

# Dual Optimization Methods for Multiuser Orthogonal Frequency Division Multiplex Systems

Wei Yu and Raymond Lui

Electrical and Computer Engineering Dept.  
University of Toronto, Canada  
{weiyu, rwmlui}@comm.utoronto.ca

Raphael Cendrillon

Electrical Engineering Dept.  
Katholieke Universiteit Leuven, Belgium  
cendrillon@ieee.org

**Abstract**—The design and optimization of orthogonal frequency division multiplex (OFDM) systems typically take the following form. The design objective is to maximize the total data rate which is the sum of individual rates in each frequency tone. The design constraints are usually linear constraints imposed across all tones. This paper shows that regardless of whether the objective and the constraints are convex, the duality gap for this class of problems is always zero in the limit as the number of frequency tones goes to infinity. As the dual problem typically decouples into many smaller per-tone problems, solving the dual problem is much more efficient. This observation leads to an efficient method to find the global optimum of non-convex optimization problems for the OFDM system. Multiuser optimal power allocation, optimal frequency planning and optimal low-complexity crosstalk cancellation for vectored DSL are used to illustrate this point.

## I. INTRODUCTION

In an orthogonal frequency-division multiplex (OFDM) system, the frequency domain is partitioned into a large number of tones. Data transmission takes place in each tone independently. The overall system throughput is the sum of individual rates in each frequency tone. The design constraints are typically linear but coupled across all the tones. The design problem involves the optimization of the overall performance subject to design constraints. For example, the optimal bit and power allocation problem is often formulated as follows. Let  $H(n)$ ,  $P(n)$  and  $N(n)$  denote the channel frequency response, the transmit power spectral density and the noise power spectral density at tone  $n$ , respectively. The optimization problem is:

$$\begin{aligned} & \text{maximize} && \sum_{n=1}^N \log \left( 1 + \frac{P(n)H^2(n)}{N(n)} \right) \\ & \text{subject to} && \sum_{n=1}^N P(n) \leq \mathbf{P} \\ & && P(n) \geq 0. \end{aligned} \quad (1)$$

The above problem has a well-known solution called “water-filling”. Efficient solution exists in this case because the objective function is concave in the optimizing variable  $P(n)$ .

<sup>1</sup>This work was supported by Bell Canada University Laboratories, Communications and Information Technology Ontario (CITO), Natural Sciences and Engineering Council (NSERC) of Canada, and by the Canada Research Chairs program.

Unfortunately, not all optimization objectives are concave. The multiuser bit and power allocation is such an example. In this case, several OFDM transmitters interfere with each other, and the sum rate maximization problem becomes:

$$\begin{aligned} & \text{max} && \sum_{k=1}^K \sum_{n=1}^N \log \left( 1 + \frac{P_k(n)H_{kk}^2(n)}{N(n) + \sum_{j \neq k} H_{jk}^2(n)P_j(n)} \right) \\ & \text{s.t.} && \sum_{n=1}^N P_k(n) \leq \mathbf{P}_k \quad k = 1, \dots, K \\ & && P_k(n) \geq 0, \quad k = 1, \dots, K \end{aligned} \quad (2)$$

where  $H_{jk}(n)$  is the channel transfer function from system  $j$  to system  $k$  in tone  $n$ ,  $P_k(n)$  is the power allocation for user  $k$  in tone  $n$ , each user has a separate power constraint. Because the objective function is not concave in  $P_k(n)$ , the optimization problem is difficult to solve. Previous approaches (e.g. iterative water-filling [1] and others [2] [3]) use heuristics to derive suboptimal solutions.

Recently, an efficient and exact “Optimal Spectrum Balancing” algorithm to solve this problem was proposed in [4]. The basically idea is as follows. Form the Lagrangian of the optimization problem (2):

$$\begin{aligned} & \text{max} && \sum_{k=1}^K \sum_{n=1}^N \log \left( 1 + \frac{P_k(n)H_{kk}^2(n)}{N(n) + \sum_{j \neq k} H_{jk}^2(n)P_j(n)} \right) \\ & && + \sum_{k=1}^K \lambda_k \left( \mathbf{P}_k - \sum_{n=1}^N P_k(n) \right) \\ & \text{s.t.} && P_k(n) \geq 0, \quad k = 1, \dots, K. \end{aligned} \quad (3)$$

