Distributed Interference Management Policies for Heterogeneous Small Cell Networks

Kartik Ahuja, Yuanzhang Xiao and Mihaela van der Schaar
Department of Electrical Engineering, UCLA, Los Angeles, CA, 90095
Email: ahujak@ucla.edu, yxiao@seas.ucla.edu and mihaela@ee.ucla.edu

Abstract

We study the problem of *distributed* interference management in a network of heterogeneous small cells with different cell sizes, different numbers of user equipments (UEs) served, and different throughput requirements by UEs. We consider the uplink transmission, where each UE determines when and at what power level it should transmit to its serving small cell base station (SBS). We propose a general framework for designing distributed interference management policies, which exploits weak interference among non-neighboring UEs by letting them transmit simultaneously (i.e., spatial reuse), while eliminating strong interference among neighboring UEs by letting them transmit in different time slots. The design of optimal interference management policies has two key steps. Ideally, we need to find all the subsets of non-interfering UEs, i.e., the maximal independent sets (MISs) of the interference graph, but this is NP-hard (non-deterministic polynomial time) even when solved in a centralized manner. Then, in order to maximize some given network performance criterion subject to UEs' minimum throughput requirements, we need to determine the optimal fraction of time occupied by each MIS, which requires global information (e.g., all the UEs' throughput requirements and channel gains). In our framework, we first propose a *distributed* algorithm for the UE-SBS pairs to find *a subset of* MISs in logarithmic time (with respect to the number of UEs). Then we propose a novel problem reformulation which enables UE-SBS pairs to determine the optimal fraction of time occupied by each MIS with only local message exchange among the neighbors in the interference graph. Despite the fact that our interference management policies are distributed and utilize only local information, we can *analytically* bound their performance under a wide range of heterogeneous deployment scenarios in terms of the competitive ratio with respect to the optimal network performance, which can only be obtained in a centralized manner with NP complexity. Remarkably, we prove that the competitive ratio is independent of the network size. Through extensive simulations, we show that our proposed policies achieve significant performance improvements (ranging from 150% to 700%) over state-of-the-art policies.

I. Introduction

Dense deployment of low-cost heterogeneous small cells (e.g. picocells, femtocells) has become one of the most effective solutions to accommodate the exploding demand for wireless spectrum [1] [2] [3]. On one hand, dense deployment of small cells significantly shortens the distances between small cell base stations (SBSs) and their corresponding user equipments (UEs), thereby boosting the network capacity. On the other hand, dense deployment also shortens the distances between neighboring SBSs, thereby potentially increasing the inter-cell interference. Hence, while the solution provided by the dense deployment of small cells is promising, its success depends crucially on interference management by the small cells. Efficient interference management is even more challenging in heterogeneous small cell networks, due to the lack of central coordinators, compared to that in traditional cellular networks.

In this paper, we propose a novel framework for designing interference management policies in the uplink of small cell networks, which specify when and at what power level each UE should transmit[1]. Our proposed design framework and the resulting interference management policies fulfill *all* the following important requirements:

- *Deployment of heterogeneous small cell networks*: Existing deployments of small cell networks exhibit significant heterogeneity such as different types of small cells (picocells and femtocells), different cell sizes, different number of UEs served, different UEs' throughput requirements etc.
- *Interference avoidance and spatial reuse*: Effective interference management policies should take into account the strong interference among neighboring UEs, as well as the weak interference among non-neighboring UEs. Hence, the policies should effectively avoid interference among neighboring UEs and use spatial reuse to take advantage of the weak interference among non-neighboring UEs.
- *Distributed implementation with local information and message exchange*: Since there is no central coordinator in small cell networks, interference management policies need to be computed and implemented by the UEs in a distributed manner, by exchanging only local information through local message exchanges among neighboring UE-SBS pairs.
- *Scalability to large networks*: Small cells are often deployed over a large scale (e.g., in a city). Effective interference management policies should scale in large networks, namely achieve efficient network performance while maintaining low computational complexity.
- *Ability to optimize different network performance criteria*: Under different deployment sce-

[1] Although we focus on uplink transmissions in this paper, our framework can be easily applied to downlink transmissions.

narios the small cell networks may have different performance criteria, e.g., weighted sum throughput or max-min fairness. The design framework should be general and should prescribe different policies to optimize different network performance criteria.

- *Performance guarantees for individual UEs*: Effective interference management should provide performance guarantees (e.g., minimum throughput guarantees) for individual UEs.

As we will discuss in detail in Section II, existing state-of-the-art policies for interference management cannot simultaneously fulfill all of the above requirements.

Next, we describe our key results and major contributions:

1. We propose a general framework for designing distributed interference management policies that maximizes the given network performance criterion subject to each UE's minimum throughput requirements. The proposed policies schedule maximal independent sets (MISs)[2] of the interference graph to transmit in each time slot. In this way, they avoid strong interference among neighboring UEs (since neighboring UEs cannot be in the same MIS), and efficiently exploit the weak interference among UEs in a MIS by letting them to transmit at the same time.

