Low-Complexity Distributed Algorithms for Spectrum Balancing in Multi-User DSL Networks

John Papandriopoulos and Jamie S. Evans
ARC Special Research Centre for Ultra-Broadband Information Networks (CUBIN),
Department of Electrical and Electronic Engineering,
University of Melbourne, VIC 3010, Australia

***Abstract*— Dynamic Spectrum Management of Digital Subscriber Lines (DSL) has the potential to dramatically increase the capacity of the aging last-mile copper access network. This paper takes an important step toward fulfilling this potential through power spectrum balancing. We derive a novel algorithm called SCALE, that provides a significant performance improvement over the existing iterative water-filling (IWF) algorithm in multi-user DSL networks, doing so with comparable low complexity. The algorithm is easily distributed through measurement and limited message-passing with the use of a Spectrum Management Center. We outline how overhead can be managed, and show that in the limit of zero message-passing, performance reduces to IWF.**

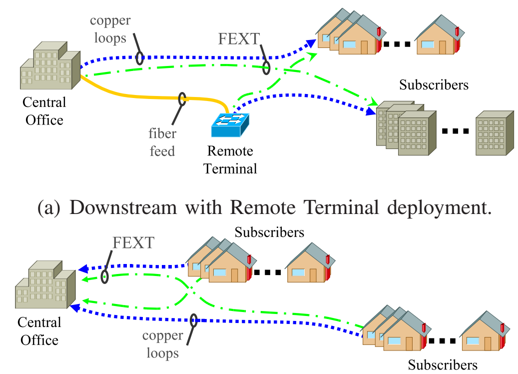
Numerical convergence of SCALE was found to be extremely fast when applied to VDSL, with performance exceeding that of iterative water-filling in just a few iterations, and to over 90% of the final rate in under 5 iterations. Lastly, we return to the problem of iterative water-filling and derive a new algorithm named SCAWF that is shown to be a very simple way to water-fill, particularly suited to the multi-user context.

I. Introduction

DIGITAL Subscriber Line (DSL) technology has helped quench our thirst for bandwidth in recent years, extending the life of existing copper twisted-pair networks that now serve over 100 million subscribers around the globe with broadband internet connectivity. Whilst DSL technology has been hugely successful, incumbent telephone operators are increasingly faced with stiff competition from the decreasing cost of optical fiber-fed leased lines, and aggressive cable television companies serving subscribers from much higher bandwidth Hybrid Fiber-Coax (HFC) cable networks. As a result, telephone companies are in desperate need for increased bit-rates on existing DSL lines in an effort to further extend the life of their aging copper plants.

Cross-talk and long loop lengths are the obstacles toward reaching these higher rates. Twisted-pairs are bundled together in groups of 25-100 lines in ducts directed toward subscribers. Lines are sufficiently close such that electromagnetic radiation induces cross-talk coupling between pairs. Near-end cross-talk (NEXT) is caused by transmitters interfering with receivers on the same side of the bundle and is often avoided by using non-overlapping transmit and receive spectra (FDD) or disjoint time intervals (TDD). Far-end cross-talk (FEXT) is caused by transmitters on opposite sides of the bundle (see Fig. 1). In some cases, this interference can be 10-20dB larger than the background noise and has been identified as the dominant source of performance degradation in DSL systems [7].

This work was supported by the Australian Research Council.
CUBIN is an affiliated program of National ICT Australia (NICTA).



(b) Upstream with varied subscriber loop lengths.

Fig. 1. Two DSL topologies where performance is significantly improved through dynamic spectrum management.

Telephone companies are increasingly shortening the loop using Remote Terminal (RT) deployments, resulting in lower signal attenuation and larger available bandwidths, see Fig. 1(a). Unfortunately this can cause other problems such as the 'near-far' effect, due to the cross-talk. Common in CDMA wireless, the near-far effect occurs when a user enjoying a good channel close to the receiver overpowers the received signal of a user further away having a worse channel and where both users transmit at a similar power level.

Two competing solutions to the cross-talk impairment are known: vectored DSL and spectrum balancing. Each falls under the umbrella of Dynamic Spectrum Management (DSM); see [11] for an overview. Vectoring treats the DSL network as a MIMO system, where modems must co-ordinate at the signal level to effectively *remove* cross-talk through successive decoding or precoding. In contrast, spectrum balancing involves a much looser level of co-ordination: much like existing systems, modems employ single-user encoding and decoding (treating interference as noise), however they may also interact on a more granular time-scale to negotiate a spectrum allocation that effectively *avoids* cross-talk as much as possible to improve the overall performance of the network.

