3, \tag{13}$$

where $B_1$, $B_2$, $B_3$ are respectively the strictly lower part, the diagonal part, and the strictly upper part of $B$ [17]. Let $\lambda_1, \ldots, \lambda_s$ be the eigenvalues of the matrix $D \triangleq (B_2 - B_1)^{-1} B_3$. The following theorem shows the sufficient condition for the convergence of ISWF.

*Theorem 4:* For any fixed scaling factors $\boldsymbol{\alpha}$, the power allocations uniquely converge under the ISWF algorithm in the $L$-user system if the following condition holds.

$$\rho(D) = \max_{1 \le i \le s} \lambda_i < 1. \tag{14}$$

*Proof:* See Appendix. ■

The above condition has the same formulation with the sufficient condition for the convergence of IWF in [17] except that $B$ matrix is constructed with scaling factors as well as channel parameters. For the two-user case, the following proposition shows more intuitive conditions.

*Proposition 1:* For any fixed scaling factors, the sufficient condition for the convergence of ISWF is as follows:

$$\frac{b_{12} b_{21}}{b_{11} b_{22}} < 1, \tag{15}$$

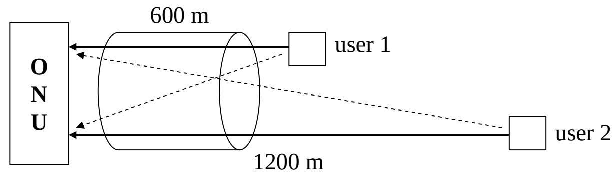



Fig. 1. Two VDSL lines connected to the optical network unit

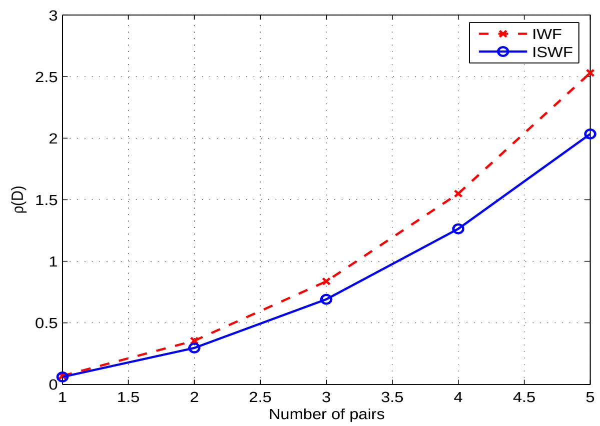



Fig. 2. $\rho(D)$ of IWF and ISWF when the number of (600m, 1200m) pairs increases

or equivalently,

$$\max_{n \in F} \left\{ \frac{\Gamma h_n^{1,2} \alpha_n^1}{h_n^{1,1}} \right\} \max_{n \in F} \left\{ \frac{\Gamma h_n^{2,1} \alpha_n^2}{h_n^{2,2}} \right\} < \max_{n \in F} \left\{ \alpha_n^1 \right\} \max_{n \in F} \left\{ \alpha_n^2 \right\}. \tag{16}$$

The sufficient condition for IWF can be obtained by setting $\boldsymbol{\alpha_1} = \boldsymbol{\alpha_2} = \mathbf{1}$, which is the same with the condition of [1]. Because $\max (a_n b_n) \le \max a_n \max b_n$ for any positive $a_n$ and $b_n$, channel parameters that satisfy the sufficient condition for IWF also satisfy that of ISWF. Therefore, the iterative algorithm is guaranteed to converge for broader range of channel parameters with scaling factors. (The scaling factors provide an opportunity to stabilize the I(S)WF in any case where the IWF might not converge.) Formally speaking, define a set $\mathbb{H}(\boldsymbol{\alpha})$ as the set of channel matrices that satisfies (15) for given $\boldsymbol{\alpha} \in \overline{\mathbb{R}}_+^N$. Then, $\mathbb{H}(\mathbf{1}) \subseteq \mathbb{H}(\boldsymbol{\alpha})$ for any fixed $\boldsymbol{\alpha} \succeq \mathbf{1}$. This sufficient condition is generally true for the line length of our interest. Fig. 2 shows the increase of $\rho(D)$ when the number of users (600m, 1200m) increases. The sufficient condition holds up to 3 pairs (6 users) and both IWF and IWSF are not guaranteed to converge for more than 3 pairs. However, it does not mean that both algorithms do not converge for those numbers, and they are actually shown to converge under various simulations.