2. We propose a distributed algorithm for the UEs to determine a subset of MISs. The subset of MISs generated ensures that each UE belongs to at least one MIS in this subset. Moreover, the subset of MISs can be generated in a distributed manner in logarithmic time (logarithmic in the number of UEs in the network) for bounded-degree interference graphs[3]. The logarithmic convergence time is significantly faster than the time (linear or quadratic in the number of UEs) required by the distributed algorithms for generating subsets of MISs in [4]–[6].

3. Given the computed subsets of MISs, we propose a distributed algorithm in which each UE determines the optimal fractions of time occupied by the MISs with only local message exchange. The message is exchanged only among the UE-SBS pairs that strongly interfere with each other, i.e. among neighbors in the interference graph. The distributed algorithm will output the optimal fractions of time for each MIS such that the given network performance criterion is

[2]Consider the interference graph of the network, where each vertex is a UE-SBS pair and each edge indicates strong interference between the two vertices. An independent set (IS) is a set of vertices in which no pair is connected by an edge. An IS is a MIS if it is not a proper subset of another IS.

[3]Bounded-degree graphs are the graphs whose maximum degree can be bounded by a constant independent of the size of the graph, i.e., $\Delta = \mathcal{O}(1)$. As we will show in Theorem 5, for the interference graphs that are not bounded-degree graphs, even the centralized solution, given all the MISs, cannot satisfy the minimum throughput requirements.

maximized subject to the minimum throughput requirements.

4. Under a wide range of conditions, we analytically characterize the competitive ratio of the proposed distributed policy with respect to the optimal network performance. Importantly, we prove that the competitive ratio is independent of the network size, which demonstrates the scalability of our proposed policy in large networks. Remarkably, the constant competitive ratio is achieved even though our proposed policy requires only local information, is distributed, and can be computed fast, while the optimal network performance can only be obtained in a centralized manner with global information (e.g., all the UEs' channel gains, maximum transmit power levels, minimum throughput requirements) and NP (non-deterministic polynomial time) complexity.

5. Through simulations, we demonstrate significant (from 160% to 700 %) performance gains over state-of-the-art policies. Moreover, we show that our proposed policies can be easily adapted to a variety of heterogeneous deployment scenarios, with dynamic entry and exit of UEs.

The rest of the paper is organized as follows. In Section II we discuss the related works and their limitations. We describe the system model in Section III. Then we formulate the interference management problem and give a motivating example in Section IV. We propose the design framework in Section **??**, and demonstrate the performance gain of our proposed policies in Section VI. Finally, we conclude the paper in Section VII.

II. Related Works

State-of-the-art interference management policies can be divided into three main categories: policies based on power control, policies based on spatial reuse, and policies based on joint spatial reuse and power control.

A. Distributed Interference Management Based on Power Control

Policies based on distributed power control, with representative references [7]–[14] have been used for interference management in both cellular and ad-hoc networks. In these policies, all the UEs in the network transmit at a *constant* power all the time (provided that the system parameters remain the same)[4]. The major limitation of policies based on power control is the difficulty in

[4]Although some power control policies [7], [8], [10] go through a transient period of adjusting the power levels before the convergence to the optimal power levels, the users maintain constant power levels after the convergence.

providing minimum throughput guarantees for each UE, especially in the presence of strong interference. Some works [7], [8], [10] use pricing to mitigate the strong interference. However, they [7], [8], [10] cannot strictly guarantee the UEs' minimum throughput requirements. Indeed, the low throughput experienced by some users, caused by strong interference, is the fundamental limitation of such power control approaches - even the optimal power control policy obtained by a central controller [15], [16] can be inefficient [5]. Since strong interference is very common in dense small cell deployments (e.g. in offices and apartments where SBSs are installed close to each other [18]), more efficient policies are required which can guarantee the individual UEs' throughput requirements. Also, there exist a different strand of work based on [19] which proposes a distributed algorithm to achieve the desired minimum throughput requirement for each UE. However, these works cannot optimize network performance criterion such as weighted sum throughput, max-min fairness etc. and hence are suboptimal.

B. Distributed Spatial Reuse Based on Maximal Independent Sets

An efficient solution to mitigate strong interference is spatial reuse, in which only a subset of UEs (which do not significantly interfere with each other) transmit at the same time. Spatial Time reuse based Time Division Multiple Access (STDMA) has been widely used in existing works on broadcast scheduling in multi-hop networks [4]–[6][6]. Specifically, these policies construct a cyclic schedule such that in each time slot an MIS of the interference graph is scheduled. The constructed schedule ensures that each UE is scheduled at least once in the cycle.

