OPTIMAL FDMA OVER WIRELESS FADING MOBILE AD-HOC NETWORKS

Alejandro Ribeiro and Georgios B. Giannakis

Dept. of ECE, University of Minnesota, Minneapolis, MN 55455, USA

ABSTRACT

We formulate a frequency-division multiple access (FDMA) networking problem for wireless mobile ad-hoc networks (MANETS) to jointly optimize end-to-end user rates, routes, link capacities, transmitted power, frequency and power allocation across subcarriers and fading states. We show that the resulting non-convex optimization problem has zero duality gap. For some types of FDMA networks this result is exploited to reformulate the original problem into a (computationally tractable) convex optimization problem. We further exploit the lack of duality gap to show that conventional layering can be optimal in FDMA wireless MANETS. Specifically, if we select Lagrange multipliers appropriately, we can decompose the original problem in smaller sub-problems associated with the conventional networking layers. The solution of these per-layer optimization problems coincides with the solution of the originally formulated cross-layer optimization problem.

Index Terms— Wireless networking, cross-layer design, optimization

1. INTRODUCTION

It is widely recognized that networking in wireless mobile ad-hoc networks (MANETS) requires joint optimization of parameters that span the different layers of the traditional network-protocol stack. This optimization, alas, is computationally complex. Consequently, advances in cross-layer optimization usually rely on either restricting attention to a subset of layers, introducing approximations to render the problem tractable, or attempting to find an approximation, e.g., a local optimum, of the optimal solution.

One of the most comprehensive approaches to the wireless networking problem is the work in [1] where suitably modified versions of the back-pressure algorithm [2] are shown to approximate solutions of various wireless networking problems. Approaching the problem from an optimization perspective the use of the Lagrangian dual function is widespread. In particular, the Lagrangian function can be used to justify the decomposition of network optimization into layers and develop distributed implementations translating a mathematical optimization to a networking protocol; see e.g., [3] and references therein. It has to be noted that since wireless networking problems are not convex, the duality gap is not null. Thus, the dual optimum is different from the primal optimum. Lyapunov's convexity theorem [4], has been recently applied to the subcarrier allocation problem in digital subscriber lines (DSL) [5]. This problem, which bears resemblance to some particular wireless networking problems, is also non-convex and computationally intractable. Nonetheless, when formulated in continuous – as opposed to discrete – frequency the problem has zero duality gap [5].

This paper formulates a frequency-division multiple access (FDMA) networking problem for wireless MANETS to jointly optimize end-to-end user rates, routes, link capacities, transmitted power, frequency and power allocation across subcarriers and fading states. This problem is computationally difficult, and in the absence of fading known to be NP-hard [6]. One would expect that fading complicates matters further. Quite surprisingly, Lyapunov's convexity theorem can be used to show that in the presence of fading there is no duality gap. Therefore, the Lagrangian dual can be solved instead of the primal problem. For some types of FDMA networks this result can be used to reformulate the FDMA networking problem into a (tractable) convex optimization problem. Perhaps more important, the zero duality gap shows that conventional layering can be optimal in FDMA wireless MANETS. Specifically, if we select Lagrange multipliers appropriately, we can decompose the original problem in smaller sub-problems associated with the conventional networking layers. The solution of these per-layer optimization problems coincides with the solution of the originally formulated cross-layer optimization problem.

Email: {aribeiro,georgios}@ece.umn.edu.

2. PROBLEM FORMULATION

Consider an ad-hoc wireless network comprising J user terminals $\{T_i\}_{i=1}^{J}$. Terminal T_i wants to deliver packets for different application level flows generically denoted by k. Flow k specifies the destination of the flow's packets, but the same destination might be associated with different flows to, e.g., accommodate different types of traffic (video, voice or data). The destination of flow k is denoted by T^k to emphasize that flow indexing is different from terminal indexing. For every flow k, packet arrivals at T_i form a stationary stochastic process with mean a_i^k.

We model network connectivity with a graph $\mathcal{G}(v,e)$ with vertices $v := \{1,\dots,J\}$ and edges $e \in \mathcal{E}$ connecting pairs of vertices (i,j) when and only when T_i and T_j can communicate with each other. The adjacency of i is denoted $n(i) := \{j : (i,j) \in \mathcal{E}\}$. Each terminal $\{T_j\}_{j\in n(i)}$ that can communicate with T_i will be referred to as a neighbor and the set of all neighbors as T_i's neighborhood. Given this model, terminals rely on multi-hop transmissions to deliver packets to the intended destination T^k of the flow k. For that matter, T_i selects an average rate r_{ij}^k for transmitting k-th flow packets to T_j. Assuming that packets are not discarded and that queues are stable throughout the network, we can write a flow conservation equation to relate rates a_i^k of exogenous packet arrivals (from the application layer) and endogenous (to the network layer) average rates r_{ij}^k of transmission to and from neighboring nodes

