Optimization of Power and Channel Allocation Using the Deterministic Channel Model

Yue Zhao, and Gregory J. Pottie
Department of Electrical Engineering
University of California, Los Angeles
Los Angeles, CA, 90095, USA
Email: yuezhao@ucla.edu, pottie@ee.ucla.edu

***Abstract*—In a multiuser interference channel, solving the optimal power and channel allocation for a weighted sum-rate maximization is a well-known non-convex problem, and has NP complexity. In this paper, we apply the recently developed deterministic channel model, and obtain a new formulation for this classic problem. Although the non-convex nature remains unavoidable, we exploit novel insights and techniques to significantly reduce the algorithm's complexity, while still guaranteeing its asymptotic optimality. For cellular structured networks with a fixed number of cells, our algorithm has a worst-case polynomial complexity. We provide simulation solutions of this non-convex optimization in a seven-cell network. The proposed algorithm also computes performance upper bounds in all simulation cases as a numerical verification of the solutions' optimality. The upper bounds demonstrate very small gaps from the maximum achieved objective values of the simulation solutions.**

I. Introduction

We consider the maximization of an arbitrarily weighted sum-rate of multiple users in multi-carrier interference channels. The main difficulty in this class of problems is that an individual user's rate is a non-convex function of all users' power. Thus, finding the optimal multi-user power and channel allocation is an NP-hard problem. In the literature, with the employment of a Gaussian interference channel model and the assumption that interference generated from Gaussian codebooks is treated as noise, various forms of this problem have been extensively studied. With a discrete number of possible power levels for each user, algorithms that use dual decomposition methods have been proposed [5] [6] [15]: For the algorithm that guarantees global optimality [6], its complexity grows exponentially as B^n, where n is the number of users, and B is the number of possible transmit power levels (usually on the scale of ten in power controlled systems.) With convexified utility functions or other convex approximations, many lower complexity and distributed algorithms are also proposed [7] [8] [11] [13].

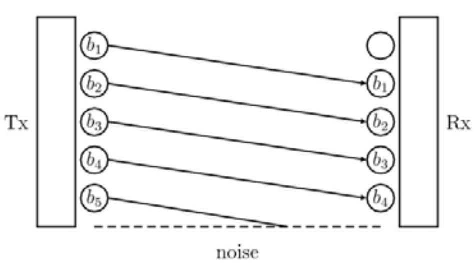

Figure 1. The Deterministic Channel Model for point to point communications [3]

Recently, a new channel model – the Deterministic Channel Model – has been developed in solving the information theoretic capacity regions of the relay channel and the interference channel to within a constant gap [1] [3] [4]. In this paper, instead of using the Gaussian interference channel model, we employ the deterministic channel model, formulate the weighted sum-rate maximization into a new form of optimization, and design novel low complexity algorithms to solve the non-convex problem of multi-user power and channel allocation. Our algorithm provides asymptotically global optimal solutions, and simultaneously provides sharp upper bounds on the global optimum as a check on the optimality of the solutions obtained. Without assuming a discrete number of power levels, our algorithm has a complexity that grows as 2^n for general networks, which is much lower compared to B^n [6]. For cellular structured networks with a fixed a number of cells, our algorithm has a strictly polynomial complexity.

Detailed explanation of the deterministic channel model can be found in [1] [3]. The key features of this model that motivate our approach in this paper are summarized in Table I.

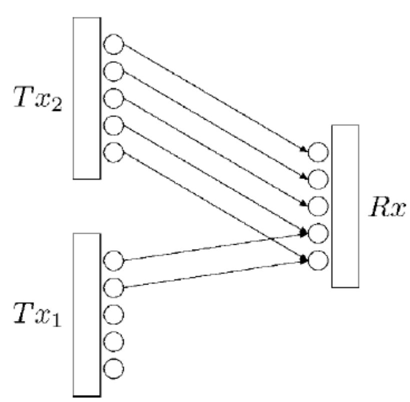

Figure 2. Interference from multiple users in the Deterministic Channel Model [3]

TABLE I. FEATURE COMPARISONS
THE DETERMINISTIC CHANNEL VS. THE GAUSSIAN CHANNEL

	Deterministic Channel	*Gaussian Channel*
1	Signal strength is represented by the number of bit levels	Signal strength is represented by its power
2	Noise is represented by a floor of truncation applied to the received bit levels	Noise is represented by its power, as added in the denominator of SINR
3	Channel gain is represented by the number of bit level shifts from the transmitter to the receiver	Channel gain is represented by the linear scaling factor (from the transmitter to the receiver) of the signal power
4	The interference from another user is represented by the bit levels of the interfering signal that overlap with those of the desired signal at the receiver	The interference from another user is represented by the received power of the interfering signal, as added in the denominator of SINR
5	The aggregate effect of interference from multiple users is represented by the *union* of the received bit levels on which at least one interferer interferes with the desired user	The aggreagate effect of interference from multiple users is represented by the *sum* of the received power of them, as added in the denominator of SINR
6	A user's rate is the number of the received bit levels (of the desired signal) that remain free of interference or noise, i.e. above the interference plus noise floor	A user's rate is calculated by log(1 + SINR)

Comparisons with the corresponding features in the Gaussian channel model are also listed.

Features 1, 2 and 3 in Table I are illustrated in Fig. 1 [3]. The channel gain is a shift of *1 bit level down*. The transmit signal strength is 5 bits above the noise floor, and the received signal strength is 4 bits above the noise floor.