In terms of performance, STDMA policies [4]–[6] cannot guarantee the minimum throughput requirement of each UE, and usually adopt a fixed scheduling (i.e. follow a fixed order in which the MISs are scheduled), which may be very inefficient depending on the given network performance criteria. For example, the policies in [6] are inefficient in terms of fairness. In terms of complexity, for the distributed generation of the subsets of MISs, the STDMA policies in [4]–[6] require an ordering of all the UEs, and have a computational complexity (in terms of the number of steps executed by the algorithm) that scales as $\mathcal{O}(|V|)$ (in [5], [6]) or $\mathcal{O}(|V||E|)$

[5]In the case of average sum throughput maximization given the minimum average throughput constraints of the UEs, the power control policies are inefficient if the feasible rate region is non-convex [17] .

[6]These works [4]–[6] do not have the exactly same model as in our setting. However, these works can be adapted to our model. Hence, we also compare with these works to have a comprehensive literature review.

(in [4]), where $|V|$ and $|E|$ are the number of vertices/UEs and the number of edges in the interference graph, respectively. Hence, in large-scale dense deployments, the complexity grows superlinearly with the number of UEs, making the policies difficult to compute. By contrast, our proposed distributed algorithm for generating subsets of MISs does not require the ordering of all the UEs, and has a complexity that scales as $\mathcal{O}(\log |V|)$, namely sublinearly with the number of the UEs, for bounded-degree graphs.[7]

Finally, the STDMA policies in [4]–[6] are designed for the MAC layer and assume that all the UEs are homogeneous at the physical layer. In practice, different UEs are heterogeneous due to their different distances from their SBSs, their different maximum transmit power levels, etc. This heterogeneity is important, and will be considered in our design framework.

C. Distributed Power Control and Spatial Reuse For Multi-Cell Networks

As we have discussed, the works in the above two categories either focus on distributed power control in the physical layer [7], [8], [10] or focus on distributed spatial reuse in the MAC layer [4]–[6]. Similar to our paper, some works (representative references [20]–[24]) adopted a cross-layer approach and proposed distributed joint power control and spatial reuse for multi-cell networks. However, although these works schedule a subset of UEs to transmit at the same time, the subset is not the MIS of the interference graph [22], [23]. For example, the policies in [22], [23] schedule one UE from each small cell at the same time, even if some UEs are from small cells very close to each other. In this case, the UEs will experience strong inter-cell interference. Hence, the works in [22], [23] cannot perfectly eliminate strong interference from neighboring cells and exploit weak interference from non-neighboring cells. Moreover, the works in [20]–[24] cannot provide minimum throughput guarantees for the UEs.

III. SYSTEM MODEL

A. Heterogeneous Network of Small Cells

We consider a heterogeneous network of K small cells operating in the same frequency band[8] (see Fig. 1), which represents a common deployment scenario considered in practice

[7] As will be shown in Theorem 5, for graphs which are not bounded degree graphs, even a centralized solution based on all the MISs cannot satisfy the minimum throughput requirements.

[8] Our solutions will be based on spatial time reuse assuming every UE uses the same frequency. Our solutions can be extended to spatial frequency reuse, where we let different MISs operate in non-overlapping frequency bands.

[10] [13] [25]. Note that the small cells can be of different types (e.g. picocells, femtocells, etc.) and thereby belong to different tiers in the heterogeneous network. Each small cell j has one SBS, (SBS-j), which serves a set of UEs under a closed access scenario [10]. Denote the set of UEs by $\mathcal{U} = \{1, ..., N\}$. We write the association of UEs to SBSs as a mapping $T : \{1, ..., N\} \to \{1, .., K\}$, where each UE-$i$ is served by SBS-$T(i)$. We focus on the uplink transmissions; the extension to downlink transmissions is straightforward when each SBS serves one UE at a time (e.g. TDMA among UEs connected to the same SBS).

Each UE-i chooses its transmit power p_i from a compact set $\mathcal{P}_i \subseteq \mathbb{R}_+$. We assume that $0 \in \mathcal{P}_i,\ \forall i \in \{1, ..., N\}$, namely any UE can choose not to transmit. The joint power profile of all the UEs is denoted by $\mathbf{p} = (p_1, ..., p_N) \in \mathcal{P} \triangleq \Pi_{i=1}^{N} \mathcal{P}_i$. Under the joint power profile $\mathbf{p}$, the signal to interference and noise ratio (SINR) of UE-i's signal, experienced at its serving SBS-$j = T(i)$, can be calculated as $\gamma_i(\mathbf{p}) = \frac{g_{ij} p_i}{\sum\limits_{k=1, k \neq i}^{N} g_{kj} p_k + \sigma_j^2}$, where g_{ij} is the channel gain from UE-i to SBS-j, and σ_j^2 is the noise power at SBS j. The UEs do not cooperate to encode their signals to avoid interference, hence, each UE-SBS pair treats the interference from other UEs as white noise. Hence, each UE-i gets the following throughput [22], $r_i(\mathbf{p}) = \log_2(1 + \gamma_i(\mathbf{p}))$[9].