$$a_i^k = \sum_{j\in n(i)} \left(r_{ij}^k - r_{ji}^k\right). \tag{1}$$

Consider now rates r_{ij}^k of all flows traversing the link $T_i \to T_j$. Denoting by c_{ij} the information capacity of this link, we ensure queue stability by requiring

$$\sum_k r_{ij}^k \le c_{ij}. \tag{2}$$

Constraints (1) and (2) are sufficient to describe traffic flow in a wireline network with fixed capacities c_{ij}. In such a case, T_i needs to determine exogenous arrival rates a_i^k and transmission rate variables r_{ij}^k to satisfy certain optimality criteria. In a wireless network however, c_{ij} is not a fixed resource given to the terminals. In fact, operating conditions are determined by a set of available frequencies $\mathcal{F}$ and power budgets $p_{\max i}$. Thus, in addition to a_i^k and r_{ij}^k, T_i has to decide how to split the power budget $p_{\max i}$ among tones $f \in \mathcal{F}$ and neighbors T_j, $j \in n(i)$. Matters are further complicated by fading as described in the next section.

2.1. Link capacities in FDMA wireless networks

For every frequency tone $f \in \mathcal{F}$ and $(i,j) \in \mathcal{E}$, let h_{ij}^f denote the channel gain from T_i to T_j. As is customary practice in wireless communications, h_{ij}^f is modeled as a random variable. The channel gains of all network links are collected in the vector $\mathbf{h}$, and all $\mathbf{h}$ realizations in the set $\mathcal{H}$.

For a given channel realization $\mathbf{h}$, let the indicator variable $\alpha_{ij}^f(\mathbf{h})$ equal 1 when T_i sends packets to T_j on tone f and 0 otherwise, i.e., $\alpha_{ij}^f(\mathbf{h})$ indicates the event that T_i chooses to transmit to T_j on the tone f when the channel vector realization is $\mathbf{h}$. When $\alpha_{ij}^f(\mathbf{h}) = 1$, $p_{ij}^f(\mathbf{h})$ denotes the power used for transmission in this link. For a given channel realization, the instantaneous total power $p_i(\mathbf{h})$ used by T_i is the sum of the power used

1-4244-1484-9/08/$25.00 ©2008 IEEE

to transmit to all selected neighbors in all selected tones, i.e.,

$$p_i(\mathbf{h}) := \sum_{j\in n(i)} \sum_{f\in\mathcal{F}} \alpha_{ij}^f(\mathbf{h}) p_{ij}^f(\mathbf{h}). \tag{3}$$

Integrating over all possible channel realizations yields the average p_i power used by T_i as

$$p_i := \mathrm{E}_{\mathbf{h}}\left[p_i(\mathbf{h})\right] = \mathrm{E}_{\mathbf{h}}\left[\sum_{j\in n(i)} \sum_{f\in\mathcal{F}} \alpha_{ij}^f(\mathbf{h}) p_{ij}^f(\mathbf{h})\right] \tag{4}$$

where $\mathrm{E}_{\mathbf{h}}[\cdot]$ denotes expectation over the channel probability distribution.

The rate of information transmission in the $T_i \to T_j$ link is a function of the power distributions $p_{ij}^f(\mathbf{h})$. In this paper, we assume that terminals avoid mutual interference (this constrains the values of the indicator variables $\alpha_{ij}^f(\mathbf{h})$ as we explain in Section 2.2) so that $c_{ij}^f(\mathbf{h})$ can be written as a function of $p_{ij}^f(\mathbf{h})$ only, i.e.,

$$c_{ij}^f(\mathbf{h}) := \alpha_{ij}^f(\mathbf{h}) C\left(h_{ij}^f p_{ij}^f(\mathbf{h})\right). \tag{5}$$

If, e.g., a capacity achieving code is used we can write $C[h_{ij}^f p_{ij}^f(\mathbf{h})] = \log[1 + h_{ij}^f p_{ij}^f(\mathbf{h})]$, where we have assumed that the noise power is 1, as can always be done, normalizing the channel gains h_{ij}^f if necessary. Another example entails the use of a finite number of adaptive modulation and coding (AMC) modes. In this case, $C(\cdot)$ is a staircase function defined by the rate of the AMC modes considered.

