Optimized Distributed Inter-Cell Interference Coordination (ICIC) Scheme Using Projected Subgradient and Network Flow Optimization

Akram Bin Sediq, *Student Member, IEEE*, Rainer Schoenen, *Senior Member, IEEE*,
Halim Yanikomeroglu, *Senior Member, IEEE*, and Gamini Senarath

Abstract—In this paper, we tackle the problem of multi-cell resource scheduling, where the objective is to maximize the weighted sum-rate through inter-cell interference coordination (ICIC). The blanking method is used to mitigate the inter-cell interference, where a resource is either used with a predetermined transmit power or not used at all, i.e., blanked. This problem is known to be strongly NP-hard, which means that it is not only hard to solve in polynomial time, but it is also hard to find an approximation algorithm with guaranteed optimality gap. In this work, we identify special scenarios where a polynomial-time algorithm can be constructed to solve this problem with theoretical guarantees. In particular, we define a dominant interference environment, in which for each user the received power from each interferer is significantly greater than the aggregate received power from all other weaker interferers. We show that the originally strongly NP-hard problem can be tightly relaxed to a linear programming problem in a dominant interference environment. Consequently, we propose a polynomial time distributed algorithm that is not only guaranteed to be tight in a dominant interference environment, but which also computes an upper bound on the optimality gap without additional computational complexity. The proposed scheme is based on the primal-decomposition method, where the problem is divided into a master-problem and multiple subproblems. We solve the master-problem iteratively using the projected-subgradient method. We also show that each subproblem has a special network flow structure. By exploiting this network structure, each subproblem is solved using the network-based optimization methods, which significantly reduces the complexity in comparison to the general-purpose convex or linear optimization methods. In comparison with baseline schemes, simulation results of the International Mobile Telecommunications-Advanced (IMT-Advanced) scenarios show that the proposed scheme achieves higher gains in aggregate throughput, cell-edge throughput, and outage probability.

I. INTRODUCTION

AGGRESSIVE reuse is inevitable in cellular networks due to the scarcity of the radio resources. Reuse-1, in which all radio resources are reused in every sector, is an example of such an aggressive reuse scheme. While reuse-1 can potentially achieve high aggregate system throughput, it jeopardizes the throughput experienced by users close to the cell[1]-edge, due to the excessive interference experienced by these users. Therefore, it is vital for the network to use robust and efficient interference mitigation techniques.

Conventionally, interference is mitigated by static resource partitioning and frequency/sector-planning, where close-by sectors are assigned orthogonal resources (clustering). A common example is reuse-3 (cluster size = 3 sectors) [2, Sec. 2.5], where adjacent sectors are assigned orthogonal channels, i.e., there is no inter-cell interference between adjacent sectors. Although such techniques can reduce inter-cell interference and improve cell-edge user throughput, they suffer from two major drawbacks. First of all, the aggregate network throughput is significantly reduced since each sector has only a fraction of the available resources, which is equal to the reciprocal of the reuse factor. Secondly, conventional frequency/sector-planning may not be possible in emerging wireless networks where new multi-tier network elements (such as relays, femto-/pico-base-stations, distributed antenna ports) are expected to be installed without prior planning in self-organizing networks (SON) [3].

To reduce the effect of the first drawback, fractional frequency reuse (FFR) schemes have been proposed. The key idea in FFR is to assign lower reuse factor for users near the cell-center and higher reuse factor for users at the cell-edge. The motivation behind such a scheme is that cell-edge users are more vulnerable to inter-cell interference than cell-center users. Soft Frequency Reuse (SFR) [4] and Partial Frequency Reuse (PFR) [5] are two examples of FFR. A comparative study between reuse-1, reuse-3, PFR, and SFR is provided in [6], [7]. While FFR schemes recover some of the throughput loss due to partitioning, they require frequency/sector planning *a priori*, which is not desirable in future cellular networks as mentioned

---

[1]The terms "cell" and "sector" are used interchangeably in this work.

earlier. As a result, developing efficient dynamic inter-cell interference coordination (ICIC) schemes is vital to the success of future cellular networks [3].

One approach to tackle the ICIC problem is to devise adaptive FFR or SFR schemes. In [8], an adaptive FFR scheme is proposed where each base-station (BS) chooses one of several reuse modes. A dynamic and centralized[2] FFR scheme is proposed in [9] that outperforms conventional FFR schemes in terms of the total system throughput. In [10], the authors propose softer frequency reuse, which is a heuristic algorithm based on modifying the proportional fair algorithm and the SFR scheme. In [11], a heuristic algorithm based on adaptive SFR is proposed for the uplink. In [12], the authors propose gradient-based distributed schemes that create SFR patterns, in an effort to achieve local maximization of the network utility. To enable gradient computation, the rate adaptation function that maps signal-to-interference-and-noise-ratio (SINR) to rate is required to be differentiable. However, this requirement is not currently feasible since rate adaption is performed in current cellular standards by using adaptive modulation and coding (AMC) with a discrete set of modulation and coding schemes [13], which results in a non-differentiable rate adaptation function.

Another approach to tackle the ICIC problem is to remove the limitations imposed by the FFR schemes and view the ICIC problem as a multi-cell scheduling problem. In [14], the authors propose a suboptimal centralized algorithm where a radio network controller is assumed to be connected to all BSs. The authors conclude that the problem is NP-hard and they resort to heuristics that show improvement in the performance. In [15], a graph-theoretic approach is taken to develop an ICIC scheme in which information about the interference experienced by each user terminal (UT) is inferred from the diversity set of that UT. In [16], a game-theoretic approach is pursued and an autonomous decentralized algorithm is developed. The proposed algorithm converges to a Nash equilibrium in a simplified cellular system. Nevertheless, a significant gap is observed between the proposed algorithm and the globally optimum one, which is computationally demanding. In [3], a partly distributed two-level ICIC scheme is proposed where a centralized entity is required to solve a binary linear optimization problem, which is generally not solvable in polynomial time.

The focus of this paper is to overcome the drawbacks of the existing schemes mentioned above by developing a distributed ICIC scheme, that can work with any AMC scheme, and yet provide theoretical guarantee. In particular, we tackle the problem of distributed multi-cell resource scheduling, where the objective is to maximize the weighted sum-rate. The blanking method is used to mitigate the inter-cell interference, where a resource is either used with a predetermined transmit power or not used at all, i.e., blanked, similar to [3]. This problem is known to be strongly NP-hard, which means that it is not only hard to solve in polynomial time, but it is also hard to find an

approximation algorithm with guaranteed optimality gap [17]. The main contributions can be summarized as follows:

1) We show that the originally strongly NP-hard problem can be tightly relaxed to a linear programming problem in a dominant interference environment, in which for each user the received power from each interferer is significantly greater than the aggregate received power from all other weaker interferers.[3] The tightness of the relaxation is shown by proving that the percentage of optimal relaxed variables that assume binary values is bounded below by $\frac{\tilde{K}(\bar{M}-1)}{(\tilde{K}+1)\bar{M}+1} \cdot 100\%$, where $\bar{M}$ is the average number of UTs per sector and $\tilde{K}$ is the number of neighboring interferers. Consequently, we devise a polynomial-time distributed algorithm that is not only guaranteed to be tight, but which also computes an upper bound on the optimality gap without additional computational complexity. The proposed algorithm does not require a central controller and it can be used with any AMC scheme, including AMC schemes that result in non-differentiable discrete rate adaption functions.

2) We demonstrate that considered optimization problem for a dominant interference environment can be transformed into an equivalent minimum-cost network flow (MCNF) optimization problem. Thus, we can use network-based algorithms which have significantly reduced complexity as compared to the general-purpose convex or linear optimization algorithms [18, p. 402]. While MCNF optimization tools have been used in the literature to solve resource allocation problem in single-cell networks, e.g., [19] and [20], as far as we know, our work is the first in literature to use these optimization tools in ICIC, i.e., in multi-cell resource allocation problems.

II. SYSTEM MODEL

We consider the network model described by the International Mobile Telecommunications-Advanced (IMT-Advanced) evaluation guidelines [21]. Based on these guidelines, the considered network consists of K sectors served by $K/3$ BSs as shown in Fig. 1. Each BS is equipped with a tri-sector antenna to serve a cell-site that consists of 3 sectors. Each BS can communicate with its neighboring BSs; this is supported in most cellular network standards, e.g., using the R8 interface in IEEE 802.16m standard and the X2 interface in Long-Term Evolution (LTE) and LTE-advanced standards. Hexagonal sectors are considered herein according to IMT-Advanced guidelines; nevertheless, the proposed scheme works also for arbitrary sector shapes. We focus on the downlink scenario in this paper. For convenience, the symbols used in this section and onward are provided in Table I.

---

[2]In this paper, an algorithm is categorized as a centralized algorithm if it requires a centralized entity to coordinate or to perform either partial or full execution of the algorithm. In contrast, a distributed algorithm is one that can be executed in each sector with limited message exchange between the sectors, without the requirement of having a centralized entity.

[3]The definition of a dominant interference environment is illustrated by the following example. Assume there are four interferers, and let $Pr^{(k)}$ denote the received power from the kth interferer. Assume further that $Pr^{(1)} > Pr^{(2)} > Pr^{(3)} > Pr^{(4)}$. Such an environment is called a dominant interference environment if and only if the following conditions are satisfied: $Pr^{(1)} \gg Pr^{(2)} + Pr^{(3)} + Pr^{(4)}$, $Pr^{(2)} \gg Pr^{(3)} + Pr^{(4)}$, and $Pr^{(3)} \gg Pr^{(4)}$.
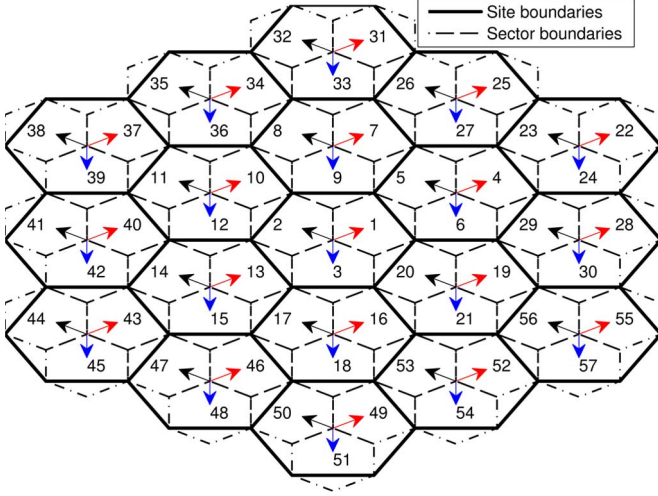
Fig. 1. The considered network layout which consists of 19 cell sites and 3 sectors per site.

TABLE I
LIST OF SYMBOLS

Symbol	Description
$M^{(k)}$	Number of UTs in sector k
N	Number of available RBs
K	Number of sectors
$\mathcal{M}^{(k)}$	Set of indices of users in sector k, i.e., $\mathcal{M}^{(k)} \triangleq \{1,\ldots,M^{(k)}\}$
$\mathcal{N}$	Set of indices of available RBs, i.e., $\mathcal{N} \triangleq \{1,\ldots,N\}$
$\mathcal{K}$	Set of indices of all sectors in the network, i.e., $\mathcal{K} \triangleq \{1,\ldots,K\}$
$\tilde{\mathcal{K}}^{(k)}$	Set of indices of the first-tier interfering sectors seen by sector k
$\tilde{K}$	The cardinality of $\tilde{\mathcal{K}}^{(k)}$, which is assumed to be the same $\forall k$
m	UT index; $m \in \mathcal{M}^{(k)}$
n	RB index; $n \in \mathcal{N}$
k	Sector index; $k \in \mathcal{K}$
$H_{m,n}^{(k,\tilde{k})}$	Channel gain from sector $\tilde{k}$ on RB n to UT m served by sector k
$R_{m,n}^{(k)}$	Achievable rate on RB n for UT m in sector k
$r_{m,n}^{(k)}$	Achievable rate on RB n for UT m in sector k, if RB n is reused by all other interfering sectors
$\tilde{r}_{m,n}^{(k,\tilde{k})}$	Additional rate on RB n for UT m in sector k, gained if RB n is blanked in sector $\tilde{k}$
$\bar{R}_m^{(k)}$	Average rate for UT m in sector k
$\Gamma_{m,n}^{(k)}$	SINR on RB n of UT m in sector k
$\gamma_{m,n}^{(k)}$	SINR on RB n of UT m in sector k, if all sectors reuse RB n
$\tilde{\gamma}_{m,n}^{(k,\tilde{k})}$	SINR on RB n of UT m in sector k, if sector $\tilde{k}$ blanks RB n
$x_{m,n}^{(k)}$	Binary variable indicating whether RB n is assigned to UT m in sector k or not
$I_n^{(k)}$	Binary variable indicating whether RB n is blanked in sector k or not

Orthogonal Frequency Division Multiple Access (OFDMA) is used as the multiple access scheme, since it is adopted in most of the contemporary cellular standards. Among the advantages offered by OFDMA is its scheduling flexibility, since users can be scheduled in both time and frequency, which can be exploited to gain time, frequency, and multi-user diversity. The time and frequency radio resources are grouped into time-frequency resource blocks (RBs). RB is the smallest radio resource block that can be scheduled to a UT and it consists of N_s OFDM symbols in the time dimension and N_f sub-carriers in the frequency dimension. The total number of RBs is denoted by N. The number of UTs in sector k is denoted by $M^{(k)}$. Both the BSs and the UTs are assumed to have single antenna each. Similar to [3], we assume that each UT estimates and reports to its serving BS the channel from its serving sector's antenna

and from the first-tier interfering sectors. The SINR observed by UT $m \in \mathcal{M}^{(k)} \triangleq \{1,\ldots,M^{(k)}\}$ in sector $k \in \mathcal{K} \triangleq \{1,\ldots,K\}$ on RB $n \in \mathcal{N} \triangleq \{1,\ldots,N\}$ is given by [3]

$$\Gamma_{m,n}^{(k)} = \frac{P_C H_{m,n}^{(k,k)}}{P_C \sum_{\tilde{k}=1,\tilde{k}\neq k}^{K} \left(1 - I_n^{(\tilde{k})}\right) H_{m,n}^{(k,\tilde{k})} + P_N} \tag{1}$$

where $I_n^{(k)}$ is a binary variable indicating whether RB n is blanked, i.e., not used, in sector k ($I_n^{(k)} = 1$) or not ($I_n^{(k)} = 0$), P_C represents the transmitted power per RB and it is assumed to be the same for all non-blanked RBs, P_N represents the thermal noise power per RB, and $H_{m,n}^{(k,\tilde{k})}$ represents the channel gain from sector $\tilde{k}$ on RB n to UT m served by sector k. Large scale channel variations (due to distance-dependant attenuation and shadowing), antenna gains, and multipath fading are all captured in $H_{m,n}^{(k,\tilde{k})}$. The achievable rate on RB n of UT m in sector k is given by

$$R_{m,n}^{(k)} = f\left(\Gamma_{m,n}^{(k)}\right) \quad \text{bit/sec,} \tag{2}$$

where $f(\cdot)$ is the AMC rate adaptation function that maps SINR to rate. The function $f(\cdot)$ is assumed to be nondecreasing with $f(0) = 0$, possibly discontinuous, which is the case for all practical AMC schemes.