Features 4 and 5 in Table I are illustrated in Fig. 2 [3]. The transmit signal strengths of the two interferers are both 5 bit levels above the noise floor. At the receiver of interest, the 1st interferer interferes 5 bit levels above the noise floor, while the 2nd interferer interferes 2 bit levels. The aggregate effect of the two interferences is the *union* of these two sets of bit levels, which is *equivalent to the effect of the interference from the 1st user only*.

Thus, feature 5 implies a distinguishing property of the deterministic model compared with the Gaussian model: the aggregate effect of the *noise plus interference* (potentially from multiple users) *is represented by the maximum*, instead of the sum, of the noise floor and all interfering signals. In wireless communications networks, this property is a good approximation of reality for the following reason. If Gaussian random codes are used, the number of interfered bit levels is the *log of the summation* of the interference powers. Since it is unlikely to have multiple users at nearly the same interference power level, the *log of the summation* is well approximated by the *log of the maximum* function.

Finally, we note that the deterministic channel model is a good approximation to the Gaussian channel model in the high SINR regime [1]. We will see in later sections that the high SINR conditions are automatically pursued by the proposed optimization algorithm.

With the above features of the deterministic channel model, we establish the system model, and formulate the optimal power and channel allocation problem into a new form in Section II. In Section III, we analyze the problem, and show that the complexity of solving the non-convex optimization can be greatly reduced with the concept of *activity matrix/vector search*. In Section IV, we design the complete low-complexity algorithm based on dual decomposition and subgradient method. In Section V, simulation results are provided, and have shown that the proposed algorithm achieves the global optimum with almost zero gaps. Conclusions are made in Section VI.

II. PROBLEM FORMULATION

A. Mathematical Model of Multiuser Deterministic Interference Channels

We consider power and channel allocation in multi-carrier interference channels with n users and m parallel channels. With the above deterministic channel model, we define the following notations:

1. n_i^j is the height of the noise floor (in other words, the number of bit levels buried below the noise floor) at user i's receiver in channel j.

2. B_i^j is the number of transmit bit levels of user i in channel j.

3. D_i^j is the direct channel gain of user i in channel j. $I_{k,i}^j$ is the interference channel gain from user k to user i in channel j. Both gains are in terms of the number of bit levels shifted down.

Notice that all n_i^j and B_i^j have *relative* values, meaning that the problem setting remains equivalent after adding a common constant to all these values. In other words, all n_i^j and B_i^j can be referenced to an *arbitrary imaginary common noise floor*. On the other hand, all D_i^j and $I_{k,i}^j$ have *absolute* values, since they characterize the difference in signal strength before and after propagation loss.

With these definitions, $B_i^j - D_i^j$ is the number of bit levels of the *desired* signal that user i receives in channel j, whereas $B_k^j - I_{k,i}^j$ is the number of bit levels of the *interfering* signal that user i receives from user k in channel j. From the implication of feature 5 in Table I, $\max\left(\max_{k\neq i}(B_k^j - I_{k,i}^j), n_i^j\right)$ is the number of bit levels buried below the *interference plus noise* at user i's receiver in channel j.

B. Formulation of the Weighted Sum-rate Maximization

Define $$\tilde{R}_i^j \triangleq (B_i^j - D_i^j) - \max\left(\max_{k\neq i}(B_k^j - I_{k,i}^j), n_i^j\right). \quad (1)$$

From feature 6 in Table I, $\tilde{R}_i^j$ is user i's rate in channel j, provided that $\tilde{R}_i^j \geq 0$. When $\tilde{R}_i^j < 0$ (i.e. the desired signal is strictly buried under the aggregate interference plus noise,) user i has zero rate. Thus, we define

$$R_i^j \triangleq \max\left(\tilde{R}_i^j, 0\right),$$

which is the actual rate of user i in channel j. Hence, the weighted sum-rate objective function for an n-user m-channel problem is $\sum_{i=1}^{n} w_i \sum_{j=1}^{m} R_i^j$. (2)

Next, we translate the transmit power constraints into constraints on transmit bit levels B_i^j. We have the Shannon channel capacity formula $B_i^j = \log_2(1 + P_i^j / N)$, where N is the imaginary noise power common to all users (as explained in part A,) and P_i^j is the transmit power of user i (or equivalently, the received power of user i with an identity channel gain.) Thus, the transmit power constraints $\sum_{j=1}^{m} P_i^j \leq P_i^{\max}$, $i = 1, 2, ..., n$ can be translated to

$$\sum_{j=1}^{m} (2^{B_i^j} - 1) \leq \frac{P_i^{\max}}{N}, \quad i = 1,\, 2,\, ...,\, n,$$

which remain as convex constraints. WLOG, we use the normalized power constraints by letting $N = 1$.

Finally, we have the following optimization problem:

$$\begin{aligned} &\max_{B_i^j \geq 0} \ \sum_{i=1}^{n} w_i \sum_{j=1}^{m} R_i^j \\ &s.t. \ \sum_{j=1}^{m} (2^{B_i^j} - 1) \leq P_i^{\max}, \quad i = 1,\, 2,\, ...,\, n \end{aligned}, \tag{3}$$

where $R_i^j \triangleq \max\left((B_i^j - D_i^j) - \max\left(\max_{k \neq i}(B_k^j - I_{k,i}^j), n_i^j\right), 0\right)$.

We define $\left\{B_i^{j*}\right\}$ and $\left\{R_i^{j*}\right\}$ to be the optimal solution and the corresponding rates achieved, $i = 1, 2,..., n$, $j = 1, 2,..., m$.

