

Multiuser Spectrum Optimization for Discrete Multitone Systems With Asynchronous Crosstalk

Vincent M. K. Chan and Wei Yu, *Member, IEEE*

Abstract—Finding computationally efficient spectrum optimization methods for a multiuser digital subscriber line (DSL) environment is a key objective of dynamic spectrum management for DSL systems. For synchronous discrete multitone (DMT) based DSL systems, the computational complexity issue can be partially addressed using a dual optimization approach that decomposes the problem in the frequency domain. However, the decomposition approach fails whenever the DMT system is asynchronous, in which case the power allocations in adjacent frequency tones interfere with each other. This paper provides a mathematical model for intercarrier interference due to symbol misalignment and proposes ways to modify the dual method for spectrum optimization of asynchronous DSL systems. The main ingredient of the algorithm is an efficient method to evaluate the minimum power required to support a given bit-allocation using a power series approximation, combined with a local gradient search.

Index Terms—Digital subscriber lines, dynamic spectrum management, intercarrier interference, spectrum balancing.

I. INTRODUCTION

SPECTRAL adaptivity is a significant benefit for discrete multitone (DMT)-based communication systems. In a DMT system, the frequency domain is partitioned into a large number of narrow tones each carrying independent bits. Frequency partition is achieved with an inverse fast Fourier transform (IFFT) at the transmitter and a fast Fourier transform (FFT) at the receiver. The Fourier basis are orthogonal at the transmitter, and by the use of a cyclic prefix, they remain orthogonal at the output of the channel, so long as the channel is linear and time-invariant and the channel memory is no longer than the cyclic prefix. The orthogonality property makes the optimization of bit and power allocation over frequency easy to implement for a single-user DMT system. When multiple DMT-based devices share the same physical medium, orthogonality among subchannels of the intended signal and the subchannels of the interference signal in different frequency tones is maintained for *synchronous* DMT systems, where the DMT symbols of the intended signal and the interfering signals are in perfect symbol synchronization with

each other at the receiver. In practice, when alignment cannot be guaranteed due to, for example, the channel time delay, cyclic suffix can be inserted to ensure orthogonality. Cyclic suffix is used in Very high bit-rate DSL (VDSL) systems. Zipper VDSL is one of the standardized VDSL systems based on the principle of cyclic suffix [1], [2].

The orthogonality of the frequency dimension simplifies the multiuser spectrum optimization problem significantly. In a synchronous system, the interference between DMT systems only occurs in a tone-by-tone basis. More specifically, let P_k^n denote the transmit power spectral density for the k th user at tone n . Symbol synchronization ensures that P_k^n interferes only with the n th tone of all coexisting DMT systems. In this case, the spectrum optimization problem that maximizes the weighted sum rate with weights w_k is formulated as

$$\begin{aligned} \max \sum_{k=1}^K w_k \sum_{n=1}^N \log_2 \left(1 + \frac{1}{\Gamma} \frac{|H_{kk}^n|^2 P_k^n}{\sigma_k^n + \sum_{j \neq k} |H_{jk}^n|^2 P_j^n} \right) \\ \text{s.t.} \sum_{n=1}^N P_k^n \leq P_{k,\text{total}} \quad k = 1, \dots, K \\ P_k^n \geq 0, \quad k = 1, \dots, K \end{aligned} \quad (1)$$

where H_{jk}^n is the channel transfer function from user j to user k in tone n , σ_k^n is the background noise power for user k in tone n , $P_{k,\text{total}}$ is the power constraint for user k , and Γ is the signal-to-noise ratio (SNR) gap. Note that the objective function is a sum of many decoupled terms, and the only coupling between the tones occurs in the power constraint. In this case, optimal spectrum balancing (OSB) algorithms can be applied [3], [4].

The main idea of OSB is that the Lagrangian dual decomposition method can be used even if the objective function is nonconvex. The dual decomposition method decouples the constraint in (1) by the use of a dual variable. The optimization problem is then decomposed into N smaller subproblems, which, in turn, can be solved individually with a lower complexity. The global optimum solution is then obtained via a subgradient search on the dual variable. Further simplification of OSB is possible, if the dual problem is solved approximately using an iterative approach. These iterative algorithms are known as iterative spectrum balancing (ISB) [4], [5].

The crucial condition that leads to the decoupling of the optimization problem, namely the synchronization condition or equivalently the insertion of cyclic suffix, is however not necessarily satisfied in practice. DMT systems deployed at present

Manuscript received July 31, 2006; revised February 12, 2007. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Mounir Ghogho. This work was supported in part by Bell Canada, by Communications and Information Technology Ontario (CITO), and in part by the Natural Science and Engineering Council (NSERC) of Canada under a collaborative research and development (CRD) Grant.

V. M. K. Chan was with the Edward S. Rogers, Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada. He is now with ViXS Systems, Inc., Toronto, ON, Canada (e-mail: vimkchan@hotmail.com).

W. Yu is with the Edward S. Rogers, Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada (e-mail: weiyu@comm.utoronto.ca).

Digital Object Identifier 10.1109/TSP.2007.896076

often do not allow an easy insertion of suffix; DMT modems belonging to different service providers are unlikely to be synchronized. In these cases, the crosstalk generated by a DMT system in one frequency tone interferes not only with other DMT systems at the same frequency tone, but also at neighboring frequency tones. This results in intercarrier interference (ICI) and leads to tone-coupling in the objective function of the spectrum optimization problem.

If OSB is applied while ignoring the presence of ICI, the resulting power spectra can lead to poor performance. This is because OSB tends to over-exploit channel characteristics. This paper represents a first effort in identifying new spectrum balancing algorithms tailored for DMT systems with asynchronous crosstalk interference. The main contributions of this paper are as follows. First, a model of the ICI for asynchronous DMT is derived. Second, several heuristic algorithms are proposed to maximize transmission rates in the presence of ICI. Finally, a complexity reduction technique for the proposed algorithm based on power series expansion is suggested.

ICI modeling for DMT systems has been done in the past for systems with an insufficient guard interval duration [6], [7]. ICI is induced when the length of the cyclic prefix is shorter than the intersymbol interference (ISI) channel impulse response. This paper introduces a different ICI mechanism resulting from an offset between the DMT frames of the direct and the interfering signals. This ICI phenomenon is commonly known in the DSL community. However, to the best of the authors' knowledge, an explicit mathematical model of ICI due to asynchronism has not appeared in the literature.

This paper proposes several novel algorithms for solving the spectrum optimization problem with ICI. The *greedy bit-adding* algorithm allocates the available power by assigning an additional bit to the tone that requires the least amount of power in each iteration starting from a low power-spectral level. In contrast, the *greedy bit-subtracting* algorithm lowers the bit assignment from a high power-spectral level. These techniques are similar to previous bit-loading methods proposed in the past for the ICI-free power control problem [8]. However, in the ICI-free case, the exact additional power can be computed in each iteration, while our proposed algorithms need to use gradient information to approximate the power adjustment when a bit is loaded or removed from the bit allocation table.

This paper also proposes a computational complexity reduction technique for the proposed algorithm by employing a first-order power series expansion of a matrix inversion. A similar technique has been applied in the past in other DSL problems for approximating the inverse of a matrix with dominant diagonal terms [9]. The proposed technique is similar but applies to block-diagonal matrices. One weakness of this approximation is that it cannot be applied in DSL configurations when the power series expansion does not converge. Nevertheless, when this method is applicable, simulation shows that this simplifying technique causes only minor performance degradation.

This paper is mainly concerned with spectrum adaptation methods assuming a fixed amount of ICI. In practical systems, it is also possible to reduce the ICI itself by transmit-windowing techniques. This paper assumes that transmit windowing is not done, as is the case for most currently deployed systems.

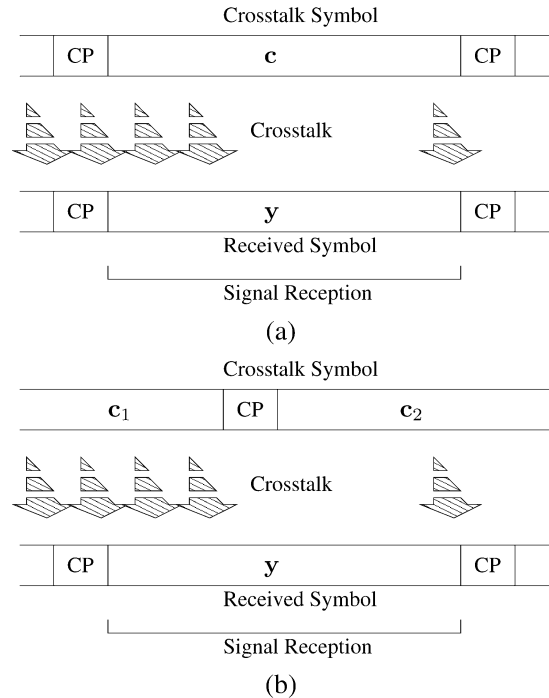


Fig. 1. (a) If two DMT symbols are synchronous at the receiver, the crosstalk channel appears to be circulant. (b) If the DMT symbols are not synchronous, the received symbol contains interference from two crosstalk symbols. The combination of the two crosstalk symbols does not have a cyclic structure. Consequently, the crosstalk channel does not appear to be circulant, which causes ICI.

The rest of the paper is organized as follows. Section II contains the new ICI model. The proposed spectrum optimization algorithms for the ICI problem are presented in Section III. Section IV contains simulation results. The paper concludes with Section V.