This paper is concerned with the balancing of users' power spectral densities (PSDs), explicitly taking cross-talk effects into account. A significant improvement in network capacity is possible by such a judicious allocation of users' power, and

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE ICC 2006 proceedings.

especially so in near-far situations as pictured in Fig. 1.

Early work in this area introduced an iterative water-filling (IWF) scheme to balance user PSDs, where each user repeatedly measures the aggregate interference received from all other users, then greedily water-pours their own power allocation without regard for the impact to be had on other users [14]. This process results in a fully distributed and autonomous algorithm having a reasonable computational complexity.

More recent efforts have focused on the underlying optimization problem that spectrum balancing aims to solve. Unfortunately this optimization (introduced in Sec. III) is a difficult nonconvex problem. As such, the Optimal Spectrum Balancing (OSB) algorithm [7] makes use of a grid-search to find the optimal power allocation to a predetermined quantization of user powers. It suffers from an exponential complexity in the number of users, and so near-optimal Iterative Spectrum Balancing (ISB) algorithms were developed that reduce complexity through a series of line-searches, avoiding the grid-search bottleneck [6], [10]. Both of these algorithms are centralized; it is unclear how well-suited they are for practical implementation.

In this paper, we return to the underlying nonconvex spectrum balancing optimization, and show that it is a NP-hard problem. We then apply a novel technique involving a series of convex relaxations to derive an algorithm called SCALE (**S**uccessive **C**onvex **A**pproximation for **L**ow-compl**E**xity). We show through numerical simulation that SCALE performs significantly better than IWF, and with comparable complexity.

An important feature of SCALE is that it may be distributed with the help of a Spectrum Management Center (SMC). The resulting method may be viewed as a distributed computation across the DSL network, in contrast to the centralized OSB and ISB schemes. Importantly, we outline how the overhead associated with this approach can be managed, and show that it degrades gracefully to the same performance as that attained by IWF when no inter-user communication is available.

Our final contribution involves a fresh look at IWF. We derive a new algorithm called SCAWF (**S**uccessive **C**onvex **A**pproximation for **W**ater-**F**illing) that simplifies existing IWF approaches and enjoys low complexity implementation.

II. System Model

We adopt a standard model for a K user xDSL system employing Discrete Multi-Tone (DMT) modulation where each user has N tones available that are used to form a set of N ISI-free orthogonal sub-channels. We make the usual assumption that users are aligned in frequency such that FEXT coupling occurs on a common tone-by-tone basis.

A fixed band-plan is assumed for simplicity, that partitions each of these tones into separate up- and down-stream bands that are the same for all users. While it is known that such a scheme is not optimal [10], partitions are a common way to avoid NEXT. The algorithms developed in this paper are easily extended to include NEXT coupling if required.

We consider continuous bit-loading and write the achievable loading on tone n, user k as

$$b_k^n(\mathbf{P}^n) = \log\left(1 + SIR_k^n(\mathbf{P}^n)\right)$$

in the units of nats, and where we define the signal-to-interference ratio (SIR) for user k on tone n as

$$SIR_k^n(\mathbf{P}^n) = \frac{G_{kk}^n P_k^n}{\sum_{j\neq k} G_{kj}^n P_j^n + \sigma_k^n}.$$

We denote by P_k^n the transmitter power of user k on tone n. For notational convenience, we write $\mathbf{P}^n = [P_1^n, P_2^n, \ldots, P_K^n]^T$ as the K-length vector of all transmitter powers on tone n. We will also make use of the notation $\mathbf{P}_k = [P_k^1, P_k^2, \ldots, P_k^N]$ as the N-length PSD vector of user k. Lastly, we will denote the $K \times N$ matrix $\mathbf{P}$ as the stacking of these vectors in the obvious way. This notation makes clear the explicit dependence of the SIR on power.

The gains G_{kj}^n model the channel power transfer on tone n from user j to the receiver of user k. For further notational convenience, we assume the gains G_{kk}^n have been normalized by an appropriate SNR-gap Γ_k^n, that depends on the coding scheme, target probability of error and noise margin [12].

Each σ_k^n models the received noise power on tone n. We assume the noise powers are constant, modeling receiver thermal noise plus any background noise injected by other co-existing systems (e.g. HDSL, ISDN, RF noise, etc.)

The achievable rate for user k is then

$$R_k(\mathbf{P}) = \sum_{n=1}^{N} b_k^n(\mathbf{P}^n) = \sum_{n=1}^{N} \log\left(1 + SIR_k^n(\mathbf{P}^n)\right)$$

nats per channel use.

