Spectrum Management for Interference-limited Multiuser Communication Systems

Shunsuke Hayashi[‡] and Zhi-Quan Luo[†]

Abstract—Consider a multiuser communication system in a frequency selective environment whereby users share a common spectrum and can interfere with each other. Assuming Gaussian signaling and no interference cancelation, we study optimal spectrum sharing strategies for the maximization of sum-rate under separate power constraints for individual users. Since the sum-rate function is non-concave in terms of the users' power allocations, there can be multiple local maxima for the sum-rate maximization problem in general. In this paper, we show that, if the normalized crosstalk coefficients are larger than a given threshold (roughly equal to $1/2$), then the optimal spectrum sharing strategy is frequency division multiple access (FDMA). In case of arbitrary positive crosstalk coefficients, if each user's power budget exceeds a given threshold, then FDMA is again sum-rate optimal, at least in a local sense. In addition, we show that the problem of finding the optimal FDMA spectrum allocation is NP-hard, implying that the general problem of maximizing sum-rate is also NP-hard, even in the case of two users. We also propose several simple distributed spectrum allocation algorithms that can approximately maximize sum-rates. Numerical results indicate that these algorithms are efficient and can achieve substantially larger sum-rates than the existing Iterative Waterfilling solutions, either in an interference-rich environment or when the users' power budgets are sufficiently high.

Index Terms—FDMA optimality, multiuser communication system, spectrum management, sum-rate maximization

I. INTRODUCTION

IN a multiuser communication system, interference mitigation is a major design and management objective. A popular approach to minimize multiuser interference is Frequency Division Multiple Access (FDMA) whereby the available spectrum is divided into multiple tones (or bands) and shared by all the users on a non-overlapping basis. Such 'orthogonal channelization' approach is well-suited for high speed structured communication in which quality of service is a major concern. However, FDMA also has major drawbacks such as high system overhead and low bandwidth utilization. The latter is due to the fact that a frequency tone, once assigned to a user, cannot be used by any other users even if it is not fully utilized.

With the proliferation of various radio devices and services, multiple wireless systems sharing a common spectrum must coexist [1]. In such scenarios, system-enforced FDMA solution may no longer be feasible or desirable, and we are naturally led to a situation whereby users can communicate simultaneously over a common spectrum, potentially causing significant interference to each other. For such a multiuser system, each user's performance depends on not only the power allocation (across spectrum) of his own, but also those of other users in the system. Thus, proper spectrum management (i.e., power control) is needed for the maximization of the overall system performance. Spectrum management problem of this type also arises in a digital subscriber line (DSL) system where multiple users communicate with a central office through separate telephone lines over a common spectrum. Due to electro-magnetic coupling, signals transmitted over different telephone wires bundled in close proximity may interfere with each other, resulting in significant signal distortion. In fact, such crosstalk is known to be the major source of signal distortion in a high speed DSL system [2]. Hence, for both wireless and wireline (DSL) applications, judicious management of spectrum among users can have a major impact on the overall system performance.

For many communication systems, a reasonable measure of overall system performance is the sum of achievable rates of all users in the system. The maximum achievable sum-rate (subject to individual power constraints) corresponds to the *social optimum* of the system. Ideally, we would like to enable the users in the system to reach the social optimum through a distributed mechanism whereby the users' power levels are adjusted only locally. Unfortunately, in a frequency selective environment and assuming Gaussian signalling, the sum-rate (in the sense of Shannon) turns out to be a non-concave function of individual user's power allocations. Consequently, the problem of maximizing the sum-rate has multiple local maxima, which in turn makes the computation of a globally optimal spectrum sharing solution difficult (indeed, NP-hard as shown in this paper). Obviously, the distributed maximization of sum-rate constitutes an even more formidable computational challenge.

Recently, some researchers proposed several spectrum management algorithms based on duality theory [3], [4], [5], [6]. In these algorithms, the authors aim to solve the Lagrangian dual relaxation problem instead of the original sum-rate maximization problem. Although the dual problem can be decomposed to lower dimensional problems and their objective functions are convex, it is difficult to solve it exactly since the evaluation of the dual objective function involves a non-concave maximization. Moreover, the duality gap can still be positive for any finite number of tones, so the optimal dual solution

can only provide an upper bound on the social optimum. In other words, the dual decomposition based algorithms as prescribed in [3], [6] cannot solve the original primal spectrum management problem due to the existence of a positive duality gap[1]. Another class of spectrum management methods (first proposed by Cioffi and Yu [2] and later studied in [8], [9], [10], [11], [12], [13]) are based on game theoretical concepts whereby users maximize their individual rates in a distributed manner using the well-known water-filling strategy. In this framework, the spectrum management problem is viewed as a non-cooperative Nash game in which each player's payoff function is his data rate, and every user greedily updates its power spectrum by treating other users' signals as Gaussian noise. When crosstalk interference is small, the resulting distributed algorithm (called iterative water-filling algorithm, or IWFA herein) is known to generate a sequence of power allocations converging to a unique Nash equilibrium point of the non-cooperative game [12], [13], [11]. When the crosstalk interference is strong, there may be multiple Nash equilibrium points and the convergence of IWFA is unknown, though empirical evidence suggests the algorithm still converges. Despite its distributed nature and simplicity, the sum-rate achieved by IWFA can be far from the social optimum; see the simulation results in Section VI.

How should users in an interference-limited communication system share spectrum in order to achieve the social optimum? The answer depends on the communication environment. Intuitively, if the crosstalk interference is absent or low, then all users should utilize the entire spectrum simultaneously. On the other hand, when the crosstalk interference is significant, the users may be much better off if no spectrum is shared, therefore giving rise to a self-induced (rather than system-enforced) FDMA solution. Our interest in FDMA type solutions is two fold. First, it is of practical interest to characterize how strong the crosstalk interference has to be before FDMA strategy becomes sum-rate optimal. The answer to this question will not only provide valuable insight into the structural property of optimal power allocation strategies, but also help simplify the spectrum management problem since it allows the users to narrow their search to FDMA type solutions only. The latter is a much simpler design problem than the general nonconvex sum-rate maximization problem. Second, there exist simple distributed algorithms that can determine the optimal FDMA type solutions, regardless of their overall optimality for the general sum-rate maximization problem. It turns out that FDMA solutions obtained this way can offer substantially higher sum-rates than the existing IWFA method in certain situations.

The structural property of optimal spectrum sharing strategies has been recently studied in [14] for a frequency flat environment. It was shown that the optimal spectrum sharing strategy is FDMA when the product of normalized crosstalk

coefficients between each pair of users is greater than 1. Moreover, when restricted to the FDMA strategies, the optimal bandwidth allocation can be computed easily using convex optimization. In this paper, we study the same problem, but for a more practical frequency selective environment. As it turns out, frequency selectivity greatly complicates the sum-rate maximization problem — it makes an otherwise computationally easy problem intractable, even if the number of users in the system is only two.

The contribution of this paper is two fold. First, for the two-user case, we show that, if the pairwise products of the normalized crosstalk coefficients at all frequency tones are larger than a certain threshold value $\frac{1}{4}(1 + \frac{1}{C-1})$, with C being the minimum number of tones used by any user), then the optimal spectrum sharing strategy for the maximization of sum-rate is FDMA. If, in addition, each normalized crosstalk coefficient is greater than $1/2$ for all users at each frequency tone, then FDMA remains sum-rate optimal for the case of arbitrary number of users. Second, we restrict ourselves to the FDMA strategies and study, for a given channel condition, how to find an optimal bandwidth allocation which maximizes the sum-rate. We show that this problem is NP-hard and propose several simple algorithms to approximately maximize the sum-rate. Numerical results indicate that these algorithms are efficient and can generate higher quality solutions than IWFA when the crosstalk coefficients are sufficiently large.

This paper is organized as follows. In Section II, we describe the system model and give some mathematical preliminaries on the first and second order necessary conditions for local optimality. In Section III, we derive a sufficient condition under which the global optimum of the sum-rate maximization problem possesses the FDMA structure. Our proof is based on an analysis of quasi-convexity using the gradient vectors and Hessian matrices of sum-rate function at each frequency tone. In Section IV, we further provide a sufficient condition for the existence of a local maxima of the sum-rate function (subject to individual power constraints) that has the FDMA structure. In Section V, we establish the NP-hardness of the sum-rate maximization problem and propose a simple distributed algorithm and two polynomial time combinatorial search algorithms for finding a FDMA solution with maximal sum-rate. Numerical results are reported in Section VI, and the concluding remarks are given in Section VII.

Throughout the paper, we use the following notations. We denote the set of frequency tones and users by $\mathcal{N}$ and $\mathcal{K}$, respectively, i.e., $\mathcal{N} := \{1, \ldots, N\}$ and $\mathcal{K} := \{1, \ldots, K\}$. Also, we use superscript n to denote the frequency tone index and subscript k to denote the user index.

II. PRELIMINARIES

We first describe the frequency selective Gaussian interference channel model and the mathematical formulation of the sum-rate maximization problem. Then we derive the first and second order conditions for sum-rate optimality.

A. Channel model and the sum-rate maximization problem

Suppose there are K users sharing a common spectrum which is divided into N frequency tones numbered by

---

[1]It is worth noting that the authors of [6] claimed that the duality gap is zero if a certain "time-sharing" property holds. They also used an intuitive (but non-rigorous) argument to show that the time-sharing property holds approximately when the tone spacing is narrow. A rigorous treatment of the issue is needed to substantiate this claim and to estimate the size of duality gap, see [7].

$\{1, 2, ..., N\}$. For notational simplicity, we assume that each user acts both as a transmitter and as a receiver[2], and we number the transmitters and receivers by the same index set $\{1, 2, ..., K\}$. In this way, a physical user may act as transmitter k and receiver l, with $l \neq k$. Let x_k^n denote the transmitted complex Gaussian signal from transmitter k at tone n, and let $S_k^n := E|x_k^n|^2$ denote its power. The received signal y_k^n is given by

$$y_k^n = \sum_{l=1}^{K} h_{lk}^n x_l^n + z_k^n, \quad n \in \mathcal{N}, \ k \in \mathcal{K},$$

where $z_k^n \sim CN(0, N_0)$ denotes the complex Gaussian channel noise with zero mean and variance N_0, and the complex scalars $\{h_{lk}^n\}$ represent channel gain coefficients. In practice, h_{lk}^n can be determined by the distance between transmitter l and receiver k. Assuming that the interference is treated as white noise, we can write transmitter k's achievable data rate R_k^n at tone n [15] as

$$R_k^n(S_1^n, \ldots, S_K^n) = \log\left(1 + \frac{|h_{kk}^n|^2 S_k^n}{N_0 + \sum_{l \neq k} |h_{lk}^n|^2 S_l^n}\right),$$

Upon normalizing the channel coefficients, we obtain

$$R_k^n(S_1^n, \ldots, S_K^n) := \log\left(1 + \frac{S_k^n}{\sigma_k^n + \sum_{l \neq k} \alpha_{lk}^n S_l^n}\right), \quad (1)$$

where $\sigma_k^n = N_0/|h_{kk}^n|^2$ denotes the normalized background noise power, and $\alpha_{lk}^n = |h_{lk}^n|^2/|h_{kk}^n|^2$ is the normalized crosstalk coefficient from transmitter l to receiver k at tone n. Due to normalization, we have $\alpha_{kk}^n = 1$ for all k.

Notice that unlike the frequency flat case considered in [14], the channel coefficients h_{lk}^n vary according to tone index n due to frequency selectivity, resulting in a non-constant normalized noise power σ_k^n across tones. As it turns out, this crucial difference greatly complicates the sum-rate maximization problem in the frequency selective case, making the computation of optimal power allocations computationally intractable; see Section V.

Throughout, we assume that transmitter k's power is bounded by $P_k > 0$, i.e.,

$$\sum_{n=1}^{N} S_k^n \leq P_k, \quad \text{for } k \in \mathcal{K}.$$

For a given power allocation $\{S_k^n\}$, transmitter k's total achievable data rate is given by $\sum_{n=1}^{N} R_k^n$ and the total sum-rate is given by $\sum_{k=1}^{K} \sum_{n=1}^{N} R_k^n$. Hence, the sum-rate maximization problem can be written as follows:

$$\underset{\{S_1^n, \ldots, S_K^n\}_{n=1}^N}{\text{maximize}} \sum_{k=1}^{K} \sum_{n=1}^{N} \log\left(1 + \frac{S_k^n}{\sigma_k^n + \sum_{l \neq k} \alpha_{lk}^n S_l^n}\right)$$

$$\text{subject to } \sum_{n=1}^{N} S_k^n \leq P_k, \ S_k^n \geq 0, \ \forall n \in \mathcal{N}, \forall k \in \mathcal{K}. \ (2)$$

[2] There is no loss of generality with this assumption since we can always create a virtual channel with zero channel gain coefficients between pair of users who do not wish to communicate.

It can be easily seen that user k's total achievable data rate $\sum_{n=1}^{N} R_k^n$ is concave for user k's power vector $(S_k^1, \ldots, S_k^N)$ when other users' power vectors are fixed. However, the total sum-rate function $\sum_{k=1}^{K} \sum_{n=1}^{N} R_k^n$ is in general non-concave even if other users' powers are fixed, since user k's power S_k^n appears in the denominators of other users' data rate function.

When interference is absent (or small), it can be easily checked [14] that signal spreading across spectrum is optimal. In other words, if the crosstalk coefficients are sufficiently small, then all frequency tones should be utilized by all users. On the other hand, if the crosstalk coefficients are large, then the communication system becomes interference limited, and spectrum sharing is no longer optimal. Intuitively, FDMA should yield a larger sum-rate in this case. Mathematically, FDMA property is defined as follows:

Definition 2.1: A feasible solution $\{S_1^n, \ldots, S_K^n\}_{n=1}^N$ of the sum-rate maximization problem (2) is said to have FDMA property, if the following implication holds for all $(n, k) \in \mathcal{N} \times \mathcal{K}$:

$$S_k^n > 0 \implies S_l^n = 0, \qquad \forall l \neq k.$$

B. First and second order optimality conditions

In this subsection, we give some necessary or sufficient conditions of local optimality for (4). Since those conditions can be derived directly from standard optimization theory [16], [17], we simply state the results without proofs.

To simplify our notations, we let $\mathbf{S}^n$, $\mathbf{S}_k$, and $\mathbf{S}$ denote the power vectors at tone n, for user k, and in the whole system, respectively, i.e.,

$$\mathbf{S}^n := (S_1^n, \ldots, S_K^n) \in \Re^K,$$
$$\mathbf{S}_k := (S_k^1, \ldots, S_k^n) \in \Re^N,$$
$$\mathbf{S} := (S_1^1, \ldots, S_K^N) \in \Re^{NK}.$$

We denote the power budget vector by $\mathbf{P}$, i.e., $\mathbf{P} := (P_1, \ldots, P_K) \in \Re^K$. Also, we denote the noise plus interference power for user k at tone n, and the sum of all users' data rates at tone n by X_k^n and f^n, respectively, i.e.,

$$X_k^n(\mathbf{S}^n) := \sigma_k^n + \sum_{l \neq k} \alpha_{lk}^n S_l^n,$$

$$f^n(\mathbf{S}^n) := \sum_{k=1}^{K} R_k^n(\mathbf{S}^n) = \sum_{k=1}^{K} \log\left(1 + \frac{S_k^n}{X_k^n}\right). \quad (3)$$

Note that X_k^n and f^n depend on $\mathbf{S}^n$ only. We adopt the following short notations for the first and second derivatives of f^n:

$$\partial_k f^n(\mathbf{S}^n) := \frac{\partial}{\partial S_k^n} f^n(\mathbf{S}^n), \ \partial_{kl} f^n(\mathbf{S}^n) := \frac{\partial^2}{\partial S_k^n \partial S_l^n} f^n(\mathbf{S}^n).$$

With these notations, we can rewrite the sum-rate maximization problem (2) as follows:

$$\underset{\mathbf{S}^1, \ldots, \mathbf{S}^N}{\text{maximize}} \sum_{n=1}^{N} f^n(\mathbf{S}^n) \quad (4)$$

$$\text{subject to } \sum_{n=1}^{N} \mathbf{S}^n \leq \mathbf{P}, \ \mathbf{S}^n \geq 0, \quad \forall n \in \mathcal{N},$$

where the vector inequalities are to be interpreted componentwise. Finally, we define the following index sets which will be convenient for describing the FDMA property of a feasible power vector.

Definition 2.2: For a feasible solution $\mathbf{S}$ of problem (4), we define the following sets.

$$\begin{aligned}
\mathcal{T}(\mathbf{S}) &:= \{(n,k) \mid S_k^n > 0\} \subseteq \mathcal{N} \times \mathcal{K}, \\
\mathcal{T}_k(\mathbf{S}_k) &:= \{n \mid S_k^n > 0\} \subseteq \mathcal{N}, \\
\mathcal{T}^n(\mathbf{S}^n) &:= \{k \mid S_k^n > 0\} \subseteq \mathcal{K}.
\end{aligned}$$

Note that $\mathcal{T}_k(\mathbf{S}_k)$ denotes the set of all tones used by user k, and $\mathcal{T}^n(\mathbf{S}^n)$ denotes the set of all users using tone n.

Definition 2.3: For a feasible solution S of problem (4), we define the tone sets by

$$\begin{aligned}
\mathcal{N}_O(\mathbf{S}) &:= \Big\{ n \,\Big|\, |\mathcal{T}^n(\mathbf{S}^n)| = 0 \Big\}, \\
\mathcal{N}_I(\mathbf{S}) &:= \Big\{ n \,\Big|\, |\mathcal{T}^n(\mathbf{S}^n)| = 1 \Big\}, \\
\mathcal{N}_{II}(\mathbf{S}) &:= \Big\{ n \,\Big|\, |\mathcal{T}^n(\mathbf{S}^n)| \geq 2 \Big\},
\end{aligned}$$

where $|\mathcal{T}^n(\mathbf{S}^n)|$ denotes the number of elements belonging to $\mathcal{T}^n(\mathbf{S}^n)$.

Since $|\mathcal{T}^n(\mathbf{S}^n)|$ implies the number of users whose transmit power is allocated to tone n, power vector $\mathbf{S}$ has FDMA property if and only if $\mathcal{N}_{II} = \emptyset$. In the ensuing discussions, we often omit the argument $\mathbf{S}$ when it is obvious from the context.

The first order necessary conditions, which are also called the Karush-Kuhn-Tucker (KKT) conditions, for problem (4) is given as follows. One can easily see that the linear independence constraint qualification always holds since $P_k > 0$ for all $k \in \mathcal{K}$.

Proposition 2.1 (Karush-Kuhn-Tucker conditions): Let $\mathbf{S}$ be a local optimum for problem (4). Then there exist nonnegative reals $\lambda_1, \ldots, \lambda_K$ such that

$$\begin{aligned}
&\lambda_k \geq 0, \ P_k - \textstyle\sum_{n=1}^N S_k^n \geq 0, \ \lambda_k\Big(P_k - \textstyle\sum_{n=1}^N S_k^n\Big) = 0, \\
&S_k^n \geq 0, \ \lambda_k - \partial_k f^n(\mathbf{S}^n) \geq 0, \ S_k^n\Big(\lambda_k - \partial_k f^n(\mathbf{S}^n)\Big) = 0,
\end{aligned} \quad (5)$$

for all $k \in \mathcal{K}$ and $n \in \mathcal{N}$.