## IV. Scaling Factor Setting

### A. Two-user case

This section explains a heuristic way to find scaling factors. Since adjacent tones are likely to have similar properties in a DSL channel, adjacent tones are grouped into a "subband", and one scaling factor is allocated to each subband. For this technique, a two-user problem in Fig.1 is considered. Assuming user 1 to be far-located, he is assumed to be a "weak" user and maximizes his rate while user 2 maintains his target rate. Therefore, user 1 is likely to load power up to the PSD masks on all tones as long as the total power constraint is satisfied. With user 1's power fixed at the PSD masks, the following bit-trade-off can be considered. If user 2 loads more bits on band 1, user 1 might lose some bits on the same band because of the increased interference from user 2. Thus, user 2 should consider loading bits on the bands where the loss of user 1 is minimized. This bit-trade-off between users is expressed as a cost table. A cost table calculated in this way significantly simplifies the optimization process since it conceals the details of bit-loading process.

Assume that $\mathbf{p}^1$ are fixed as the PSD masks $\mathbf{C}$ and $\mathbf{p}^2$ is determined to achieve user 2's rate target while maximizing user 1's rate. Tones are divided into $M$ subbands, and $B_m$ is the set of tones on band $m$. Then, a cost function is defined as follows.

$$C_m(\mathbf{p}^1, \mathbf{p}^2) = R_m^1(\mathbf{p}^1, \mathbf{0}) - R_m^1(\mathbf{p}^1, \mathbf{p}^2), \qquad (17)$$

where $R_m^1(\mathbf{p}^1, \mathbf{p}^2)$ is the number of bits of user 1 on band $m$ when each user's PSD is respectively $\mathbf{p}^1$ and $\mathbf{p}^2$. $C_m(\mathbf{p}^1, \mathbf{p}^2)$ is the decrease in user 1's bits when user 2 increases its PSD from $\mathbf{0}$ to $\mathbf{p}^2$.

With the above definition, the following cost minimization problem with two users is considered.

$$\begin{aligned}
\text{minimize} \quad & \sum_{m=1}^{M} C_m(\mathbf{p}^1, \mathbf{p}^2) \\
\text{subject to} \quad & \sum_{m=1}^{M} R_m^2(\mathbf{p}^1, \mathbf{p}^2) \geq R_{target}^2 \\
& \sum_{n=1}^{N} p_n^2 \leq P \\
& p_n^2 = \left( \frac{K}{\alpha_m^2} - \frac{\Gamma}{g_n^2} \right)^+, \quad n \in B_m,
\end{aligned} \qquad (18)$$

where the last condition means that user 2 should maintain a flat water-level $(K/\alpha_m^2)$ on each subband, as defined in the scaled water-filling. Because of the scaled water-level condition of user 2, if $R_m^2$ is fixed for $\forall m$, $\mathbf{p}^2$ is also determined. Therefore, the above minimization problem can be expressed only in terms of bits per band as follows:

$$\begin{aligned}
\text{minimize} \quad & \sum_{m=1}^{M} C_m(R_m^2) \\
\text{subject to} \quad & \sum_{m=1}^{M} R_m^2 \geq R_{target}^2,
\end{aligned} \qquad (19)$$

where $C_m(x) = C_m(\mathbf{p}^1, \mathbf{p}^2)$, such that $x = R_m^2(\mathbf{p}^1, \mathbf{p}^2)$. From this new definition of cost function, a cost table is generated for each band and each $R_m^2$. Because the table size can be prohibitively large if each entry is generated for every bit increment, incremental granularity $\Delta$ is introduced.