Solve the Lagrangian for each set of positive and fixed  $(\lambda_1, \dots, \lambda_K)$ . Then, the solution to the original problem may be found by an exhaustive search over the  $\lambda$ -space.  $\lambda_k$  is increased or decreased depending on whether  $\sum_{n=1}^N P_k(n)$  is greater or less than  $\mathbf{P}_k$ . When the process converges, either  $\lambda_k = 0$  or  $\sum_{n=1}^N P_k(n) = \mathbf{P}_k$  for each  $k$ . In this case, the Lagrangian objective is identical to the original objective, thus solving the original problem.

This Lagrangian approach works well because of the following. First, for a fixed  $\lambda_k$ , the objective decouples into  $N$  independent problems corresponding to the  $N$  frequency tones. Thus, solving the dual problem requires a much lower computational complexity as compared to the original problem. Second,  $\lambda_k$  represents the price of power for user  $k$ . A higher

price leads to a lower power usage. Thus, as a function of  $\lambda_k$ , the optimal  $\sum_{n=1}^N P_k(n)$  is monotonic in  $\lambda_k$ . An exhaustive search over the  $\lambda$ -space can then be performed using bisection on each  $\lambda_k$ . This is essentially an exhaustive search over all power usages. It leads to the global optimum regardless of whether the original problem is convex. However, with  $K$  users,  $K$  loops of bisections are involved, one for each  $\lambda_k$ . Therefore, the computational complexity of optimal spectrum balancing, although linear in  $N$ , is exponential in  $K$ . When the number of users is large, the complexity becomes prohibitive.

The purpose of this paper is first to refine the optimal spectrum balancing algorithm with an aim of eliminating the exponential complexity in Lagrangian search, and second to generalize the algorithm for other optimization problems in multiuser OFDM system design. Toward this end, we show that the optimal spectrum balancing algorithm belongs to a class of dual optimization methods. Contrary to general non-convex problems, the duality gap for multiuser OFDM optimization always tends to zero as the number of frequency tones goes to infinity regardless of whether the optimization problem is convex. This observation is inspired by the earlier work of Bertsekas *et al* [5] and it leads to more efficient  $\lambda$ -search methods. In the second part of paper, we show that the general theory is applicable to many other areas of OFDM system design. Optimal frequency planning and optimal complexity allocation in vectored digital subscriber line systems are some of these examples.

## II. DUAL OPTIMIZATION METHODS

### A. Duality Gap

Consider an optimization problem in which both the constraints and the objective function consist of a large number of individual functions, corresponding to the  $N$  frequency tones:

$$\begin{aligned} & \text{maximize} && \sum_{n=1}^N f_n(x_n) \\ & \text{subject to} && \sum_{n=1}^N h_n(x_n) \leq \mathbf{P}, \end{aligned} \quad (4)$$

where  $f_n(\cdot)$  is a scalar function which is not necessarily concave, and  $h_n(\cdot)$  is a vector-valued function that is not necessarily convex.  $\mathbf{P}$  is a vector of constraints. Also, there may be other (possibly integer) constraints implicit in the problem. The idea of the dual method is to solve (4) via its Lagrangian:

$$L(x_n, \lambda) = \sum_{n=1}^N f_n(x_n) + \lambda^T \cdot \left( \mathbf{P} - \sum_{n=1}^N h_n(x_n) \right), \quad (5)$$

where  $\lambda$  is a vector, and “ $\cdot$ ” denotes vector dot product. Note that the Lagrangian decouples into a set of  $N$  smaller problems, so optimizing the Lagrangian is much easier than solving (4). Define the dual objective  $g(\lambda)$  as the solution to the following:

$$g(\lambda) = \max_{x_n} L(x_n, \lambda) \quad (6)$$

The dual optimization problem is:

$$\begin{aligned} & \text{minimize} && g(\lambda) \\ & \text{subject to} && \lambda \geq 0. \end{aligned} \quad (7)$$

When  $f_n(x_n)$  is concave and  $h_n(x_n)$  is convex, standard convex optimization results guarantee that the primal problem (4) and the dual problem (7) have the same solution. When convexity does not hold, the dual problem provides a solution which is an upper bound to the solution of (4). The upper bound is not always tight, and the difference is called the “duality gap”.