II. ASYNCHRONOUS DMT

The effect of imperfect DMT block synchronization is investigated in this section. Crosstalk from asynchronous neighboring user is quantified. The resulting expression for ICI is incorporated into the spectrum optimization problem (1).

A. Model of Imperfect DMT Block Synchronization

The effect of imperfect block synchronization can be visualized as shown in Fig. 1. Fig. 1(a) illustrates a DMT system in which the two received DMT symbols—the intended signal and the crosstalk signal—are aligned at the receiver. In this case, the FFT operation at the receiver jointly diagonalizes both the intended signal and the crosstalk signal in the frequency domain, resulting in no ICI. Fig. 1(b) illustrates a situation in which the crosstalk symbol is not synchronized with the intended signal. In this case, the received signal contains interference from three sources: the interference from the previous symbol c_1 of the crosstalk signal, the cyclic prefix c_p , and the interference from the next symbol c_2 . The receiver FFT operates across the boundaries of all three components. Orthogonality is no longer maintained, which results in ICI.

Let \mathbf{x} be the transmit DMT symbol of the intended signal. Let \mathbf{y} represent the receive signal in the time domain. Let \mathbf{c} represent the crosstalk transmit vector. Let \mathbf{n} represent the back-

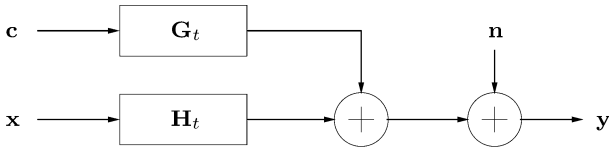


Fig. 2. DSL channel model.

ground noise vector. The crosstalk signal interferes with the transmit symbol according to the model shown in Fig. 2. The direct channel \mathbf{H}_t and the crosstalk channel \mathbf{G}_t are ISI channels, which can be represented by Toeplitz matrices, denoted here by the subscript $(\cdot)_t$.

By adding the cyclic prefix in the transmitted symbol, the Toeplitz matrices of the ISI channels are transformed into matrices with the circulant structure, denoted here by the subscript $(\cdot)_c$. Let \mathbf{H}_c represent the channel matrix for the intended signal. Let \mathbf{G}_c represent the N -by- N crosstalk channel matrix for the interfering signal. When the DMT transmission is synchronized, the received symbol can be calculated as follows:

$$\mathbf{y} = \mathbf{H}_c \mathbf{x} + \mathbf{G}_c \mathbf{c} + \mathbf{n}. \quad (2)$$

Due to the circulant nature of \mathbf{H}_c and \mathbf{G}_c , if an IFFT is performed at the transmitter and an FFT is performed at the receiver, the channel is diagonalized as follows. Let \mathbf{Q} be the N -by- N FFT matrix and \mathbf{Q}^* be the corresponding IFFT matrix. Then the transmit vectors are $\mathbf{x} = \mathbf{Q}^* \mathbf{X}$ and $\mathbf{c} = \mathbf{Q}^* \mathbf{C}$, and the receive symbols are recovered by $\mathbf{Y} = \mathbf{Q} \mathbf{y}$. Substituting these expressions into (2) yields

$$\mathbf{Y} = \mathbf{Q} \mathbf{H}_c \mathbf{Q}^* \mathbf{X} + \mathbf{Q} \mathbf{G}_c \mathbf{Q}^* \mathbf{C} + \mathbf{Q} \mathbf{n}. \quad (3)$$

The circulant nature of \mathbf{H}_c and \mathbf{G}_c results in diagonal matrices $\mathbf{Q} \mathbf{H}_c \mathbf{Q}^*$ and $\mathbf{Q} \mathbf{G}_c \mathbf{Q}^*$. Hence, interference from \mathbf{C} is injected into \mathbf{Y} on a tone-by-tone basis.

If, instead, the DMT blocks are not synchronized, (2) becomes

$$\mathbf{y} = \mathbf{H}_c \mathbf{x} + \mathbf{G}_t \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_p \\ \mathbf{c}_2 \end{bmatrix} + \mathbf{n}. \quad (4)$$

The crosstalk channel matrix \mathbf{G}_t is no longer circulant but only has a Toeplitz structure. In this case, \mathbf{c}_1 represents interference from the previous symbol. Thus, $\mathbf{c}_1 = \mathbf{Q}_l^* \mathbf{C}_1$, where \mathbf{Q}_l^* is the lower portion of \mathbf{Q}^* ; $\mathbf{c}_p = \mathbf{Q}_p^* \mathbf{C}_2$, where \mathbf{Q}_p^* is the submatrix of \mathbf{Q}^* corresponding to the cyclic prefix; $\mathbf{c}_2 = \mathbf{Q}_u^* \mathbf{C}_2$, where \mathbf{Q}_u^* is the upper portion of \mathbf{Q}^* . Combining these expressions

$$\mathbf{Y} = \mathbf{Q} \mathbf{H}_c \mathbf{Q}^* \mathbf{X} + \mathbf{Q} \mathbf{G}_t \begin{bmatrix} \mathbf{Q}_l^* \mathbf{C}_1 \\ \mathbf{Q}_p^* \mathbf{C}_2 \\ \mathbf{Q}_u^* \mathbf{C}_2 \end{bmatrix} + \mathbf{Q} \mathbf{n}. \quad (5)$$

Note that the crosstalk channel matrix is no longer diagonalized by the FFT operation at the receiver. Hence, power allocated to any tone in \mathbf{C}_1 or \mathbf{C}_2 interferes with all other tones in \mathbf{X} .

B. Quantifying Intercarrier Interference

To quantify ICI, one may think of \mathbf{Q}_l^* and \mathbf{Q}_u^* as IFFT operations multiplied by rectangular window functions in the time domain. Fig. 3 illustrates the formation of the crosstalk signal. In the figure, $C_1[n]$ and $C_2[n]$ are the two consecutive crosstalk DMT symbols, where n denotes the frequency tone index; $r_\nu[t]$ is a rectangular window function of width ν in time domain. The two time-domain signals $c_1[t]$ and $c_2[t]$ are truncated by rectangular window functions. Then, the two resulting signals are summed up to form the crosstalk signal $\tilde{c}[t]$. For simplicity, the effect of the cyclic prefix is disregarded here because the length of the cyclic prefix is usually small compared to the DMT block length.

Note that the system shown in Fig. 3 is a linear system. Also, a multiplication with $r_\nu[t]$ in the time domain is equivalent to a convolution with the FFT of $r_\nu[t]$ in the frequency domain. Thus, the system is ‘‘frequency invariant’’ in the sense that it is sufficient to examine the ICI caused by the crosstalk at each single tone individually. Superposition can then be used to determine the total ICI.

Assume that the power spectrum does not change between $C_1[n]$ and $C_2[n]$, (which is always the case for stationary DMT systems). Assume also that the power spectrum of $C_1[n]$ and $C_2[n]$ is a unit delta sequence

$$\delta[n] = \begin{cases} 1 & \text{if } n = 0 \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

where n is the frequency tone index. The power spectrum of $\tilde{c}[t]$ can then be derived analytically. The analytic expression, denoted as $\rho_\nu[n]$, is rigorously derived in Appendix A. It has the following form:

$$\rho_\nu[n] = \begin{cases} \frac{\nu^2 + (N-\nu)^2}{N^2} & \text{if } n = 0 \\ \frac{2 \sin^2(\frac{\pi}{N} n \nu)}{N^2 \sin^2(\frac{\pi}{N} n)} & \text{otherwise.} \end{cases} \quad (7)$$

The crosstalk spectrum depends on the offset ν . In practice, ν is often not known in advance. So, for the rest of the paper, the worst-case ν is considered. In this case, the crosstalk power spectrum originating from $\delta[n]$ is

$$\begin{aligned} \gamma[n] &= \max_\nu \rho_\nu[n] \\ &= \frac{2}{N^2 \sin^2(\frac{\pi}{N} n)} \quad \text{if } n \neq 0 \end{aligned} \quad (8)$$

and $\gamma[0] = 1$.

Since the truncation and summation operations are frequency invariant, applying $\delta[n-r]$ instead of $\delta[n]$ for $C_1[n]$ and $C_2[n]$ simply corresponds to a frequency shift in (8). Hence, $\gamma[n]$ represents the interference power induced on a tone that is n tones away from the crosstalk tone. Also, the interference pattern has a circular structure so that $\gamma[-n] = \gamma[N-n]$. For the rest of the paper, $\gamma[n]$ is termed as the ICI coefficients.