$$C_{m,\Delta}(i) = C_m(i\Delta), \qquad (20)$$

where $C_{m,\Delta}(i)$ is the $(i,m)^{\text{th}}$ entry of the table, which is the cost incurred when user 2 loads $i\Delta$ bits on band $m$. When user 2 can not load $i\Delta$ bits any more, $C_{m,\Delta}(i)$ is set as $\infty$. Table I shows one such example when $M = 6$ and $\Delta = 70$ for the structure of Fig. 1. For instance, if user 2 loads 210 $(= 3\Delta)$ bits in band 2, user 1 loses 8 bits in the same band. If user 2 loads $(2, 2, 3, 2, 9, 8) \cdot \Delta$ bits on each band respectively, the total incurred cost becomes $0 + 2 + 5 + 1 + 0 + 0 = 8$ bits. User 1 can not load on high frequency tones because of weak channel gains, and therefore, no cost is incurred on high frequency bands (i.e. band 5,6) although user 2 loads maximum allowed bits on those bands. Generally, a greedy algorithm does not find the optimal solution because of the non-convexity of the underlined problem. However, it can be efficiently solved by Dynamic Programming [21] as follows.

$$\begin{aligned}
f_1(i_1) &= C_{1,\Delta}(i_1) \\
f_m(i_m) &= \min_{0 \leq j_m \leq i_m} \{C_{m,\Delta}(j_m) + f_{m-1}(i_m - j_m)\},
\end{aligned} \qquad (21)$$

where $f_m(i_m)$ is the minimum cost to load $i_m\Delta$ bits from band 1 to band $m$. Therefore, the minimum cost to load $R_{target}^2$, which is the objective of (19) is $f_M(R_{target}^2/\Delta)$. The determination of $f_m(i_m)$ can be divided into $i_m + 1$ subproblems. The subproblems are to find $f_{m-1}(i_m - j_m)$ when $j_m\Delta$ bits $(0 \leq j_m \leq i_m)$ are loaded on band $m$. In this recursive way, the solution of (19) is obtained and shows that how many bits user 2 loads on each band to minimize the cost. Therefore, $\mathbf{p}^2$ can also be found, which is converted to scaling factors by Theorem 1. The resultant scaling factors are constant in each band, or one scaling factor is required per band because the water-level is assumed to be flat in each band in solving (18). The total power constraint of (18) is not used (19), and therefore, the above solution should be separately verified with the constraint before the conversion to scaling factors. If the solution violates the constraint, the entry that requires the largest power in the table is set as $\infty$, and the problem is solved again for the new table. This process repeats until the new solution satisfies the total power constraint. However, this update process is not generally required to find scaling factors with an appropriate level of accuracy. Furthermore, the actual power allocations of ISWF always satisfy the constraint even in the case when the scaling factors are converted from the power allocations that violate the power constraint. However, these could differ slightly from those found by solving (19).

### B. *User grouping*

The previous section explains how to find scaling factors for the two-user situation by the cost-table approach. However, for $L > 2$, it may not be possible to apply the method directly since it requires a multi-dimensional cost table. Thus, this section introduces the user grouping method.

TABLE I
COST TABLE FOR 600M AND 1200M VDSL LINES ($M = 6$, $\Delta = 70$)

| i | $C_{1,\Delta}$ | $C_{2,\Delta}$ | $C_{3,\Delta}$ | $C_{4,\Delta}$ | $C_{5,\Delta}$ | $C_{6,\Delta}$ |
|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 2 | 0 | 1 | 0 | 0 |
| 3 | $\infty$ | 8 | 5 | 7 | 0 | 0 |
| 4 | $\infty$ | 27 | 20 | 20 | 0 | 0 |
| 5 | $\infty$ | 67 | 58 | 48 | 0 | 0 |
| 6 | $\infty$ | 112 | 76 | 48 | 0 | 0 |
| 7 | $\infty$ | 112 | $\infty$ | $\infty$ | 0 | 0 |
| 8 | $\infty$ | $\infty$ | $\infty$ | $\infty$ | 0 | 0 |
| 9 | $\infty$ | $\infty$ | $\infty$ | $\infty$ | 0 | $\infty$ |
| 10 | $\infty$ | $\infty$ | $\infty$ | $\infty$ | $\infty$ | $\infty$ |