Note that (5) requires either $\lambda_k = 0$ or $P_k - \sum_{n=1}^N S_k^n = 0$, and $S_k^n = 0$ or $\lambda_k - \partial_k f^n(\mathbf{S}^n) = 0$. Moreover, since λ_k does not depend on tones, (5) also implies

$$n, m \in \mathcal{T}_k \implies 0 \leq \lambda_k = \partial_k f^n(\mathbf{S}^n) = \partial_k f^m(\mathbf{S}^m). \quad (6)$$

Since the objective function of problem (4) is not concave, some KKT points can be local minima or saddle points. In order to distinguish these non-optimal KKT points from local maxima, we consider the following second order necessary optimality conditions.

Proposition 2.2 (Second order necessary conditions): Suppose that $\mathbf{S}$ satisfies the KKT conditions (5), and $P_k - \sum_{n=1}^N S_k^n = 0$ for all $k \in \mathcal{K}$. If $\mathbf{S}$ is a local maximum of problem (4), then the following conditions hold: for any vector $v = (v_1^n, \ldots, v_K^n) \in \Re^{NK}$ such that

$$v_k^n = 0 \ \ \forall n \notin \mathcal{T}_k, \ \text{ and } \sum_{n \in \mathcal{T}_k} v_k^n = 0, \quad \forall k \in \mathcal{K}, \quad (7)$$

there holds

$$\sum_{n=1}^N (v^n)^T \nabla^2 f^n(\mathbf{S}^n) v^n \leq 0, \quad (8)$$

where $v^n := (v_1^n, \ldots, v_K^n) \in \Re^K$.

The second order sufficient conditions for local optimality can be described as follows.

Proposition 2.3 (Second order sufficient conditions): Suppose that $\mathbf{S}$ satisfies the KKT conditions (5), and $\lambda_k > 0$ for all $k \in \mathcal{K}$. Suppose that

$$\sum_{n=1}^N (v^n)^T \nabla^2 f^n(\mathbf{S}^n) v^n < 0 \quad (9)$$

for any vectors $v^1, \ldots, v^N \in \Re^K$ such that $v = (v^1, \ldots, v^N) \neq 0$,

$$v_k^n = 0 \ \ \forall n \notin \mathcal{T}_k, \ \text{ and } \sum_{n \in \mathcal{T}_k} v_k^n = 0, \quad \forall k \in \mathcal{K}. \quad (10)$$

Then, the power vector $\mathbf{S}$ is a local maximum of problem (4).

These optimality conditions will be used in the subsequent sections to show the optimality of FDMA-type solutions under various channel conditions.

III. SUM-RATE OPTIMALITY OF FDMA

As mentioned earlier, we expect that an FDMA-type power allocation will maximize the sum-rate when the crosstalk coefficients are sufficiently large. In this section we show the validity of this claim and derive an explicit bound on the crosstalk coefficients which will ensure the existence of an optimal FDMA type solution. We will first consider the general K-user case, and then strengthen the result in the two-user case by exploiting the quasi-convexity of the sum-rate function (4).

Notice that the first and second order derivatives of function f^n (defined by (3)) can be computed explicitly as follows:

Proposition 3.1: Denote

$$\begin{aligned}
A_k^n &:= \frac{1}{X_k^n}, \quad B_k^n := \frac{1}{X_k^n + S_k^n}, \\
P_k^n &:= A_k^n - B_k^n, \quad Q_k^n := (A_k^n)^2 - (B_k^n)^2
\end{aligned}$$

with X_k^n defined by (3). Then, for every $(n, k, l) \in \mathcal{N} \times \mathcal{K} \times \mathcal{K}$ such that $k \neq l$, we have

$$\partial_k f^n(\mathbf{S}^n) = A_k^n - \sum_{r=1}^K \alpha_{kr}^n P_r^n,$$

$$\partial_{kk} f^n(\mathbf{S}^n) = -(A_k^n)^2 + \sum_{r=1}^K (\alpha_{kr}^n)^2 Q_r^n,$$

$$\partial_{kl} f^n(\mathbf{S}^n) = -\big\{ \alpha_{lk}^n (A_k^n)^2 + \alpha_{kl}^n (A_l^n)^2 \big\} + \sum_{r=1}^K \alpha_{kr}^n \alpha_{lr}^n Q_r^n.$$

We note that the summations in the above equalities contain the case where $r = k$ and $r = l$ with $\alpha_{kk}^n = \alpha_{ll}^n = 1$. It is easy to see that $A_k^n \geq B_k^n > 0$, $P_k^n \geq 0$ and $Q_k^n \geq 0$, with $P_k^n = Q_k^n = 0$ if and only if $S_k^n = 0$.

A. Analysis of the general K-user case

Let us first introduce a condition on a feasible power allocation vector $\mathbf{S}$.

Condition 1: There holds

(a) $\min_{k\in\mathcal{K}} |\mathcal{T}_k(\mathbf{S}_k)| \geq C$, for some integer $C \geq 2$,

(b) $\sum_{n=1}^{N} \mathbf{S}^n = \mathbf{P}$.

In other words, every user uses at least two tones and exhausts his power budget. Moreover, we assume that the global maximum satisfies this condition.

Assumption A: Any global maximum of problem (4) satisfies Condition 1 for some $C \geq 2$.

Assumption A is difficult to verify since the global maximum is not known a priori. However, in a practical DSL system, the number of tones N is usually much larger than the number of users K, i.e., $K \ll N$, and the power budget for each user is sufficiently high. The following proposition shows that, when all the crosstalk coefficients are greater than or equal to $1/2$, the lower bound C of $\min_{k\in\mathcal{K}} |\mathcal{T}_k(\mathbf{S}_k)|$ can be evaluated by using the constants in the problem only.

Proposition 3.2: Suppose $\alpha_{lk}^n \geq 1/2$ for all $l, k \in \mathcal{K}$ and $n \in \mathcal{N}$. Let $C \in [2, N]$ be an arbitrary integer. If there exists another integer $m \in [C, N]$ such that

$$1 + \frac{\rho_0}{m} > \left(1 + \frac{\rho_M}{C-1}\right)^{\frac{C-1}{m}}\left(1 + \frac{K\rho_a}{N-m+1}\right), \quad (11)$$

where

$$\rho_0 := \min_{(n,k)\in\mathcal{N}\times\mathcal{K}} \frac{P_k}{\sigma_k^n}, \quad \rho_M := \max_{(n,k)\in\mathcal{N}\times\mathcal{K}} \frac{P_k}{\sigma_k^n},$$

$$\rho_a := \frac{\frac{1}{K}\sum_{k=1}^{K} P_k}{\min_{(n,k)\in\mathcal{N}\times\mathcal{K}} \sigma_k^n},$$

then $\min_{k\in\mathcal{K}} |\mathcal{T}_k(\mathbf{S}_k)| \geq C$ for any global maximizer $\mathbf{S}$ of the sum-rate maximization problem (4).

Proof: See Appendix A. ∎

In Proposition 3.2, the parameters ρ_0, ρ_M and ρ_a represent the minimum, maximum and average signal to noise ratios across all tones and among all users. In the case where $\rho_0 = \rho_M = \rho_a = \rho$ and $C = m = 2$, then the above sufficient condition (11) for $\min_{k\in\mathcal{K}} |\mathcal{T}_k(\mathbf{S}_k)| \geq 2$ simplifies to

$$1 + \frac{\rho}{2} > \left(1 + \rho\right)^{1/2}\left(1 + \frac{K\rho}{N-1}\right),$$

which can be satisfied with large ρ and small K/N.

In case of $K = 2$, the condition $\alpha_{lk}^n \geq 1/2$ in Proposition 3.2 can be relaxed to a weaker condition $\alpha_{12}^n \alpha_{21}^n > 1/4$; see the end of Section III for details. While the lower bound provided by Proposition 3.2 may not be tight, it is satisfying that such a bound for $\min_{k\in\mathcal{K}} |\mathcal{T}_k(\mathbf{S}_k)|$ can be obtained even if the global maximum $\mathbf{S}$ is not known.

We now show that, if Assumption A holds and the normalized crosstalk coefficients are sufficiently greater than $1/2$ (in the sense of (12)), then optimal spectrum sharing strategy must be FDMA. Our proofs are based on the local optimality conditions. Specifically, we will show that a feasible power allocation vector $\mathbf{S}$ cannot satisfy the second order necessary conditions for local optimality if (1) $\mathbf{S}$ is not FDMA, (2) crosstalk coefficients are larger than certain values, and (3) $\mathbf{S}$ satisfies Condition 1. This implies that every local maximum $\mathbf{S}$ of (4) satisfying (2) and (3) must be FDMA.

We first give the following proposition which characterizes a non-FDMA power allocation vector $\mathbf{S}$.

Proposition 3.3: Let $\mathbf{S} \in \Re^{NK}$ be a non-FDMA power vector (i.e., $\mathcal{N}_{II}(\mathbf{S}) \neq \emptyset$) satisfying Condition 1(a). Then, there exist an integer $M \geq 2$, a tone set $\{n_1, \ldots, n_M\} \subseteq \mathcal{N}$, and a user set $\{k_1, \ldots, k_M\} \subseteq \mathcal{K}$ such that $n_i \neq n_j$ and $k_i \neq k_j$ for any $i \neq j$, and either of the following two conditions holds:

(i) $(n_i, k_i) \in \mathcal{T}$ and $(n_i, k_{i+1}) \in \mathcal{T}$ for $i = 1, \ldots, M-1$. Moreover, $(n_M, k_M) \in \mathcal{T}$ and $(n_M, k_1) \in \mathcal{T}$.

(ii) $(n_i, k_i) \in \mathcal{T}$ and $(n_i, k_{i+1}) \in \mathcal{T}$ for $i = 1, \ldots, M-1$. Moreover, $\mathcal{N}_{II} \cap \mathcal{T}_{k_1} = \{n_1\}$ and $\mathcal{N}_{II} \cap \mathcal{T}_{k_M} = \{n_{M-1}\}$.

Proof: See Appendix B. ∎

By using Proposition 3.3, we establish sufficient conditions on crosstalk coefficients under which non-FDMA feasible solutions cannot satisfy the second order necessary conditions. We do so by considering the condition (i) and the condition (ii) of Proposition 3.3 separately.

Proposition 3.4: Suppose that

$$\alpha_{lk}^n > \frac{1}{2}$$

for all $n \in \mathcal{N}$ and $(k, l) \in \mathcal{K} \times \mathcal{K}$. Let $\mathbf{S} \in \Re^{NK}$ be an arbitrary vector satisfying Condition 1 and the condition (i) of Proposition 3.3 for some $\{n_i\}_{i=1}^{M} \subseteq \mathcal{N}$ and $\{k_i\}_{i=1}^{M} \subseteq \mathcal{K}$. Then, $\mathbf{S}$ cannot be a local maximum of the sum-rate maximization problem (4).

Proof: For any $\mathbf{S}$ satisfying the given assumptions, we can verify that the necessary condition for optimality cannot be met. So $\mathbf{S}$ cannot be a local maximum. The details of the proof are relegated to Appendix C. ∎

Now we consider the case (ii) of Proposition 3.3.

Proposition 3.5: Suppose that

$$\alpha_{lk}^n \alpha_{kl}^n > \frac{1}{4}\left(1 + \frac{1}{C-1}\right)^2 \quad (12)$$

for all $n \in \mathcal{N}$ and $(k, l) \in \mathcal{K} \times \mathcal{K}$ with $k \neq l$. Let $\mathbf{S} \in \Re^{NK}$ be an arbitrary vector satisfying Condition 1 and the condition (ii) of Proposition 3.3 for some $\{n_i\}_{i=1}^{M} \subseteq \mathcal{N}$ and $\{k_i\}_{i=1}^{M} \subseteq \mathcal{K}$. Then $\mathbf{S}$ cannot be a local maximum of the sum-rate maximization problem (4).

Proof: For any $\mathbf{S}$ satisfying the given assumptions, we can use the first and second order derivatives formula to verify that the second order necessary optimality condition must be violated. This shows that $\mathbf{S}$ cannot be a local maximum of the sum-rate maximization problem (4). See Appendix D for details. ∎

Combining the above two propositions, we obtain the following sufficient conditions under which every global maximum of sum-rate maximization problem is FDMA.

Theorem 3.1: Suppose that Assumption A holds. Then, any global maximum of problem (4) must be FDMA, provided that

$$\alpha_{lk}^n > \frac{1}{2} \quad \text{and} \quad \alpha_{lk}^n \alpha_{kl}^n > \frac{1}{4}\left(1 + \frac{1}{C-1}\right)^2$$

for all $n \in \mathcal{N}$ and $(k,l) \in \mathcal{K} \times \mathcal{K}$ with $k \neq l$.

Proof: Consider a global optimum $\mathbf{S} \in \Re^{NK}$ satisfying Condition 1 for some $C \geq 2$. If $\mathbf{S}$ is not FDMA, then there exist a tone set $\{n_i\}_{i=1}^M$ and a user set $\{k_i\}_{i=1}^M$ such that either (i) or (ii) in Proposition 3.3 holds. However, by virtue of Propositions 3.4 and 3.5, $\mathbf{S}$ cannot be a local maximum in either case. Hence, $\mathbf{S}$ must be FDMA. ∎

When C is sufficiently large, say, $C > 100$, we have $1 + \frac{1}{C-1} \approx 1$. In this case, the condition $\alpha_{lk}^n \alpha_{kl}^n > \frac{1}{4}(1 + \frac{1}{C-1})^2$ is essentially implied by the condition $\alpha_{lk}^n > 1/2$. Thus, Theorem 3.1 shows that if the normalized crosstalk coefficients are sufficiently greater than $1/2$ (in the sense of (12)), then the optimal spectrum sharing strategy must be FDMA.

B. Improved analysis for the two-user case

In this subsection, we restrict ourselves on the two-user case ($K = 2$) and show the optimality of FDMA strategy under a weaker condition than that of Theorem 3.1. Specifically, we show that the condition $\min\{\alpha_{12}^n, \alpha_{21}^n\} > 1/2$ can be dropped when $K = 2$, and the optimality of FDMA strategies is ensured under the condition $\alpha_{21}^n \alpha_{12}^n > \frac{1}{4}(1 + \frac{1}{C-1})^2$ alone. Our analysis exploits heavily the quasi-convexity property for the sum of data rate in each tone.

Since $K = 2$, function f^n can be rewritten as

$$f^n(S_1^n, S_2^n) = \log\left(1 + \frac{S_1^n}{X_1^n}\right) + \log\left(1 + \frac{S_2^n}{X_2^n}\right), \quad (13)$$

where $X_1^n = \sigma_1^n + \alpha_{21}^n S_2^n$ and $X_2^n = \sigma_2^n + \alpha_{12}^n S_1^n$. We first define the concept of quasi-convexity[3] for twice differentiable functions.

Definition 3.1: Let $g : \Omega \subseteq \Re^n \to \Re$ be a twice differentiable function. Then, g is quasi-convex on Ω if

$$v^T \nabla^2 g(x) v > 0$$

for any $x \in \Omega$ and $v \in \Re^n$ such that $\nabla g(x)^T v = 0$ and $v \neq 0$.

The following theorem gives a sufficient condition under which f^n is quasi-convex.

Theorem 3.2: If

$$\alpha_{12}^n\left(\frac{\sigma_1^n}{\sigma_2^n}\right) + \alpha_{21}^n\left(\frac{\sigma_2^n}{\sigma_1^n}\right) > 1, \quad (14)$$

then the rate function f^n defined by (13) is quasi-convex on $[0, +\infty)^2$.

Proof: The proof involves some algebraic computation verifying that the quasi-convexity condition in Definition 3.1 is satisfied. The critical step in the proof is the factorization of a certain cubic polynomial function as the product of several simpler terms whose signs can be readily verified under the given assumption. See Appendix E for details. ∎

---

[3]Our definition is stronger than the conventional notion of quasi-convexity which is defined by the convexity of level sets; see [18].

It will be interesting to see to what extent the above result (and proof) can be extended to the more general case $K \geq 3$. We expect that f^n will still be quasi-convex for $K \geq 3$ under an appropriate condition on the crosstalk coefficients α_{lk}^n and σ_k^n. However, the proof is likely to be much more complicated than the two-user case, since we do not expect the factorization trick can be applied to the $K \geq 3$ case.

The following corollary says that the assumption in Theorem 3.2 holds regardless of the background noise parameters if the crosstalk coefficients are sufficiently large.

Corollary 3.1: Let f^n be the rate function for tone n (defined by (13)). If

$$\alpha_{12}^n \alpha_{21}^n > \frac{1}{4},$$

then f^n is quasi-convex on $[0, +\infty)^2$.

Proof: We have

$$\alpha_{12}^n\left(\frac{\sigma_1^n}{\sigma_2^n}\right) + \alpha_{21}^n\left(\frac{\sigma_2^n}{\sigma_1^n}\right) \geq 2\sqrt{\alpha_{12}^n \alpha_{21}^n} > 1,$$

where the first inequality follows from the simple relation between the arithmetic and geometric means, and the last inequality is due to $\alpha_{12}^n \alpha_{21}^n > 1/4$. By Theorem 3.2, f^n is quasi-convex on $[0, +\infty)^2$. ∎

So far we have shown that for each tone n the rate function f^n is quasi-convex when the product of normalized crosstalk coefficients is greater than $1/4$. Notice that the quasi-convexity of f^n does not imply the quasi-convexity of the total sum-rate function $\sum_{n=1}^N f^n(\mathbf{S}^n)$, since the sum of two quasi-convex functions is not quasi-convex in general.

Now we proceed to establish a sufficient condition for the optimality of FDMA. Recall that $\mathcal{N}_{II}(\mathbf{S})$ denotes the set of frequency tones shared by at least two users, so $\mathbf{S}$ is FDMA if and only if $|\mathcal{N}_{II}(\mathbf{S})| = 0$. Hence, to establish the optimality of FDMA, we only need to show that any feasible solution $\mathbf{S}$ with $|\mathcal{N}_{II}(\mathbf{S})| \geq 1$ cannot be a local optimum. We will consider the two separate cases $|\mathcal{N}_{II}(\mathbf{S})| \geq 2$ and $|\mathcal{N}_{II}(\mathbf{S})| = 1$ which correspond precisely to the two cases (i) and (ii) in Proposition 3.3. Notice that the case $|\mathcal{N}_{II}(\mathbf{S})| = 1$ is already covered by Proposition 3.5 which shows that $\mathbf{S}$ cannot be a local maximum of (4) if the condition $\alpha_{12}^n \alpha_{21}^n > \frac{1}{4}(1 + \frac{1}{C-1})^2$ is satisfied. Let us now consider the remaining case $|\mathcal{N}_{II}(\mathbf{S})| \geq 2$.

Proposition 3.6: Suppose that $\mathbf{S} \in \Re^{2N}$ is a feasible solution of sum-rate maximization problem (4) such that $|\mathcal{N}_{II}(\mathbf{S})| \geq 2$. If f^n is quasi-convex for all tones $n \in \mathcal{N}$, then $\mathbf{S}$ cannot be a local maximum of the sum-rate maximization problem (4).