B. Interference Management Policies

The system is time slotted at t = 0,1,2..., and the UEs are assumed to be synchronized as in [22], [23] [26] [27].. At the beginning of each time slot t, each UE-i decides its transmit power p_i^t and obtains a throughput of $r_i(\mathbf{p}^t)$. Each UE i's strategy, denoted by $\pi_i : \mathbb{Z}_+ = \{0, 1, ..\} \to \mathcal{P}_i$, is a mapping from time t to a transmission power level $p_i \in \mathcal{P}_i$. The interference management policy is then the collection of all the UEs' strategies, denoted by $\boldsymbol{\pi} = (\pi_1, ..., \pi_N)$. The average throughput for UE i is given as $R_i(\boldsymbol{\pi}) = \lim_{T \to \infty} \frac{1}{T+1} \sum\limits_{t=0}^{T} r_i(\mathbf{p}^t)$, where $\mathbf{p}^t = (\pi_1(t), ..., \pi_N(t))$ is the power profile at time t. We assume the channel gain to be fixed over the considered time horizon as in [22] [28]–[31]. However, we will illustrate in Section VI that our framework performs well under dynamic channel conditions (due to fading, time varying channel) as well.

An interference management policy $\boldsymbol{\pi}^{const}$ is a policy based on power control [7], [8], [10] if $\boldsymbol{\pi}^{const}(t) = \mathbf{p}$ for all t. As we have discussed before, our proposed policy is based on MISs

[9]We use the Shannon capacity here. However, our analysis is general and applies to the throughput models that consider the modulation scheme used.

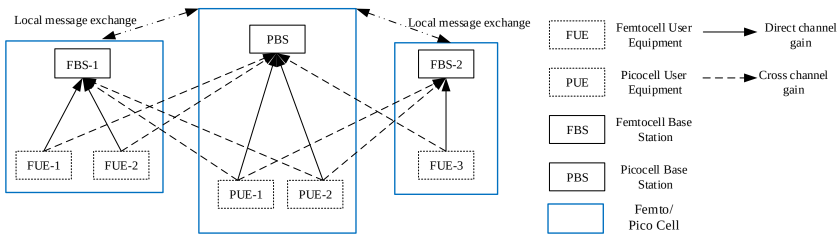

Figure 1. Illustration of a heterogeneous small cell network.

of the interference graph. The interference graph G has N vertices, each of which is one of the N UE-SBS pairs. There is an edge between two pairs/vertices if their cross interference is high (rules for deciding if interference is high will be discussed in Section V) and let there be M edges in the graph. Given an interference graph, we write $\mathbf{I} = \{I_1, ..., I_{N_{MIS}}\}$ as the set of all the MISs of the interference graph. Let $\mathbf{p}^{I_j}$ be a power profile in which the UEs in the MIS I_j transmit at their maximum power levels and the other UEs do not transmit, namely $p_k = p_k^{max} \triangleq \max \mathcal{P}_k$ if $k \in I_j$ and $p_k = 0$ otherwise. Let $\mathcal{P}^{MIS} = \{\mathbf{p}^{I_1}, ..., \mathbf{p}^{I_{N_{MIS}}}\}$ be the set of all such power profiles. Then $\boldsymbol{\pi}$ is a policy based on MIS if $\boldsymbol{\pi}(t) \in \mathcal{P}^{MIS}$ for all t. We denote the set of policies based on MISs by $\Pi^{MIS} = \{\boldsymbol{\pi} : \mathbb{Z}_+ \to \mathcal{P}^{MIS}\}$.

IV. Problem Formulation and a Motivating Example

In this section, we formulate the interference management policy design problem and give a motivating example to highlight the advantages of the proposed policy over existing policies.

A. The Interference Management Policy Design Problem

We aim to optimize a chosen network performance criterion $W(R_1(\boldsymbol{\pi}), ..., R_N(\boldsymbol{\pi}))$, defined as a function of the UEs' average throughput. We can choose any performance criterion that is concave in $R_1(\boldsymbol{\pi}), ..., R_N(\boldsymbol{\pi})$. For instance, W can be the weighted sum of all the UEs' throughput, i.e.$\sum_{i=1}^{N} w_i R_i(\boldsymbol{\pi})$ with $\sum_{i=1}^{N} w_i = 1$ and $w_i \geq 0$. Alternatively, the network performance can be max-min fairness (i.e. the worst UE's throughput) and hence W can be defined as $\min_i R_i(\boldsymbol{\pi})$. The policy design problem can be then formalized as follows:

$$\textbf{Policy Design Problem (PDP)} \qquad \max_{\boldsymbol{\pi}} \quad W(R_1(\boldsymbol{\pi}), ..., R_N(\boldsymbol{\pi}))$$
$$\text{subject to} \quad R_i(\boldsymbol{\pi}) \geq R_i^{min}, \forall i \in \{1, ..., N\}$$