In any event, the capacity c_{ij} of the wireless link $T_i \to T_j$ is obtained after integrating over all possible channel realizations:

$$c_{ij} := \mathrm{E}_{\mathbf{h}}\left[\sum_{f\in\mathcal{F}} c_{ij}^f(\mathbf{h})\right] = \mathrm{E}_{\mathbf{h}}\left[\sum_{f\in\mathcal{F}} \alpha_{ij}^f(\mathbf{h}) C\left(h_{ij}^f p_{ij}^f(\mathbf{h})\right)\right]. \tag{6}$$

The average link capacity and power expressions in (4) and (6) with the flow and rate constraints in (1) and (2) describe a generic FDMA wireless networking problem. To complete the problem description we need to constrain the indicator variables $\alpha_{ij}^f(\mathbf{h})$ in order to prevent use of the same frequency tone by neighboring terminals. This is done in the next section.

2.2. Frequency separation

To ensure frequency separation, we have to constrain the link assignment variables $\alpha_{ij}^f(\mathbf{h})$. In principle, we want to ensure that if a certain tone f is being used for transmission in a certain link, then the same tone is *not* used simultaneously by nearby nodes. Consider an arbitrary terminal T_j and recall that $\alpha_{ij}^f(\mathbf{h}) = 1$ indicates that $T_i \in n(j)$ is transmitting to T_j on the tone $f \in \mathcal{F}$ when the channel realization is $\mathbf{h} \in \mathcal{H}$. To avoid interference, a tone f cannot be used to transmit to T_j by more than one of its neighbors. In terms of indicator variables, the latter implies that if $\alpha_{i_0 j}^f(\mathbf{h}) = 1$ for some $i_0 \in n(j)$, then for any other $i \neq i_0$ we must have $\alpha_{i_0 j}^f(\mathbf{h}) = 0$. Since the indicator variables $\alpha_{ij}^f(\mathbf{h}) \in \{0,1\}$, the latter can be written as

$$\sum_{i\in n(j)} \alpha_{ij}^f(\mathbf{h}) \leq 1. \tag{7}$$

For a given channel realization $\mathbf{h}$, (7) is satisfied if either no terminal transmits to T_j on the tone f, or only one terminal T_i does so. This is not sufficient to ensure lack of interference in any $T_i \to T_j$ link, since we also have the possibility of T_j transmitting to a neighbor T_i on the same tone, i.e., the tone f being used by some outgoing transmission from T_j. To prevent this from happening we have to assure that if $\alpha_{i_0 j}^f(\mathbf{h}) = 1$ for some $i_0 \in n(j)$ then we not only have $\alpha_{ij}^f(\mathbf{h}) = 0$ for $i \neq i_0$ but $\alpha_{ji}^f(\mathbf{h}) = 0$ (note the reversal of subindexes since this indicator variables are for T_j's transmissions) for all $i \in n(j)$. This can be guaranteed if the following is satisfied

$$\sum_{i\in n(j)} \alpha_{ij}^f(\mathbf{h}) + \sum_{i\in n(j)} \alpha_{ji}^f(\mathbf{h}) \leq 1. \tag{8}$$

So far, we assured that transmissions to T_j do not interfere with each other [cf. (7)] as neither do incoming and outgoing transmissions [cf. (8)]. We have to further prevent the possibility of some neighbor of T_j using a tone f to communicate with any other node when f is used to transmit to T_j. Again, in terms of indicator variables we have that if $\alpha_{ij}^f(\mathbf{h}) = 1$ for some $i \in n(j)$, i.e., T_j is receiving information on the tone f, then for any neighbor $k \in n(j)$ we must have $\alpha_{kl}^f(\mathbf{h}) = 0$, i.e., neighbors of T_j do not use f. This yields the constraints

$$\sum_{i\in n(j)} \alpha_{ij}^f(\mathbf{h}) + \sum_{l\in n(k), l\neq j} \alpha_{kl}^f(\mathbf{h}) \leq 1, \qquad \forall k \in n(j). \tag{9}$$

The constraints in (7)-(9) state that for a given channel realization $\mathbf{h}$ one of two things happens. If f is used for transmitting to T_j, then (7) guarantees that it is used only by one node so that the sum in (7) equals 1; since this sum is also the first term of (8) and (9), (8) ensures that the tone f is not used by T_j's transmissions; while the constraints in (9) ensure that f is not used by neighboring terminals. If f is *not* used for transmitting to T_j by any of its neighbors the sum in (7) is null as are the first terms of the sums in (8) and (9); the tone f may then be used for transmission by T_j itself and/or any or many of its neighbors as long as it is compatible with the FDMA constraints associated with other nodes.