Proof: The proof consists of finding a feasible perturbation direction along which the second order necessary optimality condition (Proposition 2.2) is violated due to the quasi-convexity assumption. The details of the proof are relegated to Appendix F. ∎

We now state our main result for the two-user case.

Theorem 3.3: Suppose that $K = 2$ and Assumption A holds for some $C \geq 2$. If

$$\alpha_{12}^n \alpha_{21}^n > \frac{1}{4}\left(1 + \frac{1}{C-1}\right)^2$$

for all $n \in \mathcal{N}$, then the global maximum of sum-rate maximization problem (4) is FDMA.

Proof: We only need to argue that **S** cannot be a local maximum of (4) when $|\mathcal{N}_{II}(\mathbf{S})| \geq 1$. If $|\mathcal{N}_{II}(\mathbf{S})| = 1$, then Proposition 3.3 (ii) holds for $M = 2$. Therefore, by Proposition 3.5, **S** is not a local maximum. If $|\mathcal{N}_{II}(\mathbf{S})| \geq 2$, then Corollary 3.1 implies that the quasi-convexity condition of Propositions 3.6 is satisfied. Thus, once again **S** cannot be a local maximum. The theorem is proven. ∎

As with Theorem 3.1, the threshold for the crosstalk product in Theorem 3.3 involves the parameter C which represents the minimum number of tones used by any user. In practical situations where each user has a sufficiently high power budget and the number of available tones is large, we expect C to be large. In these cases, the threshold in Theorem 3.3 becomes essentially $1/4$. Interestingly, for the two user case, FDMA becomes sum-rate optimal whenever for each tone n the product of normalized crosstalk coefficients $\alpha_{12}^n \alpha_{21}^n$ is greater than $1/4$, even if $\min_n \{\alpha_{12}^n, \alpha_{21}^n\}$ is small. This is in contrast to the sufficient condition for the $K \geq 3$ case (Theorem 3.1) which, in addition to the condition that the pairwise product $\alpha_{kl}^n \alpha_{lk}^n$ is greater than $1/4$, also requires $\min_{n,l \neq k} \{\alpha_{kl}^n, \alpha_{lk}^n\}$ to be greater than $1/2$. It is not clear if it is possible to remove the condition $\min_n \{\alpha_{kl}^n, \alpha_{lk}^n\} > 1/2$ in Theorem 3.1. We leave this as an open question.

Before closing this section, we strengthen Proposition 3.2 by replacing the condition $\alpha_{lk}^n \geq 1/2$ with the quasi-convexity condition of Theorem 3.2.

Proposition 3.7: Suppose that $K = 2$ and (14) holds (or more strictly $\alpha_{12}^n \alpha_{21}^n > 1/4$) for all $n \in \mathcal{N}$. Let $C \in [2, N]$ be an arbitrary integer. If there exists another integer $m \in [C, N]$ such that (11) holds, then $\min_{k \in \mathcal{K}} |\mathcal{T}_k(\mathbf{S}_k)| \geq C$ for any global maximizer **S** of the sum-rate maximization problem (4).

Proof: The proof is almost identical to Appendix A except for the formulation of maximization problem (28). We need to replace (28) by the following maximization problem:

$$\text{maximize} \quad f^n(s_1^n, s_2^n)$$
$$\text{subject to} \quad s_1^n + s_2^n = U^n, \ s_1^n \geq 0, \ s_2^n \geq 0.$$

Since f^n is quasi-convex, the maximum is achieved at a vertex, i.e., $(0, U^n)$ or $(U^n, 0)$. Other parts of the proof in Appendix A remain unchanged. ∎

IV. EXISTENCE OF A LOCALLY OPTIMAL FDMA SOLUTION

In the previous section, we have shown that the optimal solution of sum-rate maximization problem is FDMA if the normalized crosstalk coefficients are sufficiently greater than $1/2$ (in the sense of (12)). This is a sufficient condition for the existence of a FDMA optimal solution. Are there other more practical situations (i.e., with weaker conditions on the crosstalk coefficients) in which FDMA strategies are optimal? Our extensive computational experiments suggest that when each user's power budget is sufficiently large, FDMA will become optimal even if the crosstalk coefficients do not satisfy the conditions in Theorems 3.1 or 3.3. The goal of this section is to derive some weaker sufficient conditions which will guarantee the existence of a FDMA type *local* maxima. Although these conditions do not guarantee the global optimum to be FDMA, the numerical results in Section VI show that, under such conditions, FDMA type power allocations often show better performance than the solutions obtained by the iterative water-filling algorithm (IWFA).

The first step of our analysis is to identify candidates for the FDMA type local maximums of the sum-rate maximization problem (4). We do so by considering the sum-rate maximization problem under the additional constraint that each user in the communication system is pre-assigned a set of fixed and mutually non-overlapping frequency tones. That is, we consider the sum-rate maximization problem under the FDMA constraint. Under this constraint, multiuser interference is no longer present and the sum-rate maximization problem decouples into a set of independent rate maximization subproblems, one per each user, whose global maximums are given by the well-known water-filling solutions. We will show later that the resulting FDMA type solutions can be local maximums for the original sum-rate maximization problem (4) when each user's power budget is sufficiently large.

Let us define the set of FDMA type frequency allocations by

$$\mathcal{FDM} := \Big\{ \mathcal{L} \ \Big| \ \min_{k \in \mathcal{K}} |\mathcal{L}_k| \geq 1, \ \cup_{k=1}^K \mathcal{L}_k = \mathcal{N},$$
$$\text{and} \ \mathcal{L}_k \cap \mathcal{L}_l = \emptyset \ (\forall k \neq l) \Big\}.$$

Here $\mathcal{L}_k$ represents the set of frequency tones allocated to user k. For any $\mathcal{L} \in \mathcal{FDM}$, we consider the following $\mathcal{L}$-restricted sum-rate maximization problem (denoted by $\mathbf{SRMP}(\mathcal{L})$):

$$\mathbf{SRMP}(\mathcal{L}) \quad
\begin{aligned}
& \underset{\{S_k^1, \ldots, S_k^N\}_{k=1}^K}{\text{maximize}} && \sum_{k=1}^K \sum_{n \in \mathcal{L}_k} \log\left(1 + \frac{S_k^n}{\sigma_k^n}\right) \\
& \text{subject to} && \sum_{n=1}^N \mathbf{S}^n \leq \mathbf{P}, \\
& && S_k^n \geq 0 \ (n \in \mathcal{L}_k), \\
& && S_k^n = 0 \ (n \notin \mathcal{L}_k), \quad (k \in \mathcal{K})
\end{aligned}$$

where the objective function is equal to the sum-rate function $\sum_{n=1}^N f^n(\mathbf{S}^n)$ since FDMA requirement implies that there is no interference among users. Notice that $\mathbf{SRMP}(\mathcal{L})$ is a concave maximization, and does not involve any crosstalk coefficient α_{lk}^n. Moreover, $\mathbf{SRMP}(\mathcal{L})$ is completely separable with respect to each user k, implying that $\mathbf{SRMP}(\mathcal{L})$ can be decomposed into the following K independent rate maximization problems:

$$\mathbf{RMP}(\mathcal{L}_k) \quad
\begin{aligned}
& \underset{\{S_k^1, \ldots, S_k^N\}}{\text{maximize}} && \sum_{n \in \mathcal{L}_k} \log\left(1 + \frac{S_k^n}{\sigma_k^n}\right) \\
& \text{subject to} && \sum_{n \in \mathcal{L}_k} S_k^n \leq P_k \\
& && S_k^n \geq 0 \ (n \in \mathcal{L}_k), \\
& && S_k^n = 0 \ (n \notin \mathcal{L}_k).
\end{aligned}$$

It is known that each user's rate maximization problem $\mathbf{RMP}(\mathcal{L}_k)$ can be solved by the water-filling procedure [8],

[9], [10], [11], [2]. To focus our analysis on the interference in the system, we make the following high signal to noise ratio assumption.

Assumption B: For all $k \in \mathcal{K}$, there holds

$$\gamma_k := \frac{P_k + \sum_{n \in \mathcal{L}_k} \sigma_k^n}{|\mathcal{L}_k|} > \max_{n \in \mathcal{L}_k} \sigma_k^n. \quad (15)$$

Under Assumption B, the water level (see [8], [9], [10], [11], [2]) is equal to γ_k, and the global maximum of $\mathbf{SRMP}(\mathcal{L})$ can be described explicitly.

$$\begin{cases} S_k^n = \gamma_k - \sigma_k^n & (\forall n \in \mathcal{L}_k) \\ S_k^n = 0 & (\forall n \notin \mathcal{L}_k) \end{cases} \quad (16)$$

for $k = 1, \ldots, K$. Moreover, $\mathcal{T}_k(S_k) = \mathcal{L}_k$ for $k = 1, \ldots, K$, and the set $\mathcal{N}_O(S)$ defined by (2.3) is empty.

We now study the conditions under which the global optimum (16) of $\mathbf{SRMP}(\mathcal{L})$ is a local optimum of the original sum-rate maximization problem (4). Needless to say, such conditions guarantee the existence of a local optimum of sum-rate maximization problem.

The following proposition gives sufficient conditions under which problem (4) has a FDMA local maximum S such that $\mathcal{T}_k(S_k) = \mathcal{L}_k$ for $k = 1, \ldots, K$.

Proposition 4.1: Let the tone allocation set be given by $\mathcal{L} \in \mathcal{FDM}$, and γ_k be defined by (15). Suppose that Assumption B holds. If

$$\frac{1}{\sigma_k^n + \alpha_{lk}^n(\gamma_l - \sigma_l^n)} - \alpha_{kl}^n\left(\frac{1}{\sigma_l^n} - \frac{1}{\gamma_l}\right) \leq \frac{1}{\gamma_k} \quad (17)$$

for any $(k, l) \in \mathcal{K} \times \mathcal{K}$ with $k \neq l$ and $n \in \mathcal{L}_l$, then the global maximum (16) of $\mathbf{SRMP}(\mathcal{L})$ is a local maximum of sum-rate maximization problem (4).

Proof: Let $\mathbf{S}$ be the global maximum of $\mathbf{SRMP}(\mathcal{L})$ given by (16). Then, we verify that $\mathbf{S}$ satisfies the KKT conditions and the second order sufficient conditions (Propositions 2.1 and 2.3) under the high SNR assumption (15). We leave the details to Appendix G. ∎

Although the condition (17) can be verified beforehand, it involves all possible combinations for k, l and n, and concerns only a given tone allocation $\mathcal{L}$. This makes it inconvenient to apply Proposition 4.1 in practice. In the following corollary result, we simplify the conditions of Proposition 4.1 so as to improve its applicability in practice.

Theorem 4.1: Let C be an arbitrary integer such that $1 \leq C \leq N/K$, and denote

$$\begin{aligned} P_M &:= \max_{k \in \mathcal{K}} P_k, & P_0 &:= \min_{k \in \mathcal{K}} P_k, \\ \sigma_M &:= \max_{(n,k) \in \mathcal{N} \times \mathcal{K}} \sigma_k^n, & \sigma_0 &:= \min_{(n,k) \in \mathcal{N} \times \mathcal{K}} \sigma_k^n, \\ \alpha_0 &:= \min_{\substack{(n,k,l) \in \mathcal{N} \times \mathcal{K} \times \mathcal{K} \\ k \neq l}} \alpha_{lk}^n, & A_0 &:= \min_{\substack{(n,k,l) \in \mathcal{N} \times \mathcal{K} \times \mathcal{K} \\ k \neq l}} \alpha_{lk}^n \alpha_{kl}^n, \\ \gamma_M &:= \frac{P_M}{C} + \sigma_M, & \gamma_0 &:= \frac{P_0}{N - (K-1)C} + \sigma_0. \end{aligned}$$

Suppose that the following inequalities hold:

$$\gamma_0 > \sigma_M, \quad (18)$$

$$A_0 \gamma_M (\gamma_0 - \sigma_M)^2 + \alpha_0(\gamma_M \sigma_0 + \gamma_0 \sigma_M)(\gamma_0 - \sigma_M)$$
$$\geq \sigma_M \gamma_0 (\gamma_M - \sigma_0). \quad (19)$$

Then, for any tone set $\mathcal{L} \in \mathcal{FDM}$ such that $\min_{k \in \mathcal{K}} |\mathcal{L}_k| \geq C$, the global maximum of $\mathbf{SRMP}(\mathcal{L})$ is a local maximum of sum-rate maximization problem (4). Moreover, if

$$P_0 \geq \left(N - (K-1)C\right)\left(\frac{1}{A_0} + \frac{1}{\sqrt{A_0}} + 1\right)\sigma_M, \quad (20)$$

then (19) holds.

Proof: The proof of this theorem involves verifying that $(18) \Rightarrow (15)$ and $(20) \Rightarrow (19) \Rightarrow (17)$. See Appendix H for details. ∎

Although condition (20) is more restrictive than (19), it is more intuitive and easier to apply in practice. Compared to our earlier results (Theorems 3.1 and 3.3), Theorem 4.1 shows the existence of a FDMA type local maxima for the sum-rate maximization problem (4) even when the crosstalk coefficients are small (but positive), so long as users' power budgets are sufficiently large.

V. FINDING AN OPTIMAL FDMA BANDWIDTH ALLOCATION

In this section, we focus our attention on the more practical issue of how to design an optimal FDMA scheme for a multiuser communication system. The latter entails allocating the available set of frequency tones to the users in the system. Let us denote the set of FDMA solutions by

$$\mathcal{S} = \left\{ \mathbf{S} \geq 0 \,\middle|\, S_k^n S_l^n = 0, \ \forall k \neq l, \forall n \right\},$$

where the condition $S_k^n S_l^n = 0$ signifies that no frequency tone can be shared by any two users. Then, the optimal FDMA frequency allocation problem can be described as follows:

$$\begin{aligned} \underset{\mathbf{S}}{\text{maximize}} \quad & \sum_{k=1}^K \sum_{n=1}^N \log\left(1 + \frac{S_k^n}{\sigma_k^n}\right) \\ \text{subject to} \quad & \mathbf{S} \in \mathcal{S}, \quad \sum_{n=1}^N S_k^n \leq P_k, \quad k \in \mathcal{K}. \quad (21) \end{aligned}$$

where $\mathbf{S}$ denotes the (NK)-dimensional vector with entries equal to S_k^n. Notice that, due to the FDMA condition, the interference term $\sum_{l \neq k} \alpha_{lk}^n S_l^n$ is absent from the sum-rate objective function. This makes the objective function concave. However, problem (21) remains a nonconvex problem due to the nonconvex constraint $\mathbf{S} \in \mathcal{S}$. The following result shows that the optimization problem (21) is NP-hard, even in the case of two users.

Theorem 5.1: For $K = 2$, the optimal bandwidth allocation problem (21) is NP-hard. Thus, the general sum-rate maximization problem (4) is also NP-hard, even in the two-user case.

Proof: The proof consists of reducing the so-called equipartition problem to (21). Specifically, given a set of N (even) positive integers, a_1, a_2, ..., a_N, the equipartition problem asks: does there exist a subset $T \subset \{1, 2, ..., N\}$ of size $|T| = N/2$ such that

$$\sum_{n \in T} a_n = \sum_{n \notin T} a_n = \frac{1}{2}\sum_{n=1}^N a_n.$$

The equipartition problem is known to be NP-complete.

For any instance of the equipartition problem, we can construct a two-user instance of (21) as well as a convex relaxation of this nonconvex problem. We can show that the two problems have the same optimal objective values if and only if the answer to the equipartition problem is 'yes'. The details are given in Appendix I. ∎

Theorem 5.1 suggests that finding either a global optimal FDMA bandwidth allocation, or a global sum-rate optimal power allocation in general is computationally intractable when the number of tones are large, even in the two-user case. Given this negative result, we are naturally led to the problem of designing efficient polynomial time algorithms which can approximately maximize the sum-rates. In what follows, we propose three simple algorithms for computing an approximately optimal FDMA bandwidth allocations. The first one is based on dual decomposition, while the other two are based on the idea of greedy local search.

Dual decomposition method

Define the bounded set $\tilde{\mathcal{S}} \subset \Re^{NK}$ by

$$\tilde{\mathcal{S}} := \mathcal{S} \cap \left\{ \mathbf{S} \, \Big| \, 0 \le S_k^n \le P_k \,\, \forall k, n \right\}.$$

Then, we can easily see that the constraint region of (21) is unchanged if $\mathcal{S}$ is replaced by $\tilde{\mathcal{S}}$. Hence, by using multipliers $\{\lambda_k\}$ to dualize the linear constraints in (21), we obtain the following dual function

$$d(\boldsymbol{\lambda})$$

$$:= \max_{\mathbf{S} \in \tilde{\mathcal{S}}} \left(\sum_{k=1}^K \sum_{n=1}^N \log\left(1 + \frac{S_k^n}{\sigma_k^n}\right) - \sum_{k=1}^K \lambda_k \left(\sum_{n=1}^N S_k^n - P_k\right) \right)$$

$$= \sum_{k=1}^K \lambda_k P_k + \max_{\mathbf{S} \in \tilde{\mathcal{S}}} \sum_{k=1}^K \sum_{n=1}^N \left(\log\left(1 + \frac{S_k^n}{\sigma_k^n}\right) - \lambda_k S_k^n \right)$$

$$= \sum_{k=1}^K \lambda_k P_k + \sum_{n=1}^N \max_{\substack{0 \le S_i^n \le P_i \\ S_i^n S_j^n = 0, \, i \ne j}} \sum_{k=1}^K \left(\log\left(1 + \frac{S_k^n}{\sigma_k^n}\right) - \lambda_k S_k^n \right) \tag{22}$$

where the last step is due to the fact that, without the power constraints, the bandwidth allocation problem decouples across tones. The inner maximization in (22) can be solved by allocating each tone to the user which can provide the maximum *shadow rate* $\log\left(1 + S_k^n/\sigma_k^n\right) - \lambda_k S_k^n$ on that tone. Simple calculation shows that the maximum shadow rate M_k^n for user k at tone n is given by

$$M_k^n(\lambda_k) := \max_{0 \le S_k^n \le P_k} \left(\log\left(1 + \frac{S_k^n}{\sigma_k^n}\right) - \lambda_k S_k^n \right) \tag{23}$$

$$= \log\left(1 + \frac{\overline{S}_k^n}{\sigma_k^n}\right) - \lambda_k \overline{S}_k^n$$

where the optimal power level is

$$\overline{S}_k^n = \begin{cases} \mathcal{P}_k(\lambda_k^{-1} - \sigma_k^n) & \text{if } \lambda_k > 0 \\ P_k & \text{if } \lambda_k \le 0 \end{cases} \tag{24}$$

Here $\mathcal{P}_k(\cdot)$ denotes the projection of a real number to the interval $[0, P_k]$. Thus, the dual function (22) can be written

analytically as

$$d(\boldsymbol{\lambda}) = \sum_{k=1}^K \lambda_k P_k + \sum_{n=1}^N \max_{k=1,\ldots,K} M_k^n(\lambda_k). \tag{25}$$

There may be more than one user attaining this maximum, in which case we simply assign the n-th tone to an arbitrary (but unique) user denoted by $k(n)$. Then a subgradient of $d(\boldsymbol{\lambda})$ is given by

$$\nabla d(\boldsymbol{\lambda}) = \left(P_1 - \sum_{n \in \mathcal{N}_1(\boldsymbol{\lambda})} \overline{S}_1^n, \ldots, \, P_K - \sum_{n \in \mathcal{N}_K(\boldsymbol{\lambda})} \overline{S}_K^n \right)^T$$

where we denote the set of tones assigned to user k by $\mathcal{N}_k(\boldsymbol{\lambda})$. Notice that the components of subgradient $\nabla d(\boldsymbol{\lambda})$ correspond to each user's unused power (or deficit power if negative).