In multiuser OFDM design, convexity often does not hold. However, it is usually the case that the following “time-sharing” property is satisfied:

*Definition 1:* An optimization problem of the form (4) satisfies the time-sharing property if the following holds: Let  $x_n$  and  $y_n$  be optimal solutions to the problem with  $\mathbf{P} = P_x$  and  $\mathbf{P} = P_y$ , respectively. Then, for any  $0 \leq \nu \leq 1$ , there exists a set of  $z_n$  such that  $\sum_n h_n(z_n) \leq \nu P_x + (1 - \nu) P_y$ , and  $\sum f_n(z_n) \geq \nu \sum f_n(x_n) + (1 - \nu) \sum f_n(y_n)$ .

This property is clearly satisfied if time-division multiplexing may be implemented. (Throughout the paper, the channels are assumed to be time invariant.) The frequency tones can then be assigned to  $x_n$  for  $\nu$  percentage of the time and  $y_n$  for  $(1 - \nu)$  percentage of the time. In practical OFDM systems in which channel conditions in adjacent tones are similar and there are a large number of frequency tones, the time-sharing property can be satisfied with frequency-sharing. This is true because time-sharing can be approximately implemented by interleaving  $x_n$  and  $y_n$  in the frequency domain. As  $N \rightarrow \infty$ , frequency-sharing is equivalent to time-sharing.

Note that the concavity of  $f_n(x_n)$  and the convexity of  $h_n(x_n)$  and all other constraints imply time-sharing but not vice versa. Time-sharing is always satisfied regardless of convexity as long as  $N$  is sufficiently large and  $f_n \cdots f_{n+k}$  are sufficiently similar for small values of  $k$  (and likewise for  $h_n \cdots h_{n+k}$ .) This is the case in almost all OFDM systems as the subchannel width in OFDM systems is chosen so that the channel response in adjacent subchannels are approximately the same.

The main result of this section is that the time-sharing property implies that the duality gap is zero.

*Theorem 1:* If an optimization problem satisfies the time-sharing property, then it has zero duality gap, i.e. the primal problem (4) and the dual problem (7) have the same solution.

*Proof:* The proof is obvious if (4) is convex. Fig. 1 illustrates the proof when convexity does not hold but time-sharing does. The first diagram illustrates a function that satisfies the time-sharing property. The solid line plots the optimal  $(\sum h_n(x_n^*), \sum f_n(x_n^*))$  as the constraint  $\mathbf{P}$  varies. The intersection of the curve with the vertical axis where  $\sum h_n(x_n^*) = \mathbf{P}$  is the optimal value of the primal objective. Clearly larger  $\mathbf{P}$  leads to higher objective value, so the

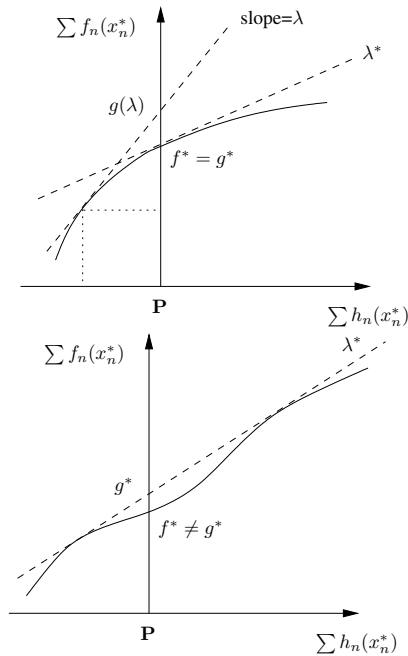


Fig. 1. Time-sharing property implies zero duality gap.

curve is increasing. More importantly, the curve is concave because of the time-sharing property. Now, consider a fixed tangent line with slope  $\lambda$ . By the definition of  $L(\lambda, x_n)$ , the intersection of the tangent line with the vertical axis is precisely  $g(\lambda)$ . This allows the minimization of the dual problem to be visualized easily. As  $\lambda$  varies,  $g(\lambda)$  achieves a minimum at exactly the maximum value of the primal objective. Thus, the duality gap is zero. (The second diagram illustrates a case where time-sharing property does not hold. In this case, the minimum  $g(\lambda)$  is strictly larger than the maximum  $\sum f_n(x_n)$ .)  $\square$