With a properly defined $\mathbf{A}$, the constraints in (7)-(9) will be henceforth denoted as $\mathbf{A}\boldsymbol{\alpha}^f(\mathbf{h}) \leq \mathbf{1}$, where $\boldsymbol{\alpha}^f(\mathbf{h})$ is a vector containing all the link indicator variables $\alpha_{ij}^f(\mathbf{h})$ for a given tone f and channel realization $\mathbf{h}$. Note that $\mathbf{A}$ is the same for all f and $\mathbf{h}$.

3. OPTIMAL FDMA NETWORKING

We have now finished relating the variables in an FDMA networking problem. For link indicator variables $\alpha_{ij}^f(\mathbf{h})$ satisfying (7)-(9), link capacities c_{ij} and power consumption p_i depend on the chosen power profiles $p_{ij}^f(\mathbf{h})$ as per (4) and (6). The average link rates r_{ij}^k are then constrained by (2) and the end-to-end flow rates a_i^k by (1). Problem variables $\alpha_{ij}^f(\mathbf{h})$, $p_{ij}^f(\mathbf{h})$, c_{ij}, p_i, r_{ij}^k and a_i^k that satisfy these equations can be supported by the network. As network designers, we want to select out of this set of feasible variables those that are optimal in some sense. We thus introduce concave $U_i^k(a_i^k)$ and convex $V_i(p_i)$ functions, respectively representing the value of rate a_i^k and the cost of power p_i. Though not required, we expect $U_i^k(a_i^k)$ and $V_i(p_i)$ to be increasing functions of their arguments. We can thus define the optimal networking problem as [cf. (1), (2), (4), (6) and (7)-(9)]

$$P = \max \sum_{i,k} U_i^k(a_i^k) - \sum_i V_i(p_i) \tag{10}$$

$$c_{ij} \leq \mathrm{E}_{\mathbf{h}}\left[\sum_{f\in\mathcal{F}} \alpha_{ij}^f(\mathbf{h}) C\left(h_{ij}^f p_{ij}^f(\mathbf{h})\right)\right] \tag{11}$$

$$p_i \geq \mathrm{E}_{\mathbf{h}}\left[\sum_{j\in n(i)} \sum_{f\in\mathcal{F}} \alpha_{ij}^f(\mathbf{h}) p_{ij}^f(\mathbf{h})\right] \tag{12}$$

$$a_i^k \leq \sum_{j\in n(i)} \left(r_{ij}^k - r_{ji}^k\right), \qquad \sum_k r_{ij}^k \leq c_{ij} \tag{13}$$

$$\mathbf{A}\boldsymbol{\alpha}^f(\mathbf{h}) \leq \mathbf{1};\ \alpha_{ij}^f(\mathbf{h}) \in \{0,1\} \tag{14}$$

where we have relaxed the constraints (1), (4) and (6), which we can do without loss of optimality. Note that all problem variables have to be non-negative, but this is left implicit in (10). We have also left implicit power constraints $p_i \leq p_{\max i}$ and $p_{ij}^f(\mathbf{h}) \leq p_{\max}$, arrival rate requirements $a_{\min i}^k \leq a_i^k \leq a_{\max i}^k$ and upper bound constraints $c_{ij} \leq c_{\max}$ and $r_{ij}^k \leq r_{\max}$ on link capacities and link flow rates. We will henceforth refer to these constraints as box constraints.

Problem (10) is difficult to solve. The function $C(\cdot)$ is not concave in general and even if we restrict attention to concave functions, the products $\alpha_{ij}^f(\mathbf{h}) C\left(h_{ij}^f p_{ij}^f(\mathbf{h})\right)$ and $\alpha_{ij}^f(\mathbf{h}) p_{ij}^f(\mathbf{h})$ still pose computational challenges. As we will see, it is possible to reformulate (11) and (12) to avoid the latter. The unsurmountable complication, however, comes from the integer constraints (14) on the indicator variables $\alpha_{ij}^f(\mathbf{h}) \in \{0,1\}$. Indeed, if we remove the expected value operators in (11) and (12), i.e., if channels are deterministic and there is only one realization $\mathbf{h}$, it has been proved that the problem (10) is NP-hard, [6]. In light of this challenge, one is well justified to look for approximate solutions. This motivates the introduction of problem relaxations outlined in the next two sections.

3.1. Constraint relaxation

A constraint relaxation of (10) is obtained by replacing the constraint $\alpha_{ij}^f(\mathbf{h}) \in \{0,1\}$ in (14) with the constraint $\alpha_{ij}^f(\mathbf{h}) \in [0,1]$; i.e., instead of $\alpha_{ij}^f(\mathbf{h})$ being 0 or 1, we let it take values over the interval [0,1].