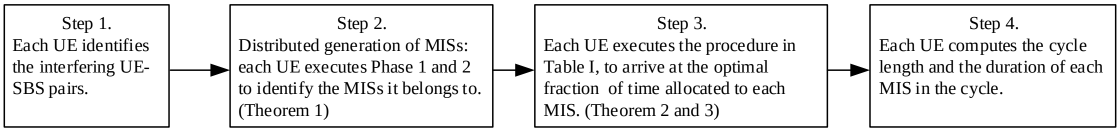

Figure 2. Steps in the Design Framework.

The above design problem is very challenging to solve even in a centralized manner (it has been shown to be NP-hard [32] even when we restrict to policies based on power control $\boldsymbol{\pi}^{const}$). Denote the optimal value of the PDP as $\mathbf{W}_{\text{opt}}$. Our goal is to develop distributed, polynomial-time algorithms to construct policies that achieve a constant competitive ratio with respect to $\mathbf{W}_{\text{opt}}$, with the competitive ratio independent of the network size. We achieve our goal by focusing on policies based on MISs Π^{MIS}, among other innovations that will be described in Section V. Next, we provide a motivating example to demonstrate the efficiency of our proposed policy.

V. Design Framework for Distributed Interference Management

A. *Proposed Design Framework*

Our proposed design framework (see Fig. 2) consists of the following four steps.

Step 1. Identification of the interfering neighbors: In Step 1, each UE-SBS pair identifies the UE-SBS pairs that strongly interfere with it. Essentially, each pair obtains a *local* view (i.e., its neighbors) of the interference graph. Note that an edge exists between two pairs if at least one of them identifies the other as a strong interferer.

Specifically, each UE-SBS pair is first informed of other pairs in the geographical proximity by managing servers (e.g., femtocell controllers/gateways) [33] [34] [29] [30]. Then each pair can decide whether another pair is strongly interfering based on various rules, such as rules based on Received Signal Strength (RSS) in the *Physical Interference Model* [33] [29] [30], and rules based on the locations in the *Protocol Model* [28]. If one pair identifies another pair as strongly interfering, its decision can be relayed by the managing servers to the latter, such that any two pairs can reach consensus of whether there exists an edge between them.

Step 2. Distributed generation of MISs that span all the UEs: In Step 2, the UE-SBS pairs generate a subset of MISs in a distributed fashion. It is important that the generated subset

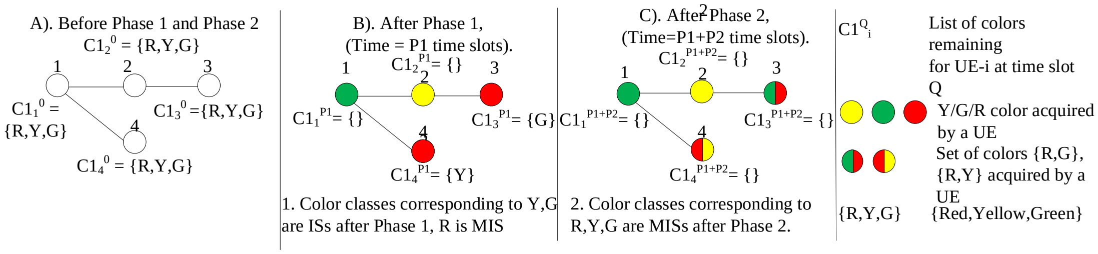

Figure 3. Illustration of the distributed generation of MISs in Step 2.

spans all the UEs, namely every UE is contained in at least one MIS in the subset. Otherwise, some UEs will never be scheduled.

The key idea is that from a given list of colors, each UE has to choose a set of colors such that the choice does not conflict with its neighbors. We should ensure that each UE has at least one color. We call the set of UEs with the same color "a color class". In addition, we should also ensure that every color class is a MIS. This step is composed of two phases: first, distributed coloring of the interference graph based on [35], and second, extension of color classes to MISs. All the UEs are synchronized and carry out their computation simultaneously. We now explain the algorithm in detail. The pseudo-codes can be found in Table II and III in the Appendix.