III. PROBLEM STATEMENT

Let us consider a generic scheduler implemented in sector k, without ICIC, such that it maximizes the weighted sum of the UTs' rates in sector k. This scheduler is implemented by solving the following optimization problem for each RB n in every sub-frame t:

$$\underset{\{x_{m,n}^{(k)}(t):m\in\mathcal{M}^{(k)}\}}{\text{maximize}} \quad \sum_{m=1}^{M^{(k)}} w_m^{(k)}(t) x_{m,n}^{(k)}(t) R_{m,n}^{(k)}(t) \tag{3a}$$

$$\text{subject to} \quad \sum_{m=1}^{M^{(k)}} x_{m,n}^{(k)}(t) = 1, \tag{3b}$$

$$x_{m,n}^{(k)}(t) \in \{0,1\}, \quad \forall m, \tag{3c}$$

where $w_m^{(k)}(t)$ is the weight of UT m in sector k, $\{x_{m,n}^{(k)}(t)\}$ are the binary decision variables such that $x_{m,n}^{(k)}(t) = 1$ when RB n is assigned to UT m in sector k, and $x_{m,n}^{(k)}(t) = 0$ otherwise. Constraint (3b) ensures that each RB is assigned to only one user in sector k.

The scheduler described above can be used to control the desired fairness-throughput tradeoff in the long-term average rates by updating the weights $\{w_m^{(k)}(t)\}$ in every sub-frame. To elaborate, let $R_m^{(k)}(t)$ denote the data rate scheduled to user m in sector k at sub-frame t, i.e., $R_m^{(k)}(t) = \sum_{n=1}^{N} x_{m,n}^{(k)}(t) R_{m,n}^{(k)}(t)$ and let $\bar{R}_m^{(k)}(t)$ denote the average rate of user m in sector k in sub-frame

t, averaged over a window of t_c sub-frames using exponentially-weighted low-pass filter [22], i.e.,

$$\bar{R}_m^{(k)}(t) = \frac{1}{t_c} \sum_{i=0}^{t} \left(1 - \frac{1}{t_c}\right)^{i-t} R_m^{(k)}(i)$$

$$= \left(1 - \frac{1}{t_c}\right) \bar{R}_m^{(k)}(t-1) + \frac{1}{t_c} R_m^{(k)}(t). \qquad (4)$$

Moreover, let $\bar{R}_m^{(k)}$ denote the long-term (steady-state) average rate of user m in sector k, i.e., $\bar{R}_m^{(k)} = \lim_{t\to\infty} \bar{R}_m^{(k)}(t)$. A common approach to achieve various tradeoffs between fairness and aggregate throughput is to maximize a concave utility $U(\bar{R}_1^{(k)}, \ldots, \bar{R}_{M^{(k)}}^{(k)})$ that incorporates both fairness and aggregate throughput. It is shown in [23] and [24] that such a maximization can be achieved by solving the instantaneous optimization problem (3) in every t with the weights $\{w_m^{(k)}(t)\}$ chosen to be the marginal utility, i.e., $w_m^{(k)}(t) = \frac{\partial U(\bar{R}_1^{(k)}(t-1),\ldots,\bar{R}_{M^{(k)}}^{(k)}(t-1))}{\partial \bar{R}_m^{(k)}(t-1)}$. Optimality in the long-term average rates is attained for sufficiently large t_c and under ergodic channels. A widely used utility in this context is the α-fair utility defined for $\alpha \geq 0$ as [25]

$$U_\alpha\left(\bar{R}_1^{(k)}, \ldots, \bar{R}_{M^{(k)}}^{(k)}\right) = \begin{cases} \sum_{m=1}^{M^{(k)}} \log \bar{R}_m^{(k)}, & \alpha = 1, \\ \frac{1}{1-\alpha} \sum_{m=1}^{M^{(k)}} \left(\bar{R}_m^{(k)}\right)^{1-\alpha}, & \alpha \neq 1. \end{cases} \quad (5)$$

Based on this choice of utility, the user weights are given by $w_m^{(k)}(t) = \left(\bar{R}_m^{(k)}(t-1)\right)^{-\alpha}$, where various degrees of fairness-throughput tradeoff can be achieved by varying the fairness exponent, α [25]. In general, as α increases, the degree of fairness increases and the total throughput decreases, and vice versa (cf. [26]). For example, max-SINR scheduler corresponds to the case where $\alpha = 0$, proportional-fair scheduler (cf. [27], [28]) corresponds to the case where $\alpha = 1$, and max-min scheduler corresponds to the case where $\alpha \to \infty$. One can also use the recently proposed utility in [29] that achieves the optimum tradeoff between total throughput and Jain's fairness index by setting the user weights as $w_m^{(k)}(t) = \beta - \bar{R}_m^{(k)}(t-1)$, where β is a non-negative parameter, such that higher β results in higher total throughput and less fairness, and vice versa (cf. [29]).

Another approach to vary the degree of fairness-throughput tradeoff is by imposing a minimum long-term average rate ($\bar{R}_{\min}$), varying $\bar{R}_{\min}$ [30], and updating $\{w_m^{(k)}(t)\}$ according to the algorithm explained in [31]. We investigate both approaches in this paper.

It can be shown that the solution to (3) is given by

$$x_{m,n}^{\star(k)}(t) = \begin{cases} 1, & m = \arg\max_m w_m^{(k)}(t) R_{m,n}^{(k)}(t), \\ 0, & m \neq \arg\max_m w_m^{(k)}(t) R_{m,n}^{(k)}(t). \end{cases} \quad (6)$$

To simplify the notation, we drop the sub-frame index (t) in the rest of the paper.

To improve the performance of the above mentioned scheduler, we seek an ICIC scheme that coordinates the scheduling in all sectors to maximize the weighted sum of the rates of all users in the network. The optimization problem for each RB n to attain this goal can be formulated as

$$\underset{\substack{\{x_{m,n}^{(k)}, I_n^{(k)}: \\ k \in \mathcal{K}, m \in \mathcal{M}^{(k)}\}}}{\text{maximize}} \quad \sum_{k=1}^{K} \sum_{m=1}^{M^{(k)}} w_m^{(k)} x_{m,n}^{(k)} R_{m,n}^{(k)} \qquad (7a)$$

$$\text{subject to} \quad \sum_{m=1}^{M^{(k)}} x_{m,n}^{(k)} = 1 - I_n^{(k)}, \quad \forall k, \qquad (7b)$$

$$x_{m,n}^{(k)}, I_n^{(k)} \in \{0,1\}, \quad \forall k, m, \qquad (7c)$$

where the binary variables $\{I_n^{(k)}\}$ are introduced such that $I_n^{(k)} = 1$ when RB n is blanked in sector k and $I_n^{(k)} = 0$ otherwise. Constraint (7b) ensures that each RB n is assigned to only one UT in sector k, given that RB n is not blanked in sector k.

The network-wide optimization problem in (7) is difficult to solve for the following reasons. First of all, this problem belongs to the class of non-linear binary combinatorial optimization problems ($R_{m,n}^{(k)}$ is a nonlinear function of $I_n^{(k)}$, cf. (1) and (2)), which are generally difficult to solve in polynomial time. Indeed, (7) is strongly NP-hard even for the case of a single user per sector [17].[4] This means that not only (7) is hard to solve in polynomial time, but it is also hard to find an approximation algorithm with guaranteed optimality gap. Moreover, the objective function is dependant on the AMC rate adaptation function that is used. Since AMC rate adaptation functions are operator dependant, it is desirable to develop an algorithm that is independent of the chosen AMC rate adaptation function. Finally, it is desirable to solve (7) in a distributed manner since most contemporary standards (such as LTE, LTE-Advanced, IEEE 802.16m) do not support a central controller. The main contribution of this work is to circumvent these difficulties and develop an efficient algorithm that obtain a near-optimal solution of (7). The development of such an algorithm is explained in the following section.

IV. PROPOSED ALGORITHM

In this section, we show the steps used to develop the proposed algorithm. We start in Section IV-A by introducing a bound and argue that this bound is a good metric for optimization, especially for dominant interference environment. Using the bound, we get a binary non-linear optimization problem and we transform it into an equivalent binary linear optimization problem in Section IV-B. Next, we relax the resulted binary linear optimization problem into a linear optimization problem in Section IV-C. Then, we devise a distributed algorithm in Section IV-D by decomposing the problem into a master problem and multiple subproblems using the primal-decomposition method. To reduce the computational complexity, we show in Section IV-E that each subproblem is amenable to powerful network flow optimization methods. We then present a pseudocode of the proposed algorithm in Section IV-F and discuss

---

[4]For instance, consider the special case explained in [17], where all cross-channel gains are either 0 or ∞ and there is one UT per sector; in this case, the optimum transmission powers are binary, yet it is strongly NP-hard to determine for each UT whether to transmit with full power or not to transmit at all.

its complexity in Section IV-G. Finally, we highlight in Section IV-H the advantages of the distributed implementation of the proposed scheme as compared to a centralized implementation of the same scheme.

A. Bound Optimization

The SINR expression in (1) can be lower-bounded as

$$\Gamma_{m,n}^{(k)} \geq \tilde{\Gamma}_{m,n}^{(k)} = \frac{P_C H_{m,n}^{(k,k)}}{P_C \sum\limits_{k=1,\tilde{k}\neq k}^{K} H_{m,n}^{(k,\tilde{k})} - \max\limits_{\tilde{k}\in\tilde{\mathcal{K}}^{(k)}} I_n^{(\tilde{k})} P_C H_{m,n}^{(k,\tilde{k})} + P_N}, \quad (8)$$

where $\tilde{\mathcal{K}}^{(k)}$ is the set of indices of the neighboring sectors of sector k, which are considered in this work to be the first-tier interfering sectors seen by sector k; e.g., in Fig. 1, $\mathcal{K}^{(3)} = \{1,2,13,17,16,20\}$.[5] The cardinality of $\tilde{\mathcal{K}}^{(k)}$ is assumed to be the same for all sectors and it is denoted by $\tilde{K}$, i.e., $|\tilde{\mathcal{K}}^{(k)}| = \tilde{K}$, $\forall k$. The bound in (8) is obtained by assuming that RB n is used in all sectors, except at most one sector. When RB n is blanked in more than one interfering sector, only the most dominant blanked interferer is considered.[6] This bound is exact if the number of blanked interferers is less or equal to one and it is tight for small number of blanked interferers. The bound is also tight for dominant interference environment, where the received power from each interferer to each user is significantly greater than the aggregate received power from all other weaker interferers.

The bound in (8) is a good metric to be used for optimization (maximization) for the following reasons. First of all, if the bound is increased by Δ, then the exact expression given by (1) will also increase by at least Δ (proof is given in Appendix A). Moreover, it is already observed in the literature that blanking a particular RB in more than two sectors can degrade the overall system performance (cf. [3]). Finally, and most importantly, based on this bound, we can develop a distributed optimization framework that is applicable to a wide range of schedulers and AMC strategies for dominant interference environments. This framework can be implemented very efficiently, and can achieve near-optimal performance, as we will see later.

Using the SINR bound given in (8), we now construct a similar bound on the rates, $\{R_{m,n}^{(k)}\}$, given in (2). To do so, we define $\gamma_{m,n}^{(k)}$ and $r_{m,n}^{(k)}$ as the SINR and the corresponding achievable rate on RB n of UT m in sector k, if all sectors use RB n, i.e.,

$$\gamma_{m,n}^{(k)} = \frac{P_C H_{m,n}^{k,k}}{P_C \sum\limits_{\tilde{k}\neq k} H_{m,n}^{(k,\tilde{k})} + P_N}, r_{m,n}^{(k)} = f\left(\gamma_{m,n}^{(k)}\right). \quad (9)$$

[5]We assume a wraparound layout so each sector has six first-tier interfering sectors; e.g., $\mathcal{K}^{(25)} = \{26,27,23,51,50,48\}$.

[6]Note that this bound considers the most dominant *blanked* interferer which is not necessarily the most dominant interferer when the most dominant interferer does not blank RB n (i.e., when it is not turned off). This differentiates (8) from the expression considered in [14], where the most dominant interferer is considered.

Similarly, we define $\tilde{\gamma}_{m,n}^{(k,\tilde{k})}$ and $\tilde{r}_{m,n}^{(k,\tilde{k})}$ as the SINR and the additional rate on RB n of UT m in sector k, if only sector $\tilde{k} \in \tilde{\mathcal{K}}^{(k)}$ blanks RB n, i.e.,

$$\tilde{\gamma}_{m,n}^{(k,\tilde{k})} = \frac{P_C H_{m,n}^{(k,k)}}{P_C \sum\limits_{\hat{k}\neq k} H_{m,n}^{k,\hat{k}} - P_C H_{m,n}^{(k,\tilde{k})} + P_N},$$

$$\tilde{r}_{m,n}^{(k,\tilde{k})} = f\left(\tilde{\gamma}_{m,n}^{(k,\tilde{k})}\right) - r_{m,n}^{(k)}. \quad (10)$$

Using the definitions of $\gamma_{m,n}^{(k)}$ and $\tilde{\gamma}_{m,n}^{(k,\tilde{k})}$ given in (9) and (10), respectively, the SINR bound given by (8) can be also written as

$$\Gamma_{m,n}^{(k)} \geq \tilde{\Gamma}_{m,n}^{(k)} = \max\left(\gamma_{m,n}^{(k)}, \max\limits_{\tilde{k}} I_n^{(\tilde{k})} \tilde{\gamma}_{m,n}^{(k,\tilde{k})}\right), \quad (11)$$

where the equivalence between (8) and (11) stems from the fact that $\{I_n^{(\tilde{k})}\}$ assume binary values.

Using (9) and (10), a bound on the rates, $\{R_{m,n}^{(k)}\}$, given by (2) is constructed based on the SINR bound given by (11) as

$$R_{m,n}^{(k)} = f\left(\Gamma_{m,n}^{(k)}\right) \geq f\left(\tilde{\Gamma}_{m,n}^{(k)}\right)$$

$$= f\left(\max\left(\gamma_{m,n}^{(k)}, \max_{\tilde{k}} I_n^{(\tilde{k})} \tilde{\gamma}_{m,n}^{(k,\tilde{k})}\right)\right)$$

$$= \max\left(r_{m,n}^{(k)}, \max_{\tilde{k}} I_n^{(\tilde{k})}\left(r_{m,n}^{(k)} + \tilde{r}_{m,n}^{(k,\tilde{k})}\right)\right)$$

$$= r_{m,n}^{(k)} + \max_{\tilde{k}\in\tilde{\mathcal{K}}^{(k)}} I_n^{(\tilde{k})} \tilde{r}_{m,n}^{(k,\tilde{k})}, \quad (12)$$

where the first equality follows from (2), the first inequality and the second equality follow from (11) and the fact that $f(\cdot)$ is assumed to be nondecreasing, the third equality follows from (9), (10), and the fact that $f(\cdot)$ is assumed to be nondecreasing with $f(0) = 0$, and the fourth equality follows from the fact that $r_{m,n}^{(k)}$ is not a function of $\tilde{k}$.

