An Efficient Cross Layer Design for OFDMA-Based Wireless Networks with Channel Reuse

Rozita Rashtchi, Ramy H. Gohary, *Senior Member, IEEE,* and Halim Yanikomeroglu, *Senior Member, IEEE*
Department of Systems and Computer Engineering, Carleton University, Ottawa, Ontario, Canada
{rozita, gohary, halim}@sce.carleton.ca

***Abstract*—This paper considers a joint design that incorporates the physical, medium access and network layers of a generic OFDMA-based wireless network with an *ad hoc* topology. The network employs channel reuse, whereby a frequency subchannel might be used simultaneously by multiple nodes. In addition to being a source and/or a destination, each node can act as a half-duplex relay to assist other nodes. The design objective is to determine the jointly optimal data routes and subchannel power allocations that maximize a weighted sum of the rates that can be reliably communicated over the network. Assuming that the signals transmitted by the nodes are Gaussian, the joint cross layer design of routing and power allocation is cast as an optimization problem. Unfortunately, this problem is non-convex, and hence difficult to solve. To circumvent this difficulty, an efficient technique based on geometric programming is developed to obtain a local solution that satisfies the Karush-Kuhn-Tucker necessary optimality conditions. Numerical results show that, despite the potential suboptimality of the obtained solution, for some network scenarios, it offers significant gains over optimal scheduling-based schemes in which a frequency band is allowed to be used by one node only at any time instant.**

***Index Terms*—Convex optimization, routing, scheduling, power allocation, KKT conditions.**

I. Introduction

The prospect of having ubiquitous wireless services is leveraged by the versatility and the portability of the communication devices that will form the nodes of future wireless networks. These devices will be able to perform various functions including sending, receiving and/or relaying data to other nodes [1]. As such, it is expected that future wireless networks will not possess a predetermined topology, but rather an *ad hoc* one. Such a model is generic and is capable of representing the key features of a wide range of networks, including current cellular ones [2].

Orthogonal Frequency Division Multiple Access (OFDMA) offers several practical advantages over other multiple access techniques. These advantages include design simplicity and resilience to interference and frequency-selective fading. In addition, OFDMA offers the potential of efficient sharing of the available frequency band by multiple nodes, depending on their channel conditions.

In order for a wireless network to be able to support the reliable communication of high data rates, the typically scarce resources available for the network must be carefully exploited. Such exploitation involves choosing the optimal propagation routes of the packets and the optimal power to be allocated by the nodes to each subchannel. Although these tasks have traditionally been performed separately to simplify the design of the network, they are interrelated and performing them in isolation may incur a significant loss in the maximum rates that the network can support. Hence, it is desirable to perform the optimization of routes and power allocation jointly, while maintaining computational cost practical.

Optimization-based techniques have been successfully employed in studying and improving the performance of various communication networks. For instance, optimal power allocations have been obtained in [3] and [4] for uplink and downlink cellular communications, respectively. A layered architecture has been considered in [5] for multicell OFDMA systems. However, further improvements have been sought by joint design of multiple network layers. For instance, joint optimization of the power allocations and binary subchannel schedules in OFDMA networks was considered in [6]. Solutions obtained therein are suboptimal and rely on the premise that each subchannel is exclusively used by one node. When the nodes are restricted to use preassigned orthogonal channels, the joint design of routes and power allocations can be cast in an efficiently solvable convex form [7]. However, this restriction was alleviated in [8] by allowing the channels to be non-orthogonal. In that case, the optimal routes and power allocations are obtained, provided that the rates are chosen from a discrete set. Cross layer designs that exploit the broadcast feature of the wireless medium were developed in [9] and [10] using superposition coding and optimization techniques, including geometric programming (GP). Techniques for jointly optimizing the power allocations, the subchannel schedules and the data routes were developed in [11], [12] for networks in which subchannel reuse among multiple nodes is not permitted.

To facilitate the design of OFDMA-based wireless networks, each subchannel is typically restricted to be exclusively used at any given instant by one of the nodes in the network. The optimal scheduling of subchannels to nodes can be incorporated with the optimization of data routes and power allocations [11]. Despite the practical advantages of using subchannel scheduling for avoiding interference, the rates provided by the networks that use this technique may be significantly inferior to those provided by networks that allow the subchannels to be used simultaneously by multiple nodes. This is especially the case when the network is composed of essentially separated clusters; subchannel reuse is optimal for clustered networks.

In contrast with scheduling-based designs, in this paper we consider a joint design that incorporates the physical, medium access and network layers of an OFDMA-based network, in which the nodes are entitled to reusing subchannels. The nodes in the considered network can assume multiple roles simultaneously including being sources, destinations and/or half-duplex decode-and-forward relays. Assuming that the channel coefficients are fixed and known *a priori*, the cross layer design objective is to determine the routes and power allocations that maximize a weighted sum of the rates injected and reliably communicated over the network. This design is cast as an optimization problem, which is, unfortunately, nonconvex and hence, difficult to solve. To overcome this difficulty, we adopt a GP-based approach [9], wherein constraints that are not compatible with the GP standard form are approximated by the first order term in the corresponding Taylor expansion in the logarithmic domain. This technique is typically referred to as monomial approximation [13]. One of the desirable features of this technique is that its successive application is guaranteed to yield a local solution of the Karush-Kuhn-Tucker (KKT) corresponding to the optimization problem [14]. The essence of this technique is to construct a convex approximation of the optimization problem in the neighbourhood of an initial feasible point. This approximation is in a GP-compatible form, which is amenable to efficient polynomial-time interior-point solvers; see e.g., [15]. The output obtained by the solver is then used to construct the subsequent approximation, and so on. Numerical results show that the performance of the proposed technique, although potentially suboptimal, can significantly outperform the performance of the optimal solution obtained by scheduling-based approaches, which restrict the subchannels to be used at most by one node at any time instant.