As previously mentioned, the most beneficial situation of spectrum balancing is when users have asymmetric locations. Otherwise, the performance gap between the IWF and OSB is not significant. Therefore, the spectrum balancing method for $L > 2$ likewise concentrates on reducing the interference from strong users to weak users to achieve better performance. For this purpose, users are grouped into a strong user group and a weak user group. In the strong user group, users who should transmit in a polite way are included, while in the weak user group, users who can maximize their own rates without hurting other users are included. That means, users in the weak user group run the IWF while users in the strong user group run the ISWF. In this way, the SMC can also use the previous approach in IV-A to compute scaling factors for more than two users. First, assume that weak users' PSDs are fixed at PSD masks. Then, one of the strong users generates a cost table. The cost function is defined as the loss of bits of all weak users when that strong user increases its PSD with other strong users's PSD fixed. Given the table, the strong user can find PSD that achieves its rate target while minimizing the cost. The next user in a strong group similarly generates the cost table and find its PSD. This sequential update of PSD repeats until the process converges. Because each strong user considers other strong users' interference as well as incurred cost, each strong user keeps its appropriate balance between politeness and selfishness.

## V. DISTRIBUTED BAND-PREFERENCE ALGORITHM

This section introduces a fully distributed Band-Preference method where scaling factors are determined in a distributed way. In the previously mentioned method, the SMC calculates scaling factors and distributes them to each user. After receiving those factors, no centralized control is required for the network. In some cases, it may not be feasible to distribute that information to each user. However, even without having that information, it is still possible to achieve performance gain over IWF using the characteristic of DSL channels. In DSL channels, direct channel gains of the low frequency region are larger than those of the high frequency region and therefore more bits are loaded in a bit-loading process. This is generally true regardless of the line length. For a multi-user case, however, strong users should load more bits in the high frequency region for the sake of weak users, which is managed through the SMC. Though fully accurate management of the spectrum is possible only with central controllers such as the SMC, the spectrum can be managed adequately without the help of the central controllers. Because of the similarity of DSL channel gain, the strong users can predict that good tones with high channel gains will also be good tones for weak users. Similarly, bad tones for strong users will also be bad ones for weak users. Thus, strong users can give up good tones for weak users if the strong users know about the existence of weak users. For this purpose, two kinds of modes are defined for distributed band-preference. The first one is the normal mode where the user runs the water-filling algorithm as in the normal IWF. The second one is the polite mode where the user is aware of the existence of weak users and gives up some good tones, thereby using bad tones while still achieving his rate target.

### A. Algorithm

This section shows the overall process of the distributed band preference algorithm. The SMC first determines a mode and a bit target for each user based on its complete knowledge about the channels. As the SMC distributes those information, each user can determine whether it should load bits greedily or politely. Because the interference from weak users is comparably small, weak users are usually allowed to load bits in a greedy manner. Meanwhile, because strong users can significantly interfere the other users' signals, they are requested to be polite to protect other users. Under normal mode, the user runs the normal bit loading process as in the IWF. Under polite mode, however, the user first runs the normal bit loading process, and then moves bits from preferred bands to less-preferred bands.

Bands are determined autonomously by each user. Consecutive tones are grouped together to form a band and the geometric means of channel gains of the bands are compared to determine how good the bands are. Then, bits are first moved from the best band to the worst band as long as the total power constraint is met. To maintain flat water-level in each band as the scaled water-filling does, one bit is removed from the tone that used the largest energy in the best band and one bit is added to the tone that will require the least energy in the worst band. This bit moving process gradually reduces the water-level of the best band and gradually increases the water-level of the worst band. Because one bit is removed and sequentially added, the total number of loaded bits remains unchanged. Once a bit can no longer be removed from the best band or a bit can no longer be added to the worst band, the second best or worst band is considered in the bit moving process. This process ends when all the bands are considered once as the best or worst band. The detailed process is described in Algorithm 2. Once the bit-moving process ends, the scaling factors can be determined based on the water-levels using (4) . Because the water-levels of the good bands are reduced and those of the bad bands are increased, the scaling factors of the good bands are resultantly smaller than those of