The dual minimization problem is given by

$$\begin{aligned} \text{minimize} \quad & d(\boldsymbol{\lambda}) \\ \text{subject to} \quad & \boldsymbol{\lambda} \ge 0. \end{aligned}$$

The standard dual descent method for this problem can now be stated as follows.

Algorithm 1:

Step 0   Choose an initial point $\boldsymbol{\lambda}^{(0)} \ge 0$ and a stepsize $\alpha^{(0)} > 0$. Set $\nu = 0$.

Step 1   For all $(n, k) \in \mathcal{N} \times \mathcal{K}$, compute

$$(\overline{S}_k^n)^{(\nu)} := \begin{cases} \mathcal{P}_k\left((\lambda_k^{(\nu)})^{-1} - \sigma_k^n\right) & \text{if } \lambda_k^{(\nu)} > 0 \\ P_k & \text{if } \lambda_k^{(\nu)} = 0, \end{cases}$$

$$(M_k^n)^{(\nu)} := \log\left(1 + \frac{(\overline{S}_k^n)^{(\nu)}}{\sigma_k^n}\right) - \lambda_k(\overline{S}_k^n)^{(\nu)}.$$

Moreover, for each $k = 1, \ldots, K$, set the FDMA tone assignment according to

$$\mathcal{N}_k(\boldsymbol{\lambda}^{(\nu)}) := \left\{ n \in \mathcal{N} \, \Big| \, (M_k^n)^{(\nu)} = \max_{k'=1,\ldots,K} (M_{k'}^n)^{(\nu)} \right\},$$

and calculate the subgradient by

$$g_k^{(\nu)} := P_k - \sum_{n \in \mathcal{N}_k(\boldsymbol{\lambda}^{(\nu)})} (\overline{S}_k^n)^{(\nu)}.$$

Step 2   Update $\boldsymbol{\lambda}^{(\nu)}$ according to

$$\lambda_k^{(\nu+1)} = \left[\lambda_k^{(\nu)} - \alpha^{(\nu)} g_k^{(\nu)}\right]_+, \quad k = 1, 2, \ldots, K,$$

where $[\cdot]_+$ denotes the positive part of a real number, and $\alpha^{(\nu)}$ is the stepsize calculated by an appropriate rule.

Step 3   Go to Step 4 if the termination criterion is satisfied. Otherwise, set $\nu := \nu + 1$, and return to Step 1.

Step 4   If $\overline{\mathbf{S}}^{(\nu)}$ is feasible for problem (21), then output it as the solution. Otherwise, choose $\overline{\nu}$ such that $\|g^{(\overline{\nu})}\| = \min\{\|g^{(0)}\|, \ldots, \|g^{(\nu)}\|\}$, and calculate the optimal power allocation $\mathbf{S}$ based on $\mathcal{N}_1(\boldsymbol{\lambda}^{(\overline{\nu})}), \ldots, \mathcal{N}_K(\boldsymbol{\lambda}^{(\overline{\nu})})$. Then, output $\mathbf{S}$ as the solution.

The above algorithm falls in the same general framework of dual decomposition method as [3], [4], [5], [6]. However, there

are two notable differences. First, we implement a projection operation in Step 1 so as to ensure the boundedness of the primal variables and to improve convergence. Second, we implement an adaptive rule to select stepsizes $\{\alpha^{(\nu)}\}$ which improves the convergence speed significantly (see Section VI). Similar to [6], the above dual descent algorithm has a per-iteration complexity of $O(NK)$. Moreover, Steps $1-3$ of the above algorithm can be implemented in a distributed manner. For example, the tone assignment step (Step 1) can be carried out using a simple carrier sensing mechanism: each user k tries to send a beacon signal on tone n after a waiting period whose length is proportional to $f((M_k^n)^{(\nu)})$, where f is any positive strictly decreasing function. In this way, the user who first sends the beacon signal over tone n gets that tone, while other users, upon hearing a beacon signal on tone n, will immediately abort their transmission of beacon signals over this tone. Also, Steps 2 and 3 require no information from other users and therefore can be performed completely locally. Step 4 requires the centralization, but it can be modified in a distributed manner as follows. If $\overline{\mathbf{S}}^{(\nu)}$ is infeasible, then each user k calculates his own power vector by using the current tone assignment $\mathcal{N}_k(\boldsymbol{\lambda}^{(\nu)})$, instead of searching the past iteration $\overline{\nu}$. Since each user k has the information of $\mathcal{N}_k(\boldsymbol{\lambda}^{(\nu)})$, σ_k^n and P_k, and $\mathcal{N}_l(\boldsymbol{\lambda}^{(\nu)}) \cap \mathcal{N}_k(\boldsymbol{\lambda}^{(\nu)}) = \emptyset$ for all $l \neq k$, he can calculate his optimal power vector easily by using the well-known water-filling strategy.

Standard convergence analysis of dual descent algorithms can be applied to the above algorithm. In particular, if stepsize $\alpha^{(\nu)}$ is chosen sufficiently small, then the distance from the iterate to the dual optimal solution set decreases monotonically (even though the objective value $d(\boldsymbol{\lambda}^{(\nu)})$ may not decrease). Moreover, every limit point of the iterate sequence $\{\boldsymbol{\lambda}^{(\nu)}\}$ solves (25). If the corresponding power sequence $\{(\overline{S}_k^n)^{(\nu)}\}$ is asymptotically feasible (i.e., satisfying individual power constraint), then each limit point of $\{(\overline{S}_k^n)^{(\nu)}\}$ is a global optimal solution of the nonconvex optimal bandwidth allocation problem (21). Notice that a similar dual descent method was considered in [6] which, like the dual algorithms proposed in this paper, may not solve the primal problem due to the existence of a duality gap.

Proposition 5.1: Let $\boldsymbol{\lambda}^* \geq 0$ and $\overline{\mathbf{S}}^* \geq 0$ be the limit points of $\{\boldsymbol{\lambda}^{(\nu)}\}$ and $\{\overline{\mathbf{S}}^{(\nu)}\}$ generated by the dual decomposition algorithm. If there holds

$$\sum_{n=1}^{N}(\overline{S}_k^n)^* \leq P_k, \quad \lambda_k^*\left(\sum_{n=1}^{N}(\overline{S}_k^n)^* - P_k\right) = 0, \quad \forall\, k \in \mathcal{K},$$

then the duality gap is zero and $\overline{\mathbf{S}}^*$ is a global optimal solution of the bandwidth allocation problem (21).

Proof: By nature of dual decomposition algorithm, we have

$$\overline{\mathbf{S}}^* = \underset{\mathbf{s} \in \tilde{\mathcal{S}}}{\operatorname{argmax}} \left(\sum_{k=1}^{K}\sum_{n=1}^{N} \log\left(1 + \frac{S_k^n}{\sigma_k^n}\right) \right.$$
$$\left. - \sum_{k=1}^{K} \lambda_k^*\left(\sum_{n=1}^{N} S_k^n - P_k\right) \right). \quad (26)$$

Then, by the hypothesis, we obtain

$$\sum_{k=1}^{K}\sum_{n=1}^{N} \log\left(1 + \frac{(\overline{S}_k^n)^*}{\sigma_k^n}\right)$$
$$= \sum_{k=1}^{K}\sum_{n=1}^{N} \log\left(1 + \frac{(\overline{S}_k^n)^*}{\sigma_k^n}\right) - \sum_{k=1}^{K}\lambda_k^*\left(\sum_{n=1}^{N}(\overline{S}_k^n)^* - P_k\right)$$
$$= \max_{\mathbf{s} \in \tilde{\mathcal{S}}}\left(\sum_{k=1}^{K}\sum_{n=1}^{N} \log\left(1 + \frac{S_k^n}{\sigma_k^n}\right) - \sum_{k=1}^{K}\lambda_k^*\left(\sum_{n=1}^{N}S_k^n - P_k\right)\right)$$
$$\geq \max_{\substack{\mathbf{s} \in \mathcal{S} \\ \sum_{n=1}^{N}S_k^n \leq P_k}}\left(\sum_{k=1}^{K}\sum_{n=1}^{N} \log\left(1 + \frac{S_k^n}{\sigma_k^n}\right) - \sum_{k=1}^{K}\lambda_k^*\left(\sum_{n=1}^{N}S_k^n - P_k\right)\right)$$
$$\geq \max_{\substack{\mathbf{s} \in \mathcal{S} \\ \sum_{n=1}^{N}S_k^n \leq P_k}}\left(\sum_{k=1}^{K}\sum_{n=1}^{N} \log\left(1 + \frac{S_k^n}{\sigma_k^n}\right)\right)$$

where the second step follows from (26), the first inequality is due to

$$\left\{\mathbf{S} \,\Big|\, \mathbf{S} \in \mathcal{S}, \ \sum_{n=1}^{N}S_k^n \leq P_k\right\} \subset \tilde{\mathcal{S}},$$

and the last step can be seen from the nonnegativity of $\boldsymbol{\lambda}^*$. This shows that $\overline{\mathbf{S}}^*$ is a global optimal solution of the bandwidth allocation problem (21) and the duality gap is zero. ∎

Proposition 5.1 implies that if the power allocations obtained in Step 1 of Algorithm 1 are asymptotically feasible, then they must be globally optimal. This provides a simple way to check the optimality of the computed solution and terminate the algorithm. The minimization of FDMA dual $d(\boldsymbol{\lambda})$ is polynomial time solvable using e.g., ellipsoid method, (i.e., finding an ϵ-optimal solution is polynomial in dimension N, K and $\log 1/\epsilon$). The NP-hardness result (Theorem 5.1) implies that the duality gap is nonzero in general, so the optimal solution of the dual is not always primal optimal. However, if the dual iterates satisfy the primal power constraints with equality in the limit, then Proposition 5.1 implies the global optimality of both primal and dual in limit. This optimality condition is usually observed in our simulation.

Local search algorithm A

We now present an efficient combinatorial local search algorithm which has an overall complexity of $O(NK)$. In this algorithm, we fix the order of tones a priori, and then sequentially allocate each tone to the user who offers the largest rate increment. This algorithm can be written as follows.

Algorithm 2:

Step 0 Permute the tones $n_1, \ldots, n_N$ arbitrarily so that $\{n_1, \ldots, n_N\} = \mathcal{N}$. Let $\mathcal{L}_k^{(0)} := \emptyset$ and $\overline{R}_k^{(0)} := 0$ for each $k = 1, \ldots, K$. Set $\nu := 0$.

Step 1 For each $k = 1, \ldots, K$, solve $\mathbf{RMP}(\mathcal{L}_k^{(\nu)} \cup \{n_{\nu+1}\})$ and obtain its optimal value $\overline{R}_k'$.

Step 2 Find a $\overline{k}$ such that

$$\overline{k} = \underset{k \in \mathcal{K}}{\operatorname{argmax}}\left(\overline{R}_k' - \overline{R}_k^{(\nu)}\right).$$

Then, define $\mathcal{L}_k^{(\nu+1)}$ and $\overline{R}_k^{(\nu+1)}$ by

$$\mathcal{L}_k^{(\nu+1)} := \begin{cases} \mathcal{L}_k^{(\nu)} \cup \{n_{\nu+1}\} & (k = \overline{k}) \\ \mathcal{L}_k^{(\nu)} & (k \neq \overline{k}), \end{cases}$$

$$\overline{R}_k^{(\nu+1)} := \begin{cases} \overline{R}'_k & (k = \overline{k}) \\ \overline{R}_k^{(\nu)} & (k \neq \overline{k}) \end{cases}$$

for each $k = 1, \ldots, K$.

Step 3 Set $\nu := \nu + 1$. If $\nu = N$, then terminate. Otherwise, return to Step 1.

In Step 1, $\overline{R}'_k$ can be obtained by the water-filling procedure. In general, the obtained solution and sum-rate depend on the initial ordering of $\{n_1, \ldots, n_N\}$.

Local search algorithm B

In Algorithm 2, we have fixed the order of tones beforehand, and then allocate a tone $n_{\nu+1}$ at the ν-th iteration. However, it is expected that the sum-rate will be improved by considering all the possible combinations of tones and users at each iteration. That is, we can consider the following three steps:

(i) For each pair of user and unallocated tone combination, we calculate the corresponding rate increment.
(ii) Find the tone $\overline{n}$ and the user $\overline{k}$ which yield the largest rate increment. Allocate tone $\overline{n}$ to user $\overline{k}$.
(iii) Remove tone $\overline{n}$ from the non-allocated tone set, and return to (i) until all tones are allocated.

A direct implementation of the above procedure will result in a computational complexity of $O(N^2 K)$. However, we note that, for any $k \in \mathcal{K}$, $\mathcal{L}_k \subseteq \mathcal{N}$, and $n, n' \in \mathcal{L}_k$, it follows $\overline{R}_k(\mathcal{L}_k \cup \{n\}) \leq \overline{R}_k(\mathcal{L}_k \cup \{n'\})$ if and only if $\sigma_k^n \geq \sigma_k^{n'}$. Therefore, by sorting the noise parameters $\{\sigma_k^n\}$ appropriately, we can reduce its complexity to $O(NK \log N)$. We describe the algorithm in the following, where $\mathcal{L}_k^{(\nu)}$, $\overline{R}_k^{(\nu)}$, and $\overline{\mathcal{N}}^{(\nu)}$ denote user k's allocated tone set, user k's temporary data rate, and unallocated tone set at the ν-th iteration.

Algorithm 3:

Step 0 For each $k = 1, \ldots, K$, sort the tone indices $\{n_1(k), \ldots, n_N(k)\} = \mathcal{N}$ so that

$$\sigma_k^{n_1(k)} \leq \cdots \leq \sigma_k^{n_N(k)}.$$

Let $\mathcal{L}_k^{(0)} := \emptyset$, $R_k^{(0)} := 0$, and $\overline{\mathcal{N}}^{(0)} := \mathcal{N}$ for all $k \in \mathcal{K}$. Set $\nu := 0$.

Step 1 For every $k = 1, \ldots, K$, perform the following steps:

Step 1-1 Find a tone $\overline{n}(k) := n_{i_-}(k)$ such that

$$i_- = \min\Big\{i \ \Big| \ n_i(k) \in \overline{\mathcal{N}}^{(\nu)}\Big\}.$$

Step 1-2 Solve $\mathbf{RMP}(\mathcal{L}_k^{(\nu)} \cup \{\overline{n}(k)\})$ and obtain its optimal value $\overline{R}'_k$.

Step 2 Find a $\overline{k} \in \mathcal{K}$ such that

$$\overline{k} = \underset{k \in \mathcal{K}}{\operatorname{argmax}} \left(\overline{R}'_k - \overline{R}_k^{(\nu)} \right).$$

Define $\mathcal{L}_k^{(\nu+1)}$ and $\overline{R}_k^{(\nu+1)}$ by

$$\mathcal{L}_k^{(\nu+1)} := \begin{cases} \mathcal{L}_k^{(\nu)} \cup \{\overline{n}(k)\} & (k = \overline{k}) \\ \mathcal{L}_k^{(\nu)} & (k \neq \overline{k}), \end{cases}$$

$$\overline{R}_k^{(\nu+1)} := \begin{cases} \overline{R}'_k & (k = \overline{k}) \\ \overline{R}_k^{(\nu)} & (k \neq \overline{k}) \end{cases}$$

for each $k \in \mathcal{K}$. Then, let $\overline{\mathcal{N}}^{(\nu+1)} := \overline{\mathcal{N}}^{(\nu)} \setminus \{\overline{n}(\overline{k})\}$.

Step 3 If $\overline{\mathcal{N}}^{(\nu+1)} = \emptyset$, then terminate. Otherwise, set $\nu := \nu + 1$ and return to Step 1.

In Step 0, the computational cost for the sort of $\{n_1(k), \ldots, n_N(k)\}$ is $O(N \log N)$ for each k. Step 1-1 implies that tone $\overline{n}(k) \in \overline{\mathcal{N}}^{(\nu)}$ is chosen so that $\sigma_k^{\overline{n}(k)} = \min\{\sigma_k^n \mid n \in \overline{\mathcal{N}}^{(\nu)}\}$. In Step 1-2, $\overline{R}'_k$ can be obtained by the water-filling procedure. One is tempted to think Algorithm 3 would always yield a better solution than Algorithm 2. While it often does, numerical results in the next section show that Algorithm 3 sometimes can lead to a worse sum-rate solution than Algorithm 2.

Both local search algorithms (Algorithms 2 and 3) can also be implemented in a distributed fashion by using the same carrier sensing mechanism stated after Algorithm 1. In the ν-th iteration, each user k sends a beacon signal after the waiting period $f(\overline{R}'_k - R_k^{(\nu)})$ with a certain strictly decreasing function f, where the value of $f(\overline{R}'_k - R_k^{(\nu)})$ can be calculated locally. Then, tone $n_{\nu+1}$ (or $\overline{n}(k)$ for Algorithm 3) is assigned to the user who first sends the beacon signal.

VI. NUMERICAL RESULTS

In this section, we consider a wireless setup and compare the performance of various spectrum management algorithms: the dual decomposition method (Algorithm 1), the local search algorithms (Algorithms 2 and 3), and the iterative water-filling algorithm (IWFA).

For the dual decomposition method, we choose the initial dual vector $\boldsymbol{\lambda}^{(0)} = (1, \ldots, 1)^T$, and consider two different stepsize rules:

Stepsize rule A $\alpha^{(\nu)} := 1/(\nu + 1)$.

Stepsize rule B $\alpha^{(\nu)} := \theta^{(\nu)}(d(\boldsymbol{\lambda}^{(\nu)}) - L^*)/\|g^{(\nu)}\|^2$, where L^* is a known lower bound of the dual function d, and $\theta^{(\nu)}$ is calculated according to the following rule: (i) $\theta^{(0)} = 2$, (ii) $\theta^{(\nu+1)} = \theta^{(\nu)}/2$ if $d(\boldsymbol{\lambda}^{(\nu)}) \geq d(\boldsymbol{\lambda}^{(\nu-10)})$ for $\nu \geq 10$, and (iii) $\theta^{(\nu+1)} = \theta^{(\nu)}$ if $d(\boldsymbol{\lambda}^{(\nu)}) < d(\boldsymbol{\lambda}^{(\nu-10)})$ or $\nu \leq 9$.

Stepsize rule A is simple and easily implementable. We note, however, that $\{\alpha^{(\nu)}\}$ converges to 0 very slowly, and hence a large number of iterations may be required. Stepsize rule B is a modification of a standard stepsize rule for maximizing a Lagrangian function [19]. If a tight lower bound L^* is available, the algorithm typically can terminate in a small number of iterations. In implementing this stepsize rule, we first calculate the sum-rate by the local search algorithm B, and then use the obtained sum-rate as the lower bound L^*. We stop the algorithm when either $\|\boldsymbol{\lambda}^{(\nu+1)} - \boldsymbol{\lambda}^{(\nu)}\| \leq 10^{-4}$ or $\nu \geq 300$.