II. System Model and Preliminaries

We consider an OFDMA-based *ad hoc* wireless network of N nodes, each with one transmit and one receive antenna, and a fixed power budget, P_n, $n \in \mathcal{N} \triangleq \{1, 2, \cdots, N\}$. In addition, each node is assumed to be capable of simultaneously transmitting, receiving and relaying data to other nodes. This assumption is generic, in the sense that constraining some nodes to perform a subset of tasks can be readily incorporated in the forthcoming formulations. For practical considerations, the relaying nodes are assumed to operate in a half-duplex mode, whereby each node uses distinct physical channels for transmission and reception. The available OFDMA frequency spectrum, W_0, is divided into K narrowband subchannels, each of bandwidth $W = \frac{W_0}{K}$. The K subchannels are assumed to remain essentially constant during the entire signalling interval.

Such a network can be represented by a weighted fully-connected directed graph with N vertices. To facilitate the enumeration of the $L = N(N-1)$ links in this graph, the link from node n to node n' will be labelled by $\ell = (N-1)(n-1)+n'-1$ if $n < n'$ and by $\ell = (N-1)(n-1)+n'$ if $n > n'$. The set of all links is denoted by $\mathcal{L}$ and the sets $\mathcal{I}(n)$ and $\mathcal{O}(n)$ represent the sets of incoming and outgoing links of node $n \in \mathcal{N}$, respectively. The connectivity of this graph can be captured by an incidence matrix, $A = [a_{n\ell}]$, where [9]

$$a_{n\ell} = \begin{cases} 1 & \text{if link } \ell \in \mathcal{O}(n), \\ -1 & \text{if link } \ell \in \mathcal{I}(n), \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

For the considered OFDMA-based scenario, each link $\ell \in \mathcal{L}$ consists of K subchannels, each with a complex random coefficient. The coefficient of the k-th subchannel of link ℓ connecting node n to node n' is denoted by $h_{nn'}^{(k)}$. The nodes receive a superposition of noise and the transmissions of all other nodes, scaled by the respective gains. Using $v_n^{(k)}$, $u_n^{(k)}$ and $y_n^{(k)}$ to respectively denote the Gaussian noise and the signals transmitted and received by node n on the k-th subchannel, we can write

$$y_n^{(k)} = \sum_{n' \in \mathcal{N} \setminus \{n\}} h_{nn'}^{(k)} u_{n'}^{(k)} + v_n^{(k)}, \qquad n \in \mathcal{N}, \quad (2)$$

where $\setminus$ denotes the set minus operation.

In contrast with [9], in the current model, it is assumed that the nodes cannot broadcast data simultaneously on the same subchannel to different destinations. However, because of the superposition, this model resembles a multiple access channel, wherein the nodes might be able to jointly decode the signals of other nodes. Such decoding involves successive detection and cancellation in a certain order, which makes the network design rather complicated. A more pragmatic approach is for each node to decode the signal received on each link separately, while treating the signals received on other links as additive interference. In this case, assuming, as before, that $\ell \in \mathcal{L}$ is the link connecting node n to node n', it can be seen from (2) that the signal-to-noise-plus-interference ratio (SNIR) observed by node n' on subchannel k of link ℓ is given by

$$\text{SNIR}(\ell, k) = \frac{p_n^{(k)} |h_{nn'}^{(k)}|^2}{WN_0 + \sum_{n'' \in \mathcal{N} \setminus \{n, n'\}} p_{n''}^{(k)} |h_{n''n'}^{(k)}|^2}, \quad (3)$$

where $p_n^{(k)}$ is the power allocated by node n to the k-th subchannel and N_0 is the variance of the additive Gaussian noise. The second term in the denominator represents the aggregate interference observed by node n' on subchannel k of link ℓ. When the nodes transmit Gaussian distributed signals, the maximum data rate that can be reliably communicated on this subchannel is given by $W \log_2(1 + \text{SNIR}(\ell, k))$.

Having described the system model, in Section III we will provide a mathematical characterization of the constraints that must be satisfied by the flows on each link, and the power and rate allocated to each transmission. This characterization will enable us to propose an efficient methodology for optimizing these design variables to maximize the weighted sum of the rates injected and reliably communicated over the network.

A. The GP Standard Form and Monomial Approximation

A GP optimization problem can be readily transformed to an efficiently solvable convex one. To provide the standard form of a GP, let $z \in \mathbb{R}^n$ be a vector of positive entries. A monomial in z is defined to be a function of the form $c_0 \prod_i z_i^{\alpha_i}$ and a posynomial in z is defined to be a function of the form $\sum_{j=1}^{J} c_j \prod_{i=1}^{n} z_i^{\alpha_{ij}}$, where $c_j > 0$, $\{\alpha_i\}$ and $\{\alpha_{ij}\}$, are arbitrary constants, $j = 0, 1, \ldots, J$, and $i = 1, \ldots, n$. A standard GP [13] is an optimization problem of the form:

$$\begin{aligned} \min_{z} \quad & f_0(z), \\ \text{subject to} \quad & f_i(z) \leq 1, \qquad i = 1, \ldots, m, \\ & g_i(z) = 1, \qquad i = 1, \ldots, p, \end{aligned} \tag{4}$$

where $\{f_i\}$ are posynomial and $\{g_i\}$ are monomials.

A monomial approximation of a differentiable function $h(z) \geq 0$ near $z^{(0)}$ is given by its first order Taylor expansion in the logarithmic domain [13]. Defining $\beta_i = \frac{z_i^{(0)}}{h(z^{(0)})} \frac{\partial h}{\partial z_i}\big|_{z=z^{(0)}}$, we have

$$h(z) \approx h(z^{(0)}) \prod_{i=1}^{n} \Big(\frac{z_i}{z_i^{(0)}} \Big)^{\beta_i}. \tag{5}$$

This approximation will be used to provide a GP approximation of the cross layer design problem.