The main consequence of Theorem 1 is that as long as the time-sharing property is satisfied, even a non-convex optimization problem can be solved by solving its dual. The dual problem is typically much easier to solve because it usually lies in a lower dimension. Further,  $g(\lambda)$  is convex regardless of the concavity of  $f_n(x_n)$ . (This is because  $L(x_n, \lambda)$  is linear in  $\lambda$  for each fixed  $x_n$ , and  $g(\lambda)$  is the maximum of linear functions and is therefore convex.) Thus, any gradient-based algorithm is guaranteed to converge. Note that the optimization of  $g(\lambda)$  requires an efficient evaluation of  $g(\lambda)$ . This usually involves an exhaustive search over the primal variables. However, as  $g(\lambda)$  is unconstrained and it decouples into  $N$  independent sub-problems, such an exhaustive search is much more manageable.

### B. Dual Methods

The optimal spectrum balancing algorithm solves  $L(x_n, \lambda)$  exhaustively for all possible values of  $\lambda$ . The multiuser spectrum optimization problem (2) consists of  $K$  constraints, and successive bisection on each component of  $\lambda$  would yield

the primal optimum. The main point of this paper is that we can take advantage of the duality relation and solve the dual objective  $g(\lambda)$  instead. By using an efficient search of  $\lambda$ , the computational efficiency of the optimal spectrum balancing can be improved drastically.

The main difficulty in deriving an efficient direction for  $\lambda$  is that  $g(\lambda)$  is not necessarily differentiable. Thus, it does not always have a gradient. Nevertheless, it is possible to find a search direction based on what is called a subgradient. A vector  $d$  is a subgradient of  $g(\lambda)$  at  $\lambda$  if for all  $\lambda'$

$$g(\lambda') \geq g(\lambda) + d^T \cdot (\lambda' - \lambda). \quad (8)$$

Subgradient is a generalization of gradient for (possibly) non-differentiable functions. Intuitively,  $d$  is a subgradient if the linear function with slope  $d$  passing through  $(\lambda, g(\lambda))$  lies entirely below  $g(\lambda)$ . For  $g(\lambda)$  defined in (6), the following choice of  $d$

$$d = \mathbf{P} - \sum_{n=1}^N h_n(x_n) \quad (9)$$

satisfies the subgradient condition (8). The subgradient search suggests that  $\lambda$  should be increased if  $\sum_{n=1}^N h_n(x_n) > \mathbf{P}$  and decreased otherwise. This is intuitively obvious as  $\lambda$  represents a price for power. Price should increase if the constraint is violated. In fact,  $\lambda$  updates can be done systematically. It is possible to prove that the following update rule

$$\lambda^{l+1} = \left[ \lambda^l - s^l \left( \mathbf{P} - \sum_n h_n(x_n) \right) \right]^+ \quad (10)$$

is guaranteed to converge to the optimal  $\lambda$  as long as  $s^l$  is chosen to be sufficiently small [6]. Here,  $s^l$  is a scalar. By Theorem 1, the minimum  $g(\lambda)$  is also equal to the maximum  $\sum f_n(x_n)$ . Thus, the solution to the dual problem immediately yields the optimal solution to the original problem.

The crucial difference between the update equation (10) and that suggested in [4] is that (10) updates all components of  $\lambda$  at the same time. Instead of doing bisection on each component individually, the subgradient method collectively finds a suitable direction for all components of  $\lambda$  at once. This eliminates the exponential complexity in  $\lambda$ -search.

However, note that the evaluation of  $g(\lambda)$  is still exponential in  $K$ . This is probably inevitable if an exact solution to the non-convex optimization problem is desired. For practical problems, however, sub-optimal methods in evaluating  $g(\lambda)$  often exist.