For IWFA, we let each user choose an initial power level randomly from the interval $[0, \max_k P_k]$, and terminate the

iteration if $\|\mathbf{S}^{(\nu+1)} - \mathbf{S}^{(\nu)}\| \le 10^{-4}$ or $\nu \ge 300$. As mentioned in Section I, IWFA maximizes each user's individual rate in a distributed manner by treating other users' signals as Gaussian noise. This can be easily implemented using the well-known water-filling strategy for a single user rate maximization. Since the FDMA concept is not considered in IWFA, the obtained power spectra are not FDMA in general.

In our simulation, we consider a multiuser wireless communication system in a frequency selective environment; see Subsection II-A for a full description. We define the channel coefficients as $h_{lk}^n := d_{lk}^{-1.8} g_{lk}^n$ where d_{lk} denotes the physical distance between transmitter l and receiver k, and g_{lk}^n is a complex normalized gaussian random variable with zero mean and unit variance. Then, the crosstalk coefficients and normalized noise power are chosen as $\alpha_{lk}^n := |h_{lk}^n|^2/|h_{kk}^n|^2$ and $\sigma_k^n := N_0/|h_{kk}^n|^2$, where the background noise level is set to $N_0 = -40$ dB. The programs were coded in MATLAB 7 and run on a machine with 3.60GHz CPU and 2GB RAM.

Experiment 1

Let there be $N = 12$ tones shared by $K = 4$ users in the system (e.g., the blue tooth setup). Then we randomly generate 4 pairs of transmitters and receivers so that each transmitter k is located in the 2-dimensional unit square and d_{kk} (the distance from transmitter k to receiver k) equals $\Delta > 0$ for all $k \in \mathcal{K}$.[4] Figure 1 shows a simple example, where the solid arrows denote the desired signal path, and all other edges in the graph (not shown) represent interferences.

We let the distance $d_{kk} = \Delta$ vary from 0.02 to 0.2, and generate 1000 test problems for each Δ. As expected, the crosstalk interference becomes stronger when the distance Δ increases. For each test problem, we choose power budget P_k randomly from the interval $[10, 16]$ (dB), and solve the corresponding spectrum management problem by the dual decomposition method with stepsize rules A and B (denoted by dual decomposition method A and B respectively), local search algorithms A and B, and IWFA. The average CPU time among 1000 trials are shown in Figure 2, which shows that the computational costs of the local search algorithms are much lower than other algorithms. Figure 3 shows the average of the obtained sum-rates for each Δ. It can be seen that, for small Δ where the crosstalk coefficients are small, IWFA yields higher sum-rate compared to our FDMA-based methods. This is expected since FDMA is strictly sub-optimal in low interference environment. However, when Δ becomes larger, our FDMA-based methods yield much higher sum-rates than IWFA, confirming the superiority of FDMA solutions under strong crosstalk conditions. Figure 4 plots the ratios of sum-rates obtained by our FDMA-based methods relative to that of the local search algorithm A. As the figure shows, the dual decomposition method B gives the highest average values. Figure 5 shows the number of times the best sum-rates are obtained among five methods. If the best sum-rate is

attained by more than one method, we duplicate counting so there can be more than 1000 times that the best sum-rate is achieved. As Figure 5 shows, the dual decomposition method B attains the best sum-rate for more than 90 % cases when $\Delta \ge 0.1$.

Experiment 2

In the second experiment, we also set $N = 12$ and $K = 4$, and generate the transmitters and receivers in a way similar to Experiment 1. However, we fix the distance to $\Delta = 0.04$ (so the crosstalk coefficients are small), and vary power budget levels. Our goal is to demonstrate that FDMA becomes optimal when power budget becomes large even in low interference environment. We choose the power budget level β from -10 dB to $+30$ dB, and generate 1000 test problems for each β. For each test problem, we choose the power budgets P_k from the interval $[\beta - 3, \beta + 3]$ (dB) randomly. Figure 6 shows the average of obtained sum-rates. Since the crosstalk coefficients are relatively small, IWFA achieves better sum-rates than our methods in the low power region. However, if the power budget level becomes higher, then our FDMA-based methods yield higher sum-rate than IWFA as predicted by our theoretical analysis (Theorem 4.1). Unfortunately, the dual decomposition method A shows worse performance than other FDMA-based methods when $\beta \ge 20$. We suppose that, in this case, the generated point $\boldsymbol{\lambda}^{(\nu)}$ is still far from the solution $\boldsymbol{\lambda}^*$ even after 300 iterations (i.e., termination criterion is satisfied). This result also shows the advantage of Stepsize rule B.

As shown in the above numerical results, the FDMA-based methods provide a much improved performance than IWFA under either the strong crosstalk or high power budget conditions. Moreover, they can be implemented in a distributed manner by introducing the aforementioned carrier sensing mechanism. (In this case, Step 4 of Algorithm 1 need to be modified so that the values of the last iteration ν are adopted instead of searching $\bar{\nu}$.) However, even in such cases, IWFA still has the following advantage. In IWFA, each user can update his power vector whenever he wishes, that is, it can be implemented in an asynchronous manner [20]. On the other hand, the carrier sensing approach needs to synchronize all users' clocks before starting the waiting period.

VII. Concluding Remarks

In this paper we have studied the structure of optimal spectrum sharing strategies for a multiuser communication system in a frequency selective environment. Our analysis and simulations show that FDMA is sum-rate optimal when either the crosstalk interference is strong or when users' power budgets are high. Unlike the frequency flat case where the sum-rate maximum FDMA solution can be found using convex optimization, the same problem in a frequency selective environment is considerably more difficult (NP-hard). To approximately solve the sum-rate maximization problem for a frequency selective environment, we have proposed several simple distributed algorithms that can find high quality sum-rate suboptimal FDMA solutions. Numerical experiments show that, if the crosstalk coefficients and/or power budgets

---

[4]More precisely, we randomly generate 4 transmitters in the 2-dimensional unit square, and then generate receiver k ($k = 1, \ldots, 4$) randomly on the circumference of the circle whose center is transmitter k ($k = 1, \ldots, 4$) and radius is Δ.

are sufficiently large, the proposed algorithms not only find higher sum-rates than those obtained by IWFA, but also enjoy faster convergence. However, when crosstalk interference is low, IWFA can deliver higher sum-rates. These results suggest a hybrid approach for the sum-rate maximization problem in practice whereby both IWFA and the proposed optimal bandwidth allocation algorithms are used, depending on the strength of crosstalk interference.

There are several issues that are worthy of further investigation. For example, the analysis in this paper shows that the sum-rate maximization problem is NP-hard even in the case of two users. However, it is not known if the same is true when we fix the number of tones and let the number of users increase to infinity. Also, this paper has not addressed the fairness issue. One popular approach to ensure user fairness [21] is to maximize the sum of proportional fair rates rather than the sum-rates; the former is defined as the logarithm of user's data rates. It will be interesting to find out under what conditions the maximization of proportional fair rates will result in FDMA solutions. Moreover, is the maximization of sum of proportional fair rates also NP-hard? Are there simple distributed algorithms which can approximately maximize the sum of proportional fair rates? The answers to these questions are of considerable value for the standardization of dynamic spectrum management technologies. Finally, one may be able to derive other sufficient conditions for FDMA optimality that are weaker than those presented in this paper (simulation results strongly suggest this). It may also be possible to design simple distributed bandwidth allocation algorithms capable of delivering sum-rates that are provably optimal up to a constant factor.
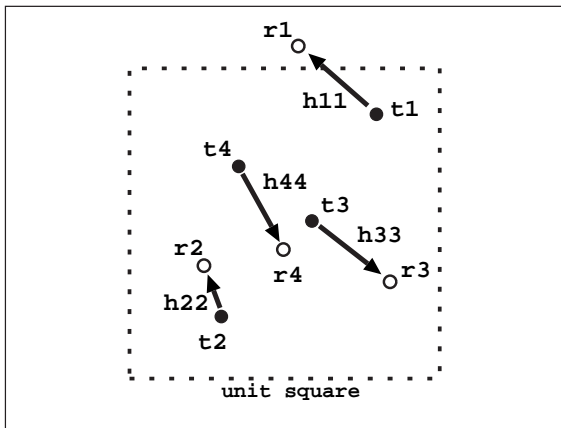

Fig. 1.   A wireless scenario

APPENDIX A
PROOF OF PROPOSITION 3.2

Let $\mathbf{S} \in \Re^{NK}$ be a globally optimal power allocation for the sum-rate maximization problem.

Assume that $\min_{k \in \mathcal{K}} |\mathcal{T}_k(\mathbf{S}_k)| \leq C - 1$ and we will derive a contradiction. Without loss of generality we can assume that user 1 uses at most $C-1$ tones in $\mathbf{S}$ and these tones are indexed tones $1, \ldots, C-1$. That is, $S_1^n \geq 0$ for $n = 1, \ldots, C-1$ and
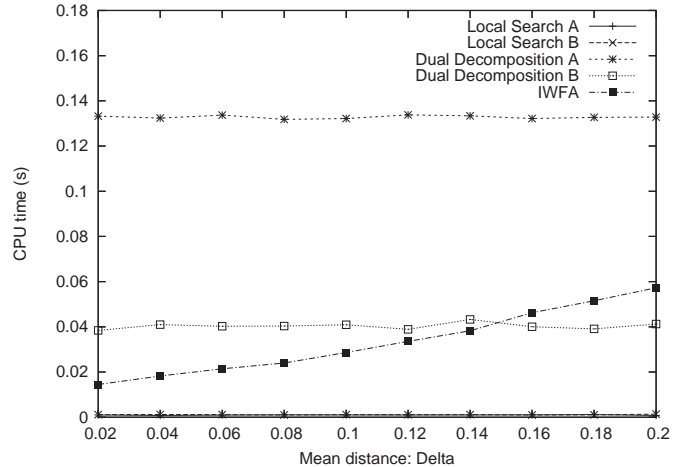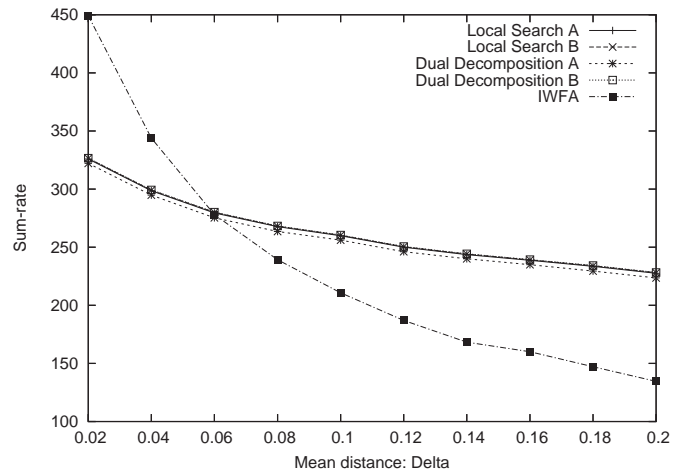

Fig. 2.   CPU time v.s. distance Δ
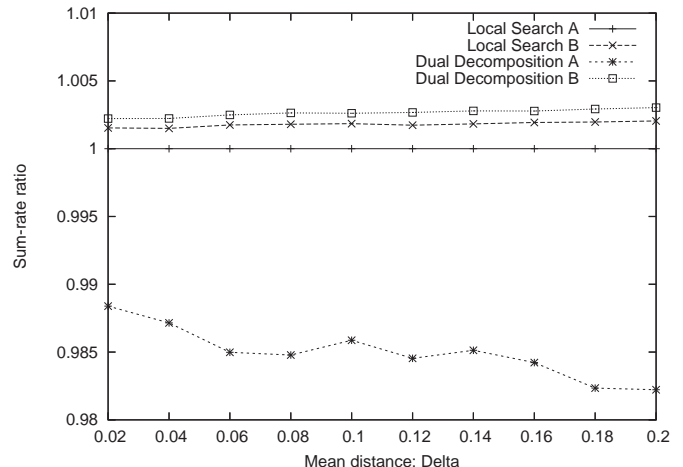

Fig. 3.   Sum-rate v.s. distance Δ

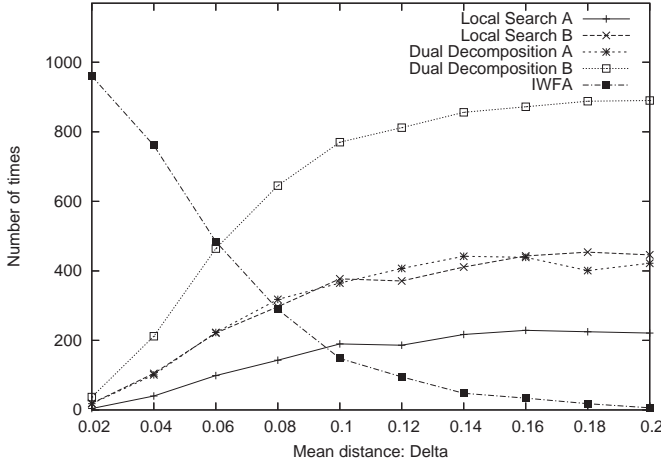Fig. 4.   Sum-rate ratios v.s. distance Δ

Fig. 5. Number of times the best sum-rates are obtained
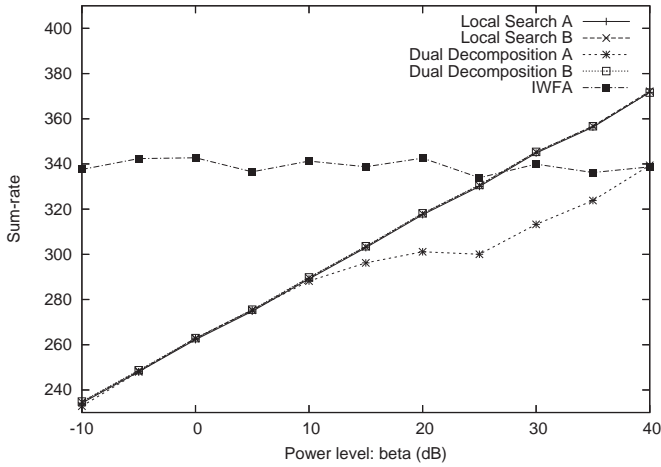

Fig. 6. Sum-rates v.s. power levels

$S_1^C = \cdots = S_1^N = 0$. Since $\sum_{n=1}^N S_k^n \le P_k$ for all $k \in \mathcal{K}$, it follows that

$$\sum_{k=1}^K \sum_{n=1}^N S_k^n \le \sum_{k=1}^K P_k.$$

Let $n_i \in \mathcal{N}$ denote the tone for which the total user power is the i-th smallest, namely,

$$\sum_{k=1}^K S_k^{n_1} \le \sum_{k=1}^K S_k^{n_2} \le \cdots \le \sum_{k=1}^K S_k^{n_N}.$$

Then we have for each $i = 1, \ldots, N$,

$$(N - i + 1) \sum_{k=1}^K S_k^{n_i} \le \sum_{j=i}^N \sum_{k=1}^K S_k^{n_j} \le \sum_{n=1}^N \sum_{k=1}^K S_k^n \le \sum_{k=1}^K P_k$$

which further implies that

$$\sum_{k=1}^K S_k^{n_i} \le \frac{1}{N - i + 1} \sum_{k=1}^K P_k, \qquad (27)$$

In what follows, we will show that it is better (in terms of achieving a higher sum-rate) to let user 1 allocate equal power to tones $n_1, \ldots, n_m$, and have other users give up using these m tones. This will imply that $\min_{k \in \mathcal{K}} |\mathcal{T}_k(\mathbf{S}_k)| \le C - 1$ cannot hold at any global sum-rate optimum since $m > C - 1$. To show this, we need to estimate the sum-rate achievable by the power vector $\mathbf{S}$ on tones $n_1, \ldots, n_m$. For this purpose, we consider the following sum-rate maximization problem for each tone $n \in \mathcal{N}$:

$$\underset{s_1^n, \ldots, s_K^n}{\text{maximize}} \quad \sum_{k=1}^K \log \left(1 + \frac{s_k^n}{\sigma_k^n + \frac{1}{2} \sum_{l \ne k} s_l^n}\right)$$

$$= \sum_{k=1}^K \log \left(1 + \frac{s_k^n}{\sigma_k^n + \frac{1}{2}(U^n - s_k^n)}\right)$$

$$\text{subject to} \quad \sum_{k=1}^K s_k^n = U^n, \ \ s_k^n \ge 0, \quad k \in \mathcal{K}, \qquad (28)$$

where $U^n := \sum_{k=1}^K S_k^n$. Notice that the crosstalk coefficients α_{ij}^n have all been reduced to $1/2$ in (28). It can be checked that the second order derivative (with respect to s_k^n) of the rate function

$$\log \left(1 + \frac{s_k^n}{\sigma_k^n + \frac{1}{2}(U^n - s_k^n)}\right)$$

is nonnegative over $s_k^n \ge 0$, implying that this rate function is convex. This shows that the objective function of (28) is convex in the feasible set. Consequently, as a maximization problem, the maximum of (28) is achieved at a vertex of the feasible set, implying

$$\max_{\substack{\sum_{k=1}^K s_k^n = U^n \\ s_k^n \ge 0}} \sum_{k=1}^K \log \left(1 + \frac{s_k^n}{\sigma_k^n + \frac{1}{2} \sum_{l \ne k} s_l^n}\right)$$

$$= \max_{k \in \mathcal{K}} \log \left(1 + \frac{U^n}{\sigma_k^n}\right) = \log \left(1 + \frac{U^n}{\sigma_0^n}\right),$$

where $\sigma_0^n := \min_{k \in \mathcal{K}} \sigma_k^n$. This gives an upper bound on the data rate achieved on tone n_i:

$$f^{n_i}(\mathbf{S}^{n_i}) = \sum_{k=1}^K \log \left(1 + \frac{S_k^{n_i}}{\sigma_k^{n_i} + \sum_{l \ne k} \alpha_{lk}^{n_i} S_l^{n_i}}\right)$$

$$\le \sum_{k=1}^K \log \left(1 + \frac{S_k^{n_i}}{\sigma_k^{n_i} + \frac{1}{2} \sum_{l \ne k} S_l^{n_i}}\right)$$

$$\le \log \left(1 + \frac{U^{n_i}}{\sigma_0^{n_i}}\right)$$

$$\le \log \left(1 + \frac{\sum_{k=1}^K P_k}{(N - i + 1)\sigma_0^{n_i}}\right)$$

where the last step follows from (27).