III. The Spectrum Balancing Problem

The spectrum balancing problem has many forms, categorized by the *Rate Adaptive* (RA) and *Fixed Margin* (FM) formulations. The RA problem seeks to maximize the data-rate of each user, subject to per-user maximum power constraints. It is inherently a multicriterion optimization problem, where one has the ability to scalarize the rates of each user, forming a weighted sum objective (see below).

The problem is written as the optimization

$$\max_{\mathbf{P}\geq 0} \sum_{k=1}^{K} \omega_k \sum_{n=1}^{N} \log\left(1 + SIR_k^n(\mathbf{P}^n)\right) \quad (1)$$

$$s.t. \sum_{n=1}^{N} P_k^n \leq P_k^{max}, \quad k = 1, \ldots, K,$$

where P_k^{max} is the maximum power constraint of user k and each ω_k is a fixed nonnegative scalarization weight that allows a trade-off between the rates allocated to each user.

Equivalently, these weights allow the system operator to place differing Qualities of Service or importance levels on each user. For example, one can weight those users having longer loop lengths more heavily, until their rate allocations are on-par with more lightly weighted users enjoying shorter loop lengths.

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE ICC 2006 proceedings.

The FM problem is concerned with finding a minimal power allocation such that each user has a minimum (or target) data-rate that is attained. These target rates must be feasible; that is, there exists a power-allocation whereby the (implicit) maximum power constraint of each user is not violated.

The set of feasible target rates is contained within the so-called rate-region. Its boundary corresponds to the Pareto optimal surface of the RA problem above, and is often explored by sweeping values of the weights ω_k and solving a sequence of such problems. Alternatively, given a set of target rates, one could solve a suitable feasibility problem to determine whether the supplied rates can be supported by the network.

In this paper, we will concentrate on the RA problem (1) for simplicity, and return to these other problems later [2].

A. Related Work

The RA and FM problems have been extensively studied for single-user $K = 1$ systems, for both continuous and discrete bit-loading. Algorithms enjoying $O(N \log N)$ or $O(N)$ complexity are well-established, their solutions usually involving some kind of water-pouring; see [5] and references therein.

For the multi-user case where $K > 1$, optimization (1) is difficult, because the objective is nonconvex in $\mathbf{P}$. More insight is gained by rewriting the problem as follows:

$$\begin{aligned} \max_{\mathbf{P}\geq 0} \quad & \sum_k \omega_k \sum_n \log\Big(\sum_j G_{kj}^n P_j^n + \sigma_k^n\Big) \\ & \qquad\qquad - \log\Big(\sum_{j\neq k} G_{kj}^n P_j^n + \sigma_k^n\Big) \\ s.t. \quad & \sum_n P_k^n \leq P_k^{max}, \quad \forall k. \end{aligned}$$

We observe that the objective is a difference of concave (d.c.) functions in $\mathbf{P}$. Such problems having d.c. structure have been of great interest to the optimization community for the past 30+ years. Unfortunately it can be shown that these problems are NP-hard [9] and often difficult to solve efficiently for the global optimum.

The IWF approach finds an approximate solution by splitting this problem into K convex sub-problems, then iterating over these until convergence. Each sub-problem concerns only the powers $\mathbf{P}_k$, fixing all other powers $\mathbf{P}_{j\neq k}$ and treating their contributions as noise (see Sec. IV-B). These sub-problems are made distributed through SIR measurements. IWF has been shown to converge to a competitive Nash equilibrium [14], and is amenable to practical implementation [8].

A very different approach is made in OSB that attempts to solve optimization (1) directly [7]. The innovation was to formulate the Lagrangian dual problem. It was then possible to iterate over N separate sub-problems for fixed Lagrangian dual variables, each sub-problem concerning only user powers $\mathbf{P}^n$ on tone $n \in [1, N]$. Each sub-problem is solved with a brute-force grid-search having $L = P^{max}/\Delta_P$ quantized power levels, requiring at least L^K operations each. An outer loop then updated the Lagrangian dual variables via bisection (or gradient-based) methods.

Although OSB has exponential complexity in the number of users, it has shown significant performance gains are possible over IWF. More recently, ISB algorithms were introduced with lower complexity, achieved by approximating the grid-search with a sequence of line-searches [6], [10]. In general, a run of K line-searches are repeated on tone n until convergence before moving on to the next: still a large computational hit!