The ICI coefficients for Asymmetric DSL (ADSL) where $N = 256$ are plotted in Fig. 4. The magnitudes of ICI are about -6.9 dB at the first adjacent tone, -20.9 dB at the fifth adjacent

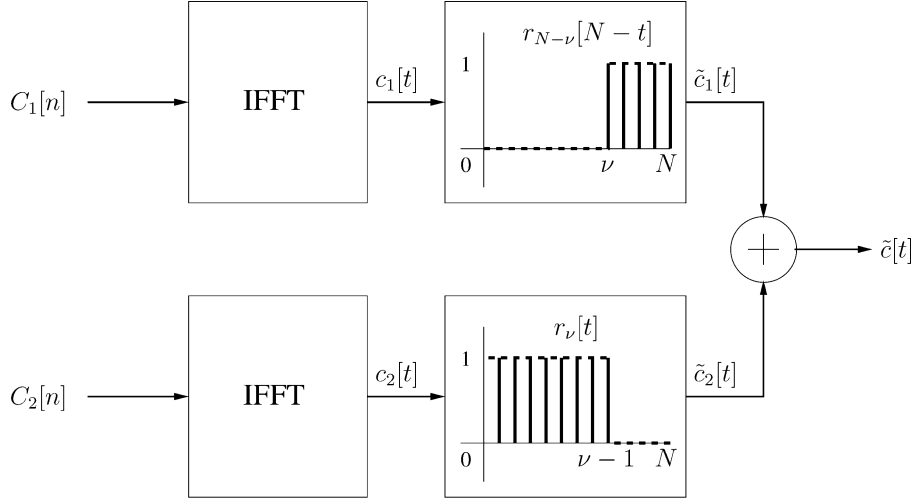
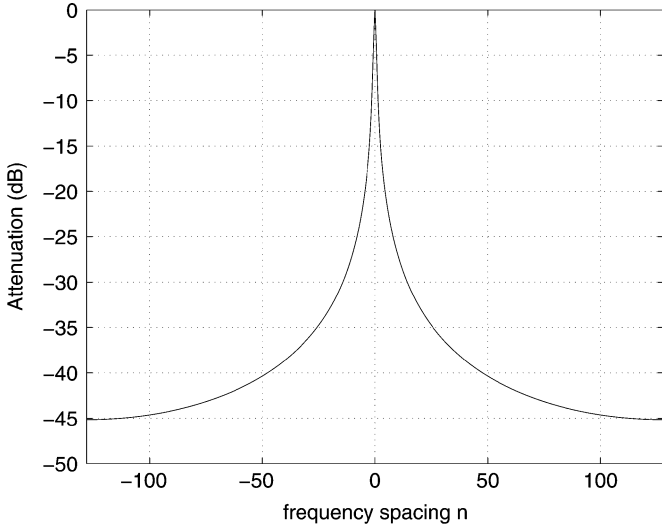


Fig. 3. Formation of intercarrier interference.

Fig. 4. Inter-carrier interference coefficients $\gamma[n]$ for ADSL ($N = 256$).

tone and -30.4 dB at the fifteenth adjacent tone, which are significant.

With ICI, the total crosstalk interference $P_{k,\text{ICI}}^n$ experienced at tone n by user k can be estimated as $P_{k,\text{ICI}}^n \approx \sum_{j \neq k} \sum_{m=0}^{N-1} \gamma[m] |H_{jk}^{(n-m)}|^2 P_j^{(n-m)}$. Again, a circular structure is assumed for the crosstalk signal power so that tone $-n$ refers to tone $N - n$.

The above ICI estimation method assumes that the time-dispersiveness of the channel can be modeled as a multiplication in the frequency domain by $|H_{jk}^{(n-m)}|^2$. This is only an approximation because the crosstalk channel is not circulant. The accuracy of such an approximation is proportional to the ratio of the cyclic prefix length and the DMT block length. Also, the cyclic prefix between $c_1[t]$ and $c_2[t]$ has also not been taken into account, which induces an additional approximation error which is proportional to the cyclic prefix length. Taking both approximations into account, in ADSL and VDSL systems where cyclic prefix is 7.8% of the DMT block length, the approximation error is at most 1 dB.

Note also that frequency offset has not been considered in this computation. In real systems where different transmitters may use slightly different clock signals, additional ICI may result. In wireless system where channel may be time-varying, yet another source of ICI may be present.

C. Spectrum Optimization in the Presence of ICI

Taking ICI into account, the spectrum optimization problem can now be set up using the results from the above section. Replacing the crosstalk power expression in (1) by the above model, the optimization problem becomes

$$\begin{aligned} \max \quad & \sum_{k=1}^K w_k \sum_{n=1}^N \log_2 \left(1 + \frac{1}{\Gamma} \cdot \frac{|H_{kk}^n|^2 P_k^n}{\sigma_k^n + P_{k,\text{ICI}}^n} \right) \\ \text{s.t.} \quad & \sum_{n=1}^N P_k^n \leq P_{k,\text{total}} \quad k = 1, \dots, K \\ & P_k^n \geq 0, \quad k = 1, \dots, K \end{aligned} \quad (9)$$

where

$$P_{k,\text{ICI}}^n = \sum_{j \neq k} \left(\sum_{m=0}^{N-1} \gamma[m] |H_{jk}^{(n-m)}|^2 P_j^{(n-m)} \right). \quad (10)$$

This is a nonconvex problem with coupled utilities.

III. SPECTRUM OPTIMIZATION ALGORITHMS FOR ASYNCHRONOUS DSL

The spectrum optimization problem with ICI is a difficult problem numerically because the problem is nonconvex, and because the problem does not decompose in a tone-by-tone basis. So, OSB is no longer applicable. Brute-force search is not feasible because its complexity is exponential in the number of tones, which is typically large. This paper proposes heuristic algorithms that have a polynomial complexity in the number of tones and the number of users.

The main idea is to first use OSB or ISB while ignoring ICI to find a reasonably good initial solution, then to perform neighborhood search using a greedy bit-adding and a greedy bit-sub-

tracting algorithm, as described below. The search direction is determined by computing the gradient of the objective function. Note that applying gradient search on a nonconvex surface does not guarantee global optimality. Extensive simulation results are used to show that the algorithms generate solutions that have a good performance.

A. Bit-Adding Algorithm

The first step of the greedy bit-adding algorithm is to apply ISB to find an initial power spectra $\{\tilde{\mathbf{P}}_1, \dots, \tilde{\mathbf{P}}_K\}$. This set of power spectra is near-optimal if the DMT system is block synchronized. It will be used as a starting point upon which the subsequent neighborhood search is based. Other choices of starting point, such as that from iterative water-filling [10], are also possible. However, simulation shows that using ISB for initialization has a better performance.

Suppose that without ICI, $\{\tilde{\mathbf{P}}_1, \dots, \tilde{\mathbf{P}}_K\}$ allows $\{\tilde{\mathbf{b}}_1, \dots, \tilde{\mathbf{b}}_K\}$ to be transmitted, i.e.

$$\tilde{b}_k^n = \left\lfloor \log_2 \left(1 + \frac{1}{\Gamma} \cdot \frac{\tilde{P}_k^n |H_{kk}^n|^2}{\sigma_k^n + \sum_{j \neq k} |H_{jk}^n|^2 \tilde{P}_j^n} \right) \right\rfloor. \quad (11)$$

Here, the $\lfloor \cdot \rfloor$ operation is used to ensure discrete bit-loading. Due to the additional asynchronous crosstalk, the actual rates transmitted in a block asynchronous system would be $\{\mathbf{b}_1^{(0)}, \dots, \mathbf{b}_K^{(0)}\}$ such that $\mathbf{b}_k^{(0)} \leq \tilde{\mathbf{b}}_k$.

Let $\mathbf{b}_k^{(i)}$ denote the bit allocated for user k at the i th iteration of the proposed algorithm. The bit allocation scheme $\{\mathbf{b}_1^{(0)}, \dots, \mathbf{b}_K^{(0)}\}$ can be calculated by

$$b_k^{n,(0)} = \left\lfloor \log_2 \left(1 + \frac{1}{\Gamma} \cdot \frac{\tilde{P}_k^n |H_{kk}^n|^2}{\sigma_k^n + \tilde{P}_{k,\text{ICI}}^n} \right) \right\rfloor. \quad (12)$$

Next, we compute the exact power required to transmit $\{\mathbf{b}_1^{(0)}, \dots, \mathbf{b}_K^{(0)}\}$, denoted by $\{\mathbf{P}_1^{(0)}, \dots, \mathbf{P}_K^{(0)}\}$. The key idea is to observe that $\{\mathbf{P}_1^{(0)}, \dots, \mathbf{P}_K^{(0)}\}$ can be determined by solving a system of equations derived from (12) as follows:

$$\text{diag} \left(\begin{bmatrix} \sigma_1 \\ \vdots \\ \sigma_K \end{bmatrix} \right) \begin{bmatrix} 2^{\mathbf{b}_1^{(i)}} - 1 \\ \vdots \\ 2^{\mathbf{b}_K^{(i)}} - 1 \end{bmatrix} \cdot \Gamma = \Theta \begin{bmatrix} \mathbf{P}_1^{(i)} \\ \vdots \\ \mathbf{P}_K^{(i)} \end{bmatrix} \quad (13)$$

where Θ is an NK -by- NK matrix

$$\Theta = \begin{bmatrix} \Theta_{11} & \cdots & \Theta_{1K} \\ \vdots & \ddots & \vdots \\ \Theta_{K1} & \cdots & \Theta_{KK} \end{bmatrix} \quad (14)$$

with Θ_{jk} defined as

$$\Theta_{jk} = \begin{cases} \text{diag}(\mathbf{H}_{kk}) & \text{for } j = k \\ -\text{diag}(2^{\mathbf{b}_j^{(i)}} - 1) \cdot \mathbf{A} \cdot \text{diag}(\mathbf{H}_{kj}) & \text{for } j \neq k \end{cases} \quad (15)$$

In the above, $\text{diag}(\mathbf{v})$ denotes the diagonal matrix formed by vector \mathbf{v} . The variables $\mathbf{H}_{kk}, \mathbf{H}_{jk}, \sigma_k$ are N -by-1 vectors whose n th element are $|H_{kk}^n|^2, |H_{jk}^n|^2, \sigma_k^n$, respectively, $\mathbf{P}_k^{(i)}$ denotes the power spectrum for user k at the i th iteration and

$$\mathbf{A} = \begin{bmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(N-1) \\ \gamma(1) & \gamma(0) & \cdots & \gamma(N-2) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(N-1) & \gamma(N-2) & \cdots & \gamma(0) \end{bmatrix}. \quad (16)$$

Equation (13) grants us means to convert any bit allocation scheme $\{\mathbf{b}_1^{(i)}, \dots, \mathbf{b}_K^{(i)}\}$ to its corresponding power allocation scheme $\{\mathbf{P}_1^{(i)}, \dots, \mathbf{P}_K^{(i)}\}$, or vice versa. This is a multitone generalization of a key step in OSB, where the bit-to-power mapping is computed in a tone-by-tone basis.