**Algorithm 2** Bit-moving process

Initialize :
Divide all tones into $M$ bands
Sort $M$ bands in descending order based on their geometric means of channel gains : $B_1, B_2, \ldots, B_M$
$i = 1, j = M$
Iteration:
**while** $i < j$ **do**
  $k \Leftarrow \arg\max_{n \in B_i} \Delta p_n(b_n)$
  $l \Leftarrow \arg\min_{n \in B_j} \Delta p_n(b_n + 1)$
  **if** $\sum_n p_n(b_n) + \Delta p_l(b_l + 1) - \Delta p_k(b_k) > P$, or $p_l(b_l + 1) = \infty$ **then**
    $j \Leftarrow j - 1$
  **else if** $b_l + 1 > b_{max}$ , or $p_l(b_l + 1) > \mathbf{C}_l$ **then**
    $\Delta p_l(b_l + 1) = \infty$
  **else if** $\Delta p_k(b_k) = 0$ **then**
    $i \Leftarrow i + 1$
  **else**
    $b_k \Leftarrow b_k - 1$, $\Delta p_k(b_k) \Leftarrow p_k(b_k) - p_k(b_k - 1)$
    $b_l \Leftarrow b_l + 1$, $\Delta p_l(b_l + 1) \Leftarrow p_l(b_l + 1) - p_l(b_l)$
  **end if**
**end while**

the bad bands. If no bits are moved during the above process, then scaling factors are identical for all bands. After one cycle of the bit-moving process and scaling-factor decision process, other users will see different interference levels and trigger their own cycles to adapt to the new interference levels. These processes will continue for all users until they converge as in IWF.

As described in the above algorithm, the only required information for each user is the mode and the rate target. Each user can autonomously determine scaling factors and those scaling factors are also used to adapt to small changes by running the bit-swapping algorithm without any aid from the SMC.

## VI. Simulation Results

The performance of the IWF, OSB, BPSM, and DBPSM are compared by simulation. Fig. 1 illustrates a two-user VDSL upstream scenario, where user 1 is distantly located and is assumed to be a weak user. The upstream signal from user 2 operates as a strong interference to user 1. In this simulation, Noise A in addition to $-140$ dBm/Hz AWGN is injected, where Noise A is a mixture of 16 ISDN, 4 HDSL and 10 ADSL disturbers [22]. Fig. 3 shows the rate region, where a large gap between the IWF and OSB can be observed. The primary reason for this gap is that the signal from the strong user induces strong interference to the weak user in the low frequency region. The OSB similarly avoids this phenomenon by allocating less power in the low frequency region of the strong user. Fig. 3 also shows that the BPSM's performance is very close to that of the OSB when $M = 15$. For example, when the strong user's rate is set as 35Mbps, the weak user achieves 4.654Mbps in the OSB, 4.477Mbps in the BPSM, 4Mbps in the DBPSM, and 1.73Mbps in the IWF. Therefore, the BPSM can achieve 96% of the OSB and 160% increase over the IWF. Interestingly, the performance of the DBPSM is also close to the performance of the BPSM, andthe DBPSM achieves 130% gain over the IWF by finding scaling factors only with given rate targets. Although the performance gain of the DBPSM decreases as the rate target of user 2 approaches the maximum rate 42.75Mbps, those points are less important than the points around 30-35Mbps, which are the most probable operating points for this two-user case. Fig. 4 shows the power spectral densities of both users when the the rate target of user 2 is 35Mbps. Because of the low direct channel gain at the high frequency region, user 1 can not transmit any bits on the high frequency region while user 2 can still transmit on that region. However, the low frequency region is also preferred by user 2 because of the high direct-channel gain, and it is also occupied by user 2 in the IWF. Meanwhile, user 2 in the OSB transmits more bits in the high frequency region than in the IWF to reduce the interference to user 1. Because user 1 does not transmit in the high frequency region, the strong interference from user 2 in the high frequency region does not reduce user 1's data rate.

## VII. Conclusion

This paper has proposed a low-overhead band preference algorithm for distributed control of modem PSDs in a DSL network. Because it utilizes power scaling factors instead of directly controlling PSDs, the proposed algorithm leverages the modems' natural adaptive capability to respond to channel and noise fluctuations. An ad-hoc algorithm for choosing band-preference parameters at the SMC was presented. When the SMC can not provide the scaling factors, a distributed way to compute scaling factors was also proposed. Because of the minimum requirement of the centralized control, these types of band preference algorithms will allow easier implementation of Level 2 DSM without significant modification of current DSL networks. Numerical simulations of these algorithms in VDSL systems show that significant gains (approaching optimal Level 2 DSM limits) can be achieved.