The physical meaning of the bound in (12) is similar to the SINR bound in (8)—the rate bound is obtained by assuming that RB n is used in all sectors, except at most one sector. When RB n is blanked in more than one interfering sector, only the most dominant blanked interferer is considered.

By substituting (12) in (7), the optimization problem in (7) can be tightly approximated for dominant interference environment as

$$\underset{\substack{\{x_{m,n}^{(k)}, I_n^{(k)}: \\ k\in\mathcal{K}, m\in\mathcal{M}^{(k)}\}}}{\text{maximize}} \quad \sum_{k=1}^{K}\sum_{m=1}^{M^{(k)}} w_m^{(k)} x_{m,n}^{(k)}\left(r_{m,n}^{(k)} + \max_{\tilde{k}\in\tilde{\mathcal{K}}^{(k)}} I_n^{(\tilde{k})} \tilde{r}_{m,n}^{(k,\tilde{k})}\right) \quad (13a)$$

$$\text{subject to} \quad \sum_{m=1}^{M^{(k)}} x_{m,n}^{(k)} = 1 - I_n^{(k)}, \quad \forall k, \quad (13b)$$

$$x_{m,n}^{(k)}, I_n^{(k)} \in \{0,1\}, \quad \forall k,m. \quad (13c)$$

The optimization problem (13) is a non-linear binary integer optimization problem, which is in general, difficult to solve. As an intermediate step to reduce the complexity of solving

(13), we convert it to an equivalent binary linear optimization problem in the following section.

B. Transforming (13) Into an Equivalent Binary Linear Optimization Problem

Our approach to tackle the binary non-linear optimization problem given by (13) is to first transform it into an equivalent binary linear optimization problem. While binary linear optimization problems are still not easy to solve in general, good approximate solutions can be computed efficiently by relaxing the constraints that restrict the variables to be either 0 or 1, into weaker constraints that restrict the variables to assume any real value in the interval [0,1]. The challenge is to construct an equivalent binary linear optimization problem that has a tight relaxation which means that the solution obtained by solving the relaxed version is very close to the one obtained by solving the binary linear optimization problem. This challenge is addressed in this section.

To convert the binary non-linear optimization problem given by (13) into an equivalent binary linear optimization problem, we need to convert the term $x_{m,n}^{(k)} \max_{\tilde{k} \in \tilde{\mathcal{K}}^{(k)}} I_n^{(\tilde{k})} \tilde{r}_{m,n}^{(k,\tilde{k})}$ into a linear term. There are two sources of non-linearity in this term: the point-wise maximum and the multiplication. The non-linear term can be written as

$$x_{m,n}^{(k)} \max_{\tilde{k} \in \tilde{\mathcal{K}}^{(k)}} I_n^{(\tilde{k})} \tilde{r}_{m,n}^{(k,\tilde{k})} = \max_{y_{m,n}^{(k,\tilde{k})} \in \mathcal{C}} \sum_{\tilde{k} \in \tilde{\mathcal{K}}^{(k)}} y_{m,n}^{(k,\tilde{k})} \tilde{r}_{m,n}^{(k,\tilde{k})}, \quad (14)$$

where the variables $\{y_{m,n}^{(k,\tilde{k})} : k \in \mathcal{K}, \tilde{k} \in \mathcal{K}^{(k)}, m \in \mathcal{M}^{(k)}, n \in \mathcal{N}\}$ are introduced as auxiliary variables that facilitate the conversion of the non-linear term into a linear term and

$$\mathcal{C} = \left\{ y_{m,n}^{(k,\tilde{k})} : \sum_{\tilde{k} \in \mathcal{K}^{(k)}} y_{m,n}^{(k,\tilde{k})} \leq 1, \quad \forall k, m, \right.$$

$$y_{m,n}^{(k,\tilde{k})} \leq x_{m,n}^{(k)}, \qquad \forall k, \tilde{k}, m,$$

$$y_{m,n}^{(k,\tilde{k})} \leq I_n^{(\tilde{k})}, \qquad \forall k, \tilde{k}, m,$$

$$\left. y_{m,n}^{(k,\tilde{k})} \in \{0,1\}, \qquad \forall k, \tilde{k}, m \right\}. \quad (15)$$

To see the equivalence, we note that since all variables assume binary values, then the point-wise maximum in the originally non-linear term is captured in (14) and the first inequality in (15). Moreover, the multiplication in the originally non-linear term is captured in the second and the third inequalities in (15).

By replacing the term $x_{m,n}^{(k)} \max_{\tilde{k} \in \tilde{\mathcal{K}}^{(k)}} I_n^{(\tilde{k})} \tilde{r}_{m,n}^{(k,\tilde{k})}$ in (13a) with (14) and (15), we get an equivalent binary linear optimization problem. Unfortunately, we found experimentally that such an equivalent optimization problem leads to loose linear relaxation, i.e., it results in solutions that are far from the binary optimal solutions. As a result, we seek tighter equivalent formulations.

A general approach to get tighter relaxation is by adding additional constraints that are called valid constraints [32, p. 585]. These valid constraints do not change the set of

feasible binary solutions; however, with proper choice of these valid constraints, one can obtain tighter relaxations. Based on this approach, we construct a set $\mathcal{C}'$ that is equivalent to $\mathcal{C}$ for the considered binary optimization problem, by adding two valid constraints:

$$\sum_{\tilde{k} \in \tilde{\mathcal{K}}^{(k)}} y_{m,n}^{(k,\tilde{k})} \leq x_{m,n}^{(k)}, \quad \forall k, m,$$

$$\sum_{m=1}^{M^{(\tilde{k})}} y_{m,n}^{(k,\tilde{k})} \leq I_n^{(\tilde{k})}, \quad \forall k, \tilde{k}. \quad (16)$$

Hence, $\mathcal{C}'$ is given by

$$\mathcal{C}' = \left\{ y_{m,n}^{(k,\tilde{k})} : \sum_{\tilde{k} \in \tilde{\mathcal{K}}^{(k)}} y_{m,n}^{(k,\tilde{k})} \leq x_{m,n}^{(k)}, \forall k, m, \right.$$

$$\sum_{m=1}^{M^{(\tilde{k})}} y_{m,n}^{(k,\tilde{k})} \leq I_n^{(\tilde{k})}, \quad \forall k, \tilde{k},$$

$$\left. y_{m,n}^{(k,\tilde{k})} \in \{0,1\}, \qquad \forall k, \tilde{k}, m \right\}. \quad (17)$$

Note that the constraints $y_{m,n}^{(k,\tilde{k})} \leq x_{m,n}^{(k)}$ and $y_{m,n}^{(k,\tilde{k})} \leq I_n^{(\tilde{k})}$, given originally in (15), are omitted from the definition of $\mathcal{C}'$ given in (17) since these constraints define a subset of the first two constraints given in (17). The equivalence of $\mathcal{C}$ and $\mathcal{C}'$ is provided in the following Lemma.

Lemma 1: The sets $\mathcal{C}$ and $\mathcal{C}'$ given by (15) and (17), respectively, are equivalent.

Proof: See Appendix B. ∎

By replacing the term $x_{m,n}^{(k)} \max_{\tilde{k}} I_n^{(\tilde{k})} \tilde{r}_{m,n}^{(k,\tilde{k})}$ in (13) with (14) and (17), we get the following binary linear optimization problem[7]:

$$\underset{\substack{\{x_{m,n}^{(k)}, y_{m,n}^{(k,\tilde{k})}, I_n^{(k)}: \\ k \in \mathcal{K}, \tilde{k} \in \mathcal{K}^{(k)}, m \in \mathcal{M}^{(k)}\}}}{\text{maximize}} \sum_{k=1}^{K} \sum_{m=1}^{M^{(k)}} w_m^{(k)} \left(x_{m,n}^{(k)} r_{m,n}^{(k)} + \sum_{\tilde{k} \in \mathcal{K}^{(k)}} y_{m,n}^{(k,\tilde{k})} \tilde{r}_{m,n}^{(k,\tilde{k})} \right)$$

$$(18a)$$

$$\text{subject to} \quad \sum_{m=1}^{M^{(k)}} x_{m,n}^{(k)} = 1 - I_n^{(k)}, \qquad \forall k, \quad (18b)$$

$$\sum_{\tilde{k} \in \mathcal{K}^{(k)}} y_{m,n}^{(k,\tilde{k})} \leq x_{m,n}^{(k)}, \qquad \forall k, m, \quad (18c)$$

$$\sum_{m=1}^{M^{(k)}} y_{m,n}^{(k,\tilde{k})} \leq I_n^{(\tilde{k})}, \qquad \forall k, \tilde{k}, \quad (18d)$$

$$x_{m,n}^{(k)}, y_{m,n}^{(k,\tilde{k})}, I_n^{(k)} \in \{0,1\}, \quad \forall k, \tilde{k}, m. \quad (18e)$$

We finally remark that many other equivalent binary optimization problems can be formulated, e.g., by replacing the term $x_{m,n}^{(k)} \max_{\tilde{k}} I_n^{(\tilde{k})} \tilde{r}_{m,n}^{(k,\tilde{k})}$ in (13) with (14) and (15). However, different equivalents will have different relaxations. Unlike many other equivalent formulations, the relaxed version of the formulation in (18) has the following advantages:

- The optimal solution is provably close to binary as we will show in Section IV-C.

---

[7]Note that the max operator in the right-hand side of (14) is omitted from the objective of the optimization problem in (18), since the optimization problem is already a maximization problem.

- It can be readily solved in distributed manner using primal decomposition as we will show in Section IV-D.
- It can be solved efficiently using network flow optimization tools as we will show in Section IV-E.

C. Linear Optimization Relaxation

An upper bound on the optimum value of (18) can be obtained by solving the relaxed version of (18) which can be constructed by replacing (18e) with the following constraints

$$x_{m,n}^{(k)}, y_{m,n}^{(k,\tilde{k})}, I_n^{(k)} \in [0,1], \quad \forall k \in \mathcal{K}, \tilde{k} \in \tilde{\mathcal{K}}^{(k)}, m \in \mathcal{M}^{(k)}. \quad (19)$$

In particular, let $p_{\text{Binary}}^{\star}$ denote the optimal value of (18), let $p_{\text{Relaxed}}^{\star}$ denote the optimal value of the relaxed version of (18), and let $\hat{p}_{\text{Relaxed}}^{\star}$ denote the value of the objective function evaluated at a rounded-solution of the relaxed problem, such that the rounded solution is a feasible binary solution of (18). Then, we have the following inequalities

$$p_{\text{Relaxed}}^{\star} \geq p_{\text{Binary}}^{\star} \geq \hat{p}_{\text{Relaxed}}^{\star}, \quad (20)$$

where the first inequality follows from the fact that the feasible set of the relaxed version is always a superset of the original problem and the second equality follows directly from the optimality of $p_{\text{Binary}}^{\star}$. We define the optimality gap, Δ_{Opt}, in percentage as

$$\Delta_{\text{Opt}} = \left(p_{\text{Binary}}^{\star} - \hat{p}_{\text{Relaxed}}^{\star}\right) / p_{\text{Binary}}^{\star} \cdot 100\%$$
$$\leq \left(p_{\text{Relaxed}}^{\star} - \hat{p}_{\text{Relaxed}}^{\star}\right) / p_{\text{Relaxed}}^{\star} \cdot 100\%, \quad (21)$$

where the inequality follows from (20). Thus, one can compute an estimate on the optimality gap in polynomial time by solving a linear optimization problem. As we will show in Section IV-A through extensive simulations, by solving the relaxed problem and rounding the solution to the closest binary feasible solution, one can obtain a solution to (18) that is near-optimal, i.e., with small Δ_{Opt}.

An important objective measure of the tightness of a relaxation is the percentage of optimal relaxed variables that assume binary values. For instance, the optimality gap goes to zero as this percentage goes to 100%. We now provide a closed-form theoretical guarantee on the this percentage. This theoretical guarantee is summarized in the following proposition.

Proposition 1: The percentage of optimal variables of the relaxed version of (18) that assume binary values is greater than or equal to $\frac{\tilde{K}(\bar{M}-1)}{(\tilde{K}+1)\bar{M}+1} \cdot 100\%$, where $\bar{M}$ is the average number of UTs per sector and $\tilde{K}$ is the number of neighboring interferers.

Proof: See Appendix C. ∎

For example, if $M^{(k)} = 20$, $\forall k$, and $\tilde{K} = 6$, then using Proposition 1 we can deduce that the percentage of optimal variables of the relaxed version of (18) that assume binary values is guaranteed to be greater than or equal to 80.8%.

However, solving the relaxed problem would require a central controller to be connected to all the BSs to solve a large linear optimization problem. Such a central controller is not supported in most contemporary cellular network standards, such as LTE, LTE-Advanced and IEEE 802.16m. Consequently,

we seek in the following section a distributed optimization method to solve the relaxed version of problem (18).

D. Primal Decomposition

The relaxed version of (18) has a special separable structure. In particular, for any set of fixed $\{I_n^{(k)}, \forall k \in \mathcal{K}\}$, it can be separated into K optimization problems, each can be solved separately in each sector. In this section, we show how this structure is exploited to develop a distributed algorithm based on the primal-decomposition method[8] [33, pp. 3–5].

To exploit the separable structure, let $\phi^{(k)}(I_n^{(1)}, \ldots, I_n^{(K)})$ denote the optimal value of the following optimization problem for given $\{I_n^{(1)}, \ldots, I_n^{(K)}\}$:

$$\underset{\substack{\{x_{m,n}^{(k)}, y_{m,n}^{(k,\tilde{k})}: \\ m \in \mathcal{M}^{(k)}, \tilde{k} \in \tilde{\mathcal{K}}^{(k)}\}}}{\text{maximize}} \sum_{m=1}^{M^{(k)}} w_m^{(k)} \left(x_{m,n}^{(k)} r_{m,n}^{(k)} + \sum_{\tilde{k} \in \tilde{\mathcal{K}}^{(k)}} y_{m,n}^{(k,\tilde{k})} \tilde{r}_{m,n}^{(k,\tilde{k})} \right) \quad (22a)$$

$$\text{subject to} \quad \sum_{m=1}^{M^{(k)}} x_{m,n}^{(k)} = 1 - I_n^{(k)}, \quad (22b)$$

$$\sum_{\tilde{k} \in \tilde{\mathcal{K}}^{(k)}} y_{m,n}^{(k,\tilde{k})} \leq x_{m,n}^{(k)}, \quad \forall m, \quad (22c)$$

$$\sum_{m=1}^{M^{(k)}} y_{m,n}^{(k,\tilde{k})} \leq I_n^{(\tilde{k})}, \quad \forall \tilde{k}, \quad (22d)$$

$$x_{m,n}^{(k)}, y_{m,n}^{(k,\tilde{k})}, \in [0,1], \quad \forall m, \tilde{k}. \quad (22e)$$

For reasons that will become apparent, we call (22) subproblem k. Using (22), the relaxed version of (18) is equivalent to

$$\underset{I_n^{(k)}, k \in \mathcal{K}}{\text{maximize}} \quad \sum_{k=1}^{K} \phi^{(k)}\left(I_n^{(1)}, \ldots, I_n^{(K)}\right) \quad (23a)$$

$$\text{subject to} \quad I_n^{(k)} \in [0,1], \quad \forall k \in \mathcal{K}. \quad (23b)$$

We call (23) the master problem. Therefore, the relaxed version of (18) has been decomposed into a master problem, given by (23), and K subproblems, each is given by (22) and can be solved separately in each sector.