Note that due to the $\lfloor \cdot \rfloor$ operation in (12) $\mathbf{b}_k^{(0)} \leq \tilde{\mathbf{b}}_k, \forall k$. It can be shown that the corresponding power vectors must also satisfy $\mathbf{P}_k^{(0)} \leq \tilde{\mathbf{P}}_k, \forall k$. This is proved in Appendix B. Now, define $\tilde{P}_k^{(i)} \triangleq P_{k,\text{total}} - \sum_n P_k^{n,(i)}$ as the slack power given by the power constraint at the i th iteration. The goal of the subsequent iterative steps is to allocate the slack power to tones where the highest rate increase can be achieved.

Given $\{\mathbf{b}_1^{(0)}, \dots, \mathbf{b}_K^{(0)}\}$ and $\{\mathbf{P}_1^{(0)}, \dots, \mathbf{P}_K^{(0)}\}$, the subsequent steps are to perform neighborhood search on $\{\mathbf{b}_1^{(i)}, \dots, \mathbf{b}_K^{(i)}\}$ and to determine $\{\mathbf{b}_1^{(i+1)}, \dots, \mathbf{b}_K^{(i+1)}\}$ based on the gradient of the objective function at $\{\mathbf{P}_1^{(i)}, \dots, \mathbf{P}_K^{(i)}\}$. Our objective is the weighted sum rate

$$J = \sum_{k=1}^K w_k \sum_{n=1}^N \log_2 \left(1 + \frac{1}{\Gamma} \cdot \frac{P_k^n |H_{kk}^n|^2}{\sigma_k^n + P_{k,\text{ICI}}^n} \right). \quad (17)$$

The gradient $\Lambda(\mathbf{P}_1, \dots, \mathbf{P}_K)$ of J at $(\mathbf{P}_1, \dots, \mathbf{P}_K)$ is

$$\Lambda(\mathbf{P}_1, \dots, \mathbf{P}_K) = \left[\frac{\partial J}{\partial P_1^1}, \dots, \frac{\partial J}{\partial P_1^N}, \dots, \frac{\partial J}{\partial P_K^1}, \dots, \frac{\partial J}{\partial P_K^N} \right] \quad (18)$$

where

$$\frac{\partial J}{\partial P_k^n} = \frac{w_k}{\ln 2} \left(\frac{|H_{kk}^n|^2}{P_k^n |H_{kk}^n|^2 + \Gamma \Phi_k^n} \right) - \sum_{j \neq k} \frac{w_j}{\ln 2} \left(\frac{\sum_{r=1}^N \frac{P_j^r |H_{jj}^r|^2 \gamma(r-n) |H_{kj}^n|^2}{\Phi_j^r \cdot (P_j^r |H_{jj}^r|^2 + \Gamma \Phi_j^r)}}{\Phi_j^r \cdot (P_j^r |H_{jj}^r|^2 + \Gamma \Phi_j^r)} \right) \quad (19)$$

with $\Phi_k^n = \sigma_k^n + \sum_{j \neq k} \sum_{l=1}^N \gamma(l-n) |H_{jk}^l|^2 P_j^l$.

For the bit-adding algorithm, the *largest* positive element in $\Lambda(\mathbf{P}_1, \dots, \mathbf{P}_K)$ is chosen and the corresponding bit in $\{\mathbf{b}_1^{(i)}, \dots, \mathbf{b}_K^{(i)}\}$ is *incremented* by one to yield $\{\mathbf{b}_1^{(i+1)}, \dots, \mathbf{b}_K^{(i+1)}\}$. Note that $\{\mathbf{P}_1^{(i)}, \dots, \mathbf{P}_K^{(i)}\}$ must be recalculated using (13) in each iteration as power expenditure increases in each iteration and the gradient vector $\Lambda(\mathbf{P}_1^{(i)}, \dots, \mathbf{P}_K^{(i)})$ also needs to be recomputed. This process of adding bit by bit according to the gradient carries on until

the slack power \tilde{P}_k is depleted for one of the users. The greedy bit-adding algorithm is summarized as follows:

Greedy Bit-Adding Algorithm

- 1) Run the iterative spectral balancing algorithm to obtain $\{\tilde{\mathbf{P}}_1, \dots, \tilde{\mathbf{P}}_K\}$.
- 2) Calculate $\{\mathbf{b}_1^{(0)}, \dots, \mathbf{b}_K^{(0)}\}$ from $\{\tilde{\mathbf{P}}_1, \dots, \tilde{\mathbf{P}}_K\}$ using (12). Find $\{\mathbf{P}_1^{(0)}, \dots, \mathbf{P}_K^{(0)}\}$ using (13).
- 3) Set $i = 0$.
- 4) Loop
 - a) Select the (k, n) pair that has the *largest* gradient value in (18) and *increment* $b_k^{n,(i)}$ to obtain $\{\mathbf{b}_1^{(i+1)}, \dots, \mathbf{b}_K^{(i+1)}\}$. Compute $\{\mathbf{P}_1^{(i)}, \dots, \mathbf{P}_K^{(i)}\}$ using (13). Compute the slack power \tilde{P}_k .
 - b) Exit loop if $\tilde{P}_k \leq 0$ for one of the users.
 - c) Set $i = i + 1$.

End loop

B. Bit-Subtracting Algorithm

Alternatively, it is also possible to use a bit-subtracting algorithm by first obtaining an upper bound on the power spectra at the initialization by setting $\{\mathbf{b}_1^{(0)}, \dots, \mathbf{b}_K^{(0)}\}$ as $\{\tilde{\mathbf{b}}_1, \dots, \tilde{\mathbf{b}}_K\}$ from (11). Then $\{\mathbf{P}_1^{(0)}, \dots, \mathbf{P}_K^{(0)}\}$ can be determined from $\{\mathbf{b}_1^{(0)}, \dots, \mathbf{b}_K^{(0)}\}$ using (13). It can be shown that $\mathbf{P}_k^{(0)} \geq \tilde{\mathbf{P}}_k$. Define $\hat{P}_k^{(i)} \triangleq \sum_n P_k^{n,(i)} - P_{k,\text{total}}$ as the extra power required to transmit $\{\mathbf{b}_1^{(i)}, \dots, \mathbf{b}_K^{(i)}\}$. The bit-subtracting algorithm finds the (k, n) pair with the *smallest* gradient and *decrements* that bit in each iteration. In the meanwhile, $\hat{P}_k^{(i)}$ will be gradually decreased.

The bit-adding and bit-subtracting algorithms differ in the terminating condition. To ensure that all users satisfy the power requirements, the bit-subtracting algorithm must continue until $\hat{P}_k^{(i)} \leq 0$ for all users. The bit-subtracting algorithm is summarized in the following:

Greedy Bit-Subtracting Algorithm

- 1) Run the iterative spectral balancing algorithm to obtain $\{\tilde{\mathbf{P}}_1, \dots, \tilde{\mathbf{P}}_K\}$.
- 2) Calculate $\{\tilde{\mathbf{b}}_1, \dots, \tilde{\mathbf{b}}_K\}$ from $\{\tilde{\mathbf{P}}_1, \dots, \tilde{\mathbf{P}}_K\}$ using (11). Set $\{\mathbf{b}_1^{(0)}, \dots, \mathbf{b}_K^{(0)}\} = \{\tilde{\mathbf{b}}_1, \dots, \tilde{\mathbf{b}}_K\}$.
- 3) Find $\{\mathbf{P}_1^{(0)}, \dots, \mathbf{P}_K^{(0)}\}$ from $\{\mathbf{b}_1^{(0)}, \dots, \mathbf{b}_K^{(0)}\}$ using (13).
- 4) Set $i = 0$.
- 5) Loop
 - a) Select the (k, n) pair that has the *smallest* gradient value in (18) and *decrement* $b_k^{n,(i)}$ to obtain $\{\mathbf{b}_1^{(i+1)}, \dots, \mathbf{b}_K^{(i+1)}\}$. Find $\{\mathbf{P}_1^{(i)}, \dots, \mathbf{P}_K^{(i)}\}$ from $\{\mathbf{b}_1^{(i)}, \dots, \mathbf{b}_K^{(i)}\}$ using (13). Compute the slack power \hat{P}_k .

- b) If $\hat{P}_k \leq 0$ for one of the users, consider that user out of the bit-subtracting process.
- c) Exit loop if $\hat{P}_k \leq 0$ for all users.
- d) Set $i = i + 1$.