## APPENDIX

### A. Proof of Theorem 4

The convergence proof of the IWF in [17] can be used to prove the sufficient condition in Theorem 4 with minor modifications. Therefore, this paper only shows the major steps of [17] with appropriate changes and the details are omitted.

In the ISWF, each user iteratively finds $\mathbf{p}$ satisfying the scaling water-filling condition (4). First, consider the scaling water-filling condition for user $i$.

$$p_n^i = \left( \frac{K}{\alpha_n^i} - \Gamma \frac{\sigma_n^2 + \sum_{j \neq i} h_n^{i,j} p_n^j}{h_n^{i,i}} \right)^+, \quad n \in E,$$
$$\sum p_n^i = P, \ K \geq 0. \qquad (22)$$

For simplicity, PSD masks are omitted here. This condition can be converted into a mixed linear complementarity problem (LCP) as follows:

$$p_n^i\left(p_n^i+\Gamma\frac{\sigma_n^2+\sum_{j\neq i}h_n^{i,j}p_n^j}{h_n^{i,i}}-\frac{K}{\alpha_n^i}\right)=0,\quad n\in E,$$
$$p_n^i\geq 0,\quad p_n^i+\Gamma\frac{\sigma_n^2+\sum_{j\neq i}h_n^{i,j}p_n^j}{h_n^{i,i}}-\frac{K}{\alpha_n^i}\geq 0,$$
$$\sum p_n^i-P=0,\ K\geq 0.\quad (23)$$

Equivalently, the above LCP can be expressed with concatenated vectors and matrices such as $\mathbf{p}^i=(p_1^i,\ldots,p_N^i)$, $M^{ij}=\Gamma\mathrm{diag}(\alpha_1^i h_1^{i,j}/h_1^{i,i},\ldots,\alpha_N^i h_N^{i,j}/h_N^{i,i})$ for $j\neq i$, $M^{ii}=\mathrm{diag}(\alpha_1^i,\ldots,\alpha_N^i)$, and $\mathbf{q}^i=\Gamma(\alpha_1^i\sigma_1^2/h_1^{i,i},\ldots,\alpha_N^i\sigma_N^2/h_N^{i,i})$.

$$(\mathbf{p}^i)^T\left(\mathbf{q}^i+\sum M^{ij}\mathbf{p}^j-K\mathbf{1}\right)=\mathbf{0},$$
$$\mathbf{p}^i\geq\mathbf{0},\quad\mathbf{q}^i+\sum M^{ij}\mathbf{p}^j-K\mathbf{1}\geq\mathbf{0},$$
$$\sum p_n^i-P=0,\ K\geq 0,\ \mathbf{p}^i\geq\mathbf{0}.\quad (24)$$

This LCP is known as the Karush-Kuhn-Tucker (KKT) condition of the following affine variational inequalities (AVI) defined in a subset $S^i$ of $\mathbb{R}^n$.

$$S^i\equiv\left\{\mathbf{p}^i\in\mathbb{R}^n:\sum p_n^i-P=0,\ \mathbf{p}^i\geq\mathbf{0}\right\},$$
$$(\mathbf{s}-\mathbf{p}^i)^T(\mathbf{q}^i+\sum M^{ij}\mathbf{p}^j)\geq 0,\forall\mathbf{s}\in S^i.\quad (25)$$

The above AVI can be solved equivalently by the following fixed-point equation with a projection to subset $S^i$.

$$\mathbf{p}^i=\left[\mathbf{p}^i-\beta(\mathbf{q}^i+\sum M^{ij}\mathbf{p}^j)\right]_{S^i}.\quad (26)$$

This fixed-point equation can be solved with an iterative method.

$$(\mathbf{p}^i)^{t+1}=\left[-\beta(\mathbf{q}^i+\sum_{j\leq i}M^{ij'}(\mathbf{p}^j)^{t+1}+\sum_{j>i}M^{ij}(\mathbf{p}^j)^t)\right]_{S^i},\quad (27)$$

where the first $\mathbf{p}^i$ term in (26) is combined in the summation and $M^{ij'}$ reflects it. Assuming that user 1 first starts its update and other users sequentially update at time $t+1$, user $i$ at time $t+1$ observes updated power $(\mathbf{p}^j)^{t+1},(j<i)$, and not-updated power $(\mathbf{p}^j)^t,(j>i)$. This sequential update is the same with the process of the ISWF. Furthermore, because (22) is equivalent to (26) as explained previously, the convergence of ISWF is equivalent to the convergence of (27). The convergence can be proved by showing a contraction in the Euclidean norm when $\rho(C)<1$ with $\beta=1/\max_n\alpha_n^i$.