Let us consider a new power allocation vector $\bar{\mathbf{S}}$:

$$\bar{S}_k^n = \begin{cases} \frac{P_1}{m}, & k = 1, \ n \in \{n_1, \ldots, n_m\} \\ 0, & k \ge 2, \ n \in \{n_1, \ldots, n_m\} \\ 0, & k = 1, \ n \in \{1, \ldots, C-1\} \setminus \{n_1, \ldots, n_m\} \\ S_k^n, & \text{else}; \end{cases}$$

It is clear that $\bar{\mathbf{S}}$ is feasible. Moreover, the sum-rate achieved by $\bar{\mathbf{S}}$ over the tones $n_1, \ldots, n_m$ is evaluated as

$$\sum_{i=1}^m f^{n_i}(\bar{\mathbf{S}}^{n_i}) = \sum_{i=1}^m \log \left(1 + \frac{P_1}{m\sigma_1^{n_i}}\right) \ge m \log \left(1 + \frac{\rho_0}{m}\right).$$

On the other hand, for any $n \in \{1, \ldots, C-1\} \setminus \{n_1, \ldots, n_m\}$, we have

$$
\begin{aligned}
&f^n(\bar{\mathbf{S}}^n) - f^n(\mathbf{S}^n) \\
&= \left[0 + \sum_{k=2}^K \log\left(1 + \frac{S_k^n}{\sigma_k^n + \sum_{l \neq 1, l \neq k} \alpha_{lk}^n S_l^n}\right)\right] \\
&\quad - \left[\log\left(1 + \frac{S_1^n}{\sigma_1^n + \sum_{l \neq 1} \alpha_{l1}^n S_l^n}\right)\right. \\
&\qquad \left. + \sum_{k=2}^K \log\left(1 + \frac{S_k^n}{\sigma_k^n + \sum_{l \neq k} \alpha_{lk}^n S_l^n}\right)\right] \\
&\geq -\log\left(1 + \frac{S_1^n}{\sigma_1^n + \sum_{l \neq 1} \alpha_{l1}^n S_l^n}\right) \geq -\log\left(1 + \frac{S_1^n}{\sigma_1^n}\right).
\end{aligned}
$$

We thus have

$$
\begin{aligned}
&\sum_{\substack{n \notin \{n_1, \ldots, n_m\} \\ n \in \{1, \ldots, C-1\}}} \left[f^n(\bar{\mathbf{S}}^n) - f^n(\mathbf{S}^n)\right] \\
&\quad \geq -\sum_{\substack{n \notin \{n_1, \ldots, n_m\} \\ n \in \{1, \ldots, C-1\}}} \log\left(1 + \frac{S_1^n}{\sigma_1^n}\right) \\
&\quad \geq -\sum_{n=1}^{C-1} \log\left(1 + \frac{S_1^n}{\sigma_1^n}\right) \\
&\quad \geq -(C-1)\log\left[\frac{1}{C-1}\sum_{n=1}^{C-1}\left(1 + \frac{S_1^n}{\sigma_1^n}\right)\right] \\
&\quad \geq -(C-1)\log\left[1 + \frac{1}{(C-1)}\sum_{n=1}^{C-1}\frac{S_1^n}{\min_{n \in \mathcal{N}}(\sigma_1^n)}\right] \\
&\quad \geq -(C-1)\log\left[1 + \frac{P_1}{(C-1)\min_{n \in \mathcal{N}}(\sigma_1^n)}\right] \\
&\quad \geq -(C-1)\log\left(1 + \frac{\rho_M}{C-1}\right),
\end{aligned}
$$

where the third inequality is due to the arithmetic-geometric mean inequality.

We can now compare the sum-rate achieved by $\mathbf{S}$ and $\bar{\mathbf{S}}$.

$$
\begin{aligned}
&\sum_{n=1}^N f^n(\bar{\mathbf{S}}^n) - \sum_{n=1}^N f^n(\mathbf{S}^n) \\
&= \sum_{n \in \{n_1, \ldots, n_m\}} \left[f^n(\bar{\mathbf{S}}^n) - f^n(\mathbf{S}^n)\right] + \sum_{\substack{n \notin \{n_1, \ldots, n_m\} \\ n \in \{1, \ldots, C-1\}}} \left[f^n(\bar{\mathbf{S}}^n) - f^n(\mathbf{S}^n)\right] \\
&\geq m\log\left(1 + \frac{\rho_0}{m}\right) - \sum_{i=1}^m \log\left(1 + \frac{\sum_{k=1}^K P_k}{(N-i+1)\sigma_0^{n_i}}\right) \\
&\quad - (C-1)\log\left(1 + \frac{\rho_M}{C-1}\right) \\
&\geq m\log\left(1 + \frac{\rho_0}{m}\right) - m\log\left(1 + \frac{K\rho_a}{N-m+1}\right) \\
&\quad - (C-1)\log\left(1 + \frac{\rho_M}{C-1}\right) > 0.
\end{aligned}
$$

This shows that the new power vector $\bar{\mathbf{S}}$ achieves a higher sum-rate than $\mathbf{S}$, contradicting the global optimality of $\mathbf{S}$.

APPENDIX B
PROOF OF PROPOSITION 3.3

Let $\mathbf{S} \in \Re^{NK}$ be an arbitrary non-FDMA power vector satisfying Condition 1 (a). We first show that (i) holds, or there exist an integer $L \geq 1$, a tone set $\{n_i\}_{i=1}^L$, and a user set $\{k_i\}_{i=1}^{L+1}$ such that $\left[n_j \neq n_{j'} \text{ and } k_j \neq k_{j'} \text{ for any } j \neq j'\right]$ and

$$
\begin{aligned}
&(n_i, k_i) \in \mathcal{T}, \ (n_i, k_{i+1}) \in \mathcal{T} \text{ for } i = 1, \ldots, L, \\
&\text{and } \mathcal{N}_{II} \cap \mathcal{T}_{k_{L+1}} = \{n_L\}.
\end{aligned} \tag{29}
$$

This can be verified according to the following procedure.

Step 0 Choose arbitrary $n_1 \in \mathcal{N}_{II}$ and $k_1 \in \mathcal{K}$ such that $(n_1, k_1) \in \mathcal{T}$. Set $i := 1$.

Step 1 Find $k_{i+1} \neq k_i$ such that $(k_{i+1}, n_i) \in \mathcal{T}$. Such a k_{i+1} exists since $n_i \in \mathcal{N}_{II}$. If $k_{i+1} \in \{k_1, \ldots, k_i\}$, then terminate since (i) holds.

Step 2 If n_i is an only element in $\mathcal{N}_{II} \cap \mathcal{T}_{k_{i+1}}$, then terminate. Otherwise, we can find $n_{i+1} \in \mathcal{N}_{II} \cap \mathcal{T}_{k_{i+1}}$ such that $n_{i+1} \neq n_i$.

Step 3 If $n_{i+1} \in \{n_1, \ldots, n_i\}$, then terminate since (i) holds. Otherwise, return to Step 1 with setting $i := i+1$.

Note that the above procedure must terminate within $\min(K, N)$ iterations since, in the i-th iteration ($i \geq 2$), we have $k_j \neq k_{j'}$ and $n_j \neq n_{j'}$ for any (j, j') with $1 \leq j < j' \leq i$. In Step 2, if the iteration terminates, then we have (29). In Step 1, we can specify the sets in (i) as follows. If $k_{i+1} = k_p$ with $1 \leq p \leq i-1$, then we have $(n_j, k_j) \in \mathcal{T}$ and $(n_j, k_{j+1}) \in \mathcal{T}$ for $j = p, \ldots, i-1$, and $(n_i, k_i) \in \mathcal{T}$ and $(n_i, k_p) \in \mathcal{T}$. Hence, renumbering the indices appropriately, we have (i). Also, in Step 3, we can see (i) as follows. If $n_{i+1} = n_q$ with $1 \leq q \leq i-1$, then we have $(n_j, k_j) \in \mathcal{T}$ and $(n_j, k_{j+1}) \in \mathcal{T}$ for $j = q+1, \ldots, i$, and $(n_q, k_{i+1}) \in \mathcal{T}$ and $(n_q, k_{q+1}) \in \mathcal{T}$. Renumbering the indices as $(k_{q+1}, \ldots, k_{i+1}) \to (k_1, \ldots, k_M)$, $(n_{q+1}, \ldots, n_i) \to (n_1, \ldots, n_{M-1})$, and $n_q \to n_M$ with $M = i-q$, we have (i).

Next we show either (i) or (ii) holds, assuming (29) holds. For the sake of convenience, we relabel the indices as $(n_L, n_{L-1}, \ldots, n_1) \to (n_1, n_2, \ldots, n_L)$ and $(k_{L+1}, k_L, \ldots, k_1) \to (k_1, k_2, \ldots, k_{L+1})$. Then, (29) is rewritten as

$$
\begin{aligned}
&(n_i, k_i) \in \mathcal{T}, \ (n_i, k_{i+1}) \in \mathcal{T} \text{ for } i = 1, \ldots, L, \\
&\text{and } \mathcal{N}_{II} \cap \mathcal{T}_{k_1} = \{n_1\}.
\end{aligned} \tag{30}
$$

Now we can carry out the above procedure again from Step 2 with $i := L$. If the iteration terminates in Step 2, then we have (ii). If it terminates in Step 1 or 3, then we have (i).

APPENDIX C
PROOF OF PROPOSITION 3.4

By reordering indices if necessary, we assume without loss of generality that $i := n_i$ and $i := k_i$ for $i = 1, \ldots, M$. Then we have

$$
(i, i) \text{ and } (i, i+1) \in \mathcal{T}, \quad \text{for } i = 1, \ldots, M,
$$

where we regard $M+1$ as 1 for convenience. Since we always write the tone index as a superscript and user index as a subscript, the notation S_{i+1}^i signifies the power value for

player $i+1$ at tone i. Also, we use $(i, i+1)$ to denote the pair of tone i and player $i+1$.

Now, let us define a vector $v \in \Re^{NK}$ as follows:

$$v_1^1 = v_2^2 = \cdots = v_M^M = 1,$$
$$v_2^1 = v_3^2 = \cdots = v_{M+1}^M (= v_1^M) = -1,$$
$$v_k^n = 0 \quad \text{(other components)}.$$

Then it can be easily seen that v satisfies (7). On the other hand, by Proposition 3.1, we have

$$\sum_{n=1}^N (v^n)^T \nabla^2 f^n(\mathbf{S}^n) v^n$$

$$= \sum_{i=1}^M \begin{pmatrix} v_i^i \\ v_{i+1}^i \end{pmatrix}^T \begin{pmatrix} \partial_{ii} f^i(\mathbf{S}^i) & \partial_{i\,i+1} f^i(\mathbf{S}^i) \\ \partial_{i\,i+1} f^i(\mathbf{S}^i) & \partial_{i+1\,i+1} f^i(\mathbf{S}^i) \end{pmatrix} \begin{pmatrix} v_i^i \\ v_{i+1}^i \end{pmatrix}$$

$$= \sum_{i=1}^M \begin{pmatrix} 1 \\ -1 \end{pmatrix}^T \left\{ \right.$$
$$- \begin{pmatrix} (A_i^i)^2 & \alpha_{i+1\,i}^i (A_i^i)^2 + \alpha_{i\,i+1}^i (A_{i+1}^i)^2 \\ \alpha_{i+1\,i}^i (A_i^i)^2 + \alpha_{i\,i+1}^i (A_{i+1}^i)^2 & (A_{i+1}^i)^2 \end{pmatrix}$$
$$+ \sum_{r=1}^K \begin{pmatrix} (\alpha_{ir}^i)^2 & \alpha_{ir}^i \alpha_{i+1\,r}^i \\ \alpha_{ir}^i \alpha_{i+1\,r}^i & (\alpha_{i+1\,r}^i)^2 \end{pmatrix} Q_r^i \left.\right\} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$= \sum_{i=1}^M \left\{ (2\alpha_{i+1\,i}^i - 1)(A_i^i)^2 + (2\alpha_{i\,i+1}^i - 1)(A_{i+1}^i)^2 \right\}$$

$$+ \sum_{i=1}^M \sum_{r=1}^K (\alpha_{ir}^i - \alpha_{i+1\,r}^i)^2 Q_r^i \ > \ 0,$$

where the inequality follows since $\alpha_{lk}^n > 1/2$, $A_k^n > 0$ and $Q_k^n \geq 0$ for all n, k and l. Since $\mathbf{S}$ does not satisfy the second order necessary conditions, $\mathbf{S}$ cannot be a local maximum.

APPENDIX D
PROOF OF PROPOSITION 3.5

By re-ordering indices if necessary, we assume $i := n_i$ for $i = 1, \ldots, M-1$ and $i := k_i$ for $i = 1, \ldots, M$. Then we have

$$(i, i) \text{ and } (i, i+1) \in \mathcal{T}, \quad \text{for } i = 1, \ldots, M-1,$$

$\mathcal{N}_{II} \cap \mathcal{T}_1 = \{1\}$, and $\mathcal{N}_{II} \cap \mathcal{T}_M = \{M-1\}$. Since $|\mathcal{T}_1|, |\mathcal{T}_M| \geq C$ (cf. Condition 1 (a)), there exist tones $n_1, \ldots, n_{C-1} \notin \{1, \ldots, M-1\}$ and $m_1, \ldots, m_{C-1} \notin \{1, \ldots, M-1\}$ such that

$$n_1, \ldots, n_{C-1} \in \mathcal{N}_I \cap \mathcal{T}_1,$$
$$m_1, \ldots, m_{C-1} \in \mathcal{N}_I \cap \mathcal{T}_M.$$

Assume, to the contrary, that $\mathbf{S}$ is a local maximum. We will derive a contradiction. Since S satisfies the KKT conditions (5), we have from (6) that

$$0 \leq \partial_1 f^1(S^1) = \partial_1 f^{n_1}(S^{n_1}) = \cdots = \partial_1 f^{n_{C-1}}(S^{n_{C-1}}). \quad (31)$$

Moreover, since $n_j \in \mathcal{N}_I \cap \mathcal{T}_1$ for all $j = 1, \ldots, C-1$, we have $S_1^{n_j} > 0$ and $S_2^{n_j} = \cdots = S_K^{n_j} = 0$. This together with Proposition 3.1 yields

$$\partial_1 f^{n_j}(S^{n_j}) = B_1^{n_j} \text{ and } \partial_{11} f^{n_j}(S^{n_j}) = -(B_1^{n_j})^2, \quad (32)$$

where we note that $P_k^n = Q_k^n = 0$ if and only if $S_k^n = 0$. Therefore, we have for each $j = 1, \ldots, C-1$

$$\partial_{11} f^{n_j}(S^{n_j}) = -(\partial_1 f^{n_j}(S^{n_j}))^2$$
$$= -(\partial_1 f^1(S^1))^2$$
$$\geq -(A_1^1)^2, \quad (33)$$

where the first equality is due to (32), the second equality holds from (31), and the inequality follows since we have $0 \leq \partial_1 f^1(S^1) = A_1^1 - \sum_{r=1}^K \alpha_{1r}^1 P_r^1 \leq A_1^1$ from (31) and Proposition 3.1. In a similar way, we also obtain

$$\partial_{MM} f^{m_j}(S^{m_j}) \geq -(A_M^{M-1})^2 \quad (34)$$

for all $j = 1, \ldots, C-1$.

Now, we can define v as

$$\begin{cases} v_i^i = \beta_i & (i = 1, \ldots, M-1) \\ v_{i+1}^i = -\beta_{i+1} & (i = 1, \ldots, M-1) \\ v_1^{n_j} = -\beta_1 (C-1)^{-1} & (j = 1, \ldots, C-1) \\ v_M^{m_j} = \beta_M (C-1)^{-1} & (j = 1, \ldots, C-1) \\ v_k^n = 0 & \text{(other components)}, \end{cases} \quad (35)$$

where

$$\begin{cases} \beta_1 = 1 + (C-1)^{-1} \\ \beta_2 = 2\alpha_{12}^1 \end{cases} \quad (M = 2), \quad (36)$$

$$\begin{cases} \beta_1 = 1 \\ \beta_{i+1} = 2\alpha_{i\,i+1}^i \beta_i & (i = 1, \ldots, M-2) \quad (M \geq 3). \quad (37) \\ \beta_M = (2\alpha_{M\,M-1}^{M-1})^{-1} \beta_{M-1} \end{cases}$$

It is easily seen that v satisfies (7). By the definition (35) of v, we have

$$\sum_{n=1}^N (v^n)^T \nabla^2 f^n(\mathbf{S}^n) v^n$$

$$= (C-1)^{-2} \sum_{j=1}^{C-1} \left(\beta_1^2 \partial_{11} f^{n_j}(S^{n_j}) + \beta_M^2 \partial_{MM} f^{m_j}(S^{m_j}) \right)$$

$$+ \sum_{i=1}^{M-1} \begin{pmatrix} \beta_i \\ -\beta_{i+1} \end{pmatrix}^T \begin{pmatrix} \partial_{ii} f^i(\mathbf{S}^i) & \partial_{i\,i+1} f^i(\mathbf{S}^i) \\ \partial_{i\,i+1} f^i(\mathbf{S}^i) & \partial_{i+1\,i+1} f^i(\mathbf{S}^i) \end{pmatrix} \begin{pmatrix} \beta_i \\ -\beta_{i+1} \end{pmatrix}$$

$$\geq (C-1)^{-2} \left\{ -(C-1)\beta_1^2 (A_1^1)^2 - (C-1)\beta_M^2 (A_M^{M-1})^2 \right\}$$

$$+ \sum_{i=1}^{M-1} \begin{pmatrix} \beta_i \\ -\beta_{i+1} \end{pmatrix}^T \left\{ \right.$$
$$- \begin{pmatrix} (A_i^i)^2 & \alpha_{i+1\,i}^i (A_i^i)^2 + \alpha_{i\,i+1}^i (A_{i+1}^i)^2 \\ \alpha_{i+1\,i}^i (A_i^i)^2 + \alpha_{i\,i+1}^i (A_{i+1}^i)^2 & (A_{i+1}^i)^2 \end{pmatrix}$$
$$+ \sum_{r=1}^K \begin{pmatrix} (\alpha_{ir}^i)^2 & \alpha_{ir}^i \alpha_{i+1\,r}^i \\ \alpha_{ir}^i \alpha_{i+1\,r}^i & (\alpha_{i+1\,r}^i)^2 \end{pmatrix} Q_r^i \left.\right\} \begin{pmatrix} \beta_i \\ -\beta_{i+1} \end{pmatrix}$$

$$\geq -(C-1)^{-1} \left\{ \beta_1^2 (A_1^1)^2 + \beta_M^2 (A_M^{M-1})^2 \right\}$$

$$+ \sum_{i=1}^{M-1} \left\{ (-\beta_i^2 + 2\beta_i \beta_{i+1} \alpha_{i+1\,i}^i)(A_i^i)^2 \right.$$
$$\left. + (-\beta_{i+1}^2 + 2\beta_i \beta_{i+1} \alpha_{i\,i+1}^i)(A_{i+1}^i)^2 \right\}, \quad (38)$$

where the first inequality follows from (33), (34) and Proposition 3.1, and the last inequality is due to $Q_k^i \geq 0$ and

$$\begin{pmatrix} \beta_i \\ -\beta_{i+1} \end{pmatrix}^T \begin{pmatrix} (\alpha_{ir}^i)^2 & \alpha_{ir}^i \alpha_{i+1\,r}^i \\ \alpha_{ir}^i \alpha_{i+1\,r}^i & (\alpha_{i+1\,r}^i)^2 \end{pmatrix} \begin{pmatrix} \beta_i \\ -\beta_{i+1} \end{pmatrix} \geq 0.$$

When $M = 2$, substituting (36) to (38), we have

$$\sum_{n=1}^{N}(v^n)^T\nabla^2 f^n(\mathbf{S}^n)v^n$$

$$\geq -(C-1)^{-1}\Big\{\beta_1^2(A_1^1)^2 + \beta_2^2(A_2^1)^2\Big\}$$

$$+\Big\{(-\beta_1^2 + 2\beta_1\beta_2\alpha_{21}^1)(A_1^1)^2 + (-\beta_2^2 + 2\beta_1\beta_2\alpha_{12}^1)(A_2^1)^2\Big\}$$

$$= (1+(C-1)^{-1})\Big\{4\alpha_{12}^1\alpha_{21}^1 - (1+(C-1)^{-1})^2\Big\}(A_1^1)^2$$

$$> 0,$$

where the last inequality follows from $\alpha_{12}^1\alpha_{21}^1 > \frac{1}{4}(1+\frac{1}{C-1})^2$. When $M \geq 3$, substituting (37) to (38), we obtain

$$\sum_{n=1}^{N}(v^n)^T\nabla^2 f^n(\mathbf{S}^n)v^n$$

$$\geq -(C-1)^{-1}\Big\{\beta_1^2(A_1^1)^2 + \beta_M^2(A_M^{M-1})^2\Big\}$$

$$+ \sum_{i=1}^{M-2}\Big\{(4\alpha_{i\,i+1}^i\,\alpha_{i+1\,i}^i - 1)\beta_i^2(A_i^i)^2\Big\}$$

$$+\Big\{(4\alpha_{M-1\,M}^{M-1}\,\alpha_{M\,M-1}^{M-1} - 1)\beta_M^2(A_M^{M-1})^2\Big\}$$

$$= \Big\{4\alpha_{12}^1\alpha_{21}^1 - (1+(C-1)^{-1})\Big\}\beta_1^2(A_1^1)^2$$

$$+\Big\{4\alpha_{M-1\,M}^{M-1}\,\alpha_{M\,M-1}^{M-1} - (1+(C-1)^{-1})\Big\}\beta_M^2(A_M^{M-1})^2$$

$$+ \begin{cases} 0 & (M=3) \\ \displaystyle\sum_{i=2}^{M-2}\Big\{(4\alpha_{i\,i+1}^i\,\alpha_{i+1\,i}^i - 1)\beta_i^2(A_i^i)^2\Big\} & (M\geq 4) \end{cases}$$

$$> 0,$$

where the last inequality follows from $\alpha_{lk}^n\alpha_{kl}^n > \frac{1}{4}(1+\frac{1}{C-1})^2 > \frac{1}{4}(1+\frac{1}{C-1}) > 1/4$, $\beta_k > 0$, and $A_k^n > 0$ for any k, l, and n. Since v does not satisfy (8), the second order necessary conditions do not hold, and hence, S cannot be a local maximum. This completes the proof.