Until now, there has been little in the way of low-complexity algorithms making use of measurement-based updates that avoid explicit line- or grid-searching. Ideally, such an algorithm would have its computation distributed, with little or no message-passing between modems.

B. *S*uccessive *C*onvex *A*pprox. for *L*ow-compl*E*xity (SCALE)

Our approach is to consider a relaxation of the nonconvex problem (1) to avoid the d.c. structure. We make use of the following lower bound

$$\alpha \log z + \beta \leq \log(1+z) \begin{cases} \alpha = \frac{z_0}{1+z_0} \\ \beta = \log(1+z_0) - \frac{z_0}{1+z_0}\log z_0 \end{cases} \quad (2)$$

that is tight with equality at a chosen value z_0 when the constants $\{\alpha, \beta\}$ are chosen as specified above.

Applying (2) to optimization (1) results in the relaxation

$$\begin{aligned} \max_{\mathbf{P}\geq 0} \quad & \sum_k \sum_n \omega_k \left(\alpha_k^n \log\left(SIR_k^n(\mathbf{P}^n)\right) + \beta_k^n\right) \\ s.t. \quad & \sum_n P_k^n \leq P_k^{max}, \quad \forall k, \end{aligned}$$

where all α_k^n and β_k^n are fixed. This maximization problem is still nonconvex, since the objective is not concave in $\mathbf{P}$. However, a transformation $\widetilde{P}_k^n = \log P_k^n$ results in an standard concave maximization problem in the new variables $\widetilde{P}_k^n$,

$$\begin{aligned} \max_{\widetilde{\mathbf{P}}} \quad & \sum_k \sum_n \omega_k \left(\alpha_k^n \log\left(SIR_k^n(e^{\widetilde{\mathbf{P}}^n})\right) + \beta_k^n\right) \\ s.t. \quad & \sum_n e^{\widetilde{P}_k^n} \leq P_k^{max}, \quad \forall k, \end{aligned} \quad (3)$$

where we denote $e^{\mathbf{x}}$ as an element-by-element operation on the vector $\mathbf{x}$. This is a standard concave maximization since each constraint is a sum of convex exponentials (thus convex), and each term in the objective sum is concave

$$\log\left(SIR_k^n(e^{\widetilde{\mathbf{P}}^n})\right) = \log G_{kk}^n + \widetilde{P}_k^n - \log\left(\sum_{j\neq k} G_{kj}^n e^{\widetilde{P}_j^n} + \sigma_k^n\right),$$

since it comprises a sum of linear and concave terms (where we note that log-sum-exp is convex [4]).

We derive an algorithm to solve the relaxation (3) in Sec. IV-A using gradient methods that are computationally efficient, and without the need for a brute-force or heuristic search of any kind. Once a solution is obtained, we may transform back to the P-space with $P_k^n = \exp(\widetilde{P}_k^n)$.

Here we are maximizing a lower-bound on the achievable system rate. It then becomes natural to improve the bound periodically, resulting in the procedure below. While the initial choice of the constants $\{\alpha, \beta\}$ is not critical, we make use of a simple high-SIR approximation with $\alpha = 1$ and $\beta = 0$.

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE ICC 2006 proceedings.

initialize iteration counter $t = 0$
initialize all $\alpha_k^{n(t)} = 1, \beta_k^{n(t)} = 0$ (high-SIR approx.)
repeat
    maximize: solve sub-problem (3) to give solution $\mathbf{P}^{(t)}$
    tighten: update $\alpha_k^{n(t+1)}, \beta_k^{n(t+1)}$ at $z_0 = SIR_k^n(\mathbf{P}^{(t)})$
    increment t
until convergence

Proposition 1 The sequence of iterates produces a monotonically increasing objective and will always converge.

Proof: Omitted due to space limitations, see [2]. ■

One consequence of Prop. 1 is that at convergence, the feasible power allocation $\mathbf{P}^*$ satisfies the Karush-Kuhn-Tucker optimality conditions of problem (1). Thus $\mathbf{P}^*$ is at least a local optimum of (1). We conjecture that the solution is also globally optimum, our statement supported by initial numerical studies showing similar significant percentage improvements over IWF as the OSB algorithm has demonstrated in [7].

Another consequence of Prop. 1 is that each sub-problem need not be maximized fully; only an improved objective is required. This lends itself toward a distributed tightening step: each user need not wait until convergence of sub-problem t, each tightens at periodic intervals and each tightening step requires only local information. A word of caution however: the relaxed lower-bound is constructed with identical slope at the tightening point as its corresponding nonconvex bounding objective. For the sequence of sub-problems to converge, each partial maximization should be 'steeper' than a simple gradient ascent step on the original nonconvex problem formulation. i.e. a number $D > 1$ (e.g. 5-10) of ascent steps should be performed before tightening.