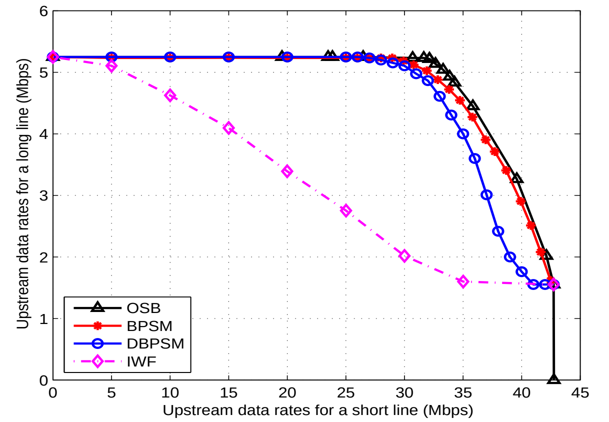



Fig. 3. Upstream data rates of OSB, BPSM, DBPSM, and IWF for VDSL with two users

## References


[1] W. Yu, G. Ginis, and J. Cioffi, "Distributed Multi-user Power Control for Digital Subscriber Lines," *IEEE JSAC*, vol. 20, no. 5, pp. 1105-1115, June 2002.

[2] R. Cendrillon, M. Moonen, W. Yu, J. Verlinden, and T. Bostoen, "Optimal Multi-user Spectrum Balancing for Digital Subscriber Lines," *IEEE Trans. On Commun.*, May 2006.

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

[4] K. Kerpez, D. Waring, S. Galli,J. Dixon, and P. Madon, "Advanced DSL management," *IEEE Commun. Mag.*, vol.41, pp.116-123, Sep. 2003.

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

[6] W. Yu and R. Lui: "Dual Methods for Nonconvex Spectrum Optimization of Multicarrier Systems", *IEEE Transactions on Communications*, vol. 54, no. 7, pp.1310-1322, July 2006.

[7] R. Lui and W. Yu, "Low-Complexity Near-Optimal Spectrum Balancing for Digital Subscriber Lines," in *IEEE ICC*, May 2005.

[8] R. Cendrillon and M. Moonen, "Iterative Spectrum Balancing for Digital Subscriber Lines," in *IEEE ICC*, May 2005.

[9] J. Papandriopoulos and J. Evans, "Low-Complexity Distributed Algorithms for Spectrum Balancing in Multi-User DSL Networks," in *IEEE ICC*, June 2006.

[10] J. Lee, S. T. Chung, and J. M. Cioffi, "Optimal discrete bit-loading for DMT-based DSL systems with equal-length loops," *IEEE Transactions on Communications*, vol. 54, pp. 1760-1764, Oct. 2006.

[11] J. Lee, R. V. Sonalkar, and J. M. Cioffi, "Multi-user bit-loading for multi-carrier systems," *IEEE Transactions on Communications*, vol. 54, pp. 1170-1174, July 2006.

[12] D. Yu, K. Seong, and J. M. Cioffi, "Multiuser discrete bit-loading for digital subscriber lines," in *IEEE ICC*, June 2007.

[13] J. Cioffi, W. Rhee, M. Mohseni, and M. Brady, "Band Preference in Dynamic Spectrum Management," *EUSIPCO*, Sep. 2004.

[14] R. Cendrillon, J. Huang M. Chiang, and M. Moonen, "Autonomous Spectrum Balancing for Digital Subscriber Lines," To appear in *IEEE Transaction on Signal Processing.*

[15] Y. Liu and Z. Su, "Distributed Dynamic Spectrum Management for Digital Subscriber Lines", *IEICE Transaction on Communications*, Mar. 2007.

[16] S. T. Chung, S. J. Kim, J. Lee, and J. M. Cioffi, "A game-theoretic approach to power allocation in frequency-selective Gaussian interference channels," in *IEEE ISIT*, June 2003.