To take care of the non-convex products in (11) and (12), we introduce the variables $q_{ij}^f(\mathbf{h}) := \alpha_{ij}^f(\mathbf{h}) p_{ij}^f(\mathbf{h})$. The power constraint is thus rewritten as [cf. (12)]

$$p_i \geq \mathrm{E}_{\mathbf{h}}\left[\sum_{j\in n(j)} \sum_{f\in\mathcal{F}} q_{ij}^f(\mathbf{h})\right] \tag{15}$$

which is a simple linear function of $q_{ij}^f(\mathbf{h})$. The link capacity constraint (11) takes the form

$$c_{ij} \leq \mathrm{E}_{\mathbf{h}}\left[\sum_{f\in\mathcal{F}} \alpha_{ij}^f(\mathbf{h}) C\left(\frac{h_{ij}^f q_{ij}^f(\mathbf{h})}{\alpha_{ij}^f(\mathbf{h})}\right)\right]. \tag{16}$$

The expression $\alpha_{ij}^f(\mathbf{h}) C\left(h_{ij}^f q_{ij}^f(\mathbf{h})/\alpha_{ij}^f(\mathbf{h})\right)$ is the perspective of the function $C(\cdot)$. Since the perspective operator preserves concavity, for concave functions $C(\cdot)$ the expression in (16) defines a convex constraint in the problem variables.

The relaxed problem defined as the maximization of the objective in (10) subject to the constraints (15), (16), (13) and $\mathbf{A}\boldsymbol{\alpha}^f(\mathbf{h}) \leq \mathbf{1}$, $\alpha_{ij}^f(\mathbf{h}) \in [0,1]$ is a convex problem that can be efficiently solved using, e.g., interior point algorithms. Since relaxing the integer constraint enlarges the set of feasible variables, the maximum $\tilde{P}$ of the relaxed problem provides an upper bound on the maximum P of the original problem (10), i.e., $\tilde{P} \geq P$. The optimal arguments of the relaxed problem will in general have $\alpha_{ij}^f(\mathbf{h}) \notin \{0,1\}$. To obtain a feasible solution, we have to project the latter so that $\alpha_{ij}^f(\mathbf{h}) \in \{0,1\}$. This might be computationally difficult and might entail a significant performance loss.

3.2. Lagrangian relaxation

Lagrangian relaxation refers to the solution D of the dual problem of (10). Since (10) is non-convex the duality gap is nonzero and we have $D \geq P$. To define the dual problem, associate Lagrange multipliers λ_{ij} with the capacity constraints in (11), μ_i with the power in (12), and ν_i^k and ξ_{ij} with the flow and rate constraints in (13). To simplify notation call $\mathbf{X}$ the set of all primal variables – i.e., $\alpha_{ij}^f(\mathbf{h})$, $p_{ij}^f(\mathbf{h})$, c_{ij}, p_i, r_{ij}^k and a_i^k – and $\boldsymbol{\Lambda}$ the set of all dual variables –i.e., λ_{ij}, μ_i, ν_i^k, ξ_{ij} – and write the Lagrangian as

$$\begin{aligned}\mathcal{L}[\mathbf{X},\boldsymbol{\Lambda}] = &\sum_{i,k} U_i^k(a_i^k) - \sum_i V_i(p_i) \\ &+ \sum_{i,j}\lambda_{ij}\left[\mathrm{E}_{\mathbf{h}}\left[\sum_{f\in\mathcal{F}} \alpha_{ij}^f(\mathbf{h}) C\left(h_{ij}^f p_{ij}^f(\mathbf{h})\right)\right] - c_{ij}\right] \\ &+ \sum_i \mu_i\left[p_i - \mathrm{E}_{\mathbf{h}}\left[\sum_{j\in n(i)}\sum_{f\in\mathcal{F}} \alpha_{ij}^f(\mathbf{h}) p_{ij}^f(\mathbf{h})\right]\right] \\ &+ \sum_{i,k}\nu_i^k\left[\sum_{j\in n(i)}\left(r_{ij}^k - r_{ji}^k\right) - a_i^k\right] + \sum_{ij}\xi_{ij}\left[c_{ij} - \sum_k r_{ij}^k\right].\end{aligned} \tag{17}$$