IV. Distributed Algorithms

A. The SCALE Protocol

In this section we outline a realization of the above procedure. The resulting algorithm then forms the basis for a scalable and distributed protocol where computation is distributed across the network. A concise summary is provided in Fig. 2.

We begin by deriving the algorithmic update of user PSDs through the dual solution to sub-problem (3). Define the Lagrangian function as

$$L_{\mathbb{S}}(\widetilde{\mathbf{P}}, \boldsymbol{\lambda}) = \sum_k \sum_n \omega_k \left(\alpha_k^n \log\left(SIR_k^n(e^{\widetilde{\mathbf{P}}^n}) \right) + \beta_k^n \right) - \sum_k \lambda_k \Big\{ \sum_n e^{\widetilde{P}_k^n} - P_k^{max} \Big\}. \quad (4)$$

The corresponding dual problem is then $\min_{\boldsymbol{\lambda}} \max_{\widetilde{\mathbf{P}}} L_{\mathbb{S}}(\widetilde{\mathbf{P}}, \boldsymbol{\lambda})$. We update dual variables λ_k through a gradient descent

$$\lambda_k^{(s+1)} = \left[\lambda_k^{(s)} + \epsilon \Big(\sum_n P_k^{n(s)} - P_k^{max} \Big) \right]^+ \quad (5)$$

for fixed $\mathbf{P}_k$, where $[\,\cdot\,]^+ = \max(0, \cdot)$, ϵ is a sufficiently small step-size and s is an iteration number for the sub-problem. Each λ_k is updated locally by each user k. As with many other Lagrangian dual formulations, the dual variables have the interpretation of price: as the power constraint is violated, the price λ_k goes up and vice-versa until the equilibrium price $\boldsymbol{\lambda}^*$ is reached that solves the dual sub-problem.

DSL Modem k Algorithm:
 Let $s_k = 0$ be a local iteration counter
 Set all $P_k^n = 0$, $\alpha_k^n = 1$
 At regular intervals:
  Receive messages $\mathcal{M}_k^n$ (8) from SMC
  repeat:
   Update PSD $\mathbf{P}_k$ using (7)
   Update power price λ_k with (5)
  until power price λ_k converges
  Measure noise to obtain $\mathcal{N}_j^n$ (9); send to SMC
  Increment s_k
  Update α_k^n with (2) if $\mod(s_k, D) = 0$
Spectrum Management Center (SMC) Algorithm:
 Repeat forever:
  Receive messages $\mathcal{N}_j^n$ (9) from each Rx j
  Calculate message $\mathcal{M}_q^n$ (8) and send to each Tx q

---

Fig. 2. The SCALE Protocol: A Summary

We solve the inner dual maximization by finding the stationary point of (4) with respect to $\widetilde{\mathbf{P}}$, and with $\boldsymbol{\lambda}$ fixed:

$$\frac{\partial L_{\mathbb{S}}}{\partial \widetilde{P}_k^n} = 0 = \omega_k \alpha_k^n - P_k^n \left(\lambda_k + \sum_{j \neq k} \omega_j \alpha_j^n G_{jk}^n \frac{SIR_j^n(\mathbf{P}^n)}{G_{jj}^n P_j^n} \right),$$

where we have transformed the partial derivative back to the P-space. From this, we can form the fixed-point equation

$$P_k^n = \frac{\omega_k \alpha_k^n}{\lambda_k + \sum_{j \neq k} G_{jk}^n \omega_j \alpha_j^n \frac{SIR_j^n(\mathbf{P}^n)}{G_{jj}^n P_j^n}}, \quad (6)$$

where we note that P_k^n also appears in the denominator of each SIR term above. Should powers be updated iteratively with (6), convergence is easily proved by showing that the right-hand side of (6) is a *standard* interference function [13]; proof omitted due to space restrictions. Convergence of the sub-problem to $\mathbf{P}^{(t)}$ then follows from convexity.

In practice, we need not fully minimize each P_k^n before updating the dual variables $\boldsymbol{\lambda}$. A single ascent step is sufficient, that equates to one iteration of (6). This power update has an elegant intuitive interpretation: the G_{jk}^n terms indicate the impact user k has on all other users j, tone n. Power is allocated in such a way so that it *takes other users into account* on a tone-by-tone basis, rather than a selfish allocation as is done in IWF. Further, should the power constraint be violated, the price of power λ_k is increased, lowering the power to a level within the power budget.