## III. APPLICATIONS

### A. Multiuser Spectrum Management

We now return to the multiuser optimal spectrum management problem. In digital subscriber line applications, electromagnetic coupling induces crosstalk between adjacent lines. The goal of optimal spectrum management is to find a set of power allocations  $(P_1(n), \dots, P_K(n))$  so that a target rate-tuple is satisfied. There is generally a trade-off between the achievable data rates of different users. Such a trade-off can be

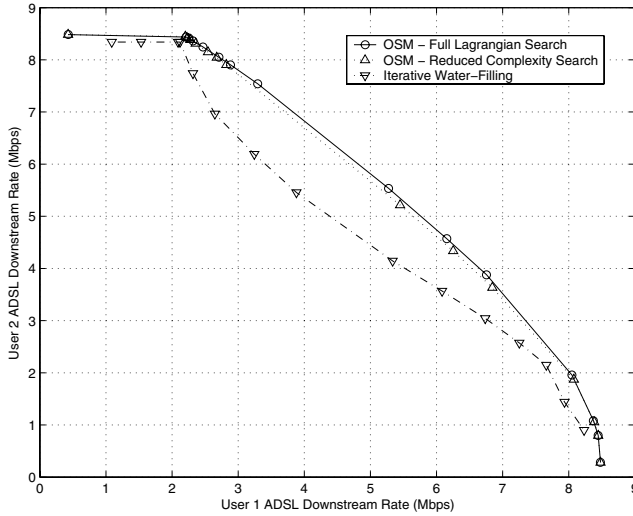


Fig. 2. Rate region for the two-user ADSL lines

represented in a rate region defined as the set of all achievable rates  $(R_1, \dots, R_K)$ . When the channel transfer function is a slow varying function of  $n$ , the spectrum optimization problem satisfies the time-sharing property.

In this section, we formulate a novel optimization problem that characterizes the boundary of the rate region. The objective is the maximization of a base rate  $R$  subject to a fixed ratio between  $R_k$  and  $R$  for each  $k = 1, \dots, K$ . More specifically, we may insist that  $R_1 : R_2 : \dots : R_K = \beta_1 : \beta_2 : \dots : \beta_K$ , where

$$R_k = \sum_{n=1}^N \log \left( 1 + \frac{P_k(n) H_{kk}^2(n)}{N(n) + \sum_{j \neq k} H_{jk}^2(n) P_j(n)} \right). \quad (11)$$

Then, the maximization problem becomes

$$\begin{aligned} \max \quad & R \\ \text{s.t.} \quad & R_k \geq \beta_k R \\ & \sum_{n=1}^N P_k(n) \leq \mathbf{P}_k, \quad k = 1, \dots, K \\ & P_k(n) \geq 0, \quad k = 1, \dots, K \end{aligned} \quad (12)$$

Here, the variables  $\beta_k$  directly represent the ratios of service rates among the different users.

The dual function for (12) can be written as follows:

$$g(\omega_1, \dots, \omega_K, \lambda_1, \dots, \lambda_K) = \max_{P_k, R} \quad (13)$$

$$R + \sum_{k=1}^K \omega_k (R_k - \beta_k R) + \sum_{k=1}^K \lambda_k \left( \mathbf{P}_k - \sum_{n=1}^N P_k(n) \right)$$

Collecting terms, we see that the maximization involves a term  $(1 - \sum \omega_k \beta_k) R$ . Since  $R$  is a free variable to be optimized, the maximization leads to  $R = \infty$  if  $(1 - \sum \omega_k \beta_k) > 0$  and  $R = 0$  if  $(1 - \sum \omega_k \beta_k) < 0$ . Thus, non-trivial solution exists only if  $(1 - \sum \omega_k \beta_k) = 0$ .

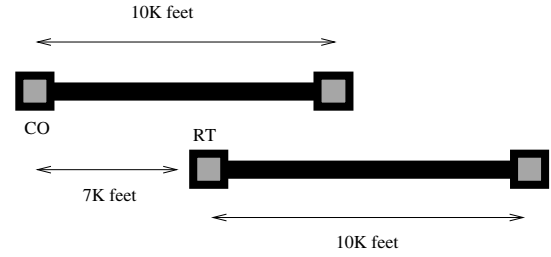


Fig. 3. Topology of the two-user ADSL lines

It is now straightforward to apply the technique developed in the previous section to derive a subgradient search for the minimization of  $g(\omega_1, \dots, \omega_K, \lambda_1, \dots, \lambda_K)$ . The idea is the following. First, solve the maximization problem (13) for a fixed set of  $(\omega_1, \dots, \omega_K, \lambda_1, \dots, \lambda_K)$  with  $(1 - \sum \omega_k \beta_k) = 0$ . This is done using exhaustive search in each tone separately and it yields a set of power allocation  $P_k(n)$  and achievable rates  $R_k$ . The maximum  $R$  can be found as  $R = \min_k R_k / \beta_k$ . The subgradient method can now be used to update  $\omega_k$  and  $\lambda_k$ :