The master problem given in (23) is a convex optimization problem with a concave objective function of the variables $\{I_n^{(1)}, \ldots, I_n^{(K)}\}$ and a convex constraint set, since the relaxed version of (18) is a linear (thus convex) optimization problem (cf. [33, p. 2]). Since the objective is not necessarily differentiable, the projected-subgradient method can be used to solve this problem iteratively [34, p. 16]. In each iteration, K subproblems are solved in each sector to evaluate $\phi^{(k)}(I_n^{(1)}, \ldots, I_n^{(K)})$, $\forall k \in \mathcal{K}$, and subgradients $[\Lambda_n^{\star(1)}, \ldots, \Lambda_n^{\star(K)}] \in \partial \sum_{k=1}^{K} \phi^{(k)}(I_n^{(1)}, \ldots, I_n^{(K)})$, where $\partial f(x)$ is the subdifferential of $f(\cdot)$ evaluated at x. To explain how the subgradients $[\Lambda_n^{\star(1)}, \ldots, \Lambda_n^{\star(K)}]$ are obtained, let $g_n^{(k)}(f(x))$ denote the kth

---

[8]Primal-decomposition is considered in this paper instead of dual-decomposition since the structure is readily separable in the primal domain. Moreover, since the objective function is not strictly concave, recovering the optimal primal variables from optimal dual variables can be challenging for dual-decomposition method [33, p. 4].

element of a subgradient of the function $f(\cdot)$ evaluated at x for RB n. Using the definition of a subgradient (cf. [35, p. 1]), it is not difficult to see that $g_n^{(k)}$ is closed under addition, and thus,

$$
\begin{aligned}
\Lambda_n^{\star(k)} &= g_n^{(k)}\left(\sum_{k=1}^K \phi^{(k)}\left(I_n^{(1)},\ldots,I_n^{(K)}\right)\right)\\
&= \sum_{k=1}^K g_n^{(k)}\left(\phi^{(k)}\left(I_n^{(1)},\ldots,I_n^{(K)}\right)\right).
\end{aligned}
\tag{24}
$$

As shown in [33, p. 5], $g_n^{(k)}(\phi^{(k)}(I_n^{(1)},\ldots,I_n^{(K)}))$ can be obtained from an optimum Lagrange multiplier (dual variable) that corresponds to the constraint where $I_n^{(K)}$ appears in its right-hand-side. Consequently, $\Lambda_n^{\star(k)}$ can be explicitly written as

$$
\Lambda_n^{\star(k)} := -\lambda_n^{\star(k)} + \sum_{\tilde{k}\in\mathcal{K}^{(k)}} \lambda_n^{\star(\tilde{k},k)},
\tag{25}
$$

where $\lambda_n^{\star(k)}$ is an optimum Lagrange multiplier corresponding to constraint (22b) and $\lambda_n^{\star(k,\tilde{k})}$, $\tilde{k}\in\mathcal{K}^{(k)}$, are optimum Lagrange multipliers corresponding to constraints (22d).

For each sector k to calculate $\Lambda_n^{\star(k)}$, it requires the knowledge of $\lambda_n^{\star(k)}$, which can be obtained locally by solving (22), and $\lambda_n^{\star(\tilde{k},k)}$, $\tilde{k}\in\mathcal{K}^{(k)}$, which can be exchanged from the neighboring sectors. In other words, each sector k sends $\lambda_n^{(k,\tilde{k})}$ for all $\tilde{k}$ sectors that are in the neighborhood of sector k, for all n. The master algorithm then updates its variables as

$$
I_n^{(k)} := I_n^{(k)} + \delta\Lambda_n^{(k)}, \qquad \forall k\in\mathcal{K},
\tag{26}
$$

where δ is the step-size which can be chosen using any of the standard methods given in [34, pp. 3–4]. Then, each $I_n^{(k)}$ is projected into the feasible set of $[0,1]$ as follows

$$
I_n^{(k)} := \begin{cases} 0, & I_n^{(k)} \le 0,\\ I_n^{(k)}, & 0 < I_n^{(k)} < 1,\\ 1, & I_n^{(k)} \ge 1. \end{cases}
\tag{27}
$$

Using (25)–(27), each sector k can compute the variables $\{I_n^{(k)},\ \forall n\in\mathcal{N}\}$ in a distributed manner without the need for a centralized entity. Then, each sector k exchanges $I_n^{(k)}$ with its neighbors and the process is repeated for N_{iter} iterations. After that, each $I_n^{(k)}$ is rounded to the nearest binary value which is denoted by $I_n^{\star(k)}$. Once $\{I_n^{\star(k)}\}$ are determined, local scheduling decision variables $\{x_{m,n}^{\star(k)}\}$ can be determined in each sector k separately as follows. For every $m\in\mathcal{M}^{(k)}$, $n\in\mathcal{N}$, $R_{m,n}^{(k)}$ is calculated using (1), (2), and $I_n^{\star(k)}$. To ensure feasibility of the resulting solution to problem (18), the scheduling decision variables are calculated as

$$
x_{m,n}^{\star(k)} = \begin{cases} 1, & m = \arg\max\limits_m w_m^{(k)} R_{m,n}^{(k)} \text{ and } I_n^{\star(k)} = 0,\\ 0, & m \neq \arg\max\limits_m w_m^{(k)} R_{m,n}^{(k)} \text{ or } I_n^{\star(k)} = 1. \end{cases}
\tag{28}
$$

Since the master problem is a convex optimization problem, the subgradient algorithm is guaranteed to converge to the optimum solution of the relaxed version of problem (18) as $N_{\text{iter}} \to \infty$ if δ is chosen properly [34, p. 6]. In this paper, we

choose δ to be square summable but not summable by setting $\delta = c/p$, where $c > 0$ is a constant and p is the iteration index. This choice of δ guarantees convergence to the optimal solution as $N_{\text{iter}} \to \infty$ [34, p. 6]. However, in practice the algorithm must terminate after a finite number of iterations which raises the following question: How different is the value obtained using finite iterations as compared to the true optimum obtained by solving (18)? We address this question in Section VI and show that few iterations are sufficient to achieve near-optimality.

The choice of the step-size, δ, also affects the rate of message exchange (overhead) indirectly, at it affects N_{iter} that is needed to achieve a particular optimality gap. Since finding an explicit relationship between N_{iter} and the step-size δ for subgradient algorithms is still an open research problem, tuning the step-size δ is performed empirically by choosing the constant c to reduce N_{iter} and thus reduce the rate of message exchange. Through extensive simulations, we found that a good choice of the parameter c is mainly dependant on the choice of the desired utility. Thus, for a given utility, the parameter c can be tuned off-line only once before executing the algorithm. The relationship between N_{iter} and the rate of message exchange required for the proposed algorithm will be discussed in Section IV-G.

Clearly, the proposed algorithm relies heavily on solving the subproblem given by (22). Hence, it is imperative to solve (22) as efficiently as possible. Interestingly, the subproblem given by (22) has a special network flow structure which can be exploited to devise efficient algorithms to solve it, as explained in the following section.

E. Transforming (22) Into an Equivalent MCNF Problem

The optimization problem (22) is a linear optimization problem which can be solved using generic simplex or interior-point methods. Nevertheless, we show in this section that (22) has a special network structure which makes it amenable to powerful network flow optimization methods that surpass conventional simplex and interior-point methods. In particular, we show that (22) can be converted into an equivalent MCNF optimization problem.

An MCNF optimization problem is defined as finding a least cost way of sending certain amount of flow over a network that is specified by a directed graph of v vertices and e edges. Such a problem has e variables, which represent the amount of flow on each arc, and v linear equality constraints, which represent the mass-balance in each vertex, such that every variable appears in exactly two constraints: one with a coefficient of $+1$ and one with a coefficient of -1 [18, p. 5]. In addition to the mass-balance constraints, constraints on the lower and upper bounds on the amount of flow on each arch are also specified. The objective function is a weighted sum of the flows in each arc, where the weight is the cost per unit flow on that arc. Thanks to the network structure of these problems, efficient combinatorial algorithms exist to solve such problems in strongly polynomial time, much faster than generic linear optimization solvers [18, Ch. 10]. For example, the enhanced capacity scaling algorithm can solve an MCNF problem with v vertices and e edges in $O(e\log v(e+v\log v))$ [18, p. 395].

Although the problem in its original form given in (22) is not an MCNF optimization problem, it can be transformed into an equivalent MCNF optimization problem as follows. If we multiply both sides of constraint (22d) with -1, we obtain a linear optimization problem with the following properties. Each variable appears in at most one constraint with a coefficient of $+1$ and at most one constraint with a coefficient of -1. According to Theorem 9.9 in [18, p. 315], a linear optimization problem with such a structure can be transformed into an equivalent MCNF optimization problem. To perform such transformation, we introduce the slack variables $\{s_m \geq 0, \forall m \in \mathcal{M}^{(k)}\}$ and surplus variables $\{\tilde{s}^{(\tilde{k})} \geq 0, \forall \tilde{k} \in \tilde{\mathcal{K}}^{(k)}\}$ to convert the inequality constraints (22c) and (22d), respectively, into equality constraints. To obtain the mass-balance constraints, we also introduce a redundant constraint by summing constraints (29b)–(29d). In addition, we convert the maximization into minimization by negating the objective function. Incorporating these transformations into (22), we get the following MCNF optimization problem:

$$\underset{\substack{\{x_{m,n}^{(k)}, y_{m,n}^{(k,\tilde{k})}, s_m, \tilde{s}^{(\tilde{k})}: \\ m \in \mathcal{M}^{(k)}, \tilde{k} \in \tilde{\mathcal{K}}^{(k)}\}}}{\text{minimize}} \quad - \sum_{m=1}^{M^{(k)}} w_m^{(k)} \left(x_{m,n}^{(k)} r_{m,n}^{(k)} + \sum_{\tilde{k} \in \mathcal{K}^{(k)}} y_{m,n}^{(k,\tilde{k})} \tilde{r}_{m,n}^{(k,\tilde{k})} \right) \quad (29a)$$

$$\text{subject to} \quad \sum_{m=1}^{M^{(k)}} x_{m,n}^{(k)} = 1 - I_n^{(k)}, \quad (29b)$$

$$\sum_{\tilde{k} \in \mathcal{K}^{(k)}} y_{m,n}^{(k,\tilde{k})} - x_{m,n}^{(k)} + s_m = 0, \quad \forall m, \quad (29c)$$

$$-\sum_{m=1}^{M^{(k)}} y_{m,n}^{(k,\tilde{k})} - \tilde{s}^{(\tilde{k})} = -I_n^{(\tilde{k})}, \quad \forall \tilde{k}, \quad (29d)$$

$$-\sum_{m=1}^{M^{(k)}} s_m + \sum_{\tilde{k} \in \mathcal{K}^{(k)}} \tilde{s}^{(\tilde{k})} = -1 + I_n^{(k)} + \sum_{\tilde{k} \in \mathcal{K}^{(k)}} I_n^{(\tilde{k})}, \quad (29e)$$

$$x_{m,n}^{(k)}, y_{m,n}^{(k,\tilde{k})}, s_m, \tilde{s}^{(\tilde{k})} \in [0,1], \quad \forall m, \tilde{k}. \quad (29f)$$

To see that (29) is indeed an MCNF optimization problem, we note that the objective function is linear and the equality constraints (29b)–(29e) represent mass-balance constraints because each variable appears in two constraints: one with a coefficient of $+1$ and one with a coefficient of -1 [18, p. 5].

F. Pseudocode

A pseudocode of the proposed algorithm to be executed in every sector k is given below.

Algorithm 1 Proposed ICIC algorithm to be executed in every sector k

---

Input: $H_{m,n}^{(k,\tilde{k})}$ and $w_m^{(k)}$, $\forall \tilde{k} \in \tilde{\mathcal{K}}^{(k)}$, $m \in \mathcal{M}^{(k)}$, $n \in \mathcal{N}$
Output: $x_{m,n}^{\star(k)}$, $I_n^{\star(k)}$, $\forall m \in \mathcal{M}^{(k)}$, $n \in \mathcal{N}$

1: Initialize $I_n^{(k)}$, $\forall n \in \mathcal{N}$
2: **Preprocessing**: Obtain $r_{m,n}^{(k)}$ and $\tilde{r}_{m,n}^{(k,\tilde{k})}$ using (9) and (10), $\forall \tilde{k} \in \tilde{\mathcal{K}}^{(k)}$, $m \in \mathcal{M}^{(k)}$, $n \in \mathcal{N}$.

3: **for** $p = 1$ to N_{iter} **do**
4:    **Solve a subproblem**: Obtain $x_{m,n}^{\star(k)}$, $y_{m,n}^{\star(k,\tilde{k})}$, $\lambda_n^{\star(k)}$, and $\lambda_n^{\star(k,\tilde{k})}$, $\forall \tilde{k} \in \tilde{\mathcal{K}}^{(k)}$, $n \in \mathcal{N}$, by solving the MCNF optimization problem (29), $\forall n \in \mathcal{N}$.
5:    **Exchange subgradients**: Send $\lambda_n^{\star(k,\tilde{k})}$ to sectors $\tilde{k} \in \tilde{\mathcal{K}}^{(k)}$, $\forall n \in \mathcal{N}$.
6:    **Update** $I_n^{(k)}$, $\forall n \in \mathcal{N}$:
7:      $\Lambda_n^{\star(k)} := -\lambda_n^{\star(k)} + \sum_{\tilde{k} \in \tilde{\mathcal{K}}^{(k)}} \lambda_n^{\star(\tilde{k},k)}$.
8:      **Subgradient step**: $I_n^{(k)} := I_n^{(k)} + \delta \Lambda_n^{\star(k)}$
9:      **Projection**: Project $I_n^{(k)}$ into the feasible set (27).
10:   **Exchange** $I_n^{(k)}$: Send $I_n^{(k)}$ to sectors $\tilde{k} \in \tilde{\mathcal{K}}^{(k)}$, $\forall n \in \mathcal{N}$.
11: **end for**
12: **Round the solution**: $I_n^{\star(k)} := \lfloor I_n^{(k)} + 0.5 \rfloor$, $\forall n \in \mathcal{N}$.
13: **Local scheduling decisions:**
14:   Obtain $R_{m,n}^{(k)}$ by substituting $H_{m,n}^{(k,\tilde{k})}$ and $I_n^{\star(k)}$ in (1) and (2), $\forall m \in \mathcal{M}^{(k)}$, $n \in \mathcal{N}$.
15:   Obtain $x_{m,n}^{\star(k)}$ by substituting $w_m^{(k)}$ and $R_{m,n}^{(k)}$ in (28), $\forall m \in \mathcal{M}^{(k)}$, $n \in \mathcal{N}$.