End loop

C. Computational Complexity

The complexity of the bit-adding/subtracting algorithms is the product of the number of iteration steps and the complexity required to execute each iteration, plus the complexity of running ISB for initialization. The complexity of initial ISB is approximately linear in the number of users and frequency tones and is small compared to the bit-addition/subtraction process. Hence, the bit-addition/subtraction process is a major component of the time complexity of the algorithm.

The complexity within each iteration is dominated by the matrix inversion (13), which determines the required power spectra for a target bit allocation (i.e., the step in which $\{\mathbf{P}_1^{(i)}, \dots, \mathbf{P}_K^{(i)}\}$ is determined from $\{\mathbf{b}_1^{(i)}, \dots, \mathbf{b}_K^{(i)}\}$). This can be done by multiplication with Θ^{-1} , which has a cubic complexity in the number of variables, NK .

The number of required iteration steps depends on the strength of crosstalk in the system. This is because the number of iteration is a function of the distance between the iterative spectrum balancing solution, which does not consider ICI, and the optimal solution with ICI taken into account. This distance is determined by the amount of ICI, which is proportional to crosstalk power in the systems. Hence, the algorithm takes a longer time to run for systems with stronger crosstalk among the users. However, it is possible to reduce the simulation time by adding or subtracting multiple bits for each iterations. This is especially useful when NK is large.

D. Complexity Reduction via Power Series Approximation

As stated earlier, the complexity of the bit-adding/subtracting algorithm is dominated by the matrix inversion (13). A complexity reduction method which approximates the matrix inversion (13) via power series expansion is proposed in this section. A convergence analysis is also presented.

In (13), $[\mathbf{P}_1^{(i)} \dots \mathbf{P}_K^{(i)}]^T$ is determined from $[\mathbf{b}_1^{(i)} \dots \mathbf{b}_K^{(i)}]^T$ by

$$\begin{bmatrix} \mathbf{P}_1^{(i)} \\ \vdots \\ \mathbf{P}_K^{(i)} \end{bmatrix} = \Theta^{-1} \text{diag} \left(\begin{bmatrix} \sigma_1 \\ \vdots \\ \sigma_K \end{bmatrix} \right) \begin{bmatrix} 2^{\mathbf{b}_1^{(i)}} - 1 \\ \vdots \\ 2^{\mathbf{b}_K^{(i)}} - 1 \end{bmatrix} \Gamma. \quad (20)$$

Since Θ is an NK -by- NK matrix, the evaluation of Θ^{-1} has a complexity $O((NK)^3)$. However, observe that ICI falls off quickly with tone spacing. So, Θ can be rearranged into a form in which the diagonal block matrices are much larger than the off-diagonal blocks.

Let \mathbf{B} represent a matrix composed of $M \times M$ blocks such that

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \cdots & \mathbf{B}_{1M} \\ \vdots & \ddots & \vdots \\ \mathbf{B}_{M1} & \cdots & \mathbf{B}_{MM} \end{bmatrix} \quad (21)$$

where each submatrix \mathbf{B}_{ij} is an L -by- L matrix. The required form is one where $\|\mathbf{B}_{ii}\| \gg \|\mathbf{B}_{ij}\|$ and $\|\mathbf{B}_{ii}\| \gg \|\mathbf{B}_{ji}\|, \forall j \neq i$.

To transform Θ into such a form, Θ can be viewed as a matrix of K -by- K blocks, where each submatrix is denoted as $\Theta_{ij}, 0 \leq i, j \leq K$. Each block Θ_{ij} is an N -by- N matrix representing the crosstalk from user j to user i . The (p, q) element in Θ_{ij} is a function of the crosstalk from tone q to tone p . ICI diminishes significantly as the difference between p and q grows. Hence, each block Θ_{ij} has its diagonal terms much larger than off-diagonal terms. Then, Θ can be rearranged into a matrix $\tilde{\Theta}$ with N -by- N blocks, where each block is a K -by- K matrix. This means that the (p, q) element in Θ_{ij} becomes the (i, j) element in $\tilde{\Theta}_{pq}$. The resulting $\tilde{\Theta}$ will have its diagonal block matrices dominating off-diagonal blocks.

We propose the following method for approximating $\tilde{\Theta}^{-1}$. Let Δ represent a block diagonal matrix in which the diagonal blocks are taken from $\tilde{\Theta}$. Let $\Pi = \tilde{\Theta} - \Delta$. Then,

$$\begin{aligned} \tilde{\Theta}^{-1} &= (\Delta + \Pi)^{-1} \\ &= (\mathbf{I}_{NK} + \Delta^{-1}\Pi)^{-1} \Delta^{-1} \\ &= (\mathbf{I}_{NK} - \Delta^{-1}\Pi + (\Delta^{-1}\Pi)^2 - \dots) \Delta^{-1} \\ &\approx (\mathbf{I}_{NK} - \Delta^{-1}\Pi) \Delta^{-1}. \end{aligned} \quad (22)$$

where \mathbf{I}_{NK} is an NK -by- NK identity matrix.

When the power series approximation is used in the bit-adding/subtracting algorithm, it is possible for the actual required power $\{\mathbf{P}_1^{(i)}, \dots, \mathbf{P}_K^{(i)}\}$ to exceed the power constraint. This problem can be solved by performing bit-subtraction with exact power-spectrum calculation after running the reduced-complexity bit-adding/subtracting algorithm. When the power series expansion results in good approximation on $\{\mathbf{P}_1^{(i)}, \dots, \mathbf{P}_K^{(i)}\}$, only a few iterations of bit subtraction with exact power calculation is required at the end.

The above method works only when the power series expansion converges. A sufficient condition for convergence is that all eigenvalues of the matrix $\Delta^{-1}\Pi$ have magnitudes that are smaller than 1 [11]. Further, if all eigenvalues are very small (e.g., less than 0.1), fast convergence of the power series is observed. Simulation experiences show that the convergence condition is satisfied for all channel configurations when the bit-adding algorithm is performed in an ADSL two-user system. For bit-subtracting, the eigenvalues of $\Delta^{-1}\Pi$ are confined in the unit circle only in the scenarios illustrated in Fig. 5. The channel lengths in the graph refers to those in Fig. 6 such that the x -axis represents the equal direct channels l_1, l_2 . The y -axis represents the crosstalk channel c_1 in Fig. 6. As shown in the graph, the domain of valid channel configurations is divided into three regions: the region of divergence, the region of slow convergence, and the region of fast convergence. The region with the largest eigenvalue of $\Delta^{-1}\Pi$ greater than 1 exists only at $l_1 \leq 11$ kft. Hence, the convergence condition is always satisfied when $l_1 > 11$ kft and partially satisfied in other cases depending on the crosstalk channel length. In general, simulation shows that convergence is more likely achieved when l_1 is large and c_1 is small.

In the cases when all eigenvalues of $\Delta^{-1}\Pi$ have magnitudes less than 0.1, the power series in (22) converges very quickly. In

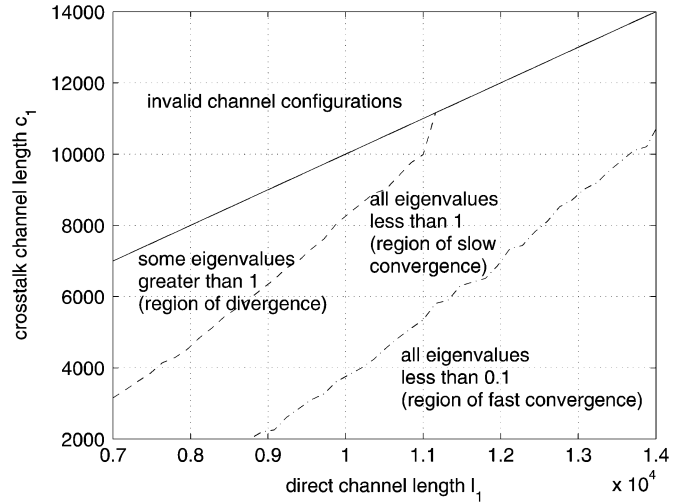


Fig. 5. Region of convergence for the reduced-complexity bit-subtracting algorithm.

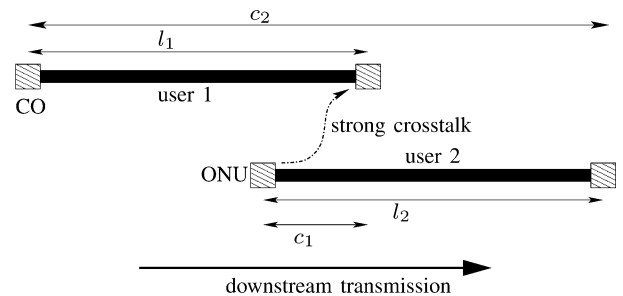


Fig. 6. Loop topology for two-user ADSL downstream, where the lines deployed from the optical network unit (ONU) produce strong crosstalk to the central office (CO) deployed lines.

this case, $\tilde{\Theta}^{-1}$ can be approximated very well by using a first-order expansion. This region of fast convergence is primarily found in cases where $l_1 \geq 9000$ and c_1 is small. Even when the first-order approximation is not entirely accurate, the power series method can still be used in practice. By adding a few iterations of bit-adding/subtracting with exact power spectra calculation at the end, the inaccuracy of the approximation can be easily corrected.

The main advantage of the above scheme is its lower complexity. This scheme replaces the inversion of a large matrix with addition and multiplication involving block diagonal matrices. The inversion of Δ and multiplication with Δ^{-1} both takes $O(N^2K^3)$ operations because Δ is block diagonal. Hence, the complexity is reduced from $O(N^3K^3)$ to $O(N^2K^3)$. Note that N is much larger than K in practice.