III. Problem Statement

In this section we will provide a characterization of the constraints that must be satisfied by the variables that represent feasible routes and power allocations. Using this characterization, we will formulate the network design as an optimization problem, which although nonconvex, will be transformed in the next section into a form that is more convenient for obtaining locally optimal solutions efficiently. For tractability, the nodes will be assumed to always have data ready for transmission [7].

A. System Constraints

Let $\mathcal{D} \triangleq \{1, 2, \cdots, D\}$ be the set of all destination nodes. Let $s_n^{(d)}$ be the rate of the data stream injected into node $n \in \mathcal{N}$ and intended for destination $d \in \mathcal{D}$, and let $x_{\ell k}^{(d)}$ be the corresponding flow on subchannel $k \in \mathcal{K}$ of link $\ell \in \mathcal{L}$. The flows, $\{x_{\ell k}^{(d)}\}$, and the injected rates, $\{s_n^{(d)}\}$, are related by the flow conservation law, which must be satisfied at each node. This law stipulates that the sum of flows intended for any destination $d \in \mathcal{D}$ at each node must be equal to zero [7]. Applying this law to the current network and using the incidence matrix in (1), it can be seen that $\{x_{\ell k}^{(d)}\}$ and $\{s_n^{(d)}\}$ must satisfy the following constraints:

$$\sum_{\ell \in \mathcal{L}} \sum_{k \in \mathcal{K}} a_{nl} x_{\ell k}^{(d)} = s_n^{(d)}, \qquad n \in \mathcal{N} \setminus \{d\}, d \in \mathcal{D}. \tag{6}$$

The flow conservation law implies that the rate of data leaving the network at $d \in \mathcal{D}$ equals the sum of the data rates injected into the network and intended for this destination. Hence, we can write $s_d^{(d)} = -\sum_{n \in \mathcal{N} \setminus \{d\}} s_n^{(d)}$.

The injected rates, $\{s_n^{(d)}\}_{n \neq d}$, are non-negative and since in our model the network will be represented by a directed graph, the flows, $\{x_{\ell k}^{(d)}\}$, must be also non-negative. Hence,

$$x_{\ell k}^{(d)} \geq 0, \qquad \ell \in \mathcal{L}, k \in \mathcal{K}, d \in \mathcal{D}, \tag{7}$$

$$s_n^{(d)} \geq 0, \qquad n \in \mathcal{N} \setminus \{d\}, \ d \in \mathcal{D}. \tag{8}$$

As mentioned in Section II, we are considering a network design problem wherein the relaying nodes operate in a half-duplex mode and the nodes cannot simultaneously broadcast to multiple destinations on the same subchannel. These requirements can be captured by introducing a new set of variables $\{q_{\ell k}\}, \ \ell \in \mathcal{L}, \ k \in \mathcal{K}$. These variables are related to the node powers by the following set of equations:

$$p_n^{(k)} = \max_{\ell \in \mathcal{O}(n)} q_{\ell k}, \qquad n \in \mathcal{N}, k \in \mathcal{K}, \tag{9a}$$

$$a_{n\ell}^{+} a_{n\ell'}^{+} q_{\ell k} q_{\ell' k} = 0, \qquad \ell, \ell' \in \mathcal{L}, \ k \in \mathcal{K}, \ n \in \mathcal{N}, \tag{9b}$$

where $a_{n\ell}^{+} = \max\{0, a_{n\ell}\}$, that is, $a_{n\ell}^{+} = 1$ if $\ell \in \mathcal{O}(n)$ and zero otherwise.

To gain a better understanding of the transformation in (9), we note that (9b) implies that for any subchannel $k \in \mathcal{K}$ and any two links $\ell, \ell' \in \mathcal{O}(n)$, at least $q_{\ell k} = 0$ or $q_{\ell' k} = 0$. In other words, this equation implies that, of all the links in $\mathcal{O}(n)$, only one element in the set $\{q_{\ell k}\}_{\ell \in \mathcal{O}(n)}, \forall \, n \in \mathcal{N} \, k \in \mathcal{K}$ can assume a strictly positive value. Now, (9a) indicates that this value is the power allocated by the node n to subchannel k. Note that using (9) will enable us to formulate the design in terms of $\{q_{\ell k}\}$ instead of $\{p_n^{(k)}\}$. In the sequel, the elements of $\{q_{\ell k}\}$ will be referred to as the link powers.

We now consider the requirement for the network to operate in a half-duplex mode. Since in the current network, no scheduling is considered, this requirement can be imposed by restricting the nodes to transmit and receive data on distinct subchannels. In other words, if a transmission intended for node n on subchannel k of an incoming link, then the power allocated by node n to subchannel k of any outgoing link must be zero. Using the link powers defined in (9), the half-duplex operation can be captured by the following set of constraints.

$$a_{n\ell}^{-} a_{n\ell'}^{+} q_{\ell k} q_{\ell' k} = 0, \qquad \ell, \ell' \in \mathcal{L}, \ k \in \mathcal{K}, \ n \in \mathcal{N}. \tag{10}$$

where $a_{n\ell}^{-} = \min\{0, a_{n\ell}\}$, that is, $a_{n\ell}^{-} = -1$ if $\ell \in \mathcal{I}(n)$ and zero otherwise. Note that (9b) and (10) are trivially satisfied if either link ℓ or ℓ' are not connected to node n.