The proposed algorithm, whose pseudocode is given in Algorithm 1, has the following intuitive interpretation. Each sector k estimates the benefit of shutting its own power on RB n to neighboring sectors by $\sum_{\tilde{k} \in \tilde{\mathcal{K}}^{(k)}} \lambda_n^{\star(\tilde{k},k)}$ and the benefit of using RB n in sector k by $\lambda_n^{\star(k)}$; if the former is more (less) than the latter, then sector k increases (decreases) its soft decision on $I_n^{(k)}$ proportionally to the difference of these benefits, and possibly set it to 1 (0).

G. Complexity Analysis of the Proposed Algorithm

In the following, we show that the worst-case time complexity of the proposed algorithm implemented in each sector k is polynomial. We start by analyzing the complexity of solving the MCNF problem given by (29). A graph that is equivalent to (29) is represented by $2 + \tilde{K} + M^{(k)}$ vertices, which is equal to the number of mass-balance constrains, and $(\tilde{K} + 2)M^{(k)} + \tilde{K}$ arcs, which is equal to the number of variables. Thus, for fixed $\tilde{K}$, (29) can be solved using the enhanced capacity scaling algorithm in $O((M^{(k)} \log M^{(k)})^2)$ [18, p. 395]. Since (29) is solved for each RB in N_{iter} iterations, then the complexity of the proposed algorithm is $O(N_{\text{iter}} N (M^{(k)} \log M^{(k)})^2)$. Since N_{iter} is a fixed constant, the complexity of the algorithm in sector k as a function of $M^{(k)}$ and N is $O(N(M^{(k)} \log M^{(k)})^2)$.

To reduce the information exchange between sectors, low complexity variants of the proposed algorithm can be implemented. One such implementation is to execute the algorithm every ρ sub-frames. Each sector has time to implement the algorithm and calculate the $\{I_n^{(k)}\}$ values during ρ sub-frames and send it to other BSs. The design parameter ρ can be adjusted to suit the practical limitations.

We now quantify the rate of message exchange required between sector k and the neighboring $\tilde{K}$ sectors, to execute the proposed algorithm. For each iteration required by the proposed

algorithm, each sector sends to each neighboring sector N subgradients values and N I values. Thus, assuming L_q bits is used to quantize the subgradients and I values, then it is not difficult to see that the rate of message exchange in bit/sec is given by

$$R_{\text{distributed exchange}} = \frac{2N_{\text{iter}}\tilde{K}NL_q}{\rho \times 1 \quad \text{ms}} \text{bit/sec} \qquad (30)$$

where the proposed algorithm is executed every ρ sub-frames, each spans 1 ms.

We remark that the number of iterations required for the algorithm to converge is a constant parameter, N_{iter}, that can be configured to obtain the desired tradeoff between complexity and the optimality gap, where increasing N_{iter} can potentially reduces the optimality gap at the expense of higher computational complexity and signalling overhead, as we will demonstrate in Section IV-A.

H. Advantages of Distributed Implementation

We conclude this section by highlighting three main advantages of the proposed distributed scheme as compared to centralized implementation of the same scheme:

I. It does not require a centralized entity to collect the channel information from all UTs in the network and make network-wide scheduling decisions. This particular advantage makes the proposed algorithm applicable to current and emerging standards such as LTE and LTE-advanced, where a centralized entity is not supported.

II. The proposed algorithm can potentially reduce the message exchange as compared to a centralized implementation. To elaborate, we start by quantifying the rate of message exchange required for centralized implementation between each sector k and a centralized entity and then compare it with the rate of message exchange for the distributed scheme given in (30). Since each sector k is required to send to the centralized entity $N \times M^{(k)}(\tilde{K}+1)$ channel coefficients and assuming L_q bits is used to quantize the channel coefficients, then it is not difficult to see that the rate of message exchange for a centralized implementation in bit/sec is given by

$$R_{\text{centralized exchange}} = \frac{N \times M^{(k)}(\tilde{K}+1)L_q}{\rho \times 1 \quad \text{ms}} \text{bit/sec}. \qquad (31)$$

Using (30) and (31), the ratio of the rate of message exchange required by the centralized scheme to the rate of message exchange by the proposed scheme is given by

$$\frac{R_{\text{centralized exchange}}}{R_{\text{distributed exchange}}} = \frac{M^{(k)}(\tilde{K}+1)}{2(\tilde{K})N_{\text{iter}}}. \qquad (32)$$

Thus, as long as $N_{\text{iter}} < \frac{1}{2}M^{(k)}(1+\frac{1}{\tilde{K}})$, then the proposed distributed implementation requires lower rate of message exchange than the centralized implementation. For instance, if $M^{(k)} = 20$ UTs, $\tilde{K} = 6$ interferers, and $N_{\text{iter}} = 5$,

then, the proposed scheme requires 2.33 times lower rate of message exchange than the centralized implementation.

III. The amount of computations that is needed to be performed in a centralized entity will scale at least linearly with the number of sectors, K. Thus, scaling the network would require to scale also the computational capabilities of the centralized entity. However, this is not an issue for the proposed distributed implementation since the computations required are distributed among the K sectors and the computations needed by each sector is independent of the number of sectors in the network.

V. SIMULATION SETUP AND PARAMETERS

The simulation parameters are based on the IMT-Advanced Urban macro-cell (UMa) scenario [21] which are provided in Table II. Based on the IMT-Advanced guidelines, a hexagonal layout with wraparound is considered with 57 hexagonal sectors and 10 UTs per sector. These sectors are served by 19 BSs, each with a tri-sector antenna to serve a 3-sector cell-site. Monte Carlo simulations are carried over 1000 sub-frames and averaged over 10 independent drops. In each drop, the average received power from all sectors are calculated for each UT; this involves the calculation of the pathloss, correlated shadowing, and antenna gains. Variations of the received signal power due to small-scale fading within each RB is negligible and thus, the channels are assumed fixed over each RB. For other RBs, the channels assume different values depending on the time-frequency correlation of the IMT-Advanced model for the UMa scenario [36]. We use the AMC strategy given in Table III [37]. Channel-aware scheduling and ICIC are done afterwards, on a sub-frame by sub-frame basis. After scheduling the last sub-frame, the time-average throughput for each user is calculated. Then, another drop commences and the process repeats. Finally, time-averaged users throughput from all drops are saved for further processing and plotting. Further details of the simulation procedure are given in [21, Sec. 7]. The simulator is validated with the UMa simulation results provided in [38].

A. Importance of Accurate Simulation

Although explaining how to do accurate simulations is not the main purpose of this paper, we nevertheless highlight an aspect in the simulation that has important effect on assessing the performance of ICIC schemes. In particular, we demonstrate that the way UTs are associated to sectors has a significant impact on the simulation results. In the following, we explain three association strategies, namely, wideband SINR-based, wideband SINR-based (excluding shadowing), and geographical-based association. Wideband SINR is defined as the ratio of the average power received from the serving sector to the sum of the average power received from all other sectors and the noise power at the UT (i.e., small scale fading is not included in the calculation of the wideband SINR). In wideband SINR-based association, each UT is associated with the antenna sector to which it has the highest wideband SINR. This association strategy resembles reality, provides the most favorable results, and it is widely used by evaluation groups. A

TABLE II
SIMULATION PARAMETERS BASED ON IMT-ADVANCED UMa SCENARIO

Parameter	Assumption
Number of sectors	57 (wraparound)
Number of UTs per sector	10
Inter-site distance	500 m
BS height	25 m
Min. dist. between a UT and a BS	25 m
UT speed	30 km/h
Bandwidth (downlink)	10 MHz
Sub-carrier spacing	15 KHz
Number of RBs (N)	50
OFDM symbol duration	66.67 μs
Number of sub-carriers per RB	12
Number of OFDM symbols per RB	7
Number of drops	10
Number of sub-frames per drop	1000
Noise power per RB (P_N)	-114.45 dBm
Carrier frequency	2.0 GHz
Total BS transmit power	46 dBm
Path loss and shadowing	Based on UMa scenario [21]
Averaging window (t_c)	100
BS antenna gain (boresight)	17 dBi
User antenna gain	0 dBi
Feeder loss	2 dB
Channel estimation delay	4 sub-frames
SINR estimation margin	6 dB
BS antenna tilt (ϕ_t)	12° [38]
BS horizontal antenna pattern	$A(\theta) = -\min\left[12(\frac{\theta}{70°})^2, 20 \text{ dB}\right]$
BS elevation antenna pattern	$A_e(\phi) = -\min\left[12(\frac{\phi-\phi_t}{15°})^2, 20 \text{ dB}\right]$
BS combined antenna pattern	$-\min\left[-(A(\theta) + A_e(\phi)), 20 \text{ dB}\right]$ [21]
Small-scale fading model	IMT-Advanced channel model [36]
Traffic model	Full buffer

TABLE III
AMC TABLE WHICH IS USED TO DETERMINE THE ACHIEVABLE DATA
RATE ON EACH RB FOR EACH USER BASED ON THE SINR. THIS
TABLE IS BASED ON THE AMC STRATEGY GIVEN IN [37]

RB SINR range (dB)	Data rate per RB (kbit/sec)	RB SINR range (dB)	Data rate per RB (kbit/sec)
$(-\infty$ -6.1]	0	(5.8, 8.5]	291.6
(-6.1, -4.1]	35.3	(8.5, 9.9]	388.4
(-4.1, -2.0]	56.4	(9.5, 12.5]	418.3
(-2.0, -0.2]	92.4	(12.5, 14.8]	544.3
(-0.2, 1.9]	131.4	14.8. 16.1]	648.1
(1.9, 3.8]	177.4	(16.1, 17.8]	721.7
(3.8, 5.8]	223.1	(17.8, ∞)	807.4

consequence of this strategy is that the coverage region of each sector is not hexagonal and it changes from drop to drop due to the different shadowing realizations. Another consequence is that a UT may not be associated to the closest sector antenna as it may experience heavy shadowing to that sector antenna. A more convenient way of doing simulation is to exclude shadowing in the calculation of the wideband SINR which leads to fixed coverage regions for each sector in all drops. Due to the directional antenna patterns, the coverage regions of each sector is not hexagonal. In geographical-based association, a UT is associated to a particular sector antenna if it resides inside the hexagonal area of that sector. In this strategy, the coverage region of each sector is hexagonal.

In Fig. 2, we plot the CDF of the wideband SINR using the three association strategies for reuse-1. For validation purposes, we also include the average CDF results produced by seven WINNER+ partners using different simulation tools for UMa scenario [38]. It is clear from the figure that wideband SINR-based association strategy produces a CDF that agree very well
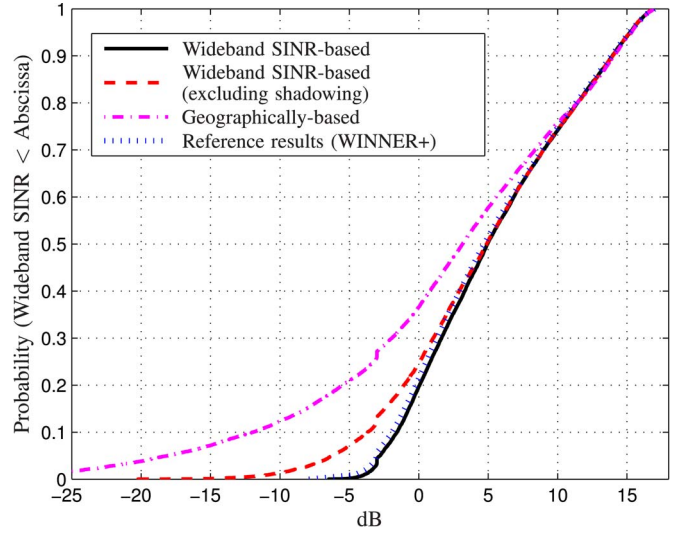

Fig. 2. CDF of the wideband SINR for different association strategies. Reference results (WINNER+) refers to the average CDF results produced by seven WINNER+ partners [38].

with the calibrated results; however, the other two association strategies produce a heavy tail which would impact the throughput of the users at the cell edge. The heavy tail is a direct consequence of associating UTs to sectors in a suboptimal manner. As a result, these two association strategies may not be suitable for assessing the performance of ICIC schemes as they tend to exaggerate the gains achieved for UTs at the cell-edge. This is the case because the cell-edge user throughput for reuse-1 for these two association strategies is very low. This makes any improvement in cell-edge user throughput to be large as compared to the very low values of reuse-1. Consequently, we use wideband SINR-based association strategy to assess the performance of ICIC schemes.

We conclude this section by highlighting the main motivations for performing SINR validation in system simulations:

1) System simulators are complex in nature which make them error-prone. As a result, careful verification of the simulation results is imperative.
2) Validation of the simulation makes it easy for other researchers to compare the performance of our schemes with other schemes in a widely-accepted scenario such as the UMa scenario specified by IMT-advanced.
3) As we observed in Fig. 2, the gains of ICIC schemes may be exaggerated if the simulation assumptions, such as the UT-BS association strategy, does not resemble reality. Validation of the SINR distribution is a good way to verify the validity of the simulation assumptions.

B. Baseline Schemes

In this paper, the followings are used as baseline schemes: reuse-1, reuse-3, PFR [5], dynamic FFR [9], and optimum FFR [39]. While reuse-1, reuse-3, and PFR are static schemes, dynamic and optimum FFR are implemented as dynamic schemes.

In PFR, RBs are divided such as 30 RBs (inner band) are used in all sectors (reuse-1) while 20 RBs (outer band) are shared

among sites in a classical reuse-3 pattern. More details about this algorithm can be found in [5].

While dynamic FFR and optimum FFR were designed to maximize the sum-rate, they can be easily modified to accommodate maximizing weighted sum-rates, and they work with differentiable and non-differentiable AMC functions, which make them good baseline schemes to be compared with the proposed scheme. For fair comparison, we convert the constraint on instantaneous minimum rate (used in dynamic and optimum FFR) to a constraint on average minimum rate, which can be implemented by updating the weights in every sub-frame according to the procedure given in [31]. We finally remark that both dynamic FFR and optimum FFR are centralized schemes.