Phase 1. Distributed coloring of the interference graph: Let H[10] be the maximum number of colors given to all SBSs at the installation and d_i be the degree (number of neighbors in the interference graph) of the i^{th} pair. The goal of this phase is to let each UE-SBS pair i choose *one* color from $C_i^0 \triangleq \{1, ...H\} \cap \{1, .., d_i + 1\}$, such that no neighbors choose the same color. The distributed coloring works as follows.

i) At the beginning of each time slot t, each UE i chooses a color from the set of remaining colors C_i^t uniformly randomly, and informs its neighbors of its tentative choice. This information can be transmitted through the back-haul network/X2 interface that is used for ICIC [34].

ii) If the tentative choice of a UE does not conflict with any of its neighbor, then it fixes its color choice and informs the neighbors of its choice. This UE does not contend for colors any

[10]The maximum number of colors H should be set to be larger than the maximum number of UE-SBS pairs interfering with any UE-SBS pair. The SBSs can determine H according to the deployment scenario. H in general will also include the number of UEs that use the same SBS who interfere with each other along with the other neighboring UEs. For example, H can be 10-15 in an office building with dense deployment of SBSs, and can be 3-5 in a residential area.

further in Phase 1. The neighbors delete the color chosen by i from their lists $C_j^{t+1}, \forall j \in \mathcal{N}(i)$, where $\mathcal{N}(i)$ is the set of i's neighbors.

iii) Otherwise, if there is a conflict, then the UE does not choose that color and repeats i) and ii) in the next time slot.

There are $\lceil c_1 \log_{\frac{4}{3}} N \rceil + 1$ time slots in Phase 1, where c_1 is the parameter given by the protocol. The number of time slots is known to the SBSs at installation. Phase 1 is successful if all the UEs acquire a color, which implies that the set of color classes (i.e., the set of UE-SBS pairs with the same color) spans all the UEs.

Phase 2. Extending color classes to the MISs: Each color class obtained at the end of Phase 1 is an independent set (IS) of the graph. In Phase 2, we extend each of these ISs to MISs and possibly generate additional MISs. After Phase 1, each UE has chosen one color and deleted some colors from its list. But there may still be remaining colors in its list that are not acquired by any of its neighbors. If the UEs can acquire these remaining colors without conflicting with its neighbors, then each color class will be a MIS. Phase 2 works as follows.

i) At each time slot in Phase 2, UE i chooses each color from the remaining colors in its list independently with probability c. Each UE i then sends the set of its tentative choices to its neighboring UEs, and receives their neighbors' choices.

ii) For any tentative choice of color, if there is a conflict with at least one neighbor, then that color is not fixed; otherwise, it is fixed.

iii) At the end of each time slot, each UE deletes its set of fixed colors from its list, and transmits this set of fixed colors to its neighbors, who will delete these fixed colors from their lists as well. Note that a UE deletes a particular color if and only if the UE itself or some of its neighbors have chosen this color. Based on this key observation, we can see that if a color is not in any UE's list, the set of UEs with this color is a MIS. If all the UEs have an empty list, then for any color in the set $\{1, ..., H\}$, the set of UEs with this color is a MIS.

There are $\lceil c_2 \log_x N \rceil + 1$ time slots in Phase 2, where $x = \frac{1}{1-(c)^H(1-c)^{H^2}}$, and c_2 is the parameter given by the protocol. The number of time slots is known to the SBSs at installation. We say that Phase 2 is successful, if it finds H MISs, or equivalently if all the UEs have an empty list.

Example: We illustrate Step 2 in a network of 4 UE-SBS pairs, whose interference graph is shown in Fig. 3. At the start, each UE-SBS pair has a list of 3 colors $\{$Red, Yellow, Green$\}$.

Phase 1 is run for $P1 = \lceil c_1 \log_{\frac{4}{3}} 5 \rceil$ time slots. At the end of Phase 1, UE 1 and UE 2 acquire Green and Yellow respectively, while UEs 3-4 acquire Red. Hence, UE 1 (UE 2) has an empty list, as Green (Yellow) is acquired by itself and Red, Yellow (Green) by its neighbors. UE 3 (UE 4) has Green (Yellow) color in its list of remaining colors. At the end of Phase 1, the Red color class is a MIS, while the Yellow and Green color classes are not. Phase 2 is run for $P2 = \lceil c_2 \log_x 5 \rceil + 1$ time slots. UE 3 (UE 4) acquires the remaining color Green (Yellow). At the end of Phase 2, the Green and Yellow color classes become MISs too.

The next theorem establishes the high success probability of Step 2.

Theorem 1. For any interference graph with the maximum degree $\Delta \leq H - 1$, the proposed algorithm in Table II and III outputs a set of H MISs that span all the UEs in $(\lceil c_1 \log_{\frac{4}{3}} N \rceil + \lceil c_2 \log_x N \rceil + 2)$ time slots with a probability no smaller than $(1 - \frac{1}{N^{c_1 - 1}})(1 - \frac{1}{N^{c_2 - 1}})$, where c_1 and c_2 are design parameters that trade-off the run time and the success probability.

See the Appendix for detailed proofs.