In a practical network, the nodes are likely to have a certain power budget which bounds the total power allocated by each node on all subchannels. This constraint can be written as

$$\sum_{k \in \mathcal{K}} p_n^{(k)} \leq P_n, \qquad n \in \mathcal{N}.$$

Using (9a), this constraint can be cast as

$$\sum_{k \in \mathcal{K}} \sum_{\ell \in \mathcal{L}} a_{n\ell}^{+} q_{\ell k} \leq P_n, \qquad n \in \mathcal{N}, \tag{11}$$

where the link powers must satisfy the following non-negativity constraints:

$$q_{\ell k} \geq 0, \qquad \ell \in \mathcal{L}, \ k \in \mathcal{K}. \tag{12}$$

To complete the network characterization, we point out that the data flows and the power allocations are coupled by the maximum rate that can be supported by the subchannels of each link. In particular, the aggregate rate $\sum_{d\in\mathcal{D}} x_{\ell k}^{(d)}$ must not exceed the capacity of the k-th subchannel of link ℓ. As mentioned in Section II, when the nodes use Gaussian signalling, this capacity is given by $W\log_2(1+\mathrm{SNIR}(\ell,k))$, where SNIR is defined in (3). For notational convenience, we will use the fact that the index of each link ℓ corresponds to a specific (n,n') pair and will use $\gamma_{\ell k}$ to denote $|h_{nn'}^{(k)}|^2$ for any two nodes $n,n'\in\mathcal{N}$. Using this notation and invoking (3), the constraint on the aggregate rate on each subchannel $k\in\mathcal{K}$ of each link $\ell\in\mathcal{L}$ can be written in terms of $\{q_{\ell k}\}$ as

$$\sum_d \frac{x_{\ell k}^{(d)}}{W} \le \log_2\Bigl(1+\frac{q_{\ell k}\gamma_{\ell k}}{WN_0+\sum_{\ell'\in\mathcal{L}\setminus\{\ell\}} q_{\ell' k}\gamma_{\ell'' k}}\Bigr), \quad (13)$$

where, here and henceforth, ℓ'' is used to denote the index of the link connecting the node at which link ℓ' originates to the node at which link ℓ ends.

B. The Cross Layer Design Problem

We are now ready to formulate the cross layer design as an optimization problem. The objective of this problem is to maximize a weighted sum of the rates injected into the network. From a practical perspective, assigning weights to the injected rates provides a convenient means for controlling the quality of service (QoS); a higher weight implies a higher priority to the corresponding rate. Such weights are typically assigned *a priori*, but can be adapted to meet variations in the QoS requirements [6]. Another advantage of considering weighted sum rates is that varying the weights over the unit simplex enables us to determine a set of rates that can be simultaneously achieved by a given network. To ensure the feasibility of the rates generated by our design, the constraints on the routes and power allocations in (6)–(13) must be satisfied. Hence, the design problem can be written as:

$$\max_{\{s_n^{(d)}\},\{x_{\ell k}^{(d)}\},\{q_{\ell k}\}} \quad \sum_{d\in\mathcal{D}}\sum_{n\in\mathcal{N}\setminus\{d\}} w_n^{(d)} s_n^{(d)}, \quad (14a)$$

subject to

$$\sum_{\ell\in\mathcal{L}}\sum_{k\in\mathcal{K}} a_{n\ell} x_{\ell k}^{(d)} = s_n^{(d)}, \qquad n\in\mathcal{N}\setminus\{d\},\, d\in\mathcal{D}, \quad (14b)$$

$$x_{\ell k}^{(d)} \ge 0, \qquad \ell\in\mathcal{L},\, k\in\mathcal{K},\, d\in\mathcal{D}, \quad (14c)$$

$$s_n^{(d)} \ge 0, \qquad n\in\mathcal{N}\setminus\{d\},\, d\in\mathcal{D}, \quad (14d)$$

$$\sum_{d\in\mathcal{D}} \frac{x_{\ell k}^{(d)}}{W} \le \log_2\Bigl(1+\frac{q_{\ell k}\gamma_{\ell k}}{WN_0+\sum_{\ell'\in\mathcal{L}\setminus\{\ell\}} q_{\ell' k}\gamma_{\ell'' k}}\Bigr), \qquad k\in\mathcal{K},\, \ell\in\mathcal{L}, \quad (14e)$$

$$q_{\ell k}\ge 0, \qquad \ell\in\mathcal{L},\, k\in\mathcal{K}, \quad (14f)$$

$$\sum_{k\in\mathcal{K}}\sum_{\ell\in\mathcal{L}} a_{n\ell}^{+} q_{\ell k} \le P_n, \qquad n\in\mathcal{N}, \quad (14g)$$

$$a_{n\ell}^{-}a_{n\ell'}^{+} q_{\ell k} q_{\ell' k} = 0, \qquad \ell,\ell'\in\mathcal{L},\, k\in\mathcal{K},\, n\in\mathcal{N}, \quad (14h)$$

$$a_{n\ell}^{+}a_{n\ell'}^{+} q_{\ell k} q_{\ell' k} = 0, \qquad \ell,\ell'\in\mathcal{L},\, k\in\mathcal{K},\, n\in\mathcal{N}, \quad (14i)$$

where $\{w_n^{(d)}\}$ are the weights assigned to the respective rates.

The optimization problem in (14) is nonconvex because the right hand side (RHS) of the capacity constraints in (14e) is the logarithm of a rational function of $\{q_{\ell k}\}$, and therefore not concave. This can be verified by showing that the corresponding Hessian matrices are non-definite. The equality constraints in (14h) and (14i) are not affine and hence, also nonconvex.

By examining (14), it can be seen that this problem shares some features with the GP standard form in (4), including the non-negativity of the optimization variables and the product form appearing in some of the constraints. To exploit this observation, in the next section we will perform a change of variables that will enable us to express the objective and all, but one set, of constraints in a GP-compatible form. The residual constraints that do not comply with the GP standard form are approximated using the monomial approximation in (5). Applying this technique iteratively is known to yield a local solution of the KKT system corresponding to (14), see e.g., [9].