The power update (6) can be distributed through a combination of measurement and message-passing. Re-write it as

$$P_k^{n(s+1)} = \frac{\omega_k \alpha_k^n}{\lambda_k^{(s)} + \mathcal{M}_k^{n(s)}} \quad (7)$$

where $\mathcal{M}_k^{n(s)} \in \mathbb{R}_+$ is a message passed to user k from the SMC, defined as

$$\mathcal{M}_k^{n(s)} = \sum_{j \neq k} G_{jk}^n \mathcal{N}_j^{n(s)}, \quad (8)$$

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE ICC 2006 proceedings.

and is formed by a weighted sum calculation at the SMC. We assume that the SMC has access to estimates of the cross-gains G_{jk}^n, obtained through measurement (see, for e.g., [15]), or with the aid of cross-talk models (for e.g., those standardized in [1]) and some knowledge of the loop topology. The terms $\mathcal{N}_j^{n(s)} \in \mathbb{R}_+$ are also messages from every other user $j \neq k$ on tone n to the SMC,

$$\mathcal{N}_j^{n(s)} = \omega_j \alpha_j^n \frac{SIR_j^n(\mathbf{P}^{n(s)})}{G_{jj}^n P_j^{n(s)}} \tag{9}$$

and is a local quantity at the receiver of user j: a simple scaled noise *measurement* on tone n.

B. A Fresh Look at Iterative Water-Filling

Consider the water-filling sub-problem for user $k \in [1, K]$ in the standard IWF procedure. We can apply the same lower-bound technique of Sec. III-B to form the relaxation

$$\begin{aligned} \max_{\mathbf{P}_k \geq 0} \quad & \sum_n \omega_k \left(\alpha_k^n \log\left(SIR_k^n(P_k^n)\right) + \beta_k^n\right) \\ s.t. \quad & \sum_n P_k^n \leq P_k^{max}, \end{aligned}$$

where again α_k^n and β_k^n are fixed constants, see (2).

This relaxation is maximized by following a similar line of development as outlined above in Sec. IV-A. That is, we again formulate an appropriate Lagrangian dual problem, this time with a single dual variable λ_k associated with the power constraint of sub-problem k. It is straightforward to show that the following solution results:

$$\lambda_k^{(s+1)} = \left[\lambda_k^{(s)} + \epsilon\Big(\sum_n P_k^{n(s)} - P_k^{max}\Big)\right]^+ \tag{10}$$

$$P_k^{n(s+1)} = \frac{\omega_k \alpha_k^n}{\lambda_k^{(s)}} \tag{11}$$

In a similar spirit to the procedure introduced earlier, we can again alternate between maximization and tightening to find a converged solution $\mathbf{P}_k^*$. It can be shown that this is exactly the unique global optimum of the k-th water-filling problem.

Comparing this solution to the SCALE algorithm, we immediately note that the λ_k update (10) is identical to (5). The power update (11) is also identical to the SCALE update (7) *when we disregard the impact user k has on other users*.

This is a significant result: SCALE degrades to IWF when no message-passing is available, or not desired, and more importantly, it motivates the use of *reduced* communication to form a hybrid SCALE-IWF scheme whereby no communication is used on tones enjoying little or no FEXT (i.e. those at low frequencies), and making full use of neighboring line conditions on tones heavily affected by FEXT to improve performance beyond IWF. As the amount of communication reduces to zero, SCALE degrades gracefully to IWF.

C. The SCAWF Algorithm: Improved Iterative Water-Filling

We can improve the convergence speed of the water-filling algorithm (10)–(11) by avoiding the gradient descent on the dual variable λ_k. Realizing that the optimal water-filling solution **always** makes use of the full transmit power P_k^{max} available, we require

$$\sum_n P_k^n = P_k^{max} \;\Rightarrow\; \lambda_k^* = \frac{\omega_k}{P_k^{max}} \sum_n \frac{SIR_k^n(P_k^n)}{1 + SIR_k^n(P_k^n)}, \tag{12}$$

where have substituted (11) and α_k^n to obtain λ_k^*.