[17] Z. Luo, and J. Pang, "Analysis of Iterative Waterfilling Algorithm for Multiuser Power Control in Digital Subscriber Lines," in *EURASIP Journal on Applied Signal Processing*, 2006.

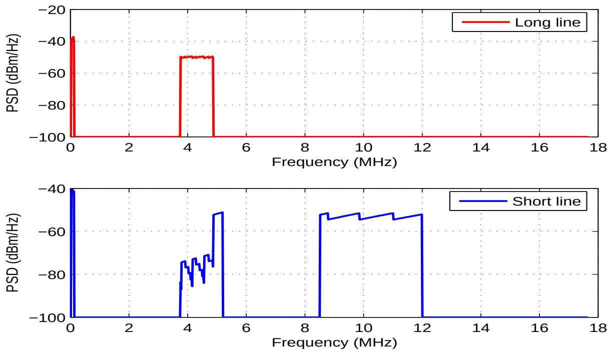
[18] S. Boyd, and L. Vandenberghe, *Convex Optimization.* Cambridge Univ. Press, 2003.

[19] J. Campello,"Optimal Discrete Bit Loading For Multicarrier Modulation Systems", in *International Symposium on Information Theory (ISIT)*, Aug. 1998.

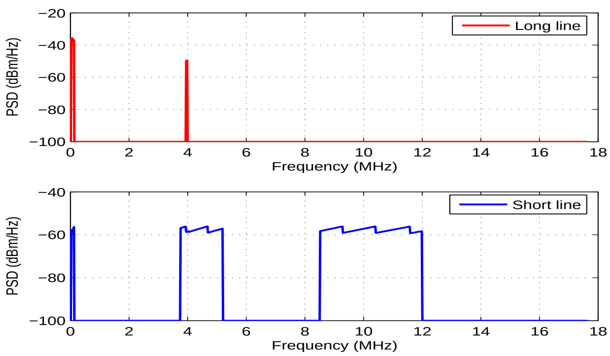
[20] B. Fox, "Discrete Optimization via Marginal Analysis", *Management Science*, 13:210-216, Nov. 1966.

[21] D. Bertsekas, *Dynamic Programming: Deterministic and Stochastic Models.* Prentice Hall, 1987.

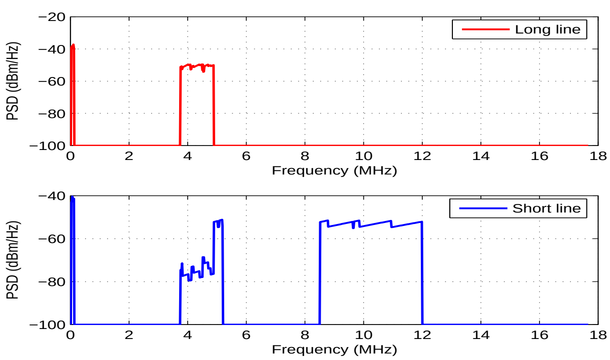
[22] M. Brady, K. Seong, J. Cioffi, and G. Ginis, "Proposed Text for Annex C in DSM Report", *NIPP-NAI-2005-217R1*, Jan. 2006.



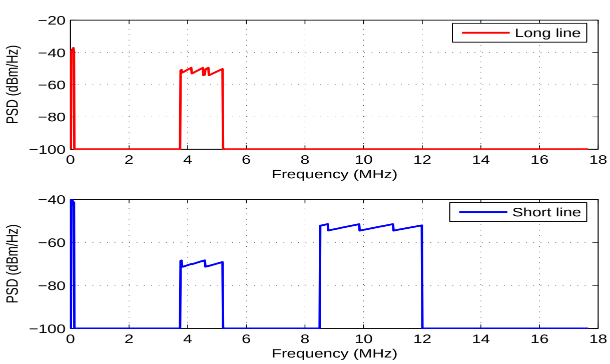

(a) OSB




(b) IWF




(c) BPSM




(d) DBPSM

Fig. 4. Power spectral densities of OSB, IWF, BPSM, and DBPSM when $R_2 = 35$Mbps (Short line)

[23] J. Cioffi, *EE379C Course Reader - Advanced Digital Communication*, Stanford University. Available at: http://www.stanford.edu/ee379c/