IV. Proposed GP-based Algorithm

Since solving the nonconvex problem in (14) directly is difficult, we will cast it in a form that is amenable to an efficient approximation technique. The structure of the constraints in (14e), (14h) and (14i) suggests GP as a candidate approach for providing a convex approximation of (14).

Let us define two new sets of variables, $\{t_n^{(d)}\}$ and $\{r_{\ell k}^{(d)}\}$, which are related to $\{s_n^{(d)}\}$ and $\{x_{\ell k}^{(d)}\}$ by the following maps:

$$s_n^{(d)} = \log_2 t_n^{(d)}, \qquad n\in\mathcal{N}\setminus\{d\},\, d\in\mathcal{D}, \quad (15)$$

$$x_{\ell k}^{(d)} = W\log_2 r_{\ell k}^{(d)}, \qquad \ell\in\mathcal{L},\, k\in\mathcal{K},\, d\in\mathcal{D}. \quad (16)$$

These maps are bijective, which renders recovering $\{s_n^{(d)}\}$ and $\{x_{\ell k}^{(d)}\}$ from $\{t_n^{(d)}\}$ and $\{r_{\ell k}^{(d)}\}$ straightforward.

Using (15) and (16), the objective in (14a) and the constraints in (14b)–(14d) can be readily expressed in a GP-compatible form. For the constraints in (14e), substituting from (16) yields the following set of equivalent constraints

$$\Bigl(WN_0+\sum_{\ell'\in\mathcal{L}\setminus\{\ell\}} q_{\ell' k}\gamma_{\ell'' k}\Bigr)\prod_{d\in\mathcal{D}} r_{\ell k}^{(d)} \le WN_0+\sum_{\ell'\in\mathcal{L}} q_{\ell' k}\gamma_{\ell'' k}, \qquad k\in\mathcal{K},\, \ell\in\mathcal{L}. \quad (17)$$

The RHS of (17) is not a monomial, which renders (17) not compatible with the GP framework in (4). To provide a cross layer design formulation that is compatible with this framework, we will invoke (5) to approximate the RHS of (17) by a monomial expression near an initial $\{q_{\ell k}^{(0)}\}$.

We now consider the constraints in (14h) and (14i). The constraints in (14h) have negative coefficients $\{a_{n\ell}^{-}\}$, but can be equivalently cast as $|a_{n\ell}^{-}|a_{n\ell'}^{+}q_{\ell k}q_{\ell' k}=0, \forall \ell,\ell',k$. Now, the RHSs of these constraints and the ones in (14i) are zero, which is not compatible with the framework in (4). This problem can be alleviated by constraining the left hand side of these constraints to be less than an arbitrary small $\epsilon>0$.

Using the transformations in (15) and (16), the monomial approximation of (17) and the relaxed versions of (14h) and (14i), the joint design of data routes and power allocations in (14) can be approximated with the following GP.

$$\max_{\{t_n^{(d)}\},\{r_{\ell k}^{(d)}\},\{q_{\ell k}\}} \quad \prod_{d\in\mathcal{D}}\prod_{n\in\mathcal{N}\setminus\{d\}} \left(t_n^{(d)}\right)^{w_n^{(d)}} \tag{18a}$$

subject to

$$\prod_{\ell\in\mathcal{L}}\prod_{k\in\mathcal{K}} \left(r_{\ell k}^{(d)}\right)^{W a_{n\ell}} = t_n^{(d)}, \quad n\in\mathcal{N}\setminus\{d\},\, d\in\mathcal{D}, \tag{18b}$$

$$r_{\ell k}^{(d)} \geq 1, \quad \ell\in\mathcal{L},\, k\in\mathcal{K},\, d\in\mathcal{D}, \tag{18c}$$

$$t_n^{(d)} \geq 1, \quad n\in\mathcal{N}\setminus\{d\}, d\in\mathcal{D}, \tag{18d}$$

$$\Big(WN_0 + \sum_{\ell'\in\mathcal{L}\setminus\{\ell\}} q_{\ell' k}\gamma_{\ell'' k}\Big) \prod_{d\in\mathcal{D}} r_{\ell k}^{(d)} \leq c_{\ell k}\prod_{\ell'\in\mathcal{L}}\Big(q_{\ell' k}/q_{\ell' k}^{(0)}\Big)^{\theta_{\ell' k}},$$
$$k\in\mathcal{K}, \ell\in\mathcal{L}, \tag{18e}$$

$$\sum_{k\in\mathcal{K}}\sum_{\ell\in\mathcal{L}} a_{n\ell}^{+} q_{\ell k} \leq P_n, \quad n\in\mathcal{N}, \tag{18f}$$

$$|a_{n\ell}^{-}| a_{n\ell'}^{+} q_{\ell k} q_{\ell' k} \leq \epsilon, \quad \ell,\ell'\in\mathcal{L},\, k\in\mathcal{K},\, n\in\mathcal{N}, \tag{18g}$$

$$a_{n\ell}^{+} a_{n\ell'}^{+} q_{\ell k} q_{\ell' k} \leq \epsilon, \quad \ell,\ell'\in\mathcal{L},\, k\in\mathcal{K},\, n\in\mathcal{N}, \tag{18h}$$

where $c_{\ell k} = \Big(WN_0 + \sum_{\ell'\in\mathcal{L}} q_{\ell' k}^{(0)}\gamma_{\ell'' k}\Big)$, and $\theta_{\ell' k} = q_{\ell' k}^{(0)}\gamma_{\ell'' k}/c_{\ell k}$. Note that the non-negativity constraints in (14f) are inherently incorporated in the GP framework.

The problem in (18) is in the form of a GP, which, using a standard exponential transformation, can be converted to a convex problem [16]. Hence, (18) enables us to efficiently solve the cross layer design problem approximately in the neighbourhood of any initial set $\{q_{\ell k}^{(0)}\}$.

Finding the global solution for the nonconvex problem in (14) is difficult, whereas solving the approximated problem in (18) is straightforward. To exploit this fact, we incorporate the formulation in (18) in an iterative algorithm, whereby the output of solving (18) for an initial set $\{q_{\ell k}^{(0)}\}$ is used as a starting point for the subsequent iteration. This technique is usually referred to as the single concentration method, e.g., [9], and under relatively mild conditions, its convergence to a solution of the KKT system corresponding to (14) is guaranteed [14].

The relaxations in (18g) and (18h) may result in infeasible power allocations that do not satisfy the constraints in (9b) and (10). To construct a feasible, but potentially suboptimal, power allocations, the elements of $\{q_{\ell k}\}$ that are less than or equal to $\sqrt{\epsilon}$ are set to zero. More specifically,

$$\textbf{If } q_{\ell k} \leq \sqrt{\epsilon}, \quad \textbf{then } q_{\ell k} = 0, \quad \forall\, k\in\mathcal{K},\ \ell\in\mathcal{L}. \tag{19}$$

After updating the link powers according to the rule in (19), the node powers can be readily obtained using (9a). The optimal routes and input rates corresponding to the (potentially suboptimal) updated node powers can be obtained through the efficiently computable linear optimization problem in (14) with power allocations being fixed.

This algorithm is described in Table I.

TABLE I
SUCCESSIVE GP-BASED ALGORITHM FOR SOLVING (18)

1- Let $U^{(0)} = 0$. Choose a feasible $\{q_{\ell k}^{(0)}\}$. Set accuracy to $\delta > 0$.
2- Solve the GP in (18). Denote the value of the objective by U.
3- **While** $U - U^{(0)} \geq \delta$,
  $\{q_{\ell k}^{(0)}\} \leftarrow \{q_{\ell k}\}$,
  $U^{(0)} \leftarrow U$,
  Solve the GP in (18). Denote the value of the objective by U,
End.
4- Update the power allocations by (19).
5- Solve the linear problem corresponding to (14) with power allocations fixed.

V. NUMERICAL EXAMPLES

In this section we provide two numerical examples that illustrate the potential advantage of the cross layer design method proposed herein. Since in this method the subchannels are allowed to be simultaneously used by multiple nodes, its advantage over scheduling-based methods is more pronounced when the network exhibits a cluster-like structure. This is because, in that case, the interference generated by any cluster has a small impact on the rest of the network, which enables efficient reuse of the subchannels.

In the numerical results reported herein, locally optimal sets of data routes and power allocations for this network are found using `MOSEK` [17], which is a powerful polynomial-time solver that converts the GP in (18) to a convex form and solves it using an interior-point method.

To maximize the sum rate, the weights $\{w_n^{(d)}\}$ are set equal to 1, $\forall n, d,\ n \neq d$ and the nodes are assumed to have identical power budgets; i.e., $P_n = P$, $n = 1, \ldots, 4$. The channels are generated using the indoor non-line-of-sight hotspot scenario of the IMT-Advanced in [18]. In this scenario, the path loss component of link ℓ connecting node n to n' is given by

$$43.3\log_{10}(r_{nn'}) + 11.5 + 20\log_{10}(f_c),$$

where $r_{nn'}$ is the distance in metres between node n and n' and f_c is the carrier frequency in GHz. The corresponding shadowing component is log-normal distributed with a 0 dB mean and a standard deviation of 4 dB. The fading component is obtained from the standard Rayleigh distribution.

In accordance with the IMT-Advanced document [18], the noise power is set to be -114 dBm/Hz, the carrier frequency is set to be $f_c = 2$ GHz and the available bandwidth is set to be 20 MHz, which is equally divided between the subchannels.

Example 1. *Illustrating the advantage of subchannel reuse in a network with $N = 4$ nodes and $K = 2$ subchannels.*

In this example, we consider maximizing the sum rate that can be reliably communicated by the quasi-static network in Figure 1. This network has $N = 4$ nodes, which were randomly dropped in a square of $100 \times 100\ m^2$. The number of links is $L = 12$, and the number of destinations is $D = 2$; i.e., $\mathcal{D} = \{1, 2\}$. Two subchannels are available for communication, i.e., $K = 2$. We obtained the realization of the channel gains represented by the matrix in (20) below. For ease of exposition, the (k, ℓ)-th entry of this matrix represents

$\log_{10}(\gamma_{\ell k}^{-1})$ instead of $\gamma_{\ell k}$.

$$\begin{bmatrix} 8.8 & 1.0 & 6.7 & 9.4 & 9.1 & 9.7 & 9.1 & 6.4 & 9.6 & 6.9 & 9.9 & 1.1 \\ 1.1 & 9.9 & 6.8 & 1.1 & 7.7 & 9.9 & 9.9 & 7.8 & 9.8 & 6.9 & 1.1 & 1.1 \end{bmatrix} \tag{20}$$

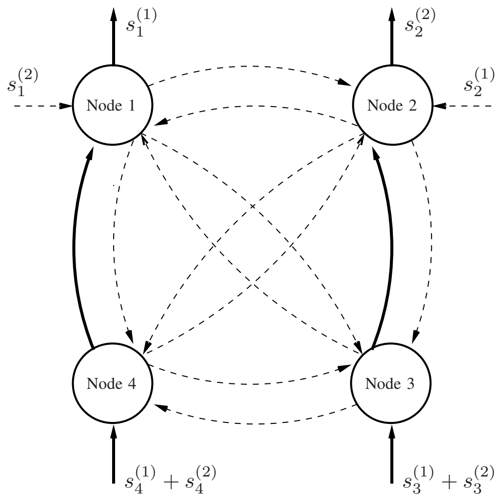

Fig. 1. Network schematic. Each link is composed of two subchannels.

Setting $\delta = 10^{-6}$, and $P = 5$ dBm, the algorithm in Table I yielded a sum rate of 37.48 bits-per-second (bps) with the 'active' links represented by the thicker solid lines in Figure 1; dashed lines represent 'inactive' links. As shown in this figure, each subchannel has been used twice during the signalling interval. It is worth noting that even for the small network considered in this example, the number of variables and constraints is relatively large; 80 variables and 172 constraints.

To demonstrate the benefit of the proposed algorithm, we compare its sum rate performance with binary and continuous scheduling-based approaches. Both approaches do not allow for subchannel reuse and ensure that, at any given time, each subchannel is used by one node only. In the binary approach, each subchannel is assigned to a certain node for the entire signalling interval. As such, binary scheduling can be regarded as a special case of the general cross layer design considered herein. In this case, the cross layer design problem corresponds to a mixed integer program [11], and can be solved using an exhaustive search over $\frac{L!}{(L-K)!}$ possible binary schedules; a computationally prohibitive approach for large networks. Performing this search for the network in Figure 1 yielded a maximum sum rate of 20.33 bps. In addition to binary scheduling, we also compare the sum rate performance of the proposed approach with that of continuous scheduling, which subsumes binary scheduling and allows each subchannel to be time-shared by multiple nodes. This type of scheduling was shown in [11] to give rise to a convex optimization problem that can be efficiently solved. Using the formulation in [11], this type of scheduling yielded a maximum sum rate of 28.47 bps. This comparison is shown in Table II.

Performing the exhaustive search required for binary scheduling at various values of P is computationally expensive. Hence, for that case, we only provide a sum rate comparison between the approach proposed herein and the one based on continuous (time-sharing) scheduling. This comparison is depicted in Figure 2, which shows that the cross layer design proposed herein can significantly outperform the scheduling-based approach. For instance, at a sum rate of 35 bps, the proposed approach provides a power gain of 15 dBm. As P increases, the effect of interference becomes significant and reuse becomes less effective. This effect would not arise had the network had a clustered structure.

TABLE II
THE SUM RATE GENERATED BY PROPOSED ALGORITHM COMPARED WITH BINARY AND CONTINUOUS SCHEDULING-BASE ALGORITHMS.

Proposed alg. with subchannel reuse	Optimal with binary sch.	Optimal with continuous sch.
37.48	20.33	28.47

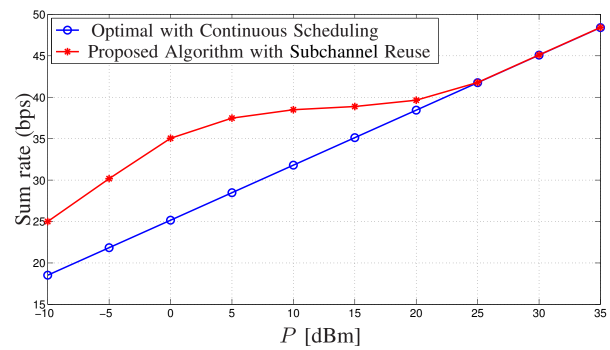

Fig. 2. Sum rate of proposed and continuous scheduling-based designs for the network considered in Example 1.

□

Example 2. *Illustrating the advantage of subchannel reuse in a network with $N = 10$ nodes and $K = 4$ subchannels.*

To further demonstrate the benefit of the proposed algorithm over scheduling-based designs even in non cluster-like structures, in this example we consider a network of $N = 10$ nodes communicating with each other, i.e., $\mathcal{D} = \mathcal{N}$. We consider the same scenario as described in Example 1, except that, now, the nodes are randomly dropped in a square of 500×500 m^2. The network has $L = 90$ links and operates over $K = 4$ subchannels. The number of variables for the considered network in this example is 4050 and the number of constraints is 10270.

Providing the subchannel gains of the considered network is not possible due to space limitations. However, since these gains are dominated by the pathloss component, in Figure 3 we provide the geographic location of the nodes. Implementing the algorithm in Table I for $P = 20$ dBm yields the routes shown in Figure 3. In this figure, the thickness of the lines are made proportional to the power allocated to the transmissions of the nodes. It can be seen from this figure that each subchannel is reused over multiple links. This is in contrast with the scheduling-based design, which allows each subchannel to be used at most once.

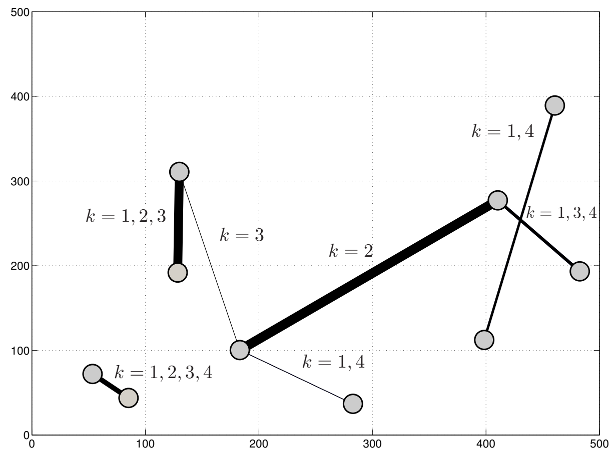

Fig. 3. Schematic of the network in Example 2, data routes, subchannels reuse and corresponding power allocations.

The sum rate achieved by the algorithm in Table I and the approach based on continuous (time-sharing) scheduling is provided in Figure 4 for various P. From this figure, it can be seen that, despite being potentially suboptimal, at a sum rate of 50 bps, the proposed design yields a power advantage in excess of 25 dBm over the design based on the optimal continuous scheduling. As the network size increases, the opportunity of reusing a subchannel also increases and therefore the advantage of using the proposed algorithm over scheduling-based techniques becomes more pronounced.

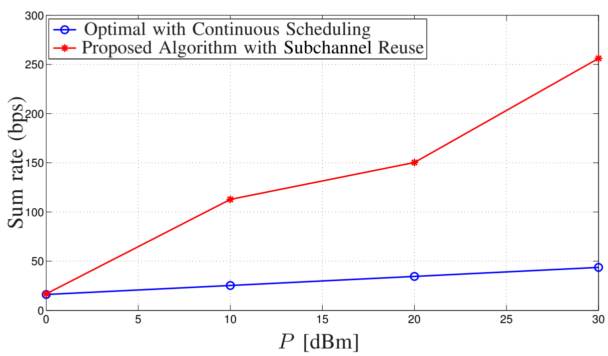

Fig. 4. Sum rate of proposed and continuous scheduling-based designs for the network considered in Example 2.

Further numerical investigations suggest that the proposed design can offer significant rate advantages over scheduling-based approaches for a large class of networks with more general structures. The results of these investigations are not presented due to space considerations. □

VI. Conclusion

We considered the joint design of data routes and power allocations in an OFDMA-based wireless network with *ad hoc* topology in which each frequency subchannel can be used by multiple nodes simultaneously. We developed an efficient iterative approach that enabled us to obtain locally optimal solutions of this nonconvex design problem in polynomial time. Although potentially suboptimal, for some network scenarios, the data routes and power allocations obtained by our technique enabled achieving significantly higher rates than those achieved by their optimal counterparts in scheduling-based cross layer designs.

Acknowledgment

The authors would like to thank Dr. C. S. Bontu and Dr. J. Womack of BlackBerry (formerly, Research In Motion (RIM)) for their support.

References

[1] R. Pabst, B. Walke, D. Schultz, P. Herhold, H. Yanikomeroglu, S. Mukherjee, H. Viswanathan, M. Lott, W. Zirwas, M. Dohler, H. Aghvami, D. Falconer, and G. Fettweis, "Relay-based deployment concepts for wireless and mobile broadband radio," *IEEE Commun. Mag.*, vol. 42, pp. 80–89, Sept. 2004.
[2] X. Bangnan, S. Hischke, and B. Walke, "The role of ad hoc networking in future wireless communications," in *Proc. Int. Conf. Commun. Tech. (ICCT)*, pp. 1353–1358, Apr. 2003.
[3] H. Inaltekin and S. V.Hanly, "Optimality of binary power control for the single cell uplink," *IEEE Trans. Inf. Theory*, vol. 58, pp. 6484–6498, Oct. 2012.
[4] J. Jang and K. B. Lee, "Transmit power adaptation for multiuser OFDM systems," *IEEE J. Select. Areas Commun.*, vol. 21, pp. 171–178, Feb. 2003.
[5] M. Moretti, A. Todini, A. Baiocchi, and G. Dainelli, "A layered architecture for fair resource allocation in multicellular multicarrier systems," *IEEE Trans. Vehicular Tech.*, vol. 60, pp. 1788–1798, May 2011.
[6] J. Huang, V. G. Subramanian, R. Agrawal, and R. A. Berry, "Downlink scheduling and resource allocation for OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 288–296, Jan. 2009.
[7] L. Xiao, M. Johansson, and S. P. Boyd, "Simultaneous routing and resource allocation via dual decomposition," *IEEE Trans. Commun.*, vol. 52, pp. 1136–1144, 2004.
[8] M. Johansson and L. Xiao, "Cross-layer optimization of wireless networks using nonlinear column generation," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 435–445, Feb. 2006.
[9] R. H. Gohary and T. J. Willink, "Joint routing and resource allocation via superposition coding for wireless data networks," *IEEE Trans. Signal Processing*, vol. 58, pp. 6387–6399, 2010.
[10] S. Shabdanov, P. Mitran, and C. Rosenberg, "Cross-layer optimization using advanced physical layer techniques in wireless mesh networks," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 1622–1631, 2012.
[11] R. Rashtchi, R. H. Gohary, and H. Yanikomeroglu, "Joint routing, scheduling and power allocation in OFDMA wireless ad hoc networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, pp. 5483–5487, June 2012.
[12] K. Karakayali, J. Kang, M. Kodialam, and K. Balachandran, "Cross-layer optimization for OFDMA-based wireless mesh backhaul networks," in *Proc. IEEE Wireless Commun. Ntwk Conf. (WCNC)*, pp. 276–281, Mar. 2007.
[13] S. Boyd, S.-J. Kim, L. Vandenberghe, and A. Hassibi, "A tutorial on geometric programming," *Optimization and Engineering*, vol. 8, pp. 67–127, 2007.
[14] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical programs," *Operations Research*, vol. 26, pp. 681–683, 1978.
[15] Y. Nesterov, A. Nemirovskii, and Y. Ye, *Interior-Point Polynomial Algorithms in Convex Programming*, vol. 13. SIAM, 1994.
[16] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.
[17] MOSEK Apps., *The MOSEK optimization toolbox for Matlab Manual, Vesion 6.0*, 2012. Available at: http://www.docs.mosek.com.
[18] Int. Telecommun. Union (ITU), *Guidelines for Evaluation of Radio Interface Technologies for IMT-Advanced*. ITU-R: TR M.21351, Dec. 2009. Available: http://www.itu.int/pub/R-REP-M.2135-1-2009.