$$\omega_k^{l+1} = [\omega_k^l - s_k^l (R_k - \beta_k R)]^+ \quad (14)$$

$$\lambda_k^{l+1} = \left[ \lambda_k^l - t_k^l \left( \mathbf{P} - \sum_{n=1}^N P_k(n) \right) \right]^+ \quad (15)$$

Note that the new  $\omega_k$  may no longer satisfy  $\sum \omega_k \beta_k = 1$ . Renormalization is needed to project  $\omega_k$  back to the proper subspace

$$\omega_k^{l+1} = \frac{\omega_k^{l+1}}{\sum_k \omega_k^{l+1} \beta_k}. \quad (16)$$

As long as  $s_k^l$  and  $t_k^l$  is sufficiently small, the subgradient algorithm is guaranteed to converge. This sub-gradient algorithm improves the computational complexity of the optimal spectrum balancing algorithm described in [4]. No bisection is needed. The number of times that  $g(\omega_k, \lambda_k)$  is evaluated is polynomial in  $K$ .

Note that the evaluation of  $g(\omega_k, \lambda_k)$ , if done exhaustively, still has a complexity exponential in  $K$ . However, for the spectrum optimization problem, experimental results suggest that lower complexity search algorithms often work well. Fig. 2 shows the rate region for a two-user ADSL system with a configuration shown in Fig. 3. Both the full implementation of optimal spectrum balancing and a reduced complexity gradient search are shown. Their performances are very similar, and both outperform iterative water-filling [1] significantly.

### B. Optimal Frequency Planning

The optimal spectrum balancing algorithm is applicable to many other areas of OFDM system design. For example, in a wireless multiuser OFDM system, different users are often allocated to different sets of tones. The optimal power and bit allocation problem is essentially the spectrum management

problem (12) with an additional constraint that only one user occupies each tone [7] [8] [9]:

$$P_k(n)P_j(n) = 0 \quad \forall k \neq j$$

Previous solutions to this problem [7] [9] [8] rely on a relaxation of the non-convex constraint. As Theorem 1 in Section II shows, this problem can instead be efficiently solved in the dual domain. The same subgradient updates as in the previous section apply here. The constraint  $P_k(n)P_j(n) = 0$  for all  $k$  and  $j$  is incorporated in the evaluation of the dual function. Theorem 1 guarantees that the dual solution is identical to the primal solution.

In fact, the complexity of this problem is strictly sub-exponential. The evaluation of the dual  $g(\omega_k, \lambda_k)$  involves an exhaustive search in  $K$  possible power allocations. Its complexity is therefore linear in  $K$ .

### C. Partial Crosstalk Cancellation in Vector DSL

Future digital subscriber line applications are expected to implement crosstalk cancellation and precoding to further improve the data rates in twisted-pair transmission. Multiple transmitters and multiple receivers at the central office can be regarded as a single entity. Crosstalk cancellation can be done in a similar way as echo cancellation.

A typical DSL bundle consists of 50 to 100 twisted pairs. Cancelling all crosstalk involves  $50 \times 50$  to  $100 \times 100$  matrix processing, which is beyond the computational complexity constraints of current digital signal processors. On the other hand, in a 50-pair DSL bundle each twisted-pair has only a limited number of nearest neighbours. Thus, we expect that the cancellation of only a few pairs would achieve most of the benefits. Furthermore, crosstalk is frequency dependent. The crosstalk level is low in low frequency bands, so cancellation in these frequency bands has limited utility. On the other hand, in very high frequency bands, the data rates are already small. Thus, as pointed out in [10], data rate improvement due to crosstalk cancellation is most noticeable in the mid-frequency range.

