Optimized Transmission for Upstream Vectored DSL Systems Using Zero-Forcing Generalized Decision-Feedback Equalizers

Chiang-yu Chen, Kibeom Seong, Rui Zhang, and John M. Cioffi
Department of Electrical Engineering
Stanford University
Stanford, CA 94305 USA
Email: {chiangyu.chen, kseong, ee.ruizhang, cioffi}@stanford.edu

***Abstract*— In upstream vectored DSL transmission, the far-end crosstalk (FEXT) can be completely cancelled by using zero-forcing generalized decision-feedback equalizers (ZF-GDFE). When the spatially correlated alien crosstalk is present, the achievable data rates of DSL lines with ZF-GDFE depend on their decoding orders at each DMT tone. Given a weighted sum-rate maximization problem, the optimal orderings for all DMT tones can be found by the Lagrange dual decomposition method. However, the computational complexity of such approach grows with the factorial of the number of users, which makes the optimal search infeasible with a large number of vectored lines.**

This paper presents a modified greedy algorithm (MGA) that performs close to the optimal search of decoding orders. The complexity of MGA is only proportional to the cube of the number of users, which is the same as it of QR decomposition. With a significant reduction of complexity, MGA is a promising technique for practical DSL systems.

I. Introduction

It has been discovered that vectored transmission in Digital Subscriber Line (DSL) systems can increase the achievable data rates dramatically [1], [2]. In upstream vectored transmission, the transmitters are distributed in different customer premises equipments (CPE) while the receivers are collocated at a central office (CO) or an optical network unit (ONU). The channel can be modeled by a multiple-access channel (MAC) and joint signal processing is allowed at the receiver. On the other hand, in downstream vectored transmission, the transmitters are collocated at a CO/ONU while the receivers are scattered. Thus, a broadcast channel (BC) is the best description of this situation.

By synchronizing the discrete multi-tone (DMT) symbols, using long enough cyclic extension and adopting frequency division duplex (FDD), near-end crosstalk (NEXT) can be completely eliminated [1]. Thus, the dominant impairment in DSL systems is FEXT that comes from the opposite end of the affected receiver. In [1], FEXT cancellation techniques utilizing QR decomposition are proposed for both upstream and downstream. The CO/ONU successively decodes the received upstream signals based on the QR decomposed channel and previous decisions. This technique is a special case of zero-forcing generalized decision-feedback equalizer (ZF-GDFE). In this case, DSL upstream channels are column-wise diagonally dominant, which makes subchannel gains independent of the orderings of the QR decompositions. Therefore, the achievable rate tuple remains constant for any orderings of QR decompositions, and the achievable rate region is a hyper-rectangle. However, in practice, alien disturbers out of the vectoring domain inject highly correlated noise among the vectored lines [3]. Since the coordination at the receiver is possible in the upstream case, data rates can be boosted further by utilizing alien noise cancellation techniques [3]. Combining with the noise whitening filter, the equivalent channel is no longer column-wise diagonally dominant. As a result, different QR orderings change subchannel gains, which further lead to various data rate tuples. For a vectored DSL system with K users and N tones, there are at most $(K!)^N$ achievable rate tuples corresponding to different sets of decoding orders. Therefore, to find the achievable rate region is a non-trivial task.

Similarly, in the downstream case, a QR precoder is used at the CO/ONU to cancel the FEXT, based on the decomposed channel and precoded downstream symbols. Unfortunately, alien noise cancellation is not possible at downstream receivers since no coordination at the receiver is possible. This observation is consistent with the typical information theoretical result of BC, where noise correlation does not change the capacity region at all [4]. Therefore, the channel matrix is row-wise diagonally dominant and the achievable rate region hardly depends on QR orderings.

Although MMSE-GDFE is the canonical receiver structure [5] and the achievable rate region can be found by solving optimization routines in [6], it requires tremendous computational complexity to obtain the energy distributions and the corresponding equalizer coefficients. Moreover, when the SNR gap is non-zero in systems such as DSL, the optimal resource allocation is still an open problem. Since the SNR is usually high for DSL channels, ZF-GDFE causes very little noise enhancement. In this paper, the achievable rate region of vectored upstream DSL using ZF-GDFE is investigated. Due to the convexity of achievable rate regions, the boundary points can be characterized by solving weighted sum-rate maximization (WSRmax) problems for all possible weight vectors. The optimal orderings can be determined in the dual domain by using Lagrange dual decomposition, which is a direct extension of [7]. Though this dual approach provides linear complexity in the number of tones, the complexity still grows with $K!$, which makes it computationally infeasible for systems with a large number of users. Thus, a modified greedy algorithm (MGA) is proposed, which provides good suboptimal solutions with substantially lower computational complexity. Simulation results show that the rate region obtained by MGA is quite close to that with the optimal exhaustive search.