Substituting (12) into (11) results in the Successive Convex Approximation for Water-Filling (SCAWF) Algorithm:

$$P_k^{n(s+1)} = P_k^{max} \frac{\frac{SIR_k^n}{1+SIR_k^n}}{\sum_{m=1}^{N} \frac{SIR_k^m}{1+SIR_k^m}}, \tag{13}$$

where the denominator sum is common to the allocation of all tones, and needs to be calculated only once (consequently, numerators are calculated for free). This algorithm is a particularly attractive alternative to the current IWF procedure where a conventional water-filling solution is computed for every user at every iteration. Such water-filling computations often require a sorting step of the channel gains or the use of complex data-structures. Instead, the SCAWF algorithm computes the partial solution (13) that adapts a user's PSD and converges together with all other users to the simultaneous multi-user water-filling solution.

The SCAWF algorithm is also extremely simple: the *SIR* on each tone is periodically measured and (13) computed to form a new power-allocation that is immediately updated. No channel sorting or complex data-structures are required.

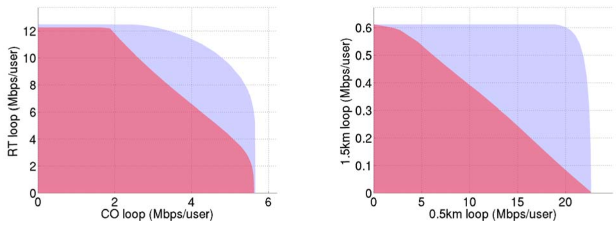
V. Performance

We compare the performance of the SCALE and SCAWF algorithms to IWF in this section. Our evaluations consider VDSL over 26-AWG (0.4mm) lines. A coding gain of 3 dB, with a 6 dB noise margin is assumed, giving a SNR-gap $\Gamma_k^n = 12.8$ dB for an error probability of 10^{-7} [12]. Each modem has maximum transmission power 11.5 dBm, and can transmit in *both* 1U and 2U upstream bands (regional-specific band; former plan 998) [1, Tab. 1] with amateur RF bands notched off [1, Tab. 17]. No other spectral masks are enforced. A DMT symbol rate of 4 kHz is assumed, with tone spacing of 4.3125 kHz. Users are subject to -140 dBm/Hz background noise and alien noises corresponding to ETSI models XA.{L,N}T.{A,D} as appropriate [1, Tab. 21–22]. The cross-gains G_{ij}^m are calculated according to [3] without FSAN combination of FEXT sources, and using standard FEXT models [12].

Our simulations consider $K = 8$ users, split into two equal groups of 4 users. The downstream topology of Fig. 1(a) has a CO-based group placed at 3km and a RT deployment at a distance of 2km, with the second RT-based user group placed 2km further along. The upstream topology of Fig. 1(b) has one group placed 0.5km from the CO, the other at 1.5km.

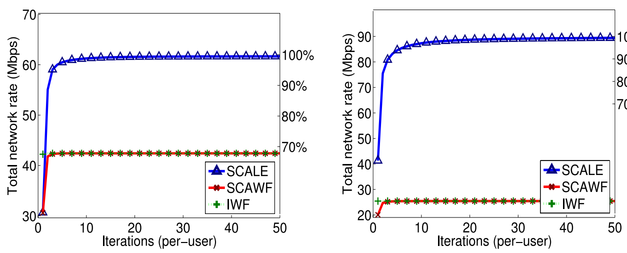
Due to the inherent symmetry in the channel models [12], the resulting rates for users having equal loop lengths end up the same. Fig. 3 then shows the rate-region between two users, one from each user group. The performance improvement of SCALE is clearly significant, where the rate-region is almost

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE ICC 2006 proceedings.



(a) Downstream (b) Upstream

Fig. 3. SCALE (lighter shade) can significantly enlarge the VDSL rate-region compared to IWF (darker shade) in both down- and up-stream topologies.



(a) Downstream (b) Upstream

Fig. 4. Convergence and performance comparison: SCALE has improved system rate with speedy convergence to rates well-above IWF and SCAWF.

doubled over IWF in the upstream direction. These gains stem from an almost disjoint frequency-division separation of the near and far user groups, negotiated automatically by a power allocation that takes other users into account on a tone-by-tone basis. In contrast, the IWF scheme overlaps the spectrum of each user group due to its selfish nature.

The SCALE region is produced by sweeping the weights ω_k and solving a complete RA problem for each tuple. A similar goal is achieved in IWF by pairing back the maximum power budget of each user k by a factor $\delta_k \leq 1$. That is, each user may use a maximum total power of $\delta_k P_k^{max}$, rather than the full power budget P_k^{max}. Selection of the scaling factors δ_k was a challenging process, as minute changes resulted in large differences of the final allocated rates.