In dynamic FFR, a centralized controller is used to determine which RBs belong to inner band (in reuse-1 pattern) and which RBs belong to outer bands (in reuse-3 pattern). The decision is made to maximize the total utility, assuming each RB is used by all the users in the region where this RB can be used, and summed for all the sectors. This assumption is critical to the development of the algorithm. Then, scheduling is done locally by each BS. More details about this algorithm can be found in [9].

In optimum FFR, the number of RBs used in inner and outer bands, and the reuse factor of the outer band are determined optimally using a centralized controller, without considering channel fading. Without channel fading, all RBs seen by a particular user have the same SINR. This makes the optimization problem tractable. However, if channel fading is considered, then the computational complexity for optimum FFR is exponential. More details about this algorithm can be found in [39].

It is common for FFR schemes (e.g., PFR and optimum FFR) to divide users into two classes: inner (cell-center) users and outer (cell-edge) users, based on SINR or distance from BS. Inner users are restricted to use inner band while outer users are restricted to use outer band. We found such restriction degrades the performance of FFR and as such, this restriction is removed to realize the full potential of FFR schemes. This restriction is already removed from dynamic FFR [9] for the same reason.

VI. SIMULATION AND ANALYTICAL RESULTS

A. Optimality Gaps

To develop an efficient algorithm to solve the difficult optimization problem given by (13) in a distributed manner, two sources of sub-optimality were introduced, namely, relaxing the integer constraints and solving the master optimization problem in finite iterations. To understand the effect of these sources of sub-optimality, we present in Fig. 3 the mean optimality gap (21), the 5th percentile, and 95th percentile of the distribution of the optimality gap obtained from simulating 22 000 instances of optimization problems for different numbers of UTs and four IMT-Advanced scenarios, namely, UMa, Urban micro-cell (UMi), Rural macro-cell (RMa), and Suburban macro-cell (SMa) scenarios [1]. In the first sub-frame, a random initial point is used. In the subsequent sub-frames, the optimal solution of the previous frame is used as an initial point. As it can be seen from this figure, the proposed algorithm converges fast
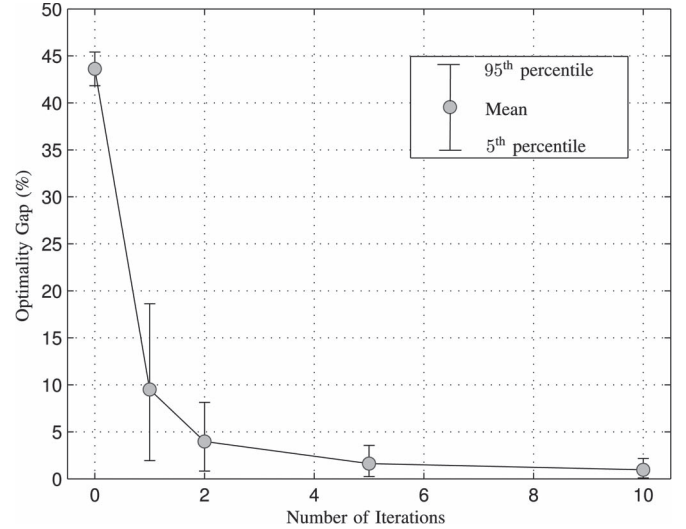

Fig. 3. Optimality gap (%) as a function of the number of iterations.

TABLE IV
MEAN AND STANDARD DEVIATION OF THE OPTIMALITY GAP (%)
COMPARED TO THE OPTIMAL VALUE OF THE ORIGINAL PROBLEM (7)

Number of runs	Mean (%)	Standard deviation (%)
1	3.8	1.7
2	2.4	0.8

to low optimality gaps; in particular, executing the algorithm in 5 iterations provided less than 1.5% mean optimality gap. As a result, the number of iterations, N_{iter}, is chosen to be $N_{\text{iter}} = 5$ throughout the paper.

Fig. 3 illustrates that our algorithm can solve the bound optimization problem (13), which is suitable in dominant interference environment, in a near-optimum manner, as expected from Proposition 1. A natural question to ask is how different is the optimal value obtained by solving the bound optimization as compared to the optimal value obtained by solving the original strongly NP-hard problem (7) using exhaustive search? Due to the exponential computational complexity of exhaustive search, simulating a system of 57 sectors is not feasible. As a result, we only show the optimality gap for a system of 12 sectors in Table IV, for proportional-fair scheduling, i.e., $\alpha = 1$. As we can see, the proposed scheme can achieve, on average, about 96% of the optimum value achieved using exhaustive search, if it is executed once (5 iterations). One can also reduce the optimality gap further by executing the algorithm more than once. That is, after finding the optimum $\{I_n^{(k)}\}$ in one run, the algorithm sets $H_{m,n}^{(\tilde{k},k)} = 0, \forall m, n, k, \tilde{k} | I_n^{(k)} = 1$, and executes the algorithm again. In this case, one can achieve about 97.6% of the optimum value achieved using exhaustive search, i.e., an incremental gain, at the expense of more computational complexity. This suggests that the bound optimization is indeed a good method to achieve near-optimality and shows that most of the gain is already captured by executing the algorithm only once.

B. Comparing the Performance of the Proposed Scheme With the Baseline Schemes

In this section, we compare the performance of the proposed scheme with the baseline schemes presented in Section V-B.
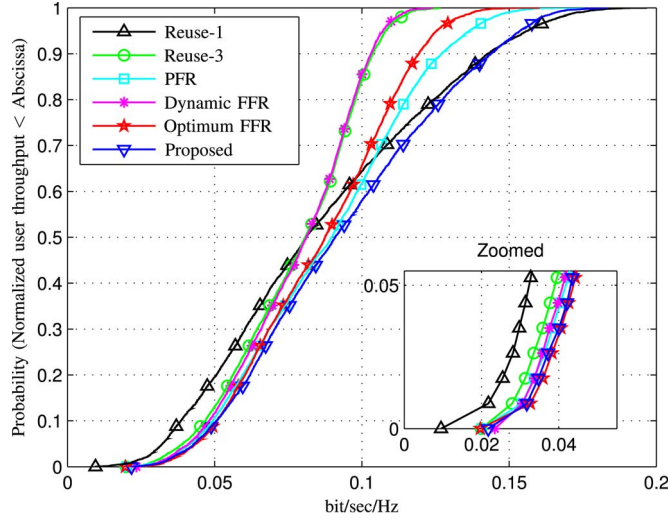
Fig. 4. CDF of the normalized user throughput of all UTs in the network for UMa scenario, $\alpha = 2$ and $\bar{R}_{min} = 0$.
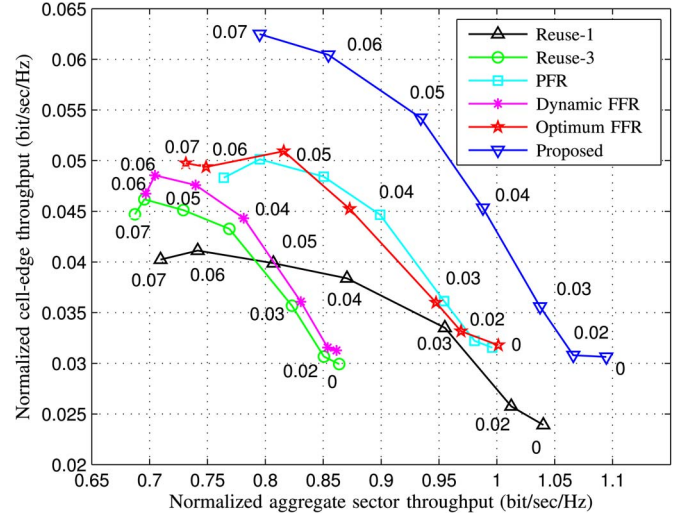

Fig. 5. Normalized cell-edge throughput versus normalized aggregate sector throughput for different schemes, for $\alpha = 1$ and $\bar{R}_{min} \in \{0, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07\}$.

In Fig. 4, we show the CDF of the normalized time-average UT throughput for all schemes. Normalization is performed by dividing the user throughput over the total downlink bandwidth, which is 10 MHz. In all schemes, α-fair scheduler is used with fairness exponent of $\alpha = 2$ (cf. Section III). To facilitate the comparison, we define the normalized cell-edge and cell-center user throughputs as the 5th and the 95th percentiles of the normalized user throughputs, respectively. It is clear from the figure that reuse-1 has the worst cell-edge performance (0.0323 bit/sec/Hz) as compared to the other three schemes, due to the excessive interference experienced at the cell-edge. Reuse-3, PFR, dynamic FFR, optimum FFR, and the proposed scheme achieve normalized cell-edge user throughputs of 0.0391, 0.0417, 0.0412, 0.0435, and 0.0430 bit/sec/Hz, respectively. However, reuse-3, PFR, dynamic FFR, and optimum FFR improve the cell-edge performance at the expense of reducing the overall throughput, especially for UTs close to the cell-center. For example, the 95th percentile achieved by reuse-3, PFR, dynamic FFR, and optimum FFR are 0.108, 0.136, 0.107, and 0.125 bit/sec/Hz, respectively, as compared to 0.156 and 0.154 bit/sec/Hz achieved by reuse-1 and the proposed scheme, respectively. Interestingly, the proposed scheme combines the advantages of all schemes, as it provides high cell-edge and cell-center throughputs simultaneously. In addition to the improvement in cell-center and cell-edge throughputs, the proposed scheme also outperforms the other schemes in terms of the normalized median throughput (50th percentile). Indeed, the gain achieved by the proposed scheme increases for higher fairness exponent, α, or higher $\bar{R}_{min}$, as we will see shortly.

To further examine the performance of the different schemes, we show in Fig. 5 the normalized cell-edge user throughput and the normalized aggregate sector throughput for all schemes and for different minimum average rate requirements $\bar{R}_{min}$, assuming proportional-fair scheduler, i.e., $\alpha = 1$. As expected, the general trend for all schemes is that as $\bar{R}_{min}$ increases, the sector throughput decreases and the cell-edge user throughput increases. For high $\bar{R}_{min}$, we observe that reuse-3, PFR,
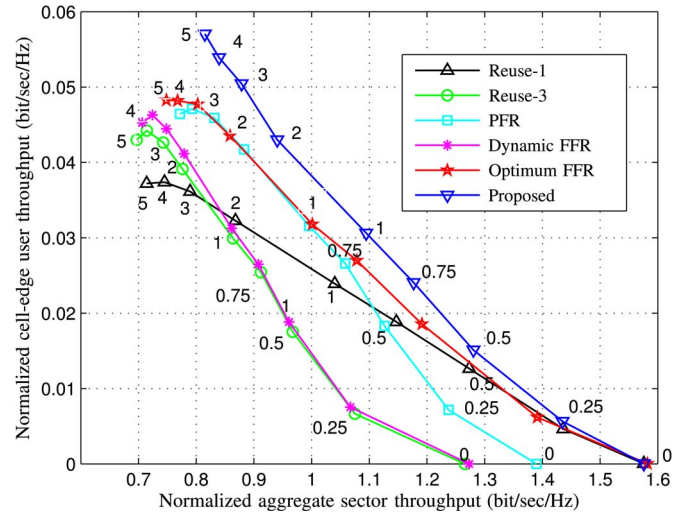
Fig. 6. Normalized cell-edge throughput versus normalized aggregate sector throughput for different schemes, for $\bar{R}_{min} = 0$ and $\alpha \in \{0, 0.25, 0.50, 0.75, 1, 2, 3, 4, 5\}$.

optimum FFR, dynamic FFR, and the proposed scheme have significantly higher cell-edge throughput than reuse-1. On the other hand, for small $\bar{R}_{min}$ it is clear that reuse-3, PFR, dynamic FFR, and optimum FFR incur significant loss in the aggregate sector throughput as compared to reuse-1. Interestingly, the proposed scheme performs very well in both the cell-edge and the sector throughput as compared to all other schemes for all values of $\bar{R}_{min}$. We also plot similar tradeoff curves in Fig. 6 by varying the fairness exponent α and fixing $\bar{R}_{min} = 0$, similar to [40]. Again, the proposed scheme outperforms other schemes in both the cell-edge and the aggregate sector throughput.

In Figs. 7 and 8, we take a closer look at the gains achieved by the different schemes as compared to reuse-1. In Fig. 7, we plot the gains in cell-edge throughput achieved for a given aggregate sector throughput. The gain in cell-edge throughput for a particular scheme for a given sector throughput x is given
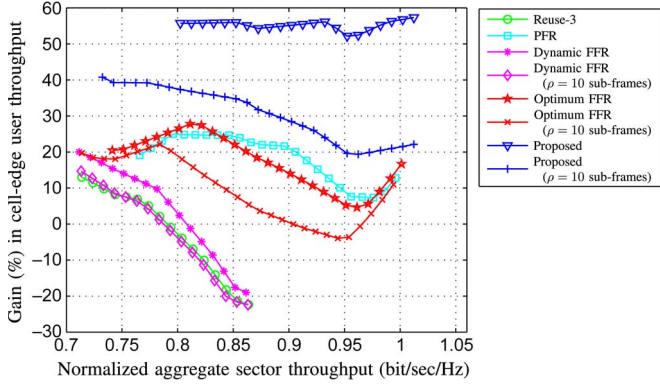
Fig. 7. Gain (%) in cell-edge throughput versus aggregate sector throughput, as compared to reuse-1.
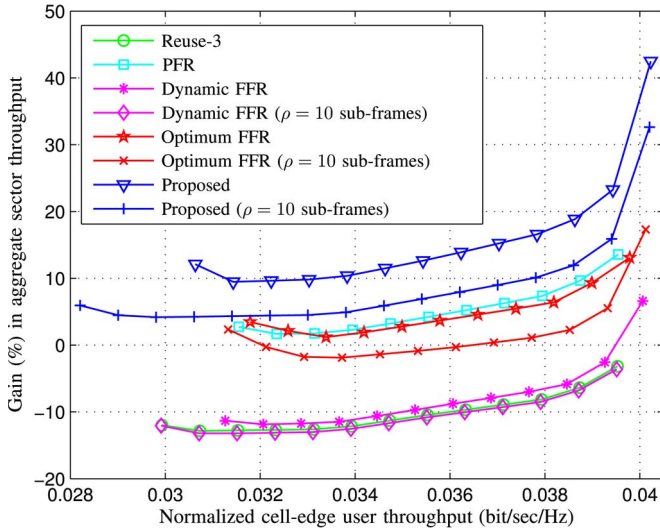

Fig. 8. Gain (%) in sector throughput versus cell-edge throughput, as compared to reuse-1.

by $\text{Gain}_{\text{scheme}}(x) = \frac{\text{CellEdge}_{\text{scheme}}(x) - \text{CellEdge}_{\text{reuse1}}(x)}{\text{CellEdge}_{\text{reuse1}}(x)} \cdot 100\%$, where $\text{CellEdge}_{\text{scheme}}(x)$ is the cell-edge throughput achieved by a particular scheme at sector throughput x, which can be obtained from the tradeoff curves given in Fig. 5. For a wide range of aggregate sector throughput, the proposed scheme achieves large gains (50% to 60%). Dynamic FFR, optimum FFR, and the proposed scheme lose some of the gain if it is executed every 10 sub-frames (as expected for any dynamic scheme); however, the gains for the proposed scheme are consistently better than the other schemes. Similarly, we plot the gains in aggregate sector throughput for a given cell-edge throughput in Fig. 8. The proposed scheme achieves consistently higher gains in aggregate sector throughput than the other schemes, especially for high cell-edge throughput.