APPENDIX E
PROOF OF THEOREM 3.2

For simplicity, we omit the superscript n (tone index) and denote $f_k := \partial_k f^n(\mathbf{S}^n)$ and $f_{kl} := \partial_{kl}f^n(\mathbf{S}^n)$ for $(k,l) \in \{1,2\}\times\{1,2\}$. Then, from Definition 3.1, we only have to show that

$$\begin{pmatrix} f_2 \\ -f_1 \end{pmatrix}^T\begin{pmatrix} f_{11} & f_{12} \\ f_{12} & f_{22} \end{pmatrix}\begin{pmatrix} f_2 \\ -f_1 \end{pmatrix}$$

$$= f_1^2 f_{22} + f_2^2 f_{11} - 2f_1 f_2 f_{12} > 0$$

for any $S_1 \geq 0$ and $S_2 \geq 0$. By Proposition 3.1, the first and second derivatives can be written as follows:

$$f_1 = \frac{1}{X_1 + S_1} - \alpha_{12}\left(\frac{1}{X_2} - \frac{1}{X_2 + S_2}\right),$$

$$f_2 = \frac{1}{X_2 + S_2} - \alpha_{21}\left(\frac{1}{X_1} - \frac{1}{X_1 + S_1}\right),$$

$$f_{11} = -\frac{1}{(X_1 + S_1)^2} + \alpha_{12}^2\left\{\frac{1}{X_2^2} - \frac{1}{(X_2 + S_2)^2}\right\},$$

$$f_{22} = -\frac{1}{(X_2 + S_2)^2} + \alpha_{21}^2\left\{\frac{1}{X_1^2} - \frac{1}{(X_1 + S_1)^2}\right\},$$

$$f_{12} = -\frac{\alpha_{21}}{(X_1 + S_1)^2} - \frac{\alpha_{12}}{(X_2 + S_2)^2}.$$

Letting $\Gamma := X_1^2 X_2^2 (X_2 + S_2)^4 (X_1 + S_1)^4 > 0$, we have

$$\Gamma f_1^2 f_{22} = \Big\{X_2(X_2 + S_2) - \alpha_{12}S_2(X_1 + S_1)\Big\}^2$$

$$\Big\{\alpha_{21}^2 S_1(2X_1 + S_1)(X_2 + S_2)^2 - X_1^2(X_1 + S_1)^2\Big\},$$

$$\Gamma f_2^2 f_{11} = \Big\{X_1(X_1 + S_1) - \alpha_{21}S_1(X_2 + S_2)\Big\}^2$$

$$\Big\{\alpha_{12}^2 S_2(2X_2 + S_2)(X_1 + S_1)^2 - X_2^2(X_2 + S_2)^2\Big\},$$

$$-2\Gamma f_1 f_2 f_{12} = 2X_1 X_2\Big\{X_2(X_2 + S_2) - \alpha_{12}S_2(X_1 + S_1)\Big\}$$

$$\Big\{X_1(X_1 + S_1) - \alpha_{21}S_1(X_2 + S_2)\Big\}$$

$$\Big\{\alpha_{12}(X_1 + S_1)^2 + \alpha_{21}(X_2 + S_2)^2\Big\}.$$

Summing up the above three equations, we have

$$\Gamma(f_1^2 f_{22} + f_2^2 f_{11} - 2f_1 f_2 f_{12})$$

$$= 2(X_1 + S_1)^2(X_2 + S_2)^2(X_1 X_2 - \alpha_{12}\alpha_{21}S_1 S_2)$$

$$(\alpha_{12}\alpha_{12}\sigma_2 S_1 + \alpha_{12}\alpha_{21}\sigma_1 S_2$$

$$+ \alpha_{12}\sigma_1^2 + \alpha_{21}\sigma_2^2 - \sigma_1\sigma_2) \quad (39)$$

$$= 2(X_1 + S_1)^2(X_2 + S_2)^2(X_1 X_2 - \alpha_{12}\alpha_{21}S_1 S_2)\,\sigma_1\sigma_2$$

$$\Big\{\alpha_{12}\alpha_{21}\Big(\frac{S_1}{\sigma_1} + \frac{S_2}{\sigma_2}\Big) + \Big(\alpha_{12}\Big(\frac{\sigma_1}{\sigma_2}\Big) + \alpha_{21}\Big(\frac{\sigma_2}{\sigma_1}\Big) - 1\Big)\Big\},$$

where the equality (39) is obtained by a straightforward factorization. Since $\Gamma > 0$ and $X_1 X_2 - \alpha_{12}\alpha_{21}S_1 S_2 > 0$ from the definitions of X_1 and X_2, we have $f_1^2 f_{22} + f_2^2 f_{11} - 2f_1 f_2 f_{12} > 0$ for any $S_1 \geq 0$ and $S_2 \geq 0$.

APPENDIX F
PROOF OF PROPOSITION 3.6

Since $|\mathcal{N}_{II}(\mathbf{S})| \geq 2$, there exist two tones n_1 and n_2 such that $S_1^{n_1} > 0$, $S_2^{n_1} > 0$, $S_1^{n_2} > 0$, and $S_2^{n_2} > 0$. Assume to the contrary that $\mathbf{S}$ is a local maximum of the sum-rate maximization problem (4). We will derive a contradiction. Then, from the KKT conditions (5) and (6), we have

$$0 \leq \lambda_1 = \partial_1 f^{n_1}(\mathbf{S}^{n_1}) = \partial_1 f^{n_2}(\mathbf{S}^{n_2}),$$

$$0 \leq \lambda_2 = \partial_2 f^{n_1}(\mathbf{S}^{n_1}) = \partial_2 f^{n_2}(\mathbf{S}^{n_2}).$$

Now, let $v \in \Re^{2N}$ be defined as

$$\begin{cases} v_1^{n_1} = \lambda_2, & v_2^{n_1} = -\lambda_1, \\ v_1^{n_2} = -\lambda_2, & v_2^{n_2} = \lambda_1, \\ v_1^n = v_2^n = 0 & (\forall n \in \mathcal{N}\setminus\{n_1, n_2\}). \end{cases}$$

Then, it is obvious that v satisfies (7). Since f^{n_1} and f^{n_2} are quasi-convex, we have

$$\sum_{n=1}^{N} (v^n)^T \nabla^2 f^n(\mathbf{S}^n) v^n$$

$$= \begin{pmatrix} \lambda_2 \\ -\lambda_1 \end{pmatrix}^T \nabla^2 f^{n_1}(\mathbf{S}^{n_1}) \begin{pmatrix} \lambda_2 \\ -\lambda_1 \end{pmatrix} + \begin{pmatrix} -\lambda_2 \\ \lambda_1 \end{pmatrix}^T \nabla^2 f^{n_2}(\mathbf{S}^{n_2}) \begin{pmatrix} -\lambda_2 \\ \lambda_1 \end{pmatrix}$$

$$= \begin{pmatrix} \partial_2 f^{n_1}(\mathbf{S}^{n_1}) \\ -\partial_1 f^{n_1}(\mathbf{S}^{n_1}) \end{pmatrix}^T \nabla^2 f^{n_1}(\mathbf{S}^{n_1}) \begin{pmatrix} \partial_2 f^{n_1}(\mathbf{S}^{n_1}) \\ -\partial_1 f^{n_1}(\mathbf{S}^{n_1}) \end{pmatrix}$$

$$+ \begin{pmatrix} -\partial_2 f^{n_2}(\mathbf{S}^{n_2}) \\ \partial_1 f^{n_2}(\mathbf{S}^{n_2}) \end{pmatrix}^T \nabla^2 f^{n_2}(\mathbf{S}^{n_2}) \begin{pmatrix} -\partial_2 f^{n_2}(\mathbf{S}^{n_2}) \\ \partial_1 f^{n_2}(\mathbf{S}^{n_2}) \end{pmatrix}$$

$$> 0,$$

which contradicts Proposition 2.2. Hence, $\mathbf{S}$ cannot be a local maximum.

APPENDIX G
PROOF OF PROPOSITION 4.1

We first show the KKT conditions (5). Choose $(n, k) \in \mathcal{N} \times \mathcal{K}$ arbitrarily, and let $\lambda_k := \gamma_k^{-1}$. From Assumption B and (16), we have $\lambda_k > 0$. Moreover, we have

$$P_k - \sum_{n=1}^{N} S_k^n = P_k - \sum_{n \in \mathcal{L}_k} (\gamma_k - \sigma_k^n)$$

$$= P_k - \left\{ |\mathcal{L}_k| \cdot \frac{P_k + \sum_{n \in \mathcal{L}_k} \sigma_k^n}{|\mathcal{L}_k|} - \sum_{n \in \mathcal{L}_k} \sigma_k^n \right\}$$

$$= 0.$$

Hence, the first part of (5) holds. Next we show the second part. If $n \in \mathcal{L}_k$, then it follows from (16) and Assumption B that $S_k^n = \gamma_k - \sigma_k^n > 0$. Therefore, we have

$$\lambda_k - \partial_k f^n(\mathbf{S}^n) = \frac{1}{\gamma_k} - \frac{\partial}{\partial S_k^n} \log \left(1 + \frac{S_k^n}{\sigma_k^n} \right)$$

$$= \frac{1}{S_k^n + \sigma_k^n} - \frac{1}{S_k^n + \sigma_k^n}$$

$$= 0,$$

where the first equality follows since tone n is used by only user k, i.e., $\mathcal{T}^n(\mathbf{S}^n) = \{k\}$, and the second equality is due to (16). If $n \notin \mathcal{L}_k$, then $S_k^n = 0$ and there exists another player $l \in \mathcal{K}$ such that $l \neq k$ and $n \in \mathcal{L}_l$ since $\mathcal{L} \in \mathcal{FDM}$. Moreover, we have

$$\lambda_k - \partial_k f^n(\mathbf{S}^n)$$

$$= \gamma_k^{-1} - (A_k^n - \alpha_{kl}^n P_l^n)$$

$$= \frac{1}{\gamma_k} - \left\{ \frac{1}{\sigma_k^n + \alpha_{lk}^n (\gamma_l - \sigma_l^n)} - \alpha_{kl}^n \left(\frac{1}{\sigma_l^n} - \frac{1}{\gamma_l} \right) \right\}$$

$$\geq 0,$$

where the first equality follows from Proposition 3.1 and $P_r^n = 0$ for all $r \neq l$, the second equality follows from $X_k^n = \sigma_k^n + \alpha_{lk}^n S_l^n$, $X_l^n = \sigma_l^n$, and $S_l^n = \gamma_l - \sigma_l^n$ from (16), and the inequality follows from (17). Hence, we obtain the second part of (5).

Next we show the second order sufficient conditions. Let $v \in \Re^{NK}$ be an arbitrary nonzero vector satisfying (10), and

choose $n \in \mathcal{N}$ arbitrarily. Then, there exists $k_n \in \mathcal{K}$ such that $\mathcal{T}^n(\mathbf{S}^n) = \{k_n\}$ from $\mathcal{L} \in \mathcal{FDM}$ and $\mathcal{N}_O(S) = \emptyset$. From (10), we have $v_k^n = 0$ for any $k \notin \mathcal{T}^n(\mathbf{S}^n)$, which together with $\mathcal{T}^n(\mathbf{S}^n) = \{k_n\}$ yields

$$(v^n)^T \nabla^2 f^n(\mathbf{S}^n) v^n = (v_{k_n}^n)^2 \partial_{k_n k_n} f^n(\mathbf{S}^n)$$
$$= -(v_{k_n}^n)^2 (B_{k_n}^n)^2, \qquad (40)$$

the last equality is due to Proposition 3.1 and $Q_k^n = (A_k^n)^2 - (B_k^n)^2 = 0$ for all $k \neq k_n$. Since $v \neq 0$ and $B_{k_n}^n > 0$, summing up (40) from $n = 1$ to N, we obtain (9).

APPENDIX H
PROOF OF THEOREM 4.1

Let $\mathcal{L} \in \mathcal{FDM}$ be an arbitrary tone allocation set such that $\min_{k \in \mathcal{K}} |\mathcal{L}_k| \geq C$. Then, we have $\max_{k \in \mathcal{K}} |\mathcal{L}_k| \leq N - (K - 1)C$ since $\min_{k \in \mathcal{K}} |\mathcal{L}_k| \geq C$ and $\sum_{k \in \mathcal{K}} |\mathcal{L}_k| = N$.

Choose $(k, l) \in \mathcal{K} \times \mathcal{K}$ with $k \neq l$ arbitrarily. Then, it can be easily seen

$$\gamma_0 \leq \gamma_k \leq \gamma_M.$$

Moreover, from (18), we have

$$\gamma_k \geq \gamma_0 > \sigma_M \geq \max_{n \in \mathcal{L}_k} \sigma_k^n,$$

that is, Assumption B holds. Hence, it suffices to show (19) $\Rightarrow$ (17) and (20) $\Rightarrow$ (19).

First we show (19) $\Rightarrow$ (17). For an arbitrary $n \in \mathcal{L}_l$, we have from (19)

$$\alpha_{lk}^n \alpha_{kl}^n \gamma_M (\gamma_0 - \sigma_M)^2 + (\alpha_{kl}^n \gamma_M \sigma_0 + \alpha_{lk}^n \gamma_0 \sigma_M)(\gamma_0 - \sigma_M)$$
$$\geq \sigma_M \gamma_0 (\gamma_M - \sigma_0),$$

which can be equivalently written as

$$\frac{1}{\sigma_0 + \alpha_{lk}^n (\gamma_0 - \sigma_M)} - \alpha_{kl}^n \left(\frac{1}{\sigma_M} - \frac{1}{\gamma_0} \right) \leq \frac{1}{\gamma_M}.$$

Hence, we have (17).

Next we show (20) $\Rightarrow$ (19). From (20), we have

$$\frac{P_0}{N - (K-1)C} \geq \left(\frac{1}{A_0} + \frac{1}{\sqrt{A_0}} + 1 \right) \sigma_M$$

$$\geq \left(\frac{1}{A_0} + \frac{1}{\sqrt{A_0}} + 1 \right) \sigma_M - \left(1 + \frac{\alpha_0}{A_0} \right) \sigma_0,$$

which is identical to

$$\frac{P_0}{N - (K-1)C} + \sigma_0 - \sigma_M \geq \frac{(\sigma_M - \alpha_0 \sigma_0) + \sigma_M \sqrt{A_0}}{A_0}.$$

Since $\gamma_0 = P_0/(N - (K-1)C) + \sigma_0$ and $2\alpha + 2\beta \geq \alpha + (\alpha^2 + 4\beta^2)^{1/2}$ for any $\alpha \geq 0$ and $\beta \geq 0$, we have

$$\gamma_0 - \sigma_M \geq \frac{(\sigma_M - \alpha_0 \sigma_0) + \sqrt{(\sigma_M - \alpha_0 \sigma_0)^2 + 4 A_0 \sigma_M^2}}{2 A_0},$$

or equivalently

$$A_0 \gamma_M (\gamma_0 - \sigma_M)^2 + \alpha_0 (\sigma_0 \gamma_M + \sigma_M \gamma_0)(\gamma_0 - \sigma_M)$$
$$\geq \sigma_M \gamma_0 (\gamma_M - \sigma_0) + \alpha_0 \gamma_0 \sigma_M (\gamma_0 - \sigma_M) + \sigma_M \sigma_0 \gamma_0.$$

Since $\alpha_0 \gamma_0 \sigma_M (\gamma_0 - \sigma_M) + \sigma_M \sigma_0 \gamma_0 \geq 0$, we have (19).