We now compare the convergence properties of each algorithm, having selected specific weightings that correspond to particular points within the rate-region.

For the downstream, our selection corresponds to a 4Mbps/user service on CO-based loops. Fig. 4(a) shows the convergence of each scheme, where iterates are shown after all users have updated their PSD. The IWF algorithm converges within two iterations. While the SCAWF algorithm takes an additional iteration to converge to the same result, it requires significantly less computation as outlined in Sec. IV-C. The SCALE algorithm also converges very quickly, and in just two iterations, far exceeds the final performance of IWF.

On the upstream, we select weightings that correspond to a 500kbps/user service on 1.5km loops. Fig. 4(b) shows the corresponding iterations. Convergence rates similar to the downstream direction are observed, where SCALE outperforms IWF after a single iteration.

Initial investigations comparing these results to the OSB and ISB algorithms are still ongoing, as they are much more complex to implement.

VI. Conclusion

Two novel algorithms for spectrum balancing in multi-user DSL networks have been introduced, each enjoying a low-complexity structure. SCALE, the first algorithm, explicitly accounts for the 'damage' a user's power allocation has on other users, resulting in higher achievable rates than the selfish competitive-optimal rates resulting from simultaneous water-filling. Through measurement and limited message-passing, SCALE is easily distributed with the help of a Spectrum Management Center. Message-passing overhead can be arbitrarily traded for performance, and in the limit of zero overhead, SCALE can perform no worse than iterative water-filling. Numerical studies have shown SCALE converging to rates far exceeding that of iterative water-filling with just 2-3 iterations, and to within 90% of the final rate in under 5 iterates.

The second algorithm, SCAWF, was shown to be an extremely simple way to water-pour that is particularly well-suited to iterative multi-user water-filling problems.

Future work includes consideration of discrete bit-loading, numerical performance evaluation against existing OSB & ISB schemes and other nonconvex solution methods well-suited to the underlying d.c. programming problem. Some of these items are currently under investigation [2].

References

[1] "ETSI TS 101 270-1 v1.3.1 Standard. Part 1: Functional Requirements," 2003-07.
[2] For latest developments, see `http://jpap.yi.org/research/`.
[3] R. Baldemair, M. Horvat, and T. Nordstöm, "Proposed method on crosstalk calculations in a distributed environment," *ETSI STC TM6 Plenary meeting*, 2003.
[4] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.
[5] J. Campello, "Practical bit loading for DMT," in *IEEE ICC*, vol. 2, June 1999, pp. 801–805.
[6] R. Cendrillon and M. Moonen, "Iterative spectrum balancing for digital subscriber lines," in *IEEE ICC*, vol. 3, May 2005, pp. 1937–1941.
[7] R. Cendrillon, W. Yu, M. Moonen, J. Verliden, and T. Bostoen, "Optimal multi-user spectrum management for digital subscriber lines," *IEEE Trans. on Comm.*, to appear.
[8] E. V. den Bogaert, T. Bostoen, J. V. Elsen, R. Cendrillon, and M. Moonen, "DSM in practice: iterative water-filling implemented on ADSL modems," in *ICASSP*, vol. 5, May 2004, pp. 337–340.
[9] R. Horst and H. Tuy, *Global optimization: deterministic approaches*, 2nd ed. Berlin; New York: Springer-Verlag, 1993.
[10] R. Lui and W. Yu, "Low-complexity near-optimal spectrum balancing for digital subscriber lines," in *IEEE ICC*, vol. 3, 2005, pp. 1947–1951.
[11] K. B. Song, S. T. Chung, G. Ginis, and J. M. Cioffi, "Dynamic spectrum management for next-generation DSL systems," *IEEE Comm. Mag.*, vol. 40, no. 10, pp. 101–109, 2002.
[12] T. Starr, J. M. Cioffi, and P. J. Silverman, *Understanding Digital Subscriber Line Technology*. NJ, USA: Prentice Hall, 1999.
[13] R. D. Yates, "A framework for uplink power control in cellular radio systems," *IEEE Journal on Selected Areas in Comm.*, vol. 13, no. 7, pp. 1341–1347, 1995.
[14] W. Yu, G. Ginis, and J. M. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE JSAC*, vol. 20, no. 5, pp. 1105–1115, June 2002.
[15] C. Zeng, C. Aldana, A. A. Salvekar, and J. M. Cioffi, "Crosstalk identification in xDSL systems," *IEEE JSAC*, vol. 19, no. 8, pp. 1488–1496, Aug. 2001.

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE ICC 2006 proceedings.