The dual function is obtained by maximizing the Lagrangian over the primal variables

$$g[\boldsymbol{\Lambda}] = \max_{\mathbf{X}} \mathcal{L}[\mathbf{X},\boldsymbol{\Lambda}] \qquad \mathbf{A}\boldsymbol{\alpha}^f(\mathbf{h}) \leq \mathbf{1};\ \alpha_{ij}^f \in \{0,1\} \tag{18}$$

where as in (10) box constraints are implicit. We finally define the dual problem as

$$D = \min_{\boldsymbol{\Lambda}\geq 0} g[\boldsymbol{\Lambda}]. \tag{19}$$

It is interesting to note that the Lagrangian of the relaxed problem, i.e., (10) with $\alpha_{ij}^f(\mathbf{h}) \in [0,1]$, is also given by (17). The dual function is different, though, since the range over which we perform the Lagrangian maximization changes, i.e.,

$$\tilde{g}[\boldsymbol{\Lambda}] = \max_{\mathbf{X}} \mathcal{L}[\mathbf{X},\boldsymbol{\Lambda}] \qquad \mathbf{A}\boldsymbol{\alpha}^f(\mathbf{h}) \leq \mathbf{1};\ \alpha_{ij}^f(\mathbf{h}) \in [0,1]. \tag{20}$$

The dual problem is defined analogously yielding the relaxed dual optimum $\tilde{D} = \min_{\boldsymbol{\Lambda}\geq 0} \tilde{g}[\boldsymbol{\Lambda}]$.

The Lagrangian in (17) depends linearly on the indicator variables $\alpha_{ij}^f(\mathbf{h})$. Therefore, we expect the maximum in (20) to be achieved at a corner of the polyhedron defined by the constraints $\mathbf{A}\boldsymbol{\alpha}^f(\mathbf{h}) \leq \mathbf{1}$ and $\alpha_{ij}^f(\mathbf{h}) \in [0,1]$. Since points with $\alpha_{ij}^f(\mathbf{h}) \in \{0,1\}$ are corners of this feasible set the question arises if these are the only possible corners so that the maximum in (20) coincides with the maximum in (18), i.e., if $\tilde{g}[\boldsymbol{\Lambda}] = g[\boldsymbol{\Lambda}]$. This is not true in general, as is well known, but for the subclass of "totally unimodular" matrices [7, p. 572]. For this subclass we have $\tilde{g}[\boldsymbol{\Lambda}] = g[\boldsymbol{\Lambda}]$ something that, in particular, holds true at the minimum value of the dual functions. This argument establishes the following result[1].

Proposition 1 *Let P denote the maximum of the primal problem* (10)*, D the minimum of its dual in* (19)*, $\tilde{P}$ the maximum of the relaxed primal problem* (10) *with $\alpha_{ij}^f(\mathbf{h}) \in [0,1]$, and $\tilde{D} = \min_{\boldsymbol{\Lambda}\geq 0} \tilde{g}[\boldsymbol{\Lambda}]$ the minimum of the dual relaxed problem with $\tilde{g}[\boldsymbol{\Lambda}]$ as in* (20)*. If the capacity function $C(\cdot)$ is convex and the matrix $\mathbf{A}$ is totally unimodular, then it holds that*

$$P \leq D = \tilde{D} = \tilde{P}. \tag{21}$$

Given the intractability of (10) we approach the problem through three, in principle, different relaxations. The optimal values obtained for totally unimodular networks are equally good as approximations of the original problem. The relaxations are different in general and yield approximating *arguments* with different properties. When solving the relaxed problem we obtain infeasible indicator variables, i.e., $\alpha_{ij}^f(\mathbf{h}) \notin \{0,1\}$. When solving the dual problems, we obtain as a byproduct primal variables that optimize the Lagrangians [cf. (18) and (20)]. The indicator variables are feasible in this case, i.e., $\alpha_{ij}^f(\mathbf{h}) \in \{0,1\}$ but other constraints in the primal problem may be violated. The usefulness of any of the relaxed problems depends on the difficulty of obtaining a primal feasible solution given the solution of the relaxed problem and the (further) loss of optimality associated with this recovery.

4. OPTIMALITY OF DUAL AND CONSTRAINT RELAXATION

The challenges in solving (10) are now clear. For deterministic channels, the problem is known to be NP-hard. The relaxations discussed in Sections 3.1 and 3.2 are certainly useful in establishing upper bounds on the achievable utility, but might or might not shed light on the optimal problem variables. One would expect that introducing fading should complicate matters further. It is thus a remarkable fact that in the presence of fading the duality gap vanishes, i.e., $P = D$. We state this result in the following theorem.

Theorem 1 *Let P denote the maximum of the primal problem* (10)*, D the minimum of its dual in* (19)*, $\tilde{P}$ the maximum of the relaxed primal problem* (10) *with $\alpha_{ij}^f(\mathbf{h}) \in [0,1]$, and $\tilde{D} = \min_{\boldsymbol{\Lambda}\geq 0} \tilde{g}[\boldsymbol{\Lambda}]$ the minimum of the dual relaxed problem with $\tilde{g}[\boldsymbol{\Lambda}]$ as in* (20)*. If the channel cumulative distribution function (cdf) is continuous, then*

$$P = D. \tag{22}$$

If in addition the matrix $\mathbf{A}$ is totally unimodular, then

$$P = D = \tilde{D} = \tilde{P}. \tag{23}$$

[1] Proofs are available in the journal version of this paper.