In Fig. 9, we plot the outage probability, which is defined as the probability of having the average UT throughput less than $\bar{R}_{\text{min}}$ for different schemes. It is clear from the figure that the proposed scheme achieves much lower outage probability as compared to other schemes. For example, at $\bar{R}_{\text{min}} = 0.05$ bits/sec/Hz, the proposed scheme has an outage probability that is at least 3.5 times less than those for PFR and optimum
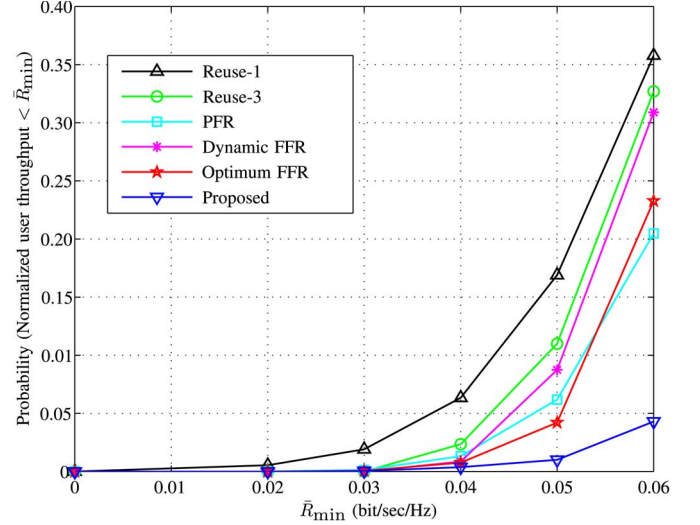

Fig. 9. Outage probability versus $\bar{R}_{\text{min}}$, which is defined as the probability that the average UT throughput less than $\bar{R}_{\text{min}}$.

FFR, and at least 6 times less than those for reuse-1, reuse-3, and dynamic FFR.

C. Comparing the Complexity of the Proposed Scheme With the Baseline Schemes

By analyzing the dynamic FFR algorithm in [9] and the optimum FFR algorithm in [39], we obtained the computational complexity and the rate of message exchange required to execute both algorithms. Due to space limitation, we only summarize the final result in Table V. In terms of computational complexity, it can be seen from Table V that the proposed algorithm scales better (worse) than dynamic and optimum FFR as N or K $(M^{(k)})$ increases. In terms of the rate of message exchange, the proposed scheme scales better than dynamic and optimum FFR as $M^{(k)}$ increases, while optimum FFR scales best as N increases.

D. Statistics of the Average Number of Blanked RBs

In Fig. 10, we plot the probability mass function of the average number of blanked RBs per sector for different fairness exponents. The proposed scheme has the flexibility to change the distribution of the blanked resources according to the desired fairness level. As α increases, cell-edge users become more important and thus more resources need to be blanked, and vice versa. The figure shows also that the proposed scheme acts as reuse-1 (no blanking) and reuse-3 (2/3 of RBs are blanking) with very small probabilities.

VII. CONCLUSION

In this paper, we tackled the problem of distributed multi-cell resource allocations with blanking where the objective is to maximize the weighted sum-rate. This problem is known to be strongly NP-hard problem; nevertheless, we demonstrated that this problem can be tightly relaxed to a linear programming problem in a dominant interference environment, in which

TABLE V
COMPARISON BETWEEN THE PROPOSED DISTRIBUTED SCHEME, DYNAMIC FFR, AND OPTIMUM FFR,
IN TERMS OF COMPUTATIONAL COMPLEXITY AND RATE OF MESSAGE EXCHANGE REQUIRED

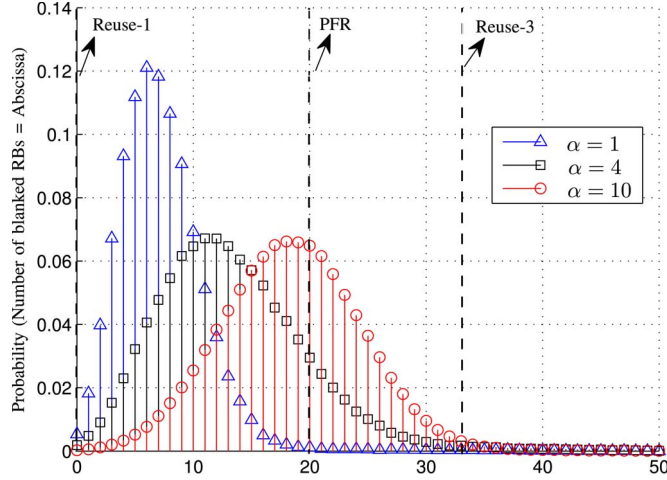| Scheme	Computational complexity		Rate of message exchange required for sector k (bit/sec)
	Centralized entity	Sector k	
Proposed Scheme	None	$O\big(N\big(M^{(k)}\log M^{(k)}\big)^2\big)$	$\dfrac{2N_{\text{iter}}\tilde{K}NL_q}{\rho \times 1\text{ms}}$
Dynamic FFR [9]	$O\big(N(N-1)K\big)$	$O\big(NM^{(k)}\big)$	$\dfrac{2NM^{(k)}L_q}{\rho \times 1\text{ms}}$
Optimum FFR [39]	$O\big(N^2\sum\limits_{k=1}^{K}M^{(k)}\big)$	$O\big(NM^{(k)}\big)$	$\dfrac{M^{(k)}L_q}{\rho \times 1\text{ms}}$

Fig. 10. Probability mass function of the average number of blanked RBs per sector (the total number of RBs is 50).

for each user the received power from each interferer is significantly greater than the aggregate received power from all other weaker interferers. Using linear relaxation, we proposed a polynomial-time distributed algorithm that is not only guaranteed to be tight for dominant interference environment, but which also computes an upper bound on the optimality gap. The proposed scheme is developed using the primal-decomposition method, which is utilized to decompose the problem into a master problem and multiple subproblems. The master problem is solved iteratively using the projected-subgradient method. We reveal that each subproblem has a network flow structure which makes it amenable to powerful MCNF algorithms and thus results in significant reduction in the computational complexity. Simulation results show that the proposed scheme achieves high gain in aggregate throughout, cell-edge throughput, and outage probability, as compared to reuse-1, reuse-3, partial frequency reuse, dynamic fractional frequency reuse, and optimum fractional frequency reuse. An interesting extension to this work is to use the decomposition algorithms proposed in [41] to devise distributed ICIC algorithms.

APPENDIX A
ON THE TIGHTNESS OF THE BOUND GIVEN IN (8)

In this appendix, we show that the SINR bound given by (8) is tight in the sense if the bound is increased by Δ, then the actual SINR will increase by at least Δ. Let

$\tilde{k}^\star = \arg\max\limits_{\tilde{k}\in\tilde{\mathcal{K}}^{(k)}} I_n^{(\tilde{k})}H_{m,n}^{(k,\tilde{k})}$, then the exact SINR expression given by (1) can be written as

$$
\begin{aligned}
\Gamma_{m,n}^{(k)} &= \frac{P_C H_{m,n}^{(k,k)}}{P_C\left(\sum\limits_{\tilde{k}\neq k}H_{m,n}^{(k,\tilde{k})}-I_n^{\tilde{k}^\star}H_{m,n}^{(k,\tilde{k}^\star)}-\sum\limits_{\tilde{k}\neq k,\tilde{k}\neq\tilde{k}^\star}I_n^{(\tilde{k})}H_{m,n}^{(k,\tilde{k})}\right)+P_N}, \\[2mm]
&= \frac{P_C H_{m,n}^{(k,k)}}{P_C\left(\sum\limits_{\tilde{k}\neq k}H_{m,n}^{(k,\tilde{k})}-I_n^{\tilde{k}^\star}H_{m,n}^{(k,\tilde{k}^\star)}\right)+P_N} \\[2mm]
&\quad\times\left(1+\frac{P_C\sum\limits_{\tilde{k}\neq k,\tilde{k}\neq\tilde{k}^\star}I_n^{(\tilde{k})}H_{m,n}^{(k,\tilde{k})}}{P_C\sum\limits_{\tilde{k}\neq k}\left(1-I_n^{(\tilde{k})}\right)H_{m,n}^{(k,\tilde{k})}+P_N}\right) \\[2mm]
&\geq \frac{P_C H_{m,n}^{(k,k)}}{P_C\left(\sum\limits_{\tilde{k}\neq k}H_{m,n}^{(k,\tilde{k})}-I_n^{\tilde{k}^\star}H_{m,n}^{(k,\tilde{k}^\star)}\right)+P_N}, \quad (33)
\end{aligned}
$$

where the right-hand side of the inequality is the bound given by (8). From the previous expression, if the bound is increased by Δ, then $\Gamma_{m,n}^{(k)}$ will increase by $\Delta \times$
$\left(1+\dfrac{P_C\sum\limits_{\tilde{k}=1,\tilde{k}\neq k,\tilde{k}\neq\tilde{k}^\star}^{K}I_n^{(\tilde{k})}H_{m,n}^{(k,\tilde{k})}}{P_C\sum\limits_{\tilde{k}=1,\tilde{k}\neq k}^{K}(1-I_n^{(\tilde{k})})H_{m,n}^{(k,\tilde{k})}+P_N}\right)$, i.e., $\Gamma_{m,n}^{(k)}$ will increase by at least Δ.

APPENDIX B
PROOF OF LEMMA 1

We proceed by proving that $\mathcal{C}\subseteq\mathcal{C}'$ and $\mathcal{C}'\subseteq\mathcal{C}$. To prove that $\mathcal{C}\subseteq\mathcal{C}'$, we assume that $y_{m,n}^{(k,\tilde{k})}\in\mathcal{C}$ and deduce that $y_{m,n}^{(k,\tilde{k})}\in\mathcal{C}'$. Since $\sum_{\tilde{k}\in\tilde{\mathcal{K}}^{(k)}}y_{m,n}^{(k,\tilde{k})}\leq 1$ and $y_{m,n}^{(k,\tilde{k})}\leq x_{m,n}^{(k)}$, then $\sum_{\tilde{k}\in\tilde{\mathcal{K}}^{(k)}}y_{m,n}^{(k,\tilde{k})}\leq x_{m,n}^{(k)}$. Moreover, since $y_{m,n}^{(k,\tilde{k})}\leq x_{m,n}^{(k)}$, then $\sum_{m=1}^{M^{(k)}}y_{m,n}^{(k,\tilde{k})}\leq\sum_{m=1}^{M^{(k)}}x_{m,n}^{(k)}\leq 1$. Since $\sum_{m=1}^{M^{(k)}}y_{m,n}^{(k,\tilde{k})}\leq 1$, $y_{m,n}^{(k,\tilde{k})}\leq I_n^{(\tilde{k})}$, and $y_{m,n}^{(k,\tilde{k})}\in\{0,1\}$, then $\sum_{m=1}^{M^{(k)}}y_{m,n}^{(k,\tilde{k})}\leq I_n^{(\tilde{k})}$. Consequently, $y_{m,n}^{(k,\tilde{k})}\in\mathcal{C}'$, which means that $\mathcal{C}\subseteq\mathcal{C}'$.

Similarly, to show that $\mathcal{C}'\subseteq\mathcal{C}$, we assume that $y_{m,n}^{(k,\tilde{k})}\in\mathcal{C}'$ and deduce that $y_{m,n}^{(k,\tilde{k})}\in\mathcal{C}$. Since $\sum_{\tilde{k}\in\tilde{\mathcal{K}}^{(k)}}y_{m,n}^{(k,\tilde{k})}\leq x_{m,n}^{(k)}$ and $x_{m,n}^{(k)},y_{m,n}^{(k,\tilde{k})}\in\{0,1\}$, then $y_{m,n}^{(k,\tilde{k})}\leq x_{m,n}^{(k)}$. Moreover, since $\sum_{m=1}^{M^{(k)}}y_{m,n}^{(k,\tilde{k})}\leq I_n^{(\tilde{k})}$ and $I_n^{(k)},y_{m,n}^{(k,\tilde{k})}\in\{0,1\}$, then $\sum_{m=1}^{M^{(k)}}y_{m,n}^{(k,\tilde{k})}\leq 1$

which means that $y_{m,n}^{(k,\tilde{k})} \leq I_n^{(\tilde{k})}$. Hence, $y_{m,n}^{(k,\tilde{k})} \in \mathcal{C}$ which means that $\mathcal{C}' \subseteq \mathcal{C}$.

APPENDIX C
PROOF OF PROPOSITION 1

We start by categorizing the constraints of the relaxed version of (18) into two groups: non-bounding constraints, which are given by (18b)–(18d), and bounding constraints, which are given by (19) (relaxed version of (18e)). It is easy to see that the number of variables that assume optimal binary values is equal to the number of bounding inequality constraints that are active at the optimal, i.e., inequalities that are satisfied with equalities at the optimal.

We now note that the optimal solution of a linear program lies on one or more vertices. For a linear program with N_{var} variables, a vertex is characterized by N_{var} equalities or active inequalities. Let $N_{\text{non-bounding}}^\star$ denote the number of non-bounding equality or active inequality constraints at the optimum, and let $N_{\text{bounding}}^\star$ denote the number of active bounding inequality constraints at the optimum. Hence, we can write the following:

$$
\begin{aligned}
N_{\text{var}} &= N_{\text{bounding}}^\star + N_{\text{non-bounding}}^\star \\
\Rightarrow N_{\text{bounding}}^\star &= N_{\text{var}} - N_{\text{non-bounding}}^\star \\
\Rightarrow N_{\text{bounding}}^\star &\geq N_{\text{var}} - \sup(N_{\text{non-bounding}}^\star) \\
&= (\tilde{K}+1)\sum_{k=1}^{K} M^{(k)} + K \\
&\quad - \left(K + \sum_{k=1}^{K} \left(M^{(k)} + \tilde{K} \right) \right) \\
&= \sum_{k=1}^{K} \tilde{K}(M^{(k)} - 1),
\end{aligned}
\tag{34}
$$

where the supremum is found by assuming that all non-bounding constraints are active. Thus, the percentage of optimal variables of the relaxed version of (18) that assume binary values can be bounded as

$$
\begin{aligned}
P_{\text{binary}} &\geq \frac{\tilde{K}\sum_{k=1}^{K} \left(M^{(k)} - 1 \right)}{(\tilde{K}+1)\sum_{k=1}^{K} M^{(k)} + K} \cdot 100\% \\
&= \frac{\tilde{K}(\bar{M} - 1)}{(\tilde{K}+1)\bar{M} + 1} \cdot 100\%,
\end{aligned}
\tag{35}
$$

where $\bar{M}$ is the average number of UTs per sector, i.e., $\bar{M} = \frac{1}{K}\sum_{k=1}^{K} M^{(k)}$.