Given a complexity constraint, how to choose the best combination of lines and tones in which to implement crosstalk cancellation is an interesting problem. This problem was first formulated in [10] and greedy algorithms were suggested. However, the solution in [10] assumes a fixed transmit spectrum level. In this section, we formulate a more realistic problem that jointly performs line/tone selection and spectrum optimization.

The basic setup is the same as the optimization problem (12) except the evaluation of  $R_k$  now takes the following form:

$$R_k = \sum_{n=1}^N \log \left( 1 + \frac{P_k(n)H_{kk}^2(n)}{N(n) + \sum_{j \neq k} G_{jk}^2(n)P_j(n)} \right). \quad (17)$$

where  $G_{kj}(n) = H_{kj}(n)$  except where crosstalk cancellation takes place, in which case  $G_{kj}(n) = 0$ . The total number of places where  $G_{kj}(n) = 0$  represents the number of crosstalk cancellation units that can be implemented. This number

is typically constrained by an implementation limit. More formally,

$$\sum_{n=1}^N \sum_{k \neq j} \mathbf{1}_{\{H_{kj}(n) \neq G_{kj}(n)\}} \leq \mathbf{C} \quad (18)$$

where  $\mathbf{1}_{\{\cdot\}}$  is an indicator function and  $\mathbf{C}$  is a constant representing the complexity constraint over all tones and all users.

Clearly (17) may be solved using the dual formulation. The complexity constraint is the same as the power constraints. As long as exhaustive search within each tone can be done with manageable complexity, the optimization over the  $N$  tones only adds a polynomial factor.

## IV. CONCLUSION

The main point of this paper is that many optimization problems in OFDM design can be decoupled in a tone-by-tone basis via the dual method. It is shown that the time-sharing property is always satisfied when the number of tones is large, and when the time-sharing property is satisfied, the duality gap becomes zero regardless of whether the original problem is convex. This observation leads to efficient dual optimization techniques such as the subgradient method. As long as the evaluation of the dual objective for each tone may be done with manageable complexity, the entire problem may be solved efficiently. This principle is applicable to a wide range of OFDM design problems. Multiuser spectrum optimization, frequency planning and line/tone selection in reduced complexity crosstalk cancellation are some of these examples.

## REFERENCES

- [1] W. Yu, G. Ginis, and J.M. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE J. Sel. Area. Comm.*, vol. 20, no. 5, pp. 1105–1115, June 2002.
- [2] K. S. Jacobsen, "Methods of upstream power backoff on very high-speed digital subscriber lines," *IEEE Comm. Mag.*, pp. 210–6, Mar. 2001.
- [3] G. Cherubini, E. Eleftheriou, and S. Olcer, "On the optimality of power back-off methods," Aug. 2000, ANSI-T1E1.4/235.
- [4] R. Cendrillon, M. Moonen, Jan Verlinden, Tom Bostoen, and W. Yu, "Optimal multi-user spectrum management for digital subscriber lines," in *IEEE Inter. Conf. Comm. (ICC)*, Paris, 2004.
- [5] D. Bertsekas, G. Lauer, N. Sandell Jr., and T. Posbergh, "Optimal short-term scheduling of large-scale power systems," *IEEE Trans. Auto. Control*, vol. 28, no. 1, pp. 1–11, Jan 1983.
- [6] D. Bertsekas, *Nonlinear programming*, Athena Scientific, 1999.
- [7] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE J. Selected Areas Comm.*, vol. 17, no. 10, pp. 1747–1758, Oct 1999.
- [8] L.M.C. Hoo, B. Halder, J. Tellado, and J.M. Cioffi, "Multiuser transmit optimization for multicarrier broadcast channels: asymptotic fdma capacity region and algorithms," *IEEE Trans. Comm.*, vol. 52, no. 6, pp. 922–930, June 2004.
- [9] W. Yu and J.M. Cioffi, "FDMA capacity of Gaussian multiple-access channels with ISI," *IEEE Trans. Comm.*, vol. 50, no. 1, pp. 102–111, Jan 2002.
- [10] R. Cendrillon, M. Moonen, G. Ginis, K. Van Acker, T. Bostoen, and P. Vandaele, "Partial crosstalk cancellation exploiting line and tone selection in upstream vdsl," in *Proc. of Sixth Baiona Workshop on Signal Processing in Communications*, Spain, September 2003.