III. REDUCED COMPLEXITY NON-CONVEX OPTIMIZATION

A. Convex Relaxation and the Sufficient Condition for its Global Optimality

We first observe that $\tilde{R}_i^j$ (1) is a concave function of all B_k^j, $k = 1, 2, ..., n$, because the maximum of linear functions is a convex function. However, $R_i^j = \max\left(\tilde{R}_i^j, 0\right)$ is not concave anymore due to the inclusion of the zero lower bound. Thus, a natural convex relaxation of the original objective (2) is $\sum_{i=1}^{n} w_i \sum_{j=1}^{m} \tilde{R}_i^j$, i.e. replacing all R_i^j with $\tilde{R}_i^j$. To close the gap between the relaxed problem and the original one, we first define an n by m *activity matrix* $\mathbf{A}$: $A_{ij} = \begin{cases} 1 & if\ R_i^j > 0 \\ 0 & otherwise \end{cases}$.

Clearly, $R_i^j = A_{ij} \cdot \tilde{R}_i^j$. We say that user i is *active* in channel j (or $\tilde{R}_i^j$ is active) if $A_{ij} = 1$, and *inactive* otherwise. $\mathbf{A}$ can also be viewed as a *channel assignment matrix*: the non-zero elements in the i^{th} row of $\mathbf{A}$ correspond to all the channels that are assigned to user i. (One channel can be assigned to multiple users.) We define $\mathbf{A}^*$ to be the *optimal channel assignment matrix* that corresponds to $\left\{B_i^{j*}\right\}$ and $\left\{R_i^{j*}\right\}$.

Now, consider the case that we already know the optimal channel assignment $\mathbf{A}^*$ in advance. In the following theorem, we show that the optimal power allocation $\left\{B_i^{j*}\right\}$ can be exactly obtained by solving the convex relaxation with an incorporation of $\mathbf{A}^*$ into the objective.

Theorem 1: Given the optimal channel assignment $\mathbf{A}^*$ of (3) in advance, the convex relaxation with $\sum_{i=1}^{n} w_i \sum_{j=1}^{m} A_{ij}^* \tilde{R}_i^j$ as the objective gives the same optimal value and solution as the original problem's (3).

Proof: First, $\mathbf{A}^*$ can be incorporated into the relaxed concave objective $\sum_{i=1}^{n} w_i \sum_{j=1}^{m} \tilde{R}_i^j$ by removing the inactive $\tilde{R}_i^j$ terms, and the modified relaxed objective $\sum_{i=1}^{n} w_i \sum_{j=1}^{m} A_{ij}^* \tilde{R}_i^j$ remains concave. We then obtain the following convex optimization:

$$\begin{aligned} &\max_{B_i^j \geq 0} \ \sum_{i=1}^{n} w_i \sum_{j=1}^{m} A_{ij}^* \tilde{R}_i^j \\ &s.t. \ \sum_{j=1}^{m} (2^{B_i^j} - 1) \leq P_i^{\max},\ i = 1,\, 2,\, ...,\, n \end{aligned} \tag{4}$$

Since $\sum_{i=1}^{n} w_i \sum_{j=1}^{m} A_{ij}^* \tilde{R}_i^j \leq \sum_{i=1}^{n} w_i \sum_{j=1}^{m} A_{ij}^* R_i^j \leq \sum_{i=1}^{n} w_i \sum_{j=1}^{m} R_i^j$, the optimal value of the original problem (3) is an *upper bound* of that of the modified convex relaxation (4). On the other hand, we show that the optimal value of (3) is also *achievable* in (4) for the following reasons. Consider the optimal solution of (3) $\left\{B_i^{j*}\right\}$ which is also a feasible solution of (4). Since $\mathbf{A}^*$ is the optimal channel assignment determined by $\left\{B_i^{j*}\right\}$, we have

$$\begin{cases} \forall A_{ij}^* = 1,\ R_i^{j*} = \max\left(\tilde{R}_i^{j*}, 0\right) > 0 \Rightarrow R_i^{j*} = A_{ij}^* \tilde{R}_i^{j*} \\ \forall A_{ij}^* = 0,\ R_i^{j*} = 0, \qquad\qquad\qquad\quad \Rightarrow R_i^{j*} = A_{ij}^* \tilde{R}_i^{j*}. \end{cases}$$

Thus, for all $i = 1, 2,\dots, n$, $j = 1, 2,\dots, m$, $R_i^{j*} = A_{ij}^* \tilde{R}_i^{j*}$, and the optimal value of (3) is achieved in (4) with $\left\{B_i^{j*}\right\}$. Therefore, the two problems (3) and (4) have the same optimal value and solution. ■

Theorem 1 shows that knowing the optimal channel assignment $\mathbf{A}^*$ is a sufficient condition under which solving a convex optimization obtains the same optimal solution and value of the original non-convex one. Thus, finding the optimal channel and power allocation can be separated into a two-step procedure: i) finding the optimal channel assignment, ii) finding the optimal power allocation that conforms to this channel assignment. From Theorem 1, step ii) is a convex optimization and can be solved with polynomial complexity. So all the NP complexity is embodied in step i). We usually do not know $\mathbf{A}^*$ in advance, and there are in total $2^{m \cdot n}$ possible $\mathbf{A}^*$ matrices. Since we can search over all $\mathbf{A}^*$ (each followed by a convex optimization of polynomial complexity,) $2^{m \cdot n}$ is an upper bound on the NP part of the complexity. In the next sub-section, we will show that the NP part of the complexity in solving this non-convex optimization can be further reduced to 2^n.

B. Dual Decomposition Method

First, we rewrite the original objective function (2) as

$$\sum_{j=1}^{m}\left(\sum_{i=1}^{n} w_i R_i^j\right). \tag{5}$$

We now consider the Lagrange dual problem of (3) [2]:

$$\min_{\boldsymbol{\lambda} \geq 0} \quad \sum_{j=1}^{m} g_j(\boldsymbol{\lambda}) + \sum_{i=1}^{n} \lambda_i (P_i^{\max} + m), \tag{6}$$

where $g_j(\boldsymbol{\lambda}) = \sup_{B_i^j \geq 0} \sum_{i=1}^{n}\left(w_i R_i^j - \lambda_i \cdot 2^{B_i^j}\right),$ (7)

$j = 1, 2, ..., m$, and $\boldsymbol{\lambda} = [\lambda_1\, \lambda_2 \ldots \lambda_n]'$.

The dual master problem (6) is a convex optimization, because $g_j(\boldsymbol{\lambda})$ is the supremum of linear functions of $\boldsymbol{\lambda}$, and hence convex in $\boldsymbol{\lambda}$ [2]. As will be shown later, a subgradient of $g_j(\boldsymbol{\lambda})$ can be obtained simultaneously while we evaluate $g_j(\boldsymbol{\lambda})$, and we apply a subgradient method to solve the dual master problem in Section IV.

In each outer iteration of updating $\boldsymbol{\lambda}$ in solving the master problem, we need to solve *m* sub-problems (7) which are non-convex optimizations because of the non-concavity of R_i^j. Since $\sum_{i=1}^{n} w_i R_i^j$ in the objective of (7) is just a single-channel case (with a small modification) of the original problem's objective (5), Theorem 1 can be applied to the sub-problems, too. We define $\mathbf{A}_j^*$ to be the *n* by 1 optimal (0,1) activity vector (a single-channel case of $\mathbf{A}^*$) that corresponds to the optimal solution of (7) $\hat{\mathbf{B}}_j^* = [\hat{B}_1^{j*}\ \hat{B}_2^{j*}, ..., \hat{B}_n^{j*}]'$, and the following corollary holds:

Corollary 1: Given $\mathbf{A}_j^*$ of (7) in advance, the convex relaxation

$$\max_{B_i^j \geq 0} \sum_{i=1}^{n}\left(w_i A_j^*(i) \cdot \tilde{R}_i^j - \lambda_i \cdot 2^{B_i^j}\right) \tag{8}$$

gives the same optimal value and solution as the original sub-problem's (7).

Because there are in total 2^n possible $\mathbf{A}_j^*$ vectors, the NP part of the complexity in solving the j^{th} sub-problem is upper bounded by 2^n. The reduction in complexity from $2^{m \cdot n}$ to 2^n comes from the fact that the dual problem is decomposed in channels: In each outer iteration, the *m* sub-problems are solved independently. Thus, the complexity of solving the dual problem is linear in *m*, and the NP part of the complexity is upper bounded by 2^n. In each channel, since the best among all possible activity vectors is selected, the optimal interference avoidance is pursued, and the low SINR conditions are automatically avoided.

As will be shown in Section IV, from solving the dual problem (6), we can obtain both primal and dual feasible solutions, which provide lower and upper bounds respectively on the primal optimal value. Because the primal problem (3) is non-convex, the duality gap between (3) and (6) is not necessarily zero. However, it has been shown in the literature that as the number of sub-channels goes to infinity, the duality gap of this problem goes to *zero* [15] [16], i.e. the dual decomposition method is asymptotically optimal. In Section V, we will show in simulations that the dual decomposition method provides solutions separated from the optimum by almost zero gaps.

C. Further Reduction of Algorithm Complexity in Cellular Structured Networks

In a cellular structured network (e.g. WLAN, cell phone system) with *c* Access Points (AP) or Base Stations (BS), every user chooses one AP with which to communicate. In this paper, we assume APs are single antenna devices. Denote the number of users communicating with the k^{th} AP, i.e. the number of users in cell *k* by n_k, $\sum_{k=1}^{c} n_k = n$.

We first consider uplink intra-cell channel allocation, i.e. the multiple access problem for the users that belong to the same cell. We have the following Theorem:

Theorem 2: For a cellular structured network, the optimal solution of problem (3) must satisfy the following condition: in any channel, in any cell, among all the users in this cell, only one (or no) user is active (i.e. transmitting).

Proof: WLOG, consider cell 1 that has users 1, 2, …, n_1, and channel 1. Denote P_i, $i = 1, 2, ..., n_1$ to be the *received* power of user *i* in channel 1 at AP 1. Let $i^* = \arg\max_{1, 2, ..., n_1}(P_i)$. From feature 6 of the deterministic channel model in Table I, $\forall i \neq i^*$, user *i*'s received power is buried below its interference plus noise floor, and user *i* has *zero* rate. In other words, in any channel, in any cell, there is at most one user that achieves a non-zero rate. Therefore, in the optimal power and channel allocation scheme, in any channel, there is at most one co-cell user active. ∎

Theorem 2 also trivially holds for the downlink scenario in any cell. The above intra-cell orthogonalization theorem is based on the basic assumption that interference is treated as noise. We note that Theorem 2 is derived for the deterministic channel model. In the low SINR case of Gaussian channels, intra-cell orthogonalization is not always optimal.

Now we apply Theorem 2 in the dual decomposition method. In a cellular structured system, for each sub-problem (7), only those activity vectors that satisfy the intra-cell orthogonalization condition are possible candidates of the *optimal* activity vector. Since the k^{th} cell can only have one or zero user (among n_k users) active, there are in total $\prod_{k=1}^{c}(n_k+1)$ candidate activity vectors $\mathbf{A}_j$. Clearly, with a fixed number of cells *c*, for any total number of users *n*,

$$\prod_{k=1}^{c}(n_k+1) \le \left(\frac{\sum_{k=1}^{c}(n_k+1)}{c}\right)^c = \left(\frac{n}{c}+1\right)^c \ll 2^n .$$

This reduces the complexity of solving (3) by the dual decomposition method to be strictly polynomial.

IV. Algorithm Design: Dual Decomposition with Activity Vector Search, Subgradient Projection and Backtracking

In this section, we provide a novel algorithm that solves the optimal power and channel allocation of (3). (Simulation results are provided in Section V.) With a dual decomposition (6), for each single channel sub-problem (7), we solve convex relaxations with an activity vector search which guarantees global optimality. For cellular structured networks, this algorithm has polynomial complexity.

A. Subgradients in the Master and Sub Problems

For the master problem (6), a basic property is that the following vector is a subgradient of the master objective function at $\boldsymbol{\lambda}$ [14]:

$$\mathbf{h} = -\sum_{j=1}^{m}[2^{\hat{B}_1^{j*}} \; 2^{\hat{B}_2^{j*}} \; ... \; 2^{\hat{B}_n^{j*}}]' + \mathbf{P}^{\max} + \mathbf{1}\cdot m \qquad (9)$$

where $\hat{\mathbf{B}}_j^* = [\hat{B}_1^{j*} \; \hat{B}_2^{j*}, ..., \hat{B}_n^{j*}]'$ is the optimal solution of the *j*th sub-problem (7) with the current $\boldsymbol{\lambda}$, and $\mathbf{P}^{\max} = [P_1^{\max} \; P_2^{\max} ... P_n^{\max}]'$. The proof can be found in e.g. [15].

To solve the sub-problem, we apply the *convex relaxation with activity vector search* as discussed in Section III. For each activity vector $\mathbf{A}_j$, the relaxed objective function is

$$\sum_{i=1}^{n}\left(w_i A_j(i)\cdot \tilde{R}_i^j - \lambda_i \cdot 2^{B_i^j}\right) , \qquad (10)$$

where $\tilde{R}_i^j = (B_i^j - D_i^j) - \max\left(\max_{k\neq i}(B_k^j - I_{k,\,i}^j), n_i^j\right)$.

Clearly, all the inactive B_i^j are kept at zero always. To compute a subgradient of $\tilde{R}_i^j$ with respect to all the active B_i^j, we apply the following lemma on the subgradient of the pointwise maximum function:

Lemma 1: For $f(x) = \max\{f_1(x), ..., f_n(x)\}$, $f_i(x)$ convex for all $i = 1,2,...,n$, define $I(x) = \{i \mid f_i(x) = f(x)\}$. Choose any $k \in I(x)$, then any subgradient of $f_k(x)$ is a subgradient of $f(x)$ at point *x*.

This is a basic result for subgradient, and its proof can be found in e.g. [12]. We apply Lemma 1 on the expression of $\tilde{R}_i^j$: For each active $\tilde{R}_i^j$, an *n* by 1 subgradient $\mathbf{g}_i$ of $\tilde{R}_i^j$ can be computed by the following procedure:

Procedure 1:

1. Initialize $\mathbf{g}_i$, such that $\mathbf{g}_i(k) = \begin{cases} 1, & if \;\; k = i \\ 0, & if \;\; k \neq i \end{cases}, \;\; k = 1, 2, ..., n$.

2. If $n_i^j \ge \max_{k\neq i}(B_k^j - I_{k,\,i}^j)$, return.

   Otherwise, choose any $k^* \neq i$ that satisfies

   $$B_{k*}^j - I_{k*,\,i}^j = \max_{k\neq i}(B_k^j - I_{k,\,i}^j),$$

   and let $\mathbf{g}_i = \mathbf{g}_i - \mathbf{e}_{k*}$, where $\mathbf{e}_{k*}$ is the elementary vector:

   $$\mathbf{e}_{k*}(k) = \begin{cases} 1, & if \;\; k = k^* \\ 0, & otherwise \end{cases}, \;\; k = 1, 2, ..., n \cdot$$

With all $\mathbf{g}_i$, a subgradient of the full objective (10) is readily obtained:

$$\mathbf{g} = \sum_{i,\; s.t. A_j(i)=1}\left(w_i \cdot \mathbf{g}_i - \lambda_i \cdot \ln 2 \cdot 2^{B_i^j} \cdot \mathbf{e}_i\right) \qquad (11)$$

This procedure to compute the subgradient (11) has a clear intuition behind it. For every active user, the 1st step of *Procedure 1* encourages this user to transmit 1 more bit, because it can increase its own rate by 1 bit (which corresponds to $B_i^j - D_i^j$ in the expression of $\tilde{R}_i^j$.) In the 2nd step, every active user pinpoints one of its *dominant interferers* – user k^*. If the dominant interference is higher than noise, this dominant interferer will be encouraged to transmit one fewer bit.

B. Subgradient Projection and Backtracking

The master problem (6) and the relaxed sub problems (8) have two common properties:

1) They are both convex optimizations with computable subgradients (as in part A.)

2) The constraints in (6) and (8) are both *non-negative orthants* $\mathbf{R}_+^n$.

With 1) and 2), the *subgradient projection* method becomes simple to apply, because the projection of a vector **x** on the non-negative orthant $\mathbf{R}_+^n$ is simply max(**x**, 0) (comparison made element-wise.)

In each inner iteration for the j^{th} sub-problem with a specified activity vector $\mathbf{A}_j$, the solution is updated by

$$\mathbf{B}_j^{(r+1)} = \max(\mathbf{B}_j^{(r)} + t_r \cdot \mathbf{g}, 0), \text{ where } \mathbf{B}_j = [B_1^j\ B_2^j\ B_3^j \dots B_n^j]', \quad (12)$$

r is the iteration index, and t_r is the r^{th} step size. There are various ways of choosing the step size t_r [12]. For this problem, using a constant step size in (12) will suffice, and it converges much faster than diminishing step sizes.

For the master problem, however, simple subgradient projection will cause severe numerical problems for the following reasons. In a projection iteration $\boldsymbol{\lambda}^{(\tilde{r}+1)} = \max(\boldsymbol{\lambda}^{(\tilde{r})} - \tilde{t}_{\tilde{r}} \cdot \mathbf{h}, 0)$, suppose $\exists k,\ \lambda_k^{(\tilde{r}+1)} = 0$, i.e. a non-trivial projection of λ_k on $\mathbf{R}_+$ is performed. Then in the next outer iteration with $\boldsymbol{\lambda}^{(\tilde{r}+1)}$, the sub-problem becomes

$$g_j(\boldsymbol{\lambda}^{(\tilde{r}+1)}) = \sup_{B_i^j \geq 0} \left(\sum_{i \neq k} \left(w_i R_i^j - \lambda_i^{(\tilde{r}+1)} \cdot 2^{B_i^j} \right) + w_k R_k^j \right).$$

Clearly, the lack of penalty on B_k^j due to $\lambda_k^{(\tilde{r}+1)} = 0$ will result in $g_j(\boldsymbol{\lambda}^{(\tilde{r}+1)}) = \infty$, since B_k^j (and hence R_k^j) is unconstrained above to infinity in the sub-problem. Therefore, the situation with any $\lambda_i = 0$ must be avoided, meaning that no simple projection to the *boundary* of $\mathbf{R}_+^n$ should be performed in the subgradient method. Thus, backtracking of $\boldsymbol{\lambda}^{(r+1)}$ to the strictly positive orthant is necessary. Various backtracking directions can be used, e.g. backtracking along the direction from the projected point, or along the subgradient direction. In our algorithm, we backtrack **λ** along the subgradient direction when the projected point is on the boundary of $\mathbf{R}_+^n$.

C. Dual Decomposition with Activity Vector Search

The complete dual decomposition algorithm with activity vector search is listed as *Algorithm 1*. It consists of outer iterations and inner iterations that solve the master and sub problems respectively. In each outer iteration, *m* single-channel sub-problems are solved. For the j^{th} sub-problem, (j = 1, 2,…, *m*,) all activity vectors $\mathbf{A}_j$ that are not ruled out to be optimal are traversed. With each $\mathbf{A}_j$, a subgradient projection method that consumes inner iterations is applied to solving the relaxed sub-problem (8). In general networks, there are 2^n candidate $\mathbf{A}_j$, whereas in cellular structured networks, there are $\prod_{k=1}^{c}(n_k+1)$ candidates. After traversing all candidate activity vectors in each of the *m* channels, a dual achievable objective value with the current **λ** is obtained. If the corresponding $\{B_i^j\}$ is a primal feasible solution, we update the best primal solution and the best primal achievable value. At the end of an outer iteration, **λ** is updated with the computed subgradient (9) using backtracking method. After the algorithm terminates, the dual optimal value that serves as an upper bound on the optimal value of the original primal problem (3) is obtained. Meanwhile, the best achieved primal feasible solution is also obtained. The gap between these two is asymptotically zero.

Algorithm 1

Dual Decomposition with Activity Vector Search, Subgradient Projection and Backtracking

Initialize $\boldsymbol{\lambda} = [\lambda_1\ \lambda_2\ \lambda_3 \dots \lambda_n]'$

Initialize $\{B_i^j\}$ ($\mathbf{B}_1, \mathbf{B}_2, \dots, \mathbf{B}_m$)

Repeat (Outer iterations)

  For j = 1 to m (traverse all channels)

    For l = 1 to $\prod_{k=1}^{c}(n_k+1)$ (or 2^n, traverse all $\mathbf{A}_j$)

      Apply the next candidate activity vector $\mathbf{A}_j$

      Repeat (Inner iterations)

      Compute the sub-problem objective value

      Compute the sub-problem subgradient **g** (11)

      Update $\mathbf{B}_j$ with **g** using subgradient projection (12)

      Until $\mathbf{B}_j = [B_1^j\ B_2^j\ B_3^j\ \dots\ B_n^j]'$ converges

    End

  End

  Update the lowest dual achievable value, i.e. the upper bound on the primal optimal value.

  If $\{B_i^j\}$ are primal feasible, update the best primal feasible solution and the best primal achievable value.

  Compute the master problem subgradient **h** (9)

  Update **λ** with **h** using subgradient method with backtracking

Until **λ** converges

V. Simulation and Results

In this section, we provide simulation results of the proposed algorithm applied to cellular structured networks. Several comments are made on the performance observed.

A. Simulation Scenario

1. Wireless Channel Model

We consider path loss (PL), shadowing and multipath fading while computing the wireless channel gains. All simulated channel gains are translated into numbers of bit level shifts, to be applied in the deterministic channel model. We employ a simplified path loss model [10]: $P_r = P_t K \left(\frac{d_0}{d}\right)^{-\gamma}$. We consider an indoor propagation scenario with $\gamma = 3,\ d_0 = 2m$, and normalize $K = 1$. Since the channel gains are represented by the number of bit-level shifts (by taking a base 2 Log on the actual power gain,) a numerical example of the above PL setting is that when d = 20*m*, $PL = 10^{-3} \approx 2^{-10}$ which corresponds to a *10 bit-level shift*. We assume an indoor Log-Normal shadowing with a variance of 3dB [9]. It naturally corresponds to a $N(0,1)$ Gaussian random variable added to the channel gains in terms of the number of bit-level shifts. Finally, since we are interested in optimizations of multi-channel frequency selective problems, we assume independent Rayleigh fading in all parallel channels.

2. Geometric Setting and Optimization Parameters

We setup a *seven-cell* scenario, in which the positions of the 7 APs form a hexagon. The distance between adjacent APs is set to be 20*m*. We simulate multiple independent realizations of 15 users uniformly scattered within a circle with a radius of 30*m*. Each user communicates with the AP to which it has the shortest distance. A typical realization of the above setting is plotted in Fig. 3.

We consider 20 parallel channels that are available to all users. We assume a common receiver noise floor on bit-level *zero*. In our simulation, we assume that all users have a common power constraint of 16 transmit bit-levels per channel in 20 channels (or equivalently, 20.3 transmit bit levels with only one channel occupied.) In other words, if a user's distance to its AP is 20*m* with no shadowing and fading, it can transmit 120 bits if it occupies all 20 channels (or 10.3 bits if it occupies one channel,) interference free. Finally, we use an equal weight on all users' rates in the objective in our simulation.

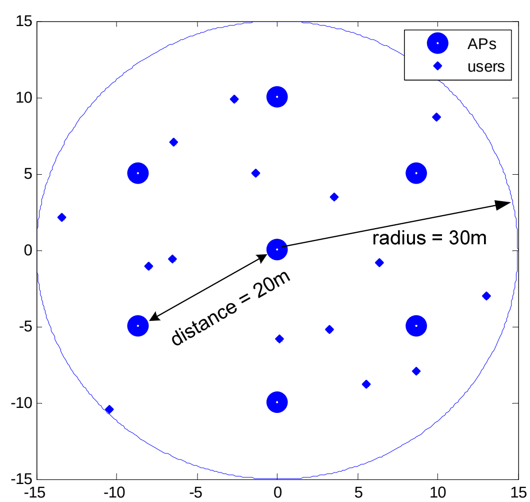

Figure 3. The geometric setting of the simulation. 7 cells, 15 users.

B. Simulation Results

With 20 parallel channels, 20 independent realizations of 15 users scattered in 7 cells are simulated, and results are averaged. As the number of outer iterations grows, the best sum-rate achieved and the sharpest upper bound obtained are plotted in Fig. 4. A closer look into the gaps between the best sum-rate achieved and the *final* upper bound (since we know it a-posteriori) is given in Fig. 5, in which the ratio between the gap and the optimum is plotted. We make several interesting observations. First, within less than 15 iterations, the gap between the best sum-rate achieved to the final upper bound falls exponentially to a very small level. Second, there is a non-zero floor for this gap, essentially due to the small but still non-

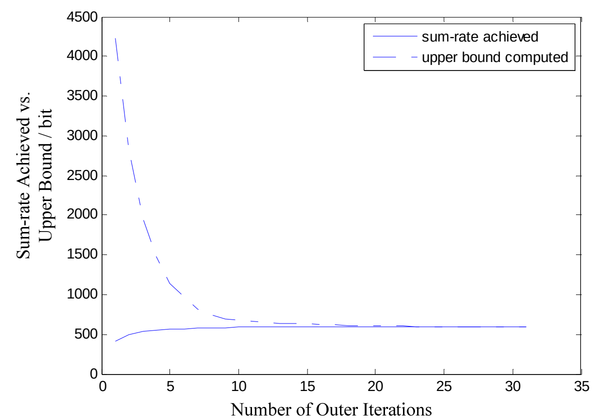

Figure 4. Primal feasible sum-rates achieved and the upper bounds obtained from dual feasible solutions, averaged over 20 independent realizations.

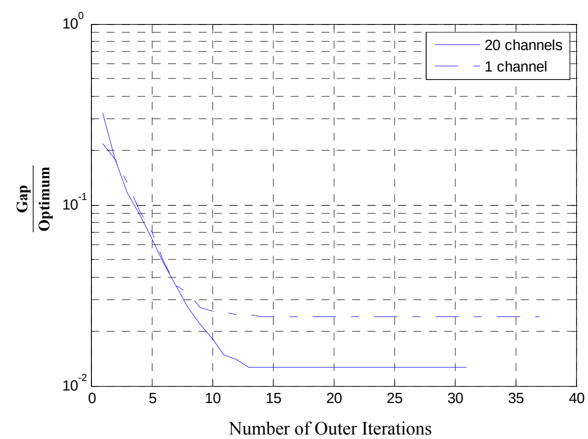

Figure 5. Gaps between the sum-rates achieved and the final upper bound, divided by the maximum achieved sum-rate, plotted in Log Scale. 20 channels vs. 1 channel

zero duality gap from the non-convexity of the problem. Third, after we plot the simulation results of the *single*-channel case as a comparison to the 20-channel case (Fig. 5), we see a similar exponential decay of the performance gap, but a relatively higher floor ($Gap/Optimum = 0.0242$) than the 20 channel case ($Gap/Optimum = 0.0127$). This is consistent with the intuition from the asymptotic zero duality gap result in the literature (as pointed out in Section III.B.): the non-zero duality gap goes to zero as the number of sub-channels goes to infinity.

VI. CONCLUSIONS

We designed a low complexity power and channel allocation algorithm that approaches the optimal throughput performance using the recently developed deterministic channel model. We formulated the weighted sum-rate maximization into a new form. Although it is still a non-convex problem, we proved that knowing the optimal channel assignment (i.e. the activity matrix/vector) is sufficient for solving a convex optimization to get to the original global optimal power allocation scheme. Applying this activity vector search idea with the dual decomposition method, we reduced the complexity of solving the non-convex optimal power and channel allocation to 2^n . We further show that for cellular structured networks with a fixed number of cells, this complexity can be reduced to be strictly polynomial. We designed the complete algorithm applying subgradient methods. While our algorithm provides primal feasible solutions which converge to the optimum, it also provides upper bounds from dual feasible solutions, acting as a check of the performance gap from the achieved to the optimum. Simulation results have shown that the proposed algorithm achieves the global optimum with almost zero gaps.

The polynomial complexity of our algorithm enables solving the global optimal solutions of the non-convex problem of power and channel allocation for relatively large networks. It thus can serve as a benchmark in performance evaluations, especially on how far from optimality distributed and lower complexity algorithms perform. However, the complexity of $\left(n/c+1\right)^c$ is still quite high while the number of cells c is large. Future research to further reduce the complexity of solving such problems with performance guarantees remains very interesting.

REFERENCES

[1] A.S. Avestimehr, S.N. Diggavi, and D. Tse, “A deterministic approach to wireless relay networks,” *Proceedings of Allerton Conference on Communication, Control, and Computing, Illinois, USA,* September 2007.

[2] Boyd, S.P., Vandenberghe, L., "Convex Optimization," Cambridge University Press, 2004.

[3] Bresler, G. and Tse, D.N.C., "The Two-User Gaussian Interference Channel: A Deterministic View", vol 19, *European Transactions in Telecommunications*, pp. 333-354, April 2008.

[4] Bresler, G., Parekh, A. and Tse, D.N.C., “The approximate capacity of the many-to-one and one-to-many Gaussian interference channels,” arXiv:0809.3554v1, 2008.

[5] Cendrillon, R.; Jianwei Huang; Mung Chiang; Moonen, M., "Autonomous Spectrum Balancing for Digital Subscriber Lines," *Signal Processing, IEEE Transactions on* , vol.55, no.8, pp.4241-4257, Aug. 2007.

[6] Cendrillon, R.; Wei Yu; Moonen, M.; Verlinden, J.; Bostoen, T., "Optimal multiuser spectrum balancing for digital subscriber lines," *Communications, IEEE Transactions on* , vol.54, no.5, pp. 922-933, May 2006.

[7] Mung Chiang, "Balancing transport and physical Layers in wireless multihop networks: jointly optimal congestion control and power control," *Selected Areas in Communications, IEEE Journal on* , vol.23, no.1, pp. 104-116, Jan. 2005

[8] Mung Chiang; Chee Wei Tan; Palomar, D.P.; O'Neill, D.; Julian, D., "Power Control By Geometric Programming," *Wireless Communications, IEEE Transactions on* , vol.6, no.7, pp.2640-2651, July 2007

[9] Ghassemzadeh, S.S.; Greenstein, L.J.; Kavcic, A.; Sveinsson, T.; Tarokh, V., "UWB indoor path loss model for residential and commercial buildings," *Vehicular Technology Conference, 2003. VTC 2003-Fall. 2003 IEEE 58th* , vol.5, no., pp. 3115-3119 Vol.5, 6-9 Oct. 2003.

[10] Goldsmith, A., “Wireless Communications,” Cambridge University Press, 2005.

[11] Jianwei Huang; Berry, R.A.; Honig, M.L., "Distributed interference compensation for wireless networks," *Selected Areas in Communications, IEEE Journal on* , vol.24, no.5, pp. 1074-1084, May 2006

[12] N. Z. Shor. “Minimization Methods for Non-di erentiable Functions,” Springer Series in Computational Mathematics. Springer, 1985.

[13] Chee Wei Tan; Palomar, D.P.; Mung Chiang, "Solving nonconvex power control problems in wireless networks: low SIR regime and distributed algorithms," *Global Telecommunications Conference, 2005. GLOBECOM '05. IEEE* , vol.6, no., pp. 6 pp.-, 28 Nov.-2 Dec. 2005.

[14] Vandenberghe, L. “Gradient Projection,” (April 2009). Lecture notes on Optimization Methods for Large-Scale Systems. [Online]. Available http://www.ee.ucla.edu/~vandenbe/236C/lectures/gprojection.pdf

[15] Wei Yu; Lui, R., "Dual methods for nonconvex spectrum optimization of multicarrier systems," *Communications, IEEE Transactions on* , vol.54, no.7, pp. 1310-1322, July 2006.

[16] Yue Zhao and Gregory J. Pottie, "Optimal Spectrum Management in Multiuser Interference Channels," *Information Theory, 2009. ISIT 2009. IEEE International Symposium on,* vol., no., pp.2266-2270, June 28 – July 3, 2009.