IV. SIMULATIONS

The performances of the greedy bit-adding and bit-subtracting algorithms are examined in this section. The proposed algorithms are compared with existing algorithms, including iterative water-filling [10] and the OSB algorithm [3], via extensive simulation.

A. Two-User Block-Asynchronous ADSL

The problem of spectrum optimization in the presence of ICI in an ADSL environment with $N = 256$ is used as the

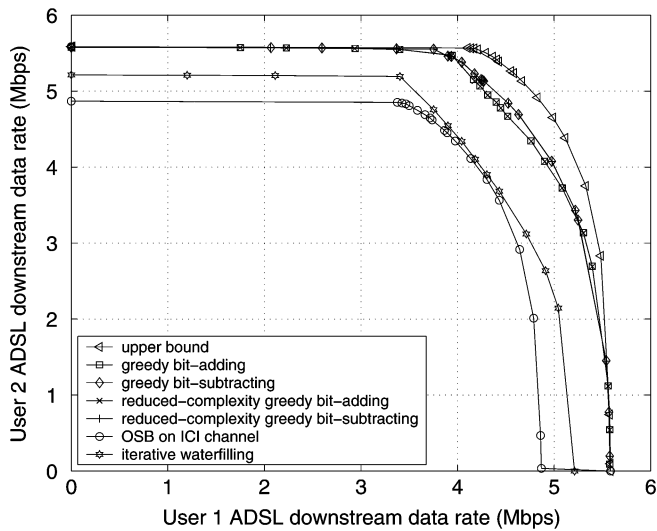


Fig. 7. Achievable rate region for two-user asynchronous ADSL downstream with 12 Kft direct channels and 9 kft crosstalk channel.

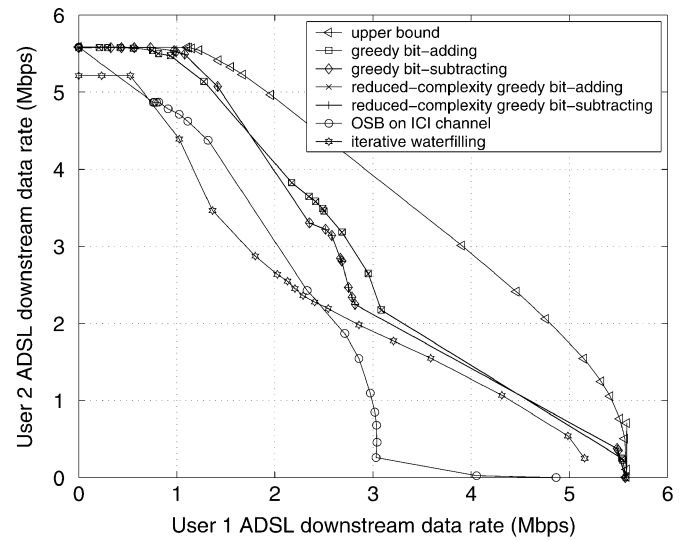


Fig. 9. Achievable rate region for two-user asynchronous ADSL downstream with 12 Kft direct channels and 3 kft crosstalk channel.

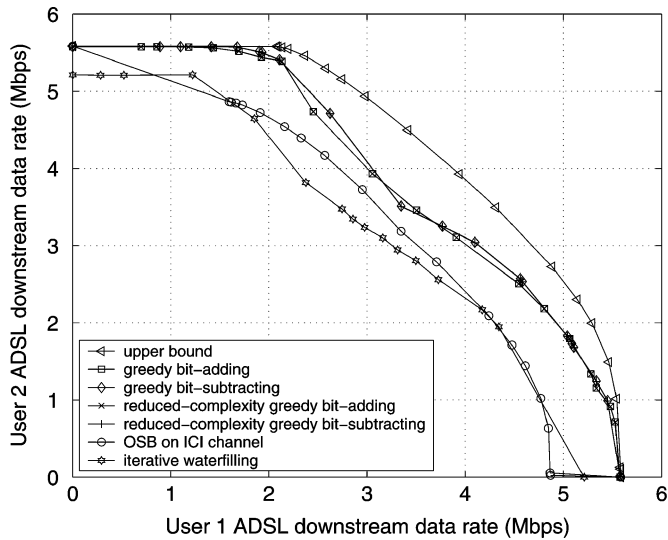


Fig. 8. Achievable rate region for two-user asynchronous ADSL downstream with 12 Kft direct channels and 6 kft crosstalk channel.

first example. Fig. 6 shows a two-user scenario where the direct channel lengths are denoted by l_1 , l_2 and the length of the crosstalk channel from user 2 to 1 is denoted by c_1 . Due to imperfect synchronization on the transmitted DMT blocks, the ICI model of Fig. 4 is used. For all simulations, a target error probability of 10^{-7} with a 3 dB coding gain and 6 dB noise margin are used. Discrete bit-loading is assumed for all cases. The channel models are generated using transmission line models for 26-gauge twisted pairs [12]. The crosstalk transfer function is obtained from empirical formula (used in standardization groups) in which the magnitude of crosstalk depends on f^2 (see, e.g., [13]).

Figs. 7–9 show a comparison of the rate regions achieved by the bit-adding and bit-subtracting algorithms with those of iterative water-filling and OSB. The rate regions are plotted by varying the weights w_k from $(w_1, w_2) = (0, 1)$ to $(w_1, w_2) = (1, 0)$ in the rate maximization problem. The direct

channel lengths are 12 kft for all three cases. The crosstalk channel lengths are 9, 6, and 3 kft for Figs. 7, 8, and 9, respectively. The three scenarios represent cases of weak, medium, and strong interference between users.

The upper bounds are obtained by applying the OSB algorithm in an ICI-free environment. The performance curves cannot exceed this upper bound because OSB is optimal in an ICI-free environment and the rate region with ICI must be smaller than the rate region without ICI. The iterative water-filling and OSB algorithms are designed for systems without ICI. Both schemes exhibit poor performance when ICI is present. With ICI, OSB is no longer optimal and it can even be inferior to iterative water-filling in some cases. The performance degradation of OSB is represented by the gap between the upper bound and the OSB rate region. It can be observed that the proposed bit-adding/subtracting algorithms significantly narrow down this gap, especially when the crosstalk signal is relatively weak.

When the crosstalk channel is strong, a sizable gap still exists between the best achievable rate region and the ICI-free upper bound. In general, the ICI-free upper bound is not expected to be achievable. However, the extent to which this gap can be partially closed is not yet known. Note that since ISB is used as a starting point for greedy search, both bit-adding and bit-subtracting algorithms always achieve higher rates than ISB.

The rate improvement offered by the proposed algorithms as compared to the OSB solution can be explained by examining the resulting power spectra. The power spectra determined by various algorithms for the case of 6 kft crosstalk channel are shown in Fig. 10. It can be observed that the optimal spectrum for an ICI-free environment contains a frequency-division multiplex (FDM) region [from 580 to 980 kHz in Fig. 10(a)] where the frequency usage alternates between the users. Alternating power allocation between users in an environment with ICI can be costly. Hence, the bit-adding and bit-subtracting algorithms attempt to avoid switching power between users as much as possible, which results in a better overall performance.

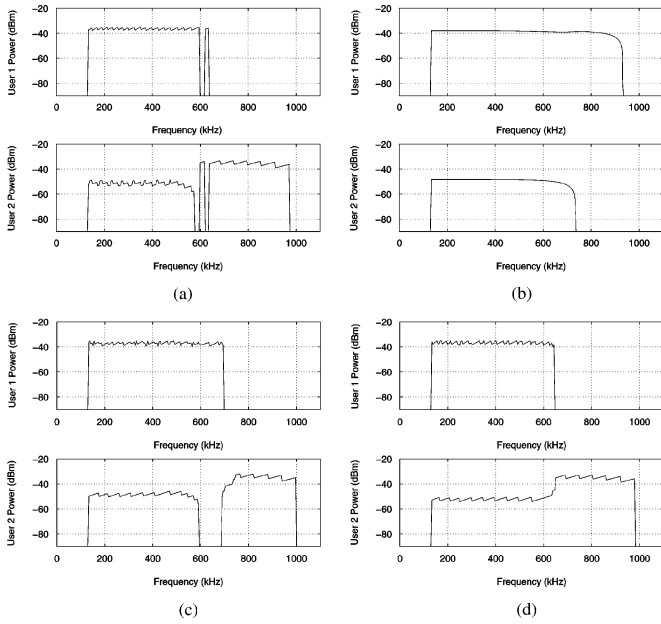


Fig. 10. Power allocation derived from various spectrum management algorithms. (a) Optimal spectrum balancing. (b) Iterative water-filling. (c) Greedy bit-adding. (d) Greedy bit-subtracting.

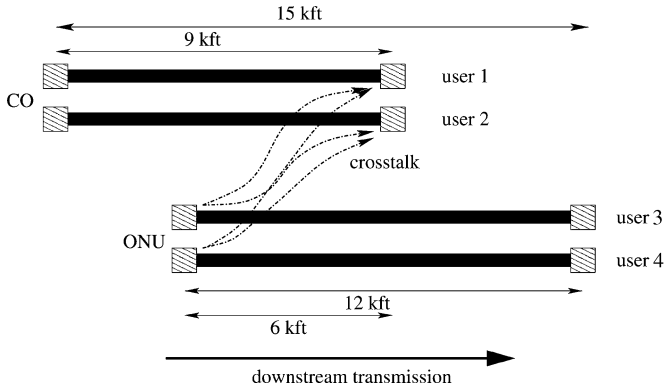


Fig. 11. Loop topology for a four-user ADSL downstream configuration.