II. System Model and Problem Formulation

In upstream vectored transmission, each of the K loops has one transmit spatial dimension. The transmitted signal can be observed at all K receivers, either by direct channel or crosstalk. Synchronized DMT symbols divide the whole bandwidth into N parallel subchannels. The subchannel on tone n can be represented by a K-by-K real matrix H_n, where $n \in \{1, 2, \cdots, N\}$. The relationship between input and output signal vectors can be written as:

$$\mathbf{y}_n = H_n \mathbf{x}_n + \mathbf{z}_n, \tag{1}$$

where $\mathbf{x}_n$ and $\mathbf{y}_n$ are K-dimensional vectors denoting transmitted and received signals on tone n. $\mathbf{z}_n$ is additive Gaussian alien noise PSD with a non-singular covariance matrix R_n. This alien noise can be cancelled by a noise whitening filter at the receiver, which is information theoretically lossless. After noise whitening, the equivalent channel becomes $\tilde{H}_n = R_n^{-1/2} H_n$ and the channel model can be expressed as:

$$\tilde{\mathbf{y}}_n = R_n^{-1/2}\mathbf{y}_n = \tilde{H}_n \mathbf{x}_n + \tilde{\mathbf{z}}_n, \tag{2}$$

where $\tilde{\mathbf{z}}_n = R_n^{-1/2}\mathbf{z}_n \sim \mathcal{N}(0, I)$.

Since the equivalent channel matrix is square, the ZF-GDFE can be obtained by applying a QR decomposition to $\tilde{H}_n$. Define $\boldsymbol{\pi}_n = [\pi_{n1}\ \pi_{n2}\ \cdots\ \pi_{nK}]$ as a K-dimensional vector denoting the permutation on $\{1, 2, \cdots, K\}$. To be more specific, π_{nk} indicates the input dimension to be decoded $(K+1-k)$-th in the order on tone n. Clearly, $\pi_{nk} \in \{1, \cdots, K\}$, and $\pi_{ni} \neq \pi_{nj}$, for all $i \neq j$. Also, each $\boldsymbol{\pi}_n$ defines a permutation matrix $M_{\boldsymbol{\pi}_n}$, of which (i,j)-th element $M_{\boldsymbol{\pi}_n}^{(i,j)}$ is equal to 1 if $\pi_{nj} = i$. Otherwise $M_{\boldsymbol{\pi}_n}^{(i,j)} = 0$. For example, $M_{\boldsymbol{\pi}_n}$ is an identity matrix if $\boldsymbol{\pi}_n = [1\ 2\ \cdots\ K]$.

The QR decomposition given an order then becomes $\tilde{H}_n M_{\boldsymbol{\pi}_n} = Q_{n,\boldsymbol{\pi}_n} D_{n,\boldsymbol{\pi}_n} G_{n,\boldsymbol{\pi}_n}$, where $Q_{n,\boldsymbol{\pi}_n}$ is an unitary matrix, $D_{n,\boldsymbol{\pi}_n}$ is a diagonal matrix, and $G_{n,\boldsymbol{\pi}_n}$ is a monic upper triangular matrix. The QR decomposition can be computed by a sequence of $K-1$ Householder transforms ([8]). In each Householder transform, a symmetric and unitary Householder matrix is found. To make the notations simple, we remove the tone index and ordering from Householder matrices and define them as $Q^1, \cdots, Q^{K-1}$. Each Q^i is used to reflect or to rotate the equivalent channel so that zero elements are produced in the lower triangle. Therefore, the equivalent channel matrix is iteratively transformed to an upper triangular one. Consequently, the QR decomposition can be written as:

$$\begin{aligned} Q_{n,\boldsymbol{\pi}_n} &= Q^1 Q^2 \cdots Q^{K-1}, \\ D_{n,\boldsymbol{\pi}_n} G_{n,\boldsymbol{\pi}_n} &= Q^{K-1} Q^{K-2} \cdots Q^1 \tilde{H}_n M_{\boldsymbol{\pi}_n}. \end{aligned} \tag{3}$$

It also should be noted that the complexity of computing QR decomposition is $\mathcal{O}(K^3)$.