Theorem 1 characterizes the performance of our proposed algorithm, in terms of the run time of the algorithm and the lower bound of the success probability. When the parameters c_1 and c_2 are larger, the lower bound of the success probability increases at the expense of a longer run time. When the maximum degree of the interference graph is larger, we need to set a higher H, which results in a longer run time. This is reasonable, because it is harder to find coloring and MISs when the number of interfering neighbors is higher. Finally, we can see that the lower bound of the successful probability is very high even under smaller c_1 and c_2, especially if the number of UEs is large. Note that the exact successful probability should depend on the probability c in Phase 2, while the lower bound in Theorem 1 does not. Hence, our lower bound is robust to different system parameters. Note also that the interference graph here is a bounded-degree graph since the maximum degree is bounded by a given constant, $H - 1$. The algorithms in [4] [6] (require ordering of the vertices, work sequentially and have a higher complexity) can be used to output the MISs spanning all the UEs for arbitrary graphs. However, we will show in Theorem 5, that the restriction to bounded-degree graphs is a must to ensure that the minimum throughput requirement of each UE is satisfied for any MIS based policy.

Step 3. Distributed computation of the optimal fractions of time for each MIS: Let the set of MISs generated in Step 2 be $\{I_1^{'}, ..., I_H^{'}\}$. In Step 3, the UE-SBS pairs compute the fractions of time allocated to each MIS in a distributed manner.

When an MIS is scheduled, the UEs in this MIS transmit at their maximum power levels, and the other UEs do not transmit. Define R_i^k as the *instantaneous* throughput obtained by UE i in the MIS $I_k^{'}$, which can be calculated as $\log_2(1+\frac{g_{iT(i)}p_i^{I_k^{'}}}{\sum_{r=1,r\neq i}^{N} g_{rT(i)}p_r^{I_k^{'}}+\sigma_{T(i)}^2})$, where $p_i^{I_k^{'}}=p_i^{max}$ if $i\in I_k^{'}$ and $p_i^{I_k^{'}}=0$ otherwise. To determine R_i^k, the UE needs to know the total interference it experiences when transmitting in $I_k^{'}$. This can be measured by having an initial cycle of transmissions of UEs in each MIS in the order of the indices of MISs/colors.

From now on, we assume that the network performance criterion $W(\boldsymbol{y})$ is concave in $\mathbf{y}$ and is separable, namely $W(y_1,...y_N)=\sum_{i=1}^{N} W_i(y_i)$. Examples of separable criteria include weighted sum throughput and proportional fairness. Our framework can also deal with max-min fairness$\min_i R_i(\boldsymbol{\pi})$, although it is not separable (see the discussion in the Appendix) The problem of computing the optimal fractions of time for the MISs is given as follows:

$$\textbf{Coupled Problem (CP)} \qquad \max_\alpha \quad \sum_{i=1}^{N} W_i\left(\sum_{k=1}^{H}\alpha^k R_i^k\right)$$

$$\text{subject to} \quad \sum_{k=1}^{H}\alpha^k R_i^k \geq R_i^{min},\ \forall i\in\{1,..,N\}$$

$$\sum_{k=1}^{H}\alpha^k=1, \alpha^k\geq 0,\ \forall k\in\{1,..,H\}$$

Each UE i knows only its own utility function W_i and minimum throughput requirement R_i^{min}. Hence, it cannot solve the above problem by itself. We will first reformulate the above problem into a decoupled problem and then show that the reformulated problem can be solved in a distributed manner. Let each UE i have a local estimate β_i^k of the fractions of time allocated to each MIS $I_k^{'}$ (including those MISs that UE i does not belong to). We impose an additional constraint that all the UEs' local estimates are the same. Note that this constraint will be satisfied by our solution, and is not an assumption. Such a constraint is still global, because any two UEs, even if they are not neighbors, need to have the same local estimate. Hence, global message exchange among any pair of UEs is still needed to solve this problem with local estimates and global constraints[11]. To avoid global message exchange, we reformulate the CP into a decoupled

[11]If the UEs could exchange messages globally, i.e. broadcast messages to all the UEs in the network, and if the network performance criterion is strictly concave, we could use standard dual decomposition with augmented Lagrangian in [36] to derive a distributed algorithm. However, in large networks, the UEs cannot exchange messages globally with other UEs, and the network performance criterion may not be strictly concave (e.g., the weighted sum throughput is linear).

problem (DP) that involves only local coupling among the neighbors and can be solved with local message exchange using Alternating Direction Method of Multipliers (ADMM) [37].