The proof relies on Lyapunov's convexity theorem [4] first used in the context of physical layer spectrum management for deterministic DSL channels in [5]. Recall that the link capacity function $C(\cdot)$ is not necessarily concave in Theorem 1. Even if $C(\cdot)$ is concave, the optimization problem is still non-convex. The duality gap, however, is null. Continuity of the channel cdf ensures that no channel realization has positive probability. This is satisfied by commonly used channel models including Rayleigh, Rice and Nakagami fading.

When $\mathbf{A}$ is totally unimodular any set of primal variables $\mathbf{X}$ feasible for the problem (10) is also feasible for the relaxed problem with $\alpha_{ij}^{f}(\mathbf{h}) \in [0,1]$. This is in particular true for the arguments $\mathbf{X}^*$ that solve (10). But since $P = \tilde{P}$, $\mathbf{X}^*$ is also a solution of the relaxed problem. This implies that the optimal arguments $\tilde{\mathbf{X}}^*$ of the relaxed problem satisfy $\alpha_{ij}^{f}(\mathbf{h}) \in \{0,1\}$, and are thus a feasible point of (10). Consequently, the convex (thus computationally tractable) relaxed problem can be solved instead of the (difficult) primal problem.

Another implication of Theorem 1 is the optimality of conventional layering in wireless FDMA networking problems. As is usually the case, the Lagrangian exhibits a separable structure in the sense that it can be written as a sum of terms that depend on a few primal variables. Rearranging terms in (17) and assuming that the optimal dual argument $\mathbf{\Lambda}^*$ is available, we can write

$$\begin{aligned}\mathcal{L}[\mathbf{X},\mathbf{\Lambda}^*] = &\sum_{i,k}\Big(U_i^k(a_i^k) - \nu_j^{k*}a_j^k\Big) + \sum_i \Big(\mu_i^* p_i - V_i(p_i)\Big) \\ &+ \sum_{i,j}\Big(\xi_{ij}^* - \lambda_{ij}^*\Big)c_{ij} + \sum_{i,j,k}\Big(\nu_i^{k*} - \nu_j^{k*} - \xi_{ij}^*\Big)r_{ij}^k \\ &+ \sum_{i,j,f} \mathrm{E}_{\mathbf{h}}\left[\alpha_{ij}^f(\mathbf{h})\left[\lambda_{ij}^* C\left(h_{ij}^f p_{ij}^f(\mathbf{h})\right) - \mu_i^* p_{ij}^f(\mathbf{h})\right]\right].\end{aligned} \tag{24}$$

The zero duality gap implies that if $\mathbf{\Lambda}^*$ is known we can, instead of solving (10), solve the (separable) problem

$$P = D = g[\mathbf{\Lambda}^*] = \max_{\mathbf{X}} \mathcal{L}[\mathbf{X},\mathbf{\Lambda}^*] \tag{25}$$

where the maximization is constrained to the $\alpha_{ij}^f(\mathbf{h})$ satisfying $\mathbf{A}\boldsymbol{\alpha}^f(\mathbf{h}) \leq \mathbf{1}$ and $\alpha_{ij}^f(\mathbf{h}) \in \{0,1\}$.

Because primal variables are decoupled in the Lagrangian $\mathcal{L}[\mathbf{X},\mathbf{\Lambda}^*]$ [cf. (24)], the maximization in (25) can be split into smaller maximization problems. This separability can be used to prove the following theorem.

Theorem 2 *Let λ_{ij}^*, μ_i^*, ν_i^{k*}, and ξ_{ij}^* denote the optimal dual variables that solve* (19)*. Consider the sub-problems*

$$P(a_i^k) = \max_{a_{\min i}^k \leq a_i^k \leq a_{\max i}^k} \Big[U_i^k(a_i^k) - \nu_i^{k*} a_i^k\Big] \tag{26}$$

$$P(r_{ij}^k) = \max_{0 \leq r_{ij}^k \leq r_{\max}} \Big[\Big(\nu_i^{k*} - \nu_j^{k*} - \xi_{ij}^*\Big) r_{ij}^k\Big] \tag{27}$$

$$P(c_{ij}) = \max_{0 \leq c_{ij} \leq c_{\max}} \big[\left(\xi_{ij}^* - \lambda_{ij}^*\right) c_{ij}\big] \tag{28}$$

$$P(p_i) = \max_{0 \leq p_i \leq p_{\max i}} \left[\mu_i^* p_i - V_i(p_i)\right]. \tag{29}$$