With these representations, the overall feedforward filter becomes $D_{n,\boldsymbol{\pi}_n}^{-1} Q_{n,\boldsymbol{\pi}_n}^* R_n^{-1/2}$, and the feedback filter is $G_{n,\boldsymbol{\pi}_n} M_{\boldsymbol{\pi}_n}^T$. $(.)^*$ and $(.)^T$ are notations for Hermitian and transpose operations. Let $g_{nk}^{(\boldsymbol{\pi}_n)}$ and p_{nk} be the subchannel gain and power spectral density (PSD) of user k on tone n respectively, it can be shown that $[g_{n1}^{(\boldsymbol{\pi}_n)}\ g_{n2}^{(\boldsymbol{\pi}_n)}\ \cdots\ g_{nK}^{(\boldsymbol{\pi}_n)}] = \text{diag}(D_{n,\boldsymbol{\pi}_n}^2) M_{\boldsymbol{\pi}_n}^T$, where diag$(.)$ represents a row vector that extracts the diagonal elements from the input matrix.

Since the user decoded later has chances to see a cleaner channel, the user index that appears former in $\boldsymbol{\pi}_n$ indicates the corresponding user has higher priority in decoding. Therefore, the π_{n1}-th user is given the highest priority on tone n, because its signal is decoded after all FEXT is cancelled. Moreover, its subchannel gain achieves its maximum possible value, which is the same as the square of norm of the π_{n1}-th column of $\tilde{H}_n$. If an SNR gap Γ is considered, the data rate of user k on tone n is determined by $\log_2(1 + p_{nk} g_{nk}^{(\boldsymbol{\pi}_n)}/\Gamma)$ per complex dimension. Note that for each user there is an individual power constraint, i.e. $\sum_{n=1}^N p_{nk} \leq P_k$ for all k.

The number of possible decoding orders for each tone is $K!$, resulting in $K!$ different receiver realizations. After QR decomposition and successive decoding, each user's channel is completely decoupled. Thus, once the orderings $\boldsymbol{\pi_n}$ are fixed, the weighted sum rate is maximized by simply doing single-user rate-adaptive waterfilling for each user. The achievable rate region of ZF-GDFE can be characterized by all the possible $(K!)^N$ rate tuples together with the time-sharing concept. Since the rate region is convex, it is sufficient to solve the maximum weighted sum-rate for all possible weight $\boldsymbol{\mu} = [\mu_1,\ \mu_2,\ \cdots,\ \mu_K]$. This problem can be formulated as:

$$\begin{aligned} \text{maximize} \quad & \max_{\boldsymbol{\pi}_1, \boldsymbol{\pi}_2, \cdots, \boldsymbol{\pi}_n} \sum_{k=1}^K \sum_{n=1}^N \mu_k \log_2 \left(1 + \frac{p_{nk} g_{nk}^{(\boldsymbol{\pi}_n)}}{\Gamma}\right) \\ \text{subject to} \quad & \sum_{n=1}^N p_{nk} \leq P_k, \quad \forall k \\ & p_{nk} \geq 0 \quad \forall n, k. \end{aligned} \tag{4}$$

(4) is not a convex optimization problem because taking the maximum of concave functions is not necessarily concave. The complexity of an exhaustive search of the optimal orderings is $\mathcal{O}((K!)^N)$. A more efficient search using Lagrange dual decomposition is shown in the next section.

III. Optimal and Modified Greedy Algorithms via Lagrange Dual Decomposition

A. Optimal Ordering

In [7], an efficient algorithm applying Lagrange dual decomposition is developed to maximize the sum rate of the downstream vectored DSL. The WSRmax problem for MAC can also be solved by extending this approach. First, the Lagrangian associated with the optimization problem (4) is defined over the domain $\mathcal{D} = \{p_{nk} : p_{nk} \geq 0, \forall n, k\}$ as

$$\begin{aligned} \mathcal{L}(\{p_{nk}\}, \boldsymbol{\lambda}, \{\boldsymbol{\pi}_n\}) \quad = \quad & \sum_{k=1}^K \sum_{n=1}^N \mu_k \log_2 \left(1 + \frac{p_{nk} g_{nk}^{(\boldsymbol{\pi}_n)}}{\Gamma}\right) \\ & - \sum_{k=1}^K \lambda_k \left(\sum_{n=1}^N p_{nk} - P_k\right), \end{aligned} \tag{5}$$

where $\boldsymbol{\lambda} = [\lambda_1, \ \cdots, \ \lambda_K]$ denotes the Lagrange multipliers with non-negative values. The Lagrange dual function is:

$$f(\boldsymbol{\lambda}) = \max_{\{p_{nk}\}, \boldsymbol{\pi}_n} \mathcal{L}(\{p_{nk}\}, \boldsymbol{\lambda}, \{\boldsymbol{\pi}_n\}) = \sum_{n=1}^{N} f_n(\boldsymbol{\lambda}) + \sum_{k=1}^{K} \lambda_k P_k, \tag{6}$$

where

$$f_n(\boldsymbol{\lambda}) = \max_{\{p_{nk}\}, \boldsymbol{\pi}_n} \sum_{k=1}^{K} \left(\mu_k \log_2 \left(1 + \frac{p_{nk} g_{nk}^{(\boldsymbol{\pi}_n)}}{\Gamma} \right) - \lambda_k p_{nk} \right). \tag{7}$$

(6) provides upper bounds on the weighted sum-rate for any $\lambda_k \geq 0$. For a fixed $\boldsymbol{\lambda}$, the maximization of (6) is equivalent to optimizing N subproblems in (7) separately. Given an ordering $\boldsymbol{\pi}_n$, the objective in (7) is concave in $\{p_{nk}\}$. Thus, the PSD allocation on user k's tone n can be represented as the following waterfilling equation.

$$p_{nk} = \left(\gamma_k - \frac{\Gamma}{g_{nk}^{(\boldsymbol{\pi}_n)}} \right)^+, \tag{8}$$

where $\gamma_k = \mu_k / (\log 2 \cdot \lambda_k)$ denotes the water-level for the k-th user and $(x)^+ = \max(x, 0)$. By assigning this power distribution, the solution to (7) can be obtained by searching all $K!$ orderings. That is,

$$f_n(\boldsymbol{\lambda}) = \max_{\boldsymbol{\pi}_n} \left\{ \sum_{k=1}^{K} \left(\mu_k \left(\log_2 \left(\frac{\gamma_k g_{nk}^{(\boldsymbol{\pi}_n)}}{\Gamma} \right) \right)^+ - \lambda_k \left(\gamma_k - \frac{\Gamma}{g_{nk}^{(\boldsymbol{\pi}_n)}} \right)^+ \right) \right\}. \tag{9}$$

Therefore, the optimal ordering on tone n for a fixed $\boldsymbol{\lambda}$ is

$$\boldsymbol{\pi}_{n,\text{opt}} = \arg\max_{\boldsymbol{\pi}_n} \left\{ \sum_{k=1}^{K} \left(\mu_k \left(\log_2 \left(\frac{\gamma_k g_{nk}^{(\boldsymbol{\pi}_n)}}{\Gamma} \right) \right)^+ - \lambda_k \left(\gamma_k - \frac{\Gamma}{g_{nk}^{(\boldsymbol{\pi}_n)}} \right)^+ \right) \right\}. \tag{10}$$

After solving (9) for all n, $f(\boldsymbol{\lambda})$ can be derived from (6). Finally, the dual optimal solution is obtained by minimizing $f(\boldsymbol{\lambda})$ over non-negative λ_k's. Though $f(\boldsymbol{\lambda})$ is convex, a search method based on gradient is infeasible since the dual function is not differentiable. However, the search direction for non-differentiable functions can be found by using subgradient-type methods. A vector $\mathbf{d}$ is a subgradient of $f(\boldsymbol{\lambda})$ at $\boldsymbol{\lambda}$ if for any $\boldsymbol{\lambda}' \succeq 0$, $f(\boldsymbol{\lambda}') \geq f(\boldsymbol{\lambda}) + \sum_{k=1}^{K} d_k(\lambda_k' - \lambda_k)$. Suppose $\boldsymbol{\lambda}^*$ minimize $f(\boldsymbol{\lambda})$, the sub-gradient at any $\boldsymbol{\lambda}$ indicates that $\boldsymbol{\lambda}^*$ can not lie in the half-space $\{\boldsymbol{\lambda}' : \sum_{k=1}^{K} d_k(\lambda_k' - \lambda_k) \geq 0\}$. The next proposition shows a sub-gradient for the above problem.

Proposition 1: For WSRmax with a dual objective $f(\boldsymbol{\lambda})$ in (6), the following choice of $\mathbf{d}$ is a subgradient for $f(\boldsymbol{\lambda})$:

$$d_k = P_k - \sum_{n=1}^{N} p_{nk}^* \qquad k = 1, \cdots, K, \tag{11}$$

where $\{p_{nk}^*\}$ and $\{\boldsymbol{\pi}_n^*\}$ optimize the maximization problem in the definition of $f(\boldsymbol{\lambda})$.

Proof: Since $\{p_{nk}^*\}$ and $\{\boldsymbol{\pi}_n^*\}$ are already in $\mathcal{D}$, for any $\boldsymbol{\lambda}' \succeq 0$,

$$\begin{aligned} f(\boldsymbol{\lambda}') &\geq \mathcal{L}(\{p_{nk}^*\}, \boldsymbol{\lambda}', \{\boldsymbol{\pi}_n^*\}) \\ &= f(\boldsymbol{\lambda}) + \sum_{k=1}^{K} \left(P_k - \sum_{n=1}^{N} p_{nk}^* \right) (\lambda_k' - \lambda_k). \end{aligned} \tag{12}$$

■

The ellipsoid method is one of efficient sub-gradient search methods for updating $\boldsymbol{\lambda}$. This method is shown to converge in $\mathcal{O}(m^2)$ iterations where m is the number of variables [9]. If each user's converged sum power is equal to individual power constraint, the duality gap is zero and the obtained dual solution is in fact globally optimal. According to Theorem 1 in [10], in multi-tone applications, as the number of tones goes to infinity, many resource allocation problems such as WSRmax satisfy the time-sharing property and the duality gap converges to zero. [11] corroborates this argument by showing that for the downlink orthogonal frequency division multiple access (OFDMA) systems, the duality gap for WSRmax is virtually zero with only tens of tones. Thus, in DSL systems with hundreds or thousands of tones, the duality gap is expected to completely vanish, which validates the dual approach.

B. Modified Greedy Algorithm

[12] presents a greedy search algorithm for sum-rate maximization in MIMO BC using precoding based on zero-forcing dirty paper coding (ZF-DPC). ZF-DPC is also based on QR decomposition of the channel matrix and can be considered as the dual of ZF-GDFE for MAC. To efficiently find orderings maximizing the sum-rate, greedy algorithm reduces the search domain by successively selecting the user with the best channel SNR. However, the extension of greedy algorithm to the general weighted sum-rate maximization problem is a nontrivial open problem. In this subsection, we propose a modified greedy algorithm (MGA), which very efficiently finds a good suboptimal solution to WSRmax in upstream vectored DSL systems using ZF-GDFE. By combining the greedy algorithm with Lagrange dual decomposition, MGA significantly reduces complexity of the subproblem (7).

In MGA, the ordering is determined successively, i.e. from $\pi_{n1,\text{MGA}}$ to $\pi_{nK,\text{MGA}}$, based on which user can help the maximization of (7) most. At the beginning of MGA, the user with highest priority (last in decoding) in the QR decomposition is selected. The criterion is to find the user with the largest $\mu_k \log_2 \left(1 + p_{nk} g_{nk}^{(\boldsymbol{\pi}_n)} / \Gamma \right) - \lambda_k p_{nk}$. Since the waterfilling solution (8) is also applicable, the selection becomes:

$$\pi_{n1,\text{MGA}} = \arg\max_{k} \left\{ \mu_k \left(\log_2 \left(\frac{\gamma_k \left(g_{nk}^{(\boldsymbol{\pi}_n)} |_{\pi_{n1}=k} \right)}{\Gamma} \right) \right)^+ - \lambda_k \left(\gamma_k - \frac{\Gamma}{\left(g_{nk}^{(\boldsymbol{\pi}_n)} |_{\pi_{n1}=k} \right)} \right)^+ \right\}. \tag{13}$$

Note that $g_{nk}^{(\boldsymbol{\pi}_n)}|_{\pi_{n1}=k}$ is completely determined by knowing $\pi_{n1} = k$ and is equal to the square of norm of the k-th column of the equivalent channel matrix $\tilde{H}_n$.

After finding the user with the highest priority, the Householder matrix Q^1 corresponding to this user can be computed. The subchannel gains for all the users become the square of the norm of Q^1 times the equivalent channel. The user with the second highest priority is then determined by the similar argument. In general, at the i-th step, Q^{i-1} and the following subchannel gains have to be calculated before $\pi_{ni,\text{MGA}}$ is determined:

$$\begin{aligned} g_{nk}^{\text{MGA}(i)} &= g_{nk}^{(\boldsymbol{\pi}_n)}|_{\substack{\pi_{ni}=k \\ \pi_{nj}=\pi_{nj,\text{MGA}}\forall j<i}}, \\ &= \left|\left|A^{(i:K,k)}\right|\right|^2 \\ &\forall \quad k \notin \{\pi_{n1,\text{MGA}}, \cdots, \pi_{n(i-1),\text{MGA}}\}, \end{aligned} \quad (14)$$

where $A = Q^{i-1} \cdots Q^1 \tilde{H}_n$ and $A^{(i:K,k)}$ represents the vector formed by the i-th to K-th elements on the k-th column of A.

Thus, the ordering and PSD is assigned by:

$$\pi_{ni,\text{MGA}} = \arg\max_k \left\{ \mu_k \left(\log_2 \left(\frac{\gamma_k g_{nk}^{\text{MGA}(i)}}{\Gamma} \right) \right)^+ - \lambda_k \left(\gamma_k - \frac{\Gamma}{g_{nk}^{\text{MGA}(i)}} \right)^+ \right\}, \quad (15)$$

and

$$p_{nk} = \left(\gamma_k - \frac{\Gamma}{g_{nk}^{(\boldsymbol{\pi}_{n,\text{MGA}})}} \right)^+. \quad (16)$$

Furthermore, $Q_{n,\boldsymbol{\pi}_{n,\text{MGA}}}$, $D_{n,\boldsymbol{\pi}_{n,\text{MGA}}}$, and $G_{n,\boldsymbol{\pi}_{n,\text{MGA}}}$ are obtained by (3) after the search.

To summarize, the whole problem can be solved in a suboptimal sense by the following algorithm:

```
Select an initial λ
While PSD for individual user does not converge, do
    For n = 1 to N
        For i = 1 to K
            Calcualte Householder matrix Q^{i-1} if i ≠ 1
            Find g_nk^{MGA(i)} according to (14)
            Determine π_{ni,MGA} and p_nk according to
                (15) and (16)
    Update λ by the ellipsoid method based on
        the sub-gradient in (11)
End While
```

At each iteration, the PSD assignments, $\{p_{nk}\}$, given by the suboptimal decoding orders, $\boldsymbol{\pi}_{n,\text{MGA}}$, deviate from those given by the optimal decoding orders, $\boldsymbol{\pi}_{n,\text{opt}}$. As a result, the sub-gradient in MGA is only an approximation of the true sub-gradient in (11). Fortunately, in all of the examples tested, it is observed that the weighted sum-rate achieved by MGA is very close to the optimal and all individual power constraints are active as well.

It also worth mentioning that instead of selecting one best user, it is also possible for MGA to select L users at a time. At i-th iteration, $[\pi_{n,(i-1)L+1,\text{MGA}}, \cdots, \pi_{n,iL,\text{MGA}}]$ is determined by searching all $\frac{(K-L(i-1))!}{(K-Li)!}$ possible permutations of the users with unassigned orderings.

C. Complexity

The complexity of the ellipsoid method grows with the square of the number of variables, so $\mathcal{O}(K^2)$ iterations are needed no matter optimal search or MGA is used. Since N subproblems (7) have to be solved, the overall computational complexity is $\mathcal{O}(NK^2)$ times the complexity of solving (7).

For optimal search (or equivalently MGA with $L = K$), each subproblem requires a search of $K!$ orderings, and for each ordering a summation of K terms is done. Therefore, the overall complexity of solving this problem becomes $\mathcal{O}(NK^3K!)$. Moreover, the search in (9) requires the information of all $g_{nk}^{(\boldsymbol{\pi}_n)}$. All subchannel gains are stored in memory and be retrieved at every iteration. The desired memory size is then equal to $BNKK!$ bits, where B is the number of bits required to store a subchannel gain. If one chooses to reserve the memory by calculating QR decompositions inside the subproblem repeatedly at every iteration, this would produce K^3 more computational complexity. These properties make the implementation of this search infeasible, especially when K is large.

On the other hand, when $L = 1$, selecting the best user in MGA does not increase the order of the complexity. Hence, the complexity of MGA is the same as that of QR decomposition, $\mathcal{O}(K^3)$. Considering all tones and outer loop, $\mathcal{O}(NK^5)$ executions are required to find the suboptimal solution. The significant reduction of computational complexity makes MGA a practical scheme for vectored DSL. Moreover, no reservation of memory is needed for storing the subchannel gains since the QR decomposition is done inside the subproblem. As L increases, the weighted sum-rate also increases but so does the computational complexity. An example in the next section shows how this tradeoff works.

IV. Simulation Results

In Fig. 1, the achievable rate regions for a very-high-bit-rate DSL (VDSL) system with two users are simulated. The loop length of the first user is 300 feet while it of the second user is 3 Kft. 2048 complex tones each with subchannel bandwidth 4.3125kHz, upstream transmit power 14.5 dBm, carrier mask, and PSD mask are assigned according to [13]. Although the PSD mask introduce N more constraints to the optimization problem, it can be applied to the search domain $\mathcal{D}$ and the PSD in (8), without changing the main body of the algorithm. One T1 disturber is introduced as the common alien noise source. To get the noise covariance matrices, the correlation coefficients are assigned to 0.99 for all tones. -140dBm/Hz white noise is also added to all users. And an SNR gap of 12dB, which may combine coding gain and margin, is assumed. There are total 4 regions in Fig. 1. Three of them are obtained by ZF-GDFE with different strategies of ordering

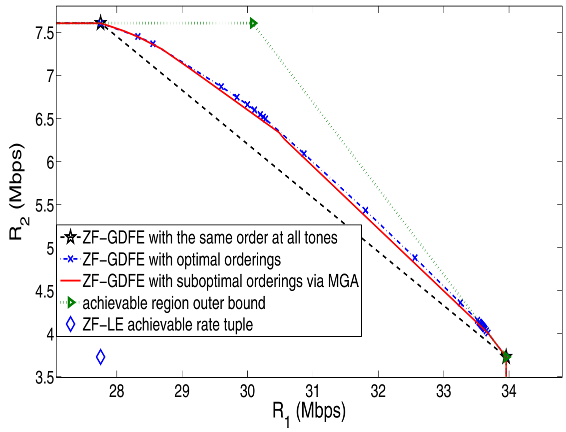

Fig. 1. Achievable rate region for ZF-GDFE

search. The dashed curve shows the rate region of using the same ordering for all tones, i.e. $\boldsymbol{\pi}_1 = \boldsymbol{\pi}_2 = \cdots = \boldsymbol{\pi}_N$. Since there are only 2! rate-tuples, time-sharing is used to find the achievable rate region. The remaining two ZF-GDFE rate regions are obtained by applying optimal search and MGA ($L = 1$), respectively. Other than those 3 achievable rate regions, an outer bound containing three line segments, where R_1, R_2, and sum rate are maximized, is provided as in [1]. It can be seen that the region achieved by MGA is very close to the optimal ZF-GDFE achievable rate region, and both are strictly larger than the one using time-sharing of rates with uniform ordering. The outer bound is loose in general, but it is assured that in this environment the bound is tight when R_1 is greater than 33 Mbps and can be achieved by ZF-GDFE whenever a good orderings is chosen. The achievable rate tuple for zero-forcing linear equalizer (ZF-LE) [14] is also shown in the figure. When noise is correlated, ZF-LE is far from optimal and feeding back previous decisions is important.

In Table I, a four-user system with loop lengths 300 feet, 1 Kft, 2 Kft, and 3 Kft is simulated. The parameters are selected the same as in the previous example. The weight for WSRmax is selected at random. And one of the 4! rate tuples using the same ordering for all tones is selected for comparison. As L and the computational complexity increases, the weighted sum-rate also increases. Thus, the tradeoff between complexity and weighted sum-rate can be verified.

In addition to higher weighted sum-rate, MGA can support much larger number of rate tuples than only $K!$ rate tuples achieved by the same ordering on every tone. Though the number of rate tuples can grow by time-sharing or frequency-sharing multiple rate tuples, these methods substantially increase the computational complexity. In order to obtain a new rate tuple by time-sharing, $NK!$ executions of QR decomposition and $KK!$ single-user waterfillings are required. Moreover, possibly up to $K!$ linear equations need to be solved in determining each user's time-sharing portion. Therefore, higher achievable rates, a simpler implementation and more flexible rate choices make MGA an advantageous scheme.

$\boldsymbol{\mu}$	data rate	$L=1$	$L=2$	$L=4$	unform ordering	ZF-LE
μ_1 = 0.424	R_1	35.380	35.302	32.557	27.742	27.742
μ_2 = 0.338	R_2	27.681	27.377	25.157	28.567	23.156
μ_3 = 0.566	R_3	17.974	18.055	17.600	14.244	11.412
μ_4 = 0.621	R_4	4.052	4.256	8.477	11.331	3.734
	WSR	37.047	37.083	37.533	36.516	28.367

TABLE I

SIMULATION RESULTS FOR A 4-USER SYSTEM

V. CONCLUSION

In upstream vectored DSL systems with ZF-GDFE, different decoding order at each tone can substantially affect each user's achievable rate. For a given weighted sum-rate maximization problem, this paper employs Lagrange dual decomposition to achieve linear complexity in the number of tones. To find the optimal orderings, exhaustive search over all possible $K!$ orderings is required at each maximization of Lagrangian. In this paper, proposed MGA avoids the exhaustive search by maximizing individual terms of the subproblem iteratively. The search complexity is significantly reduced, while the loss on the weighted sum-rate is very small. With large number of vectored DSL lines in the upstream, MGA emerges as a promising algorithm for its high numerical efficiency.

REFERENCES

[1] G. Ginis and J.M. Cioffi "Vectored transmission for digital subscriber line systems," *IEEE Journal on Selected Areas in Communications*, Vol. 20, No.5, pp.1085-1104, June 2002.

[2] T. Starr, M. Sorbara, J. M. Cioffi, and P. J. Silverman, *DSL Advances*, Prentice Hall, 2003.

[3] G. Ginis and C.-N. Peng, "Alien crosstalk cancellation for multi-pair digital subscriber line systems," in *EURASIP Special Issue on Digital Subscriber Line* , Volume 2006, Article ID 16828, 2006.

[4] T. Cover and J. Thomas, *Elements of Information Theory*, New York: John Wiley and Sons Inc., 1997.

[5] J. M. Cioffi and G. D. Forney, "Generalized decision-feedback equalization for packet trasmission with ISI and Gaussian noise," in *Communication, Computation, Control and Signal Processing (A tribute to Thomas Kailath),* A. Paulraj, V. Roychowdhury, and C. Schaper, Eds. Boston, MA: Kluwer, 1997, Ch.4, pp. 79-127.

[6] M. Mohseni, R. Zhang, and J. M. Cioffi, "Optimized transmission of fading multiple-access and broadcast channels with multiple antennas," in *IEEE J. Select. Areas Comm.*, vol. 24, no. 8, pp.1627-1639, Aug. 2006.

[7] M. Mohseni and J. M. Cioffi, "Sub-optimal throughput maximization schemes for Guassian vector broadcast channels," in *Proceedings of the 2005 IEEE International Conference on Communications, ICC 2005*, Vol.1, pp. 709–714, Seoul, May 2005.

[8] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Combridge University Press, 1985.

[9] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2003.

[10] W. Yu and R. Lui, "Dual methods for non-convex spectrum optimization of multi-carrier systems," accepted in *IEEE Transactions on Communications*, 2006.

[11] K. Seong, M. Mohseni, and J. Cioffi, "Optimal resource allocation for OFDMA downlink systems," in *Proc. IEEE International Symposium on Information Theory*, Seattle, WA, pp. 1394-1398, July 2006.

[12] Z. Tu and R. S. Blum, "Multiuser diversity for a dirty paper approach," *IEEE Communication Letters*, Vol. 7, No. 8, August 2003.

[13] S. Schelstraete, "Very-high-bit-rate digital subscriber line (VDSL) metallic interface," ANSI T1E1.4/2003-210R1, August 2003.

[14] R. Cendrillon, G. Ginis, E. Van Den Bogaert, and M. Moonen, "A near-optimal linear crosstalk canceler for upstream VDSL," *IEEE Transactions on Signal Processing*, Vol. 54, No.8, pp.3136-3146, Aug. 2006.