Now we reformulate the CP into a decoupled problem (DP) that involves only local coupling among the neighbors and that can be solved by Alternating Direction Method of Multipliers (ADMM) [37]. If UE i and l are connected by an edge (i,l) then for each set $I_k^{'}$ define $\theta_{(i,l)i}^k = \beta_i^k$ and $\theta_{(i,l)l}^k = -\beta_l^k$, note that these auxiliary variables are introduced to formulate the problem into the ADMM framework [37]. Define a polyhedron for each i, $\mathcal{T}_i = \{\boldsymbol{\beta_i} | \text{s.t. } \mathbf{1^t}\boldsymbol{\beta_i} = \mathbf{1}, \boldsymbol{\beta_i} \geq \mathbf{0},\, \boldsymbol{R_i^{'}}\boldsymbol{\beta_i} \geq R_i^{min}\}$, here $\boldsymbol{\beta_i} = (\beta_i^1, ..., \beta_i^H)$ and $\boldsymbol{R_i} = (R_i^1, ..., R_i^H)$ and $()^{'}$corresponds to the transpose. Let $\boldsymbol{\beta} = \boldsymbol{(\beta_1, ..., \beta_N)} \in \boldsymbol{\mathcal{T}}$, where $\boldsymbol{\mathcal{T}} = \prod_{i=1}^N \mathcal{T}_i$ and $\prod$ corresponds to the Cartesian product of the sets. Also, let $\boldsymbol{\beta^k} = (\beta_1^k, ..., \beta_N^k)$, $\forall k \in \{1, .., H\}$. Define another polyhedron $\Theta_{(i,l)}^k = \{(\theta_{(i,l)i}^k, \theta_{(i,l)l}^k) : \theta_{(i,l)i}^k + \theta_{(i,l)l}^k = 0,\, -1 \leq \theta_{(i,l)s}^k \leq 1, \forall s \in \{i,l\}\}$, $\boldsymbol{\Theta^k} = \prod_{(i,l)\in E} \Theta_{(i,l)}^k$ here $E = (e_1, ..e_M)$ is the set of all the M edges in the interference graph. A vector $\boldsymbol{\theta^k} \in \boldsymbol{\Theta^k}$ is written as $\boldsymbol{\theta^k} = (\theta_{e_1,z(e_1)}^k, \theta_{e_1,t(e_1)}^k, .., \theta_{e_M,z(e_M)}^k, \theta_{e_M,t(e_M)}^k)$, here $z(e_i)$, $t(e_i)$ correspond to the vertices in the edge, e_i. Similarly define, $\boldsymbol{\theta} = \boldsymbol{(\theta^1, ..., \theta^H)} \in \boldsymbol{\Theta}$, where $\boldsymbol{\Theta} = \prod_{k=1}^H \boldsymbol{\Theta^k}$.

Decoupled Problem (DP) $\quad \min_{\boldsymbol{\beta}\in\boldsymbol{\mathcal{T}}, \boldsymbol{\theta}\in\boldsymbol{\Theta}} -\sum_{i=1}^N \boldsymbol{W_i(R_i}^{'}\boldsymbol{\beta}_i)$

subject to $\boldsymbol{D^k\beta^k - \theta^k = 0}$, $\forall k \in \{1, .., H\}$

Here, $\boldsymbol{D^k} \in \mathbb{R}^{2M\times N}$, is a matrix in which each row has exactly one non-zero element which is 1 or -1. Each element of the matrix, D_{vj}^k is evaluated as follows, the index v can be uniquely expressed in terms of quotient q and the remainder w as $v = 2q + w$, and if $j \neq z(e_{q+1})$, $j \neq t(e_{q+1})$ then $D_{vj}^k = 0$. If $w = 1$, $j = z(e_{q+1})$, then $D_{vj}^k = 1$ else if $w = 0$, $j = z(e_{q+1})$ then $D_{vj}^k = 0$. Also, if $w = 0$, $j = t(e_{q+1})$, then $D_{vj}^k = -1$ else if $w = 1$, $j = t(e_{q+1})$ then $D_{vj}^k = 0$.

Theorem 2: For any connected interference graph, the coupled problem (CP) is equivalent to the decoupled problem (DP).

The above theorem shows that the original problem (CP), which requires global information and global message exchange to solve, is transformed into an equivalent problem (DP), which as we will show, can be solved in a distributed manner with local message exchange

We denote the optimal solution to the DP by $\mathbf{W}_{\text{distributed}}^G$. We associate with each constraint $D_{eq}^k\beta_q^k = \theta_{eq}^k$ a dual variable λ_{eq}^k. The augmented Lagrangian for DP is $L_y\left(\{\boldsymbol{\beta}_i\}_i, \{\theta_{eq}^k\}_{k,e,q}, \{\lambda_{eq}^k\}_{k,e,q}\right) = -\sum_{i=1}^N W_i(\boldsymbol{\beta}_i^T\boldsymbol{R}_i) + \sum_{k=1}^H \sum_{e\in E} \sum_{q\in e} \left[\lambda_{eq}^k\left(D_{eq}^k\beta_q^k - \theta_{eq}^k\right) + \frac{y}{2}\left(D_{eq}^k\beta_q^k - \theta_{eq}^k\right)^2\right]$. In the ADMM procedure (see Table IV in the Appendix), each UE i solves for its optimal local estimates $\boldsymbol{\beta}_i(t)$