Comparing the two proposed algorithms, the differences between the bit-adding and bit-subtracting algorithms are small in all three scenarios. Also, the reduced complexity method using power series approximation is shown to achieve performance almost identical to the original methods. Only first-order expansion, as described in previous section, is used in all cases.

B. Four-User Block-Asynchronous ADSL

The bit-adding/subtracting algorithms can be readily extended to systems with more than two users. A four-user ADSL downstream scenario, as shown in Fig. 11, is examined in this section. In this case, data transmissions by users 3 and 4 induce crosstalk interference to users 1 and 2. The greedy bit-adding and bit-subtracting algorithms are compared with the iterative water-filling and iterative spectrum balancing algorithms.

In the four-user case, it is difficult to represent the rate tuple (R_1, R_2, R_3, R_4) in a single graph because the rate region is 4-D. In this simulation, the maximum common rate such that $R_1 = R_2 = R_3 = R_4$ is found and used as a comparison measure. Table I shows the rates achievable for all users when the

TABLE I
PERFORMANCE COMPARISON FOR THE FOUR-USER CHANNEL

common rate achievable for all users (Mbps)	
iterative water-filling	5.1535
iterative spectrum balancing	5.0343
greedy bit-adding	5.3652
greedy bit-subtract	5.2878
reduced-complexity bit-adding	5.3628
reduced-complexity bit-subtracting	5.2611
upper bound	5.8336

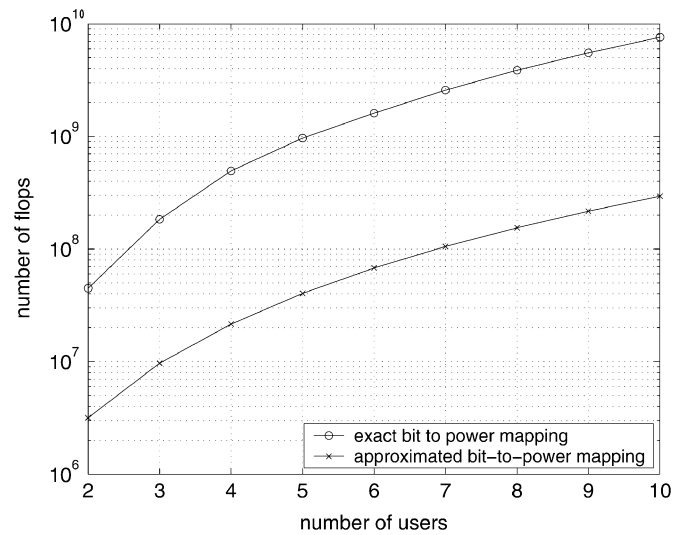


Fig. 12. Algorithmic complexity reduction of approximated bit-power mapping using first-order power series approximation for ADSL.

various algorithms are applied. As shown in the table, the bit-adding algorithm increases the rate of each user by 6.6% from OSB and 4.1% from iterative water-filling. The reduced-complexity methods using power series approximation again provide a performance close to the original methods.

C. Complexity Reduction Using Power Series Approximation

Fig. 12 illustrates the amount of complexity reduction achieved by the approximated algorithm using first-order power series expansion for an ADSL system with 256 tones. The number of MATLAB flops is used as the measure of complexity. Complexity of the bit-adding/subtracting algorithms is dominated by the computation of $\{\mathbf{P}_1^{(i)}, \mathbf{P}_2^{(i)}\}$ from $\{\mathbf{b}_1^{(i)}, \mathbf{b}_2^{(i)}\}$ using (13). The number of flops required to execute a single bit-to-power mapping function is shown in Fig. 12. Using the power series approximation, the complexity is reduced by an order of magnitude. For this reason, power series approximation is recommended whenever it is applicable. Note that the proposed algorithms are always more complex than ISB (since ISB is used as a first step), but are less complex than OSB (since the latter has an exponential complexity with respect to the number of users.)

V. CONCLUSION

Intercarrier interference is a performance-limiting factor to existing DSL systems where the DMT frames are not perfectly aligned. This paper characterizes the amount of intercarrier interference in an asynchronous DMT system and proposes two heuristics algorithms, the greedy bit-adding and bit-subtracting algorithms, for determining a set of transmit power spectra that reduces the degradation caused by intercarrier interference. Simulation results show that the proposed algorithms offer better performance than existing schemes. To enhance the potential of implementing the greedy bit-adding/subtracting algorithms, a method for reducing the complexity using power series approximation is further proposed in this paper. Simulation shows that this approximation method gives a performance which is very close to that of the full-blown algorithms.

APPENDIX

A. Derivation of ICI Coefficients

The derivation of the expressions for ICI coefficients in (7) and (8) are presented here. The ICI coefficients represent ICI generated by a crosstalk source with a single tone $\delta[n]$ in the frequency domain. Referring to Fig. 3, the received crosstalk is

$$c[t] = \tilde{c}_1[t] + \tilde{c}_2[t] \quad (23)$$

where $\tilde{c}_1[t]$ and $\tilde{c}_2[t]$ represent crosstalk symbols $c_1[t]$ and $c_2[t]$ which have an offset ν with respect to the received symbol. Assume that the power spectra of $c_1[t]$ and $c_2[t]$ are both given by $\delta[n]$, i.e.

$$|C_1[n]|^2 = |C_2[n]|^2 = \delta[n], \quad (24)$$

where n is the frequency tone index. The goal of this section is to find the power spectrum density of $c[t]$.

The crosstalk symbols $\tilde{c}_1[t]$ and $\tilde{c}_2[t]$ can be represented as truncated versions of $c_1[t]$ and $c_2[t]$. The truncation is equivalent to a multiplication with a rectangular window in the time domain. Hence

$$\tilde{c}_1[t] = c_1[t]r_{N-\nu}[N-t] \quad (25)$$

$$\tilde{c}_2[t] = c_2[t]r_\nu[t] \quad (26)$$

where $r_\nu[t]$ is a rectangular window of width ν . Now suppose that $R_\alpha[n]$, $R_\beta[n]$, and $\tilde{C}_i[t]$ are the N -point discrete Fourier transforms of $r_\nu[t]$, $r_{N-\nu}[N-t]$, and $\tilde{c}_i[t]$, respectively. Then,

$$\tilde{C}_1[n] = \frac{1}{N} \sum_{l=0}^{N-1} C_1[n-l]R_\beta[l] \quad (27)$$

$$\tilde{C}_2[n] = \frac{1}{N} \sum_{l=0}^{N-1} C_2[n-l]R_\alpha[l] \quad (28)$$

where, if $n \neq 0$,

$$\begin{aligned} R_\alpha[n] &= \sum_{t=0}^{N-1} r_\nu[t]e^{-j(2\pi/N)nt} \\ &= \sum_{t=0}^{\nu-1} e^{-j(2\pi/N)nt} \\ &= \frac{1 - e^{-j(2\pi/N)n\nu}}{1 - e^{-j(2\pi/N)n}} \\ &= \frac{\sin((\pi/N)n\nu)}{e^{-j(\pi/N)n(\nu-1)} \sin((\pi/N)n)} \end{aligned} \quad (29)$$

and

$$\begin{aligned} R_\beta[n] &= \sum_{t=0}^{N-1} r_{N-\nu}[N-t]e^{-j(2\pi/N)nt} \\ &= \sum_{t=\nu}^{N-1} e^{-j(2\pi/N)nt} \\ &= e^{-j(2\pi/N)n\nu} \frac{1 - e^{-j(2\pi/N)n(N-\nu)}}{1 - e^{-j(2\pi/N)n}} \\ &= e^{-j(\pi/N)n(\nu+1)} \frac{\sin((\pi/N)n)(N-\nu)}{\sin((\pi/N)n)}. \end{aligned} \quad (30)$$

And, $R_\alpha[0] = \nu$, $R_\beta[0] = N - \nu$.

Let $P_{\tilde{c}_1}[n]$ and $P_{\tilde{c}_2}[n]$ denote the power spectrum density of $\tilde{c}_1[t]$ and $\tilde{c}_2[t]$, respectively. Then, if $n \neq 0$

$$\begin{aligned} P_{\tilde{c}_1}[n] &= \left| \frac{1}{N} \sum_{l=0}^{N-1} C_1[n-l]R_\beta[l] \right|^2 \\ &= \left| \frac{1}{N} \sum_{l=0}^{N-1} \delta[n-l]R_\beta[l] \right|^2 \\ &= \frac{1}{N^2} |R_\beta[n]|^2 \\ &= \frac{1}{N^2} \frac{\sin^2((\pi/N)n(N-\nu))}{\sin^2((\pi/N)n)} \end{aligned} \quad (31)$$

$$\begin{aligned} P_{\tilde{c}_2}[n] &= \left| \frac{1}{N} \sum_{l=0}^{N-1} C_2[n-l]R_\alpha[l] \right|^2 \\ &= \left| \frac{1}{N} \sum_{l=0}^{N-1} \delta[n-l]R_\alpha[l] \right|^2 \\ &= \frac{1}{N^2} |R_\alpha[n]|^2 \\ &= \frac{1}{N^2} \frac{\sin^2((\pi/N)n\nu)}{\sin^2((\pi/N)n)} \end{aligned} \quad (32)$$

and $P_{\tilde{c}_1}[0] = \nu^2/N^2$, $P_{\tilde{c}_2}[0] = (N-\nu)^2/N^2$.

Now, since the data transmitted in blocks \mathbf{c}_1 and \mathbf{c}_2 are independent, their powers add together to form the crosstalk interference. Therefore, $\rho_\nu[n]$, which is the power spectrum densities of $\tilde{c}[t]$, can be expressed as

$$\begin{aligned} \rho_\nu[n] &= P_{\tilde{c}_1}[n] + P_{\tilde{c}_2}[n] \\ &= \frac{\sin^2((\pi/N)n(N-\nu)) + \sin^2((\pi/N)n\nu)}{N^2 \sin^2((\pi/N)n)} \\ &= \frac{2 \sin^2((\pi/N)n\nu)}{N^2 \sin^2((\pi/N)n)}, \quad \text{if } n \neq 0. \end{aligned} \quad (33)$$

When $n = 0$

$$\rho_\nu[0] = P_{z_1}[0] + P_{z_2}[0] = \frac{\nu^2 + (N - \nu)^2}{N^2} \quad (34)$$

This coincides with (7). The last equality in (33) follows because:

$$\sin\left(\pi n - \frac{\pi n \nu}{N}\right) = \begin{cases} \sin\left(\frac{\pi n \nu}{N}\right) & \text{if } n \text{ is odd} \\ -\sin\left(\frac{\pi n \nu}{N}\right) & \text{if } n \text{ is even.} \end{cases} \quad (35)$$

Thus, $\sin^2((\pi/N)n(N - \nu)) = \sin^2((\pi/N)n\nu)$ when n is an integer.

Since the offset ν is not known in advance, the definition of ICI coefficients used in this paper accounts for the worst possible offset ν . Following (33), the expression for ICI coefficients $\gamma[n]$ can be derived as follows: when $n = 0$, the worst ν is $\nu = 0$, in which case

$$\gamma[0] = 1 \quad (36)$$

when $n \neq 0$, the worst ν depends on n . It can be shown that

$$\begin{aligned} \gamma[n] &= \max_{\nu} \frac{2 \sin^2((\pi/N)n\nu)}{N^2 \sin^2((\pi/N)n)} \\ &= \frac{2}{N^2 \sin^2\left(\frac{\pi}{N}n\right)}. \end{aligned} \quad (37)$$

For each n , which represent the frequency spacing between the crosstalk tone and the subject tone, the maximum of $\sin^2((\pi/N)n\nu)$ is attained when $\nu = N/2n$. The approximation $\max_{\nu} \sin^2((\pi/N)n\nu) = 1$ for integer ν is accurate when N is large, as is the case for most DSL systems.

B. Monotonicity of the Bit-Power Relation

Lemma 1: Let $\mathbf{P} = \{P_k^n\}$ and $\mathbf{b} = \{b_k^n\}, k = 1, \dots, K, n = 1, \dots, N$ be related by

$$b_k^n = \log_2 \left(1 + \frac{1}{\Gamma} \cdot \frac{P_k^n |H_{kk}|^2}{\sigma_k^n + P_{k,ICI}^n} \right) \quad (38)$$

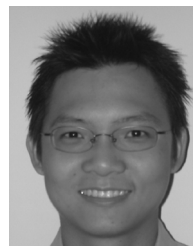
where $P_{k,ICI}^n$ is as defined in (10). Let \mathbf{P}' and \mathbf{b}' be related similarly. Suppose that $\mathbf{b}' \leq \mathbf{b}$ component-wise. Then, $\mathbf{P}' \leq \mathbf{P}$ component-wise.

Proof: Fix \mathbf{b} and \mathbf{P} as related in (38). Let $\mathbf{b}' \leq \mathbf{b}$. The idea is to use the following algorithm to find \mathbf{P}' starting from \mathbf{P} , so that \mathbf{P}' and \mathbf{b}' are also related as in (38).

Start by setting $\mathbf{P}^{(0)} = \mathbf{P}$. Now, fix the interference terms in the denominator of (38) as in $\mathbf{P}^{(0)}$ and reduce the P_k^n terms in the numerators so that (38) is satisfied with equality for \mathbf{b}' . Call the newly reduced power $\mathbf{P}^{(1)}$. Clearly, $\mathbf{P}^{(1)} \leq \mathbf{P}^{(0)}$ component-wise, since $\mathbf{b}' \leq \mathbf{b}$ component-wise. Next, replace the interference terms in (38) by that of $\mathbf{P}^{(1)}$ and resolve for the numerators of (38) with equality for \mathbf{b}' . Since the interference is reduced, the power needed, which we call $\mathbf{P}^{(2)}$, must further decrease component-wise. This process can go on and must eventually converge since the iteration is monotonic component-wise for $\mathbf{P}^{(k)}$. Let $\mathbf{P}' = \lim_{k \rightarrow \infty} \mathbf{P}^{(k)}$. At the point of convergence, we must have $\mathbf{P}' \leq \mathbf{P}$ and have \mathbf{b}' and \mathbf{P}' related exactly by (38). ■

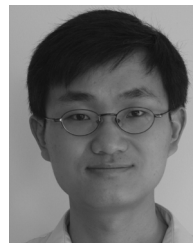
REFERENCES

- [1] F. Sjöberg, M. Isaksson, R. Nilsson, P. Ödling, S. K. Wilson, and P. O. Börjesson, "Zipper: A duplex method for VDSL based on DMT," *IEEE Trans. Commun.*, vol. 47, no. 8, pp. 1245–52, Aug. 1999.
- [2] F. Sjöberg, R. Nilsson, M. Isaksson, P. Ödling, and P. O. Börjesson, "Asynchronous zipper," in *IEEE Int. Conf. Commun. (ICC)*, Jun. 1999, vol. 1, pp. 231–235.
- [3] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen, "Optimal multiuser spectrum management for digital subscriber lines," *IEEE Trans. Commun.*, vol. 54, no. 5, pp. 922–933, May 2007.
- [4] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Commun.*, vol. 54, no. 6, pp. 1310–1322, Jun. 2007.
- [5] R. Cendrillon and M. Moonen, "Iterative spectrum balancing for digital subscriber lines," in *IEEE Int. Conf. Commun. (ICC)*, Seoul, Korea, May 2005, vol. 3, pp. 1937–1941.
- [6] T. Pollet and M. Peeters, "Performance degradation of multi-carrier systems caused by an insufficient guard interval duration," in *Proc. Int. Work. Copper Wire Access Syst. (CWA)*, Budapest, Hungary, Oct. 1997, pp. 125–127.
- [7] W. Henkel, G. Tauböck, P. Ödling, P. O. Börjesson, and N. Petersson, "The cyclic prefix of OFDM/DMT—An analysis," *Int. Zurich Seminar Broadband Commun.: Access, Transmission, Netw.*, pp. 22.1–22.3, Feb. 2002.
- [8] J. Lee, R. V. Sonalkar, and J. M. Cioffi, "A multi-user power control algorithm for digital subscriber lines," *IEEE Commun. Lett.*, vol. 9, pp. 193–195, Mar. 2005.
- [9] A. Leshem and L. Youming, "A low complexity coordinated fext cancellation for vdsl," in *Proc. 11th IEEE Int. Conf. Elect. Circuits Syst. (ICECS)*, Dec. 2004, pp. 338–341.
- [10] W. Yu, G. Ginis, and J. M. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 5, pp. 1105–1115, Jun. 2002.
- [11] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: The Johns Hopkins Univ. Press, 1996.
- [12] D. G. Messerschmitt, LINEMOD—A Transmission Line Modeling Program [Online]. Available: <http://www.stanford.edu/group/cioffi/linemod/linemod.html>
- [13] G. Ginis and J. M. Cioffi, "Vectored transmission for digital subscriber line systems," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 5, pp. 1085–1104, Jun. 2002.



Vincent M. K. Chan received the B.A.Sc. degree in computer engineering from the University of Waterloo, Waterloo, Ontario, Canada, in 2003, and the M.A.Sc degree from University of Toronto, Ontario, in 2005.

Through the co-op program of the University of Waterloo, he worked at Genesis Microchip Inc., Broadcom Corp., and International Business Machines Corporation (IBM). He is currently with ViXS Systems Inc., Toronto, as an embedded system engineer working on video processing products.



Wei Yu (S'97–M'02) received the B.A.Sc. degree in computer engineering and mathematics from the University of Waterloo, Waterloo, Ontario, Canada, in 1997 and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1998 and 2002, respectively.

Since 2002, he has been an Assistant Professor with the Electrical and Computer Engineering Department, University of Toronto, Toronto, Ontario, where he currently holds a Tier II Canada Research Chair in Digital Communications. His main research

interests include multiuser information theory, optimization, wireless communications, and broadband access networks.

Prof. Wei Yu received an Early Researcher Award from the Province of Ontario in 2006. He was an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2004 to 2007. He was a Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS Special Issue on "Non-linear Optimization of Communications Systems," and a Guest Editor of the EURASIP Journal on Applied Signal Processing Special Issue on "Advanced Signal Processing for Digital Subscriber Lines." He is a registered Professional Engineer (P.Eng.) in the Province of Ontario.