ACKNOWLEDGMENT

REFERENCES

[1] A. Bin Sediq, R. Schoenen, H. Yanikomeroglu, G. Senarath, and Z. Chao, "A novel distributed inter-cell interference coordination scheme based on projected subgradient and network flow optimizations," in *Proc. IEEE Int. Symp. PIMRC*, Toronto, ON, Canada, Sep. 2011, pp. 1595–1600.

[2] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1996.

[3] M. Rahman and H. Yanikomeroglu, "Enhancing cell-edge performance: A downlink dynamic interference avoidance scheme with inter-cell coor-

[4] dination," *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, pp. 1414–1425, Apr. 2010.

[4] "Soft frequency reuse scheme for UTRAN LTE," Cedex, France, Tech. Rep. 3GPP Project Doc. R1-050 507, May 2005.

[5] M. Sternad, T. Ottosson, A. Ahlen, and A. Svensson, "Attaining both coverage and high spectral efficiency with adaptive OFDM downlinks," in *Proc. IEEE VTC-Fall*, Orlando, FL, USA, Oct. 2003, pp. 2486–2490.

[6] Y. Xiang, J. Luo, and C. Hartmann, "Inter-cell interference mitigation through flexible resource reuse in OFDMA based communication," in *Proc. 13th Eur. Wireless Conf.*, Paris, France, Apr. 2007.

[7] A. Simonsson, "Frequency reuse and intercell interference co-ordination in E-UTRA," in *Proc. IEEE VTC-Spring*, Dublin, Ireland, Apr. 2007, pp. 3091–3095.

[8] "Further discussion on adaptive fractional frequency reuse," Cedex, France, Tech. Rep. 3GPP Project Doc. R1-072 762, Jun. 2007.

[9] S. H. Ali and V. C. M. Leung, "Dynamic frequency allocation in fractional frequency reused OFDMA networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 4286–4295, Aug. 2009.

[10] X. Zhang, C. He, L. Jiang, and J. Xu, "Inter-cell interference coordination based on softer frequency reuse in OFDMA cellular systems," in *Proc. ICNNSP*, Zhenjiang, China, Jun. 2008, pp. 270–275.

[11] X. Mao, A. Maaref, and K. H. Teo, "Adaptive soft frequency reuse for inter-cell interference coordination in SC-FDMA based 3GPP LTE uplinks," in *Proc. IEEE GLOBECOM*, New Orleas, LA, USA, Dec. 2008, pp. 1–6.

[12] A. Stolyar and H. Viswanathan, "Self-organizing dynamic fractional frequency reuse for best-effort traffic through distributed inter-cell coordination," in *Proc. IEEE INFOCOM*, Rio de Janeiro, Brazil, Apr. 2009, pp. 1287–1295.

[13] D. Bültmann, T. Andre, and R. Schoenen, "Analysis of 3GPP LTE-Advanced cell spectral efficiency," in *Proc. IEEE Int. Symp. PIMRC*, Istanbul, Turkey, Sep. 2010, pp. 1876–1881.

[14] G. Li and H. Liu, "Downlink radio resource allocation for multi-cell OFDMA system," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3451–3459, Dec. 2006.

[15] R. Y. Chang, Z. Tao, J. Zhang, and C. C. J. Kuo, "Multicell OFDMA downlink resource allocation using a graphic framework," *IEEE Trans. Veh. Technol.*, vol. 58, no. 12, pp. 3494–3507, Sep. 2009.

[16] J. Ellenbeck, C. Hartmann, and L. Berlemann, "Decentralized inter-cell interference coordination by autonomous spectral reuse decisions," in *Proc. 14th European Wireless Conf.*, Prague, Czech Republic, Jun. 2008, pp. 1–7.

[17] Z.-Q. T. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 57–73, Feb. 2008.

[18] R. Ahuja, T. Magnanti, and J. Orlin, *Network Flows: Theory, Algorithms, Applications*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.

[19] M. Tao and Y. Liu, "A network flow approach to throughput maximization in cooperative OFDMA networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 3, pp. 1138–1148, Mar. 2013.

[20] A. N. Zaki and A. O. Fapojuwo, "Optimal and efficient graph-based resource allocation algorithms for multiservice frame-based OFDMA networks," *IEEE Trans. Mobile Comput.*, vol. 10, no. 8, pp. 1175–1186, Aug. 2011.

[21] International Telecommunication Union (ITU). (Dec. 2009). Guidelines for evaluation of radio interface technologies for IMT-Advanced, Geneva, Switzerland, Tech. Rep. ITU-R M.2135-1. [Online]. Available: http://www.itu.int/pub/R-REP-M.2135-1-2009

[22] H. Kushner and P. Whiting, "Convergence of proportional-fair sharing algorithms under general conditions," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1250–1259, Jul. 2004.

[23] A. L. Stolyar, "On the asymptotic optimality of the gradient scheduling algorithm for multiuser throughput allocation," *Oper. Res.*, vol. 53, no. 1, pp. 12–25, Jan. 2005.

[24] G. Song and Y. Li, "Cross-layer optimization for OFDM wireless networks-part II: Algorithm development," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 625–634, Mar. 2005.

[25] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. Netw.*, vol. 8, no. 5, pp. 556–567, Oct. 2000.

[26] T. Lan, D. Kao, M. Chiang, and A. Sabharwal, "An axiomatic theory of fairness in network resource allocation," in *Proc. IEEE INFOCOM*, San Diego, CA, USA, Mar. 2010, pp. 1–9.

[27] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan, "Rate control for communication networks: Shadow prices, proportional fairness and stability," *J. Oper. Res. Soc.*, vol. 49, no. 3, pp. 237–252, 1998.

[28] P. Viswanath, D. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1277–1294, Jun. 2002.

[29] A. Bin Sediq, R. H. Gohary, R. Schoenen, and H. Yanikomeroglu, "Optimal tradeoff between sum-rate efficiency and Jain's fairness index in resource allocation," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3496–3509, Jul. 2013.

[30] H. T. Cheng and W. Zhuang, "An optimization framework for balancing throughput and fairness in wireless networks with QoS support," *IEEE Trans. Wireless Commun.*, vol. 7, no. 2, pp. 584–593, Feb. 2008.

[31] M. Andrews, L. Qian, and A. Stolyar, "Optimal utility based multi-user throughput allocation subject to throughput constraints," in *Proc. IEEE INFOCOM*, Miami, FL, USA, Mar. 2005, pp. 2415–2424.

[32] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA, USA: Athena Scientific, 2008.

[33] S. Boyd, L. Xiao, A. Mutapcic, and J. Mattingley, *Notes on Decomposition Methods*, *Lecture notes for EE364B*. Stanford, CA, USA: Stanford Univ. Press, Oct. 2003.

[34] S. Boyd and A. Mutapcic, *Subgradient Methods*, *Lecture notes for EE364b*. Stanford, CA, USA: Stanford Univ. Press, Oct. 2003.

[35] S. Boyd, *Subgradients*, *Lecture notes for EE364b*. Stanford, CA, USA: Stanford Univ. Press, Apr. 2011.

[36] J. Nystrom, Software implementation of IMT.Eval channel model, WINNER+, Munich, Germany, Tech. Rep. CELTIC/CP5-026 Project WINNER+ Doc. 5D/478-E. [Online]. Available: http://www.itu.int/oth/R0A06000022/en

[37] R. Schoenen and C. Teijeiro, "System level performance evaluation of LTE with MIMO and relays in reuse-1 IMT-Advanced scenarios," in *Proc. IEEE WiCom*, Chengdu, China, Sep. 2010, pp. 1–5.

[38] "Calibration for IMT-Advanced Evaluations," Munich, Germany, Tech. Rep. CELTIC/CP5-026 Project WINNER+, May 2010.

[39] M. Assaad, "Optimal fractional frequency reuse (FFR) in multicellular OFDMA system," in *Proc. IEEE VTC-Fall*, Marina Bay, Singapore, Sep. 2008, pp. 1–5.

[40] R. Schoenen, A. Bin Sediq, H. Yanikomeroglu, G. Senarath, and Z. Chao, "Fairness analysis in cellular networks using stochastic Petri nets," in *Proc. IEEE Int. Symp. PIMRC*, Toronto, ON, Canada, Sep. 2011, pp. 1983–1988.

[41] G. Scutari, F. Facchinei, P. Song, D. P. Palomar, and J.-S. Pang, "Decomposition by partial linearization: Parallel optimization of multi-agent systems," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 641–656, Feb. 2014.

Akram Bin Sediq (S'04) received the B.Sc. degree in electrical engineering (*summa cum laude*) from the American University of Sharjah (AUS), Sharjah, United Arab Emirates, in 2006 and the M.A.Sc. degree in electrical engineering and the Ph.D. degree in electrical and computer engineering from Carleton University, Ottawa, ON, Canada, in 2008 and 2013, respectively.

In Winter 2012, he was a course instructor at the Department of Systems and Computer Engineering at Carleton University. During 2013–2014, he was a system engineer at BLiNQ Networks, Kanata, ON, Canada, where he was responsible for developing novel algorithms for non-line-of-sight small-cell backhaul networks, including, dual-carrier joint scheduling, interference avoidance, beam selection, and IEEE 1588 clock synchronization algorithms. He has recently joined the R&D Group at Ericsson Canada, Ottawa, Canada. His research interests include inter-cell interference coordination, radio resource allocations, heterogenous networks, constellation design, cooperative communications, clock synchronization, and applications of optimization in wireless networks.

Dr. Bin Sediq is a recipient of the IEEE Antennas and Propagation Society Undergraduate Scholarship in 2004, the IEEE Microwave Theory and Techniques Society Undergraduate/Pre-Graduate Scholarship in 2005, the third prize in IEEE IAS Myron Zucker Undergraduate Student Design Contest in 2005, the President Cup for graduating with the highest GPA at the Bachelor's level at AUS in 2006, the Ontario Graduate Scholarship for international students for three years in a row during 2007–2010, two Senate Medals for outstanding academic achievement from Carleton University at the Masters and Ph.D. levels in 2008 and 2013, respectively, the EDC Teaching Assistant Outstanding Award at Carleton University in 2013, and NSERC Industrial R&D Fellowship (IRDF), 2013–2015.

Rainer Schoenen (SM'13) received the German Diplom-Ingenieur and Ph.D. degrees in electrical engineering from RWTH Aachen University, Aachen, Germany, in 1995 and 2000, respectively. His research interests include stochastic Petri nets and queuing systems, ATM, TCP/IP, switching, flow control, QoS, tariffs, Userin-the-loop (UIL), wireless resource and packet scheduling and the MAC layer of 4+5G systems. His Ph.D. thesis was "System Components for Broadband Universal Networks with QoS Guarantee" with the ISS group of Prof. Heinrich Meyr at RWTH Aachen University from 1995 to 2000. He started working self-employed in 2000. Dr. Schoenen was a senior researcher at the Communication Networks (ComNets) Research Group, RWTH Aachen with Professor Walke from 2005 to 2009, working on computer networks, queuing theory, Petri nets, LTE-Advanced, FDD relaying, scheduling, OSI layer 2 (MAC) and IMT-Advanced Evaluation within WINNER+. From 2010 to 2014 he worked at the Department of Systems and Computer Engineering, Carleton University, Ottawa, Canada, as a project manager supporting Professor Halim Yanikomeroglu. Since 2015 he is a professor at HAW Hamburg, Germany.

Halim Yanikomeroglu (S'96–M'98–SM'12) was born in Giresun, Turkey, in 1968. He received the B.Sc. degree in electrical and electronics engineering from the Middle East Technical University, Ankara, Turkey, in 1990, and the M.A.Sc. degree in electrical engineering (now ECE) and the Ph.D. degree in electrical and computer engineering from the University of Toronto, Toronto, ON, Canada, in 1992 and 1998, respectively.

During 1993–1994, he was with the R&D Group of Marconi Kominikasyon A.S., Ankara, Turkey. Since 1998 he has been with the Department of Systems and Computer Engineering at Carleton University, Ottawa, Canada, where he is now a Full Professor. His research interests cover many aspects of wireless technologies with a special emphasis on cellular networks. He coauthored over 60 IEEE journal papers, and has given a high number of tutorials and invited talks on wireless technologies in the leading international conferences. In recent years, his research has been funded by Huawei, Blackberry, Samsung, Communications Research Centre of Canada (CRC), Telus, and Nortel. This collaborative research resulted in about 20 patents (granted and applied). Dr. Yanikomeroglu has been involved in the organization of the IEEE Wireless Communications and Networking Conference (WCNC) from its inception, including serving as Steering Committee Member as well as the Technical Program Chair or Co-Chair of WCNC 2004 (Atlanta), WCNC 2008 (Las Vegas), and WCNC 2014 (Istanbul). He was the General Co-Chair of the IEEE Vehicular Technology Conference Fall 2010 held in Ottawa. He has served in the editorial boards of the IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and IEEE COMMUNICATIONS SURVEYS & TUTORIALS. He was the Chair of the IEEE's Technical Committee on Personal Communications (now called Wireless Technical Committee). He is a Distinguished Lecturer for the IEEE Vehicular Technology Society since 2010.

Dr. Yanikomeroglu is a recipient of the IEEE Ottawa Section Outstanding Educator Award in 2014, Carleton University Faculty Graduate Mentoring Award in 2010, the Carleton University Graduate Students Association Excellence Award in Graduate Teaching in 2010, and the Carleton University Research Achievement Award in 2009. Dr. Yanikomeroglu spent the 2011–2012 academic year at TOBB University of Economics and Technology, Ankara, Turkey, as a Visiting Professor. He is a registered Professional Engineer in the province of Ontario, Canada.

Gamini Senarath received the B.Sc. degree in electrical and electronics engineering from the Moratuwa University, Sri Lanka, in 1980, the Master of Electronics Engineering degree from the Eindhoven (Phillips) International Institute, the Netherlands, in 1986 and the Ph.D. degree in telecommunications engineering from the University of Melbourne, Australia, in 1996, respectively.

During 1980–1990, he was with the public telecommunication provider of the Government of Sri Lanka, as a Radio Transmission Engineer and a Regional Telecommunication Manager. During 1996 to 2009 he was with the Advanced Wireless Technology Labs, Nortel Networks, Ottawa, Canada, where he was involved in many next generation wireless system research projects and contributed to various 3G and 4G wireless standard development activities. The main focus was on propagation, cell planning, relaying and MAC/RRM areas. Since 2009, he has been with 5G Research Lab, Huawei Technologies Canada. His current research work involves software defined networks, virtual network embedding, power and the resource management of 3rd party infrastructure provided by multiple operators for the 5G wireless systems. His work resulted in more than 50 granted patents in wireless communication systems. Dr. Senarath is also the project manager for the Huawei's external collaboration project with Carleton University, Telus and Communications Research Centre of Canada (CRC).