APPENDIX I
PROOF OF THEOREM 5.1

Given an even integer N and a set of N positive integers $a_1, a_2, ..., a_N$, we construct a two-user communication system as follows: let there be a total of N frequency tones, and let the channel noise powers for the two users be $\sigma_1^n = \sigma_2^n = a_n$, for $n = 1, 2, ..., N$. We also set the crosstalk coefficients $\alpha_{12}^n = \alpha_{21}^n = 1.01$ for all n, and let $P_1 = P_2 = P := (N+1)^3 \sigma_M$, with $\sigma_M := \max_n a_n$. In this case, problem (21) is reduced to the following:

$$\underset{\mathbf{S}=(\mathbf{S}_1,\mathbf{S}_2)\in\Re^{2N}}{\text{maximize}} \quad \sum_{n=1}^{N} \log\left(1 + \frac{S_1^n}{a_n}\right) + \sum_{n=1}^{N} \log\left(1 + \frac{S_2^n}{a_n}\right)$$

$$\text{subject to} \quad \mathbf{S} \in \mathcal{S}, \ \sum_{n=1}^{N} S_1^n \leq P, \ \sum_{n=1}^{N} S_2^n \leq P, \quad (41)$$

where we denote the optimal value by R_{fdma}. Let us consider a convex relaxation of (41) by dropping the nonconvex FDMA constraint $\mathbf{S} \in \mathcal{S}$, and by combining the two separate power constraints as a single one:

$$\underset{(S^1,S^2,...,S^N)\in\Re^N}{\text{maximize}} \quad \sum_{n=1}^{N} \log\left(1 + \frac{S^n}{a_n}\right)$$

$$\text{subject to} \quad \sum_{n=1}^{N} S^n \leq 2P, \ \ S^n \geq 0, \ \ \forall n, \quad (42)$$

where we denote the optimal value by R_{relax}. Notice that the relaxed problem (42) is a standard single user sum-rate maximization problem, so R_{relax} can be evaluated easily using the water-filling procedure (or the classical Karush-Kuhn-Tucker optimality condition). Specifically, there exists some $\gamma > 0$ (water level) such that

$$\sum_{n=1}^{N} S^n = 2P, \ \ S^n + a_n = \gamma, \quad \forall n,$$

where we have used the fact that the power level $2P = 2(N+1)^3 \max_n a_n$ is large enough so that $S^n > 0$ for all n. The above condition further implies

$$2P + \sum_{n=1}^{N} a_n = N\gamma. \quad (43)$$

Thus, we obtain

$$R_{relax} = \sum_{n=1}^{N} \log\left(1 + \frac{S^n}{a_n}\right) = \sum_{n=1}^{N} \log\left(\frac{a_n + S^n}{a_n}\right)$$

$$= \sum_{n=1}^{N} \log\left(\frac{\gamma}{a_n}\right) = N\log\gamma - \sum_{n=1}^{N} \log a_n$$

$$= N\log\left(\frac{2P + \sum_{n=1}^{N} a_n}{N}\right) - \sum_{n=1}^{N} \log a_n, \quad (44)$$

where the last step (44) follows from (43).

Our main claim is that

$$R_{fdma} \leq R_{relax} = N\log\left(\frac{2P + \sum_{n=1}^{N} a_n}{N}\right) - \sum_{n=1}^{N} \log a_n$$

and the equality holds if and only if the equipartition problem has a 'yes' answer. This claim implies the NP-hardness of the optimal bandwidth allocation problem (21) in the two-user case. To establish this claim, let us consider a globally optimal solution $(\mathbf{S}_1, \mathbf{S}_2) \in \Re^{2N}$ of (41). Due to the FDMA constraint, the sets $\mathcal{T}_1(\mathbf{S}_1)$ and $\mathcal{T}_2(\mathbf{S}_2)$ form a partition of $\mathcal{N} = \{1, 2, ..., N\}$, where $\mathcal{T}_1(\mathbf{S}_1) := \{n \,|\, S_1^n > 0\} \subseteq \mathcal{N}$ and $\mathcal{T}_2(\mathbf{S}_2) := \{n \,|\, S_2^n > 0\} \subseteq \mathcal{N}$. Since $(\mathbf{S}_1, \mathbf{S}_2)$ is a sum-rate optimal solution of (41), it follows that the subvectors $\{S_1^n\}_{n\in\mathcal{T}_1(\mathbf{S}_1)}$, $\{S_2^n\}_{n\in\mathcal{T}_2(\mathbf{S}_2)}$ must be the water-filling solutions of corresponding sum-rate maximization subproblems over the tones in $\mathcal{T}_1(\mathbf{S}_1)$ and $\mathcal{T}_2(\mathbf{S}_2)$ respectively. In other words, there exist two positive constants (water levels) γ_1, γ_2 such that

$$\sum_{n\in\mathcal{T}_k(\mathbf{S}_k)} S_k^n = P, \ \ S_k^n + a_n = \gamma_k, \ \ \forall n \in \mathcal{T}_k(\mathbf{S}_k) \ \text{ and } \ k = 1, 2,$$

where we have used the fact that the power level $P = (N+1)^3 \sigma_M = (N+1)^3 \max_n a_n$ is sufficiently large so that $S_k^n > 0$ for all $n \in \mathcal{T}_k(\mathbf{S}_k)$ and $k = 1, 2$. Simple algebraic manipulations of above condition show

$$P + \sum_{n\in\mathcal{T}_k(\mathbf{S}_k)} a_n = N_k\gamma_k, \quad k = 1, 2, \quad (45)$$

so that

$$2P + \sum_{n=1}^{N} a_n = N_1\gamma_1 + N_2\gamma_2, \quad (46)$$

where $N_k := |\mathcal{T}_k(\mathbf{S}_k)|$ for $k = 1, 2$. Thus, we obtain

$$R_{fdma} = \sum_{n\in\mathcal{T}_1(\mathbf{S}_1)} \log\left(\frac{a_n + S_1^n}{a_n}\right) + \sum_{n\in\mathcal{T}_2(\mathbf{S}_2)} \log\left(\frac{a_n + S_2^n}{a_n}\right)$$

$$= \sum_{n\in\mathcal{T}_1(\mathbf{S}_1)} \log\left(\frac{\gamma_1}{a_n}\right) + \sum_{n\in\mathcal{T}_2(\mathbf{S}_2)} \log\left(\frac{\gamma_2}{a_n}\right)$$

$$= N_1 \log\gamma_1 + N_2 \log\gamma_2 - \sum_{n=1}^{N} \log a_n$$

$$\leq N\log\left(\frac{N_1\gamma_1 + N_2\gamma_2}{N}\right) - \sum_{n=1}^{N} \log a_n \quad (47)$$

$$= N\log\left(\frac{2P + \sum_{n=1}^{N} a_n}{N}\right) - \sum_{n=1}^{N} \log a_n = R_{relax},$$

where the last equality follows from (46), while the inequality (47) is due to the (strict) concavity of the $\log x$ function. Moreover, the inequality (47) holds with equality if and only if $\gamma_1 = \gamma_2$.

It remains to show that the condition $\gamma_1 = \gamma_2$ holds for some global optimal solution $(\mathbf{S}_1, \mathbf{S}_2)$ of (41) is equivalent to the equipartition problem having a 'yes' answer. Clearly, if the equipartition problem is a 'yes' instance, then there exists some $T \subset \{1, 2, ..., N\}$ with $|T| = N/2$ such that

$$\sum_{n\in T} a_n = \sum_{n\notin T} a_n = \frac{1}{2}\sum_{n=1}^{N} a_n. \quad (48)$$

Let us define

$$S_1^n = \begin{cases} \gamma - a_n, & \text{if } n \in T \\ 0, & \text{if } n \notin T, \end{cases} \quad S_2^n = \begin{cases} \gamma - a_n, & \text{if } n \notin T \\ 0, & \text{if } n \in T, \end{cases}$$

where $\gamma = (2P + \sum_{n=1}^{N} a_n)/N$. From these definitions and using (48), we can verify that $\mathcal{T}_1(\mathbf{S}_1) = T$, $\mathcal{T}_2(\mathbf{S}_2) = T^c = \mathcal{N}/T$, and $(\mathbf{S}_1, \mathbf{S}_2)$ is a feasible solution of (41) with an objective value of

$$\sum_{n \in T} \log\left(\frac{a_n + S_1^n}{a_n}\right) + \sum_{n \notin T} \log\left(\frac{a_n + S_2^n}{a_n}\right)$$
$$= N \log\left(\frac{2P + \sum_{n=1}^{N} a_n}{N}\right) - \sum_{n=1}^{N} \log a_n = R_{relax}.$$

By the fact that $R_{fdma} \leq R_{relax}$, we can conclude that $(\mathbf{S}_1, \mathbf{S}_2)$ is a global optimal solution of (41). Moreover, $\gamma_1 = \gamma_2 = \gamma$ in this case.

Conversely, if $\gamma_1 = \gamma_2$ for some global optimal solution $(\mathbf{S}_1, \mathbf{S}_2)$ of the optimal bandwidth allocation problem (41), then condition (45) shows

$$\frac{P + \sum_{n \in \mathcal{T}_1(\mathbf{S}_1)} a_n}{P + \sum_{n \in \mathcal{T}_2(\mathbf{S}_2)} a_n} = \frac{N_1}{N_2}.$$

We claim $N_1 = N_2$. Suppose this is not true, then we can assume without loss of generality that $N_1 < N/2 < N_2$, implying

$$\frac{P + \sum_{n \in \mathcal{T}_1(\mathbf{S}_1)} a_n}{P + \sum_{n \in \mathcal{T}_2(\mathbf{S}_2)} a_n} = \frac{N_1}{N_2} < 1 - \frac{1}{N}. \quad (49)$$

On the other hand, since $P = (N+1)^3 \sigma_M = (N+1)^3 \max_n a_n$, it follows that

$$\frac{P + \sum_{n \in \mathcal{T}_1(\mathbf{S}_1)} a_n}{P + \sum_{n \in \mathcal{T}_2(\mathbf{S}_2)} a_n} \geq \frac{(N+1)^3 \sigma_M}{(N+1)^3 \sigma_M + N \sigma_M} \geq 1 - \frac{N}{(N+1)^3},$$

which clearly contradicts with (49) for $N \geq 1$. Thus, we must have $N_1 = N_2$, which together with (45) and the fact $\gamma_1 = \gamma_2$ further imply

$$\sum_{n \in \mathcal{T}_1(\mathbf{S}_1)} a_n = \sum_{n \in \mathcal{T}_2(\mathbf{S}_2)} a_n.$$

This shows that the equipartition problem has a 'yes' answer with $T = \mathcal{T}_1(\mathbf{S}_1)$, which establishes the NP-hardness of the optimal bandwidth allocation problem (21) in the two-user case.

Finally, we argue that the sum-rate maximization problem (4) is NP-hard in the two-user case. To do so, we consider the same two-user communication system defined in the beginning of this proof, and show that the two optimization problems (4) and (21) are equivalent in this case. Let $(\mathbf{S}_1, \mathbf{S}_2)$ be an optimal power allocation for the corresponding sum-rate maximization problem (4). We first show that in this case $C := \min\{\mathcal{T}_1(\mathbf{S}_1), \mathcal{T}_2(\mathbf{S}_2)\} \geq 2$, that is, the minimum number of tones used by each user must be at least 2. Suppose the contrary and let us assume (without loss of generality) that user 1 uses only one frequency tone 1. In this case, user 2 is better off not to use frequency tone 1 since the noise plus interference power for user 2 at tone 1 is $a_1 + 1.01P$, which is greater than user 2's optimal water level $\gamma_2 = (P + \sum_{n=2}^{N} a_n)/(N-1)$ where $P = (N+1)^3 \max_n a_n$. Thus, the maximum achievable

sum-rate is upper bounded by

$$\log\left(1 + \frac{P}{a_1}\right) + \sum_{n=2}^{N} \log\left(1 + \frac{S_2^n}{a_n}\right)$$
$$= \log(a_1 + P) + (N-1) \log \frac{P + \sum_{n=2}^{N} a_n}{N-1} - \sum_{n=1}^{N} \log a_n$$
$$\leq N \log(P + N\sigma_M) - (N-1)\log(N-1) - \sum_{n=1}^{N} \log a_n$$
$$\leq N \log P - (N-1)\log(N-1) - \sum_{n=1}^{N} \log a_n + O(1/N), \quad (50)$$

where we have used the fact $P = (N+1)^3 \sigma_M$ and the water-filling property of $(S_2^2, S_2^3, ..., S_2^N)$. On the other hand, if we allocate the first $N/2$ tones to user 1 and the remaining $N/2$ tones to user 2, then the users achieve a sum-rate of

$$\sum_{n=1}^{N/2} \log\left(1 + \frac{S_1^n}{a_n}\right) + \sum_{n=1+N/2}^{N} \log\left(1 + \frac{S_2^n}{a_n}\right)$$
$$= \frac{N}{2} \log \frac{P + \sum_{n=1}^{N/2} a_n}{N/2} + \frac{N}{2} \log \frac{P + \sum_{n=1+N/2}^{N} a_n}{N/2}$$
$$- \sum_{n=1}^{N} \log a_n$$
$$\geq N \log P - N \log \frac{N}{2} - \sum_{n=1}^{N} \log a_n$$

which is strictly greater the rate (50) for large N. Thus, letting user 1 uses exactly one tone cannot be sum-rate optimal, implying $C \geq 2$ for our problem.

Since $C \geq 2$ and $\alpha_{12}^n = \alpha_{21}^n = 1.01$ for all n (by definition), it follows that

$$\alpha_{12}^n \alpha_{21}^n > 1 = \frac{1}{4}(1+1)^2 \geq \frac{1}{4}\left(1 + \frac{1}{C-1}\right)^2, \quad \forall n,$$

so Theorem 3.3 shows that the optimal power allocation strategy is FDMA. Thus, for the two-user communication system defined in the beginning of this proof, the sum-rate maximization problem (4) is equivalent to the optimal bandwidth allocation problem (21) which is NP-hard. This shows the NP-hardness of the sum-rate maximization (4) in the two-user case. The proof is complete.

REFERENCES

[1] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal Selected Areas in Communications*, vol. 23, pp. 201–220, 2005.
[2] W. Yu, G. Ginis, and J. M. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE Journal on Selected Areas in Communications*, vol. 20, pp. 1105–1115, 2002.
[3] R. Cendrillon, W. Yu, M. Moonen, J. Verliden, and T. Bostoen, "Optimal multi-user spectrum management for digital subscriber lines," *IEEE Transactions on Communications*, to appear.

[4] V. M. K. Chan and W. Yu, "Joint multiuser detection and optimal spectrum balancing for digital subscriber lines," *EURASIP Journal on Applied Signal Processing*, vol. 2006, Article ID 80941, 13 pages, 2006.

[5] R. Lui and W. Yu, "Low complexity near optimal spectrum balancing for digital subscriber lines," 2005, IEEE International Conference on Communications (ICC), Seoul, Korea.

[6] W. Yu, R. Lui, and R. Cendrillon, "Dual optimization methods for multiuser orthogonal frequency division multiplex systems," in *Proceedings of IEEE Global Communications Conference (Globecom)*, Dallas, USA, 2004, pp. 225–229.

[7] Z.-Q. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE Journal of Selected Topics in Signal Processing*, Special Issue on Signal Processing and Networking for Dynamic Spectrum Access, vol. 2, pp. 57–73, 2008.

[8] K. B. Song, S. T. Chung, G. Ginis, and J. M. Cioffi, "Dynamic spectrum management for next-generation DSL systems," *IEEE Communications Magazine*, vol. 40, pp. 101–109, 2002.

[9] S. T. Chung, S. J. Kim, J. Lee, and J. M. Cioffi, "A game-theoretic approach to power allocation in frequency-selective Gaussian interference channels," in *Proceedings of 2003 IEEE International Symposium on Information Theory*, Yokohama, Japan, 2003.

[10] Z.-Q. Luo and J.-S. Pang, "Analysis of iterative waterfilling algorithm for multiuser power control in digital subscriber lines," *EURASIP Journal on Applied Signal Processing*, vol. 2006, Article ID 24012, 10 pages, 2006.

[11] N. Yamashita and Z.-Q. Luo, "A nonlinear complementarity approach to multiuser power control for digital subscriber lines," *Optimization Methods and Software*, vol. 19, pp. 633–652, 2004.

[12] G. Scutari, D. Palomar, and S. Barbarossa, "Optimal linear precoding/multiplexing for wideband multipoint-to-multipoint systems based on game theory-part I: Nash equilibria," *IEEE Transactions on Signal Processing*, vol. 56, pp. 1230–1249, 2008.

[13] ——, "Optimal linear precoding/multiplexing for wideband multipoint-to-multipoint systems based on game theory-part II: algorithms," *IEEE Transactions on Signal Processing*, vol. 56, pp. 1250–1267, 2008.

[14] R. Etkin, A. Parekh, and D. Tse, "Spectrum sharing for unlicensed bands," in *Proceedings of First IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks*, 2005, pp. 251–258.

[15] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, 1991.

[16] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming*: *Theory and Algorithms*, 2nd ed. New York: John Wiley & Sons, 1993.

[17] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Athena Scientific, 1999.

[18] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[19] E. Kalvelagen, "Lagrangian relaxation with GAMS," GAMS Development Corporation, technical note, 2002, http://www.gams.com/~erwin/lagrel/lagrange.pdf.

[20] G. Scutari, D. Palomar, and S. Barbarossa, "Asynchronous iterative water-filling for Gaussian frequency-selective interference channels: a unified framework," in *Proceedings of Information Theory and Applications (ITA) Workshop*, San Diego, USA, 2007.

[21] F. P. Kelly, A. Maulloo, and D. Tan, "Rate control in communication networks: shadow prices, proportional fairness and stability," *Journal of the Operational Research Society*, vol. 49, pp. 237–252, 1998.

Zhi-Quan (Tom) Luo (F'07) received his B.Sc. degree in Applied Mathematics in 1984 from Peking University, Beijing, China. Subsequently, he was selected by a joint committee of American Mathematical Society and the Society of Industrial and Applied Mathematics to pursue Ph.D study in the United States. After an one-year intensive training in mathematics and English at the Nankai Institute of Mathematics, Tianjin, China, he studied in the Operations Research Center and the Department of Electrical Engineering and Computer Science at MIT, where he received a Ph.D degree in Operations Research in 1989. From 1989 to 2003, Dr. Luo held a faculty position with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada, where he eventually became the department head and held a Canada Research Chair in Information Processing. Since April of 2003, he has been with the Department of Electrical and Computer Engineering at the University of Minnesota (Twin Cities) as a full professor and holds an endowed ADC Chair in digital technology. His research interests include optimization algorithms, data communication and signal processing.

Dr. Luo serves on the IEEE Signal Processing Society Technical Committees on Signal Processing Theory and Methods (SPTM), and on the Signal Processing for Communications (SPCOM). He is a recipient of the 2004 IEEE Signal Processing Society's Best Paper Award, and has held editorial positions for several international journals including Journal of Optimization Theory and Applications, SIAM Journal on Optimization, Mathematics of Computation, and IEEE Transactions on Signal Processing. He currently serves on the editorial boards for a number of international journals including Mathematical Programming and Mathematics of Operations Research.
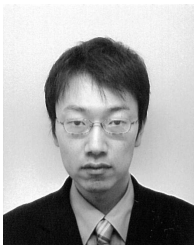
Shunsuke Hayashi received his B. Eng. degree from Faculty of Engineering, the University of Tokyo, Japan, in 2000, and M. Inf. and D. Inf. degrees from Graduate School of Informatics, Kyoto University, Japan, in 2002 and 2005, respectively.

He was a postdoctoral fellow at University of Minnesota, USA, from April to September 2005, and at National Cheng Kung University, Taiwan, from October to December 2005. He has been with Graduate School of Informatics, Kyoto University as an assistant professor from December 2005 to present. His research interests are in optimization algorithms, game theory, and dynamic spectrum management.

Recently, Dr. Hayashi has received the Best Presentation Award from the Operations Research Society of Japan (ORSJ) during the Seminar on Advanced Computation and Optimization (S@CO) in Tsukuba, 2008.