Define further the optimal power allocation problem

$$\phi_{ij}^f(\mathbf{h}) = \max_{0 \leq p_{ij}^f(\mathbf{h}) \leq p_{\max}} \Big[\lambda_{ij}^* C\left(h_{ij}^f p_{ij}^f(\mathbf{h})\right) - \mu_i^* p_{ij}^f(\mathbf{h})\Big] \tag{30}$$

and the optimal frequency allocation problem

$$\begin{aligned} P(\mathbf{h}) = \max &\left[\sum_{i,j,f} \alpha_{ij}^f(\mathbf{h}) \phi_{ij}^f(\mathbf{h})\right] \\ &\mathbf{A}\boldsymbol{\alpha}^f(\mathbf{h}) \leq \mathbf{1},\ \ \alpha_{ij}^f(\mathbf{h}) \in \{0,1\}.\end{aligned} \tag{31}$$

Then, the optimal utility yield P in (10) *is given by*

$$P = \sum_{i,k} P(a_i^k) + \sum_{i,j,k} P(r_{ij}^k) + \sum_{i,j} P(c_{ij}) + \sum_i P(p_i) + \mathrm{E}_{\mathbf{h}}\left[P(\mathbf{h})\right] \tag{32}$$

i.e., the primal problem (10) *can be separated into the (sub-) problems* (26)-(31) *without loss of optimality.*

The rate problem in (26) dictates the amount of traffic allowed into the network. It therefore solves the flow control problem at the transport layer. Likewise, (27) represents the network layer routing problem, (28) determines link-level capacities at the data link layer and (29) is the (physical layer) average power control problem. Eqs. (30) and (31) represent resource allocation at the physical layer. Therefore, it is a consequence of Theorem 2 that layering, in the sense of problem separability as per (32) is optimal in faded FDMA wireless networks. Furthermore, (30) and (31) dictate that the FDMA resource allocation separates into subproblems that depend on the instantaneous channel realization only.

We remark that Theorem 2 assumes availability of the optimal Lagrange multipliers λ_{ij}^*, μ_i^*, ν_i^{k*}, and ξ_{ij}^*. Finding them, while possible, is a non-trivial problem that we will address in forthcoming contributions. However, it has to be appreciated that Theorem 2 establishes two fundamental properties of wireless FDMA networks in the presence of fading: i) the decomposition of the problem into the traditional networking layers can be optimal; and ii) the separability of the resource allocation problem into per-fading-state subproblems is possible. None of these properties applies to static wireless networks with deterministic channels.

5. CONCLUSIONS

We investigated the optimal FDMA networking problem defined in (10). Among other important questions we set up the following ones: i) how difficult is it to solve (10), ii) how can one find the optimal solution, iii) how can this solution be used to design FDMA wireless networking protocols, and iv) what does the solution say about fundamental properties of FDMA wireless networking.

In this paper, we tackled questions (i) and (iv). We have shown that in the presence of fading (10) can sometimes be solved by algorithms with manageable complexity. This is a remarkable property given the fact that in deterministic channels the same problem is NP-hard. We further established that (10) can be decomposed into layers and fading states without loss of optimality. We will address (ii) and (iii) in forthcoming papers[2].

6. REFERENCES

[1] L. Georgiadis, M. J. Neely, and L. Tassiulas, "Resource allocation and cross-layer control in wireless networks," *Foundations and Trends in Networking*, vol. 1, no. 1, pp. 1–144, 2006.

[2] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," *IEEE Trans. Autom. Control*, vol. 37, no. 12, pp. 1936–1948, December 1992.

[3] B. Johansson, P. Soldati, and M. Johansson, "Mathematical decomposition techniques for distributed cross-layer optimization of data networks," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1535–1547, August 2006.

[4] Aleksei Andreevich Lyapunov, "Sur les fonctions-vecteur complètement additives," *Bull. Acad. Sci. URSS. Sèr. Math.*, vol. 4, pp. 465–478, 1940.

[5] Z.-Q Luo and S. Zhang, "Dynamic spectrum management: complexity and duality," *Tech. report, Dept. of ECE, Univ. of Minnesota (submitted for publication)*, 2007.

[6] S. Hayashi and Z.-Q Luo, "Spectrum management for interference-limmited multiuser communication systems," *Tech. report, Dept. of ECE, Univ. of Minnesota (submitted for publication)*, 2006.

[7] Dimitri Bertsekas, *Nonlinear Programming*, Athena Scientific, 2nd edition, 1999.

[2] Work in this paper was prepared through collaborative participation in the Communications and Networks Consortium sponsored by the U. S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0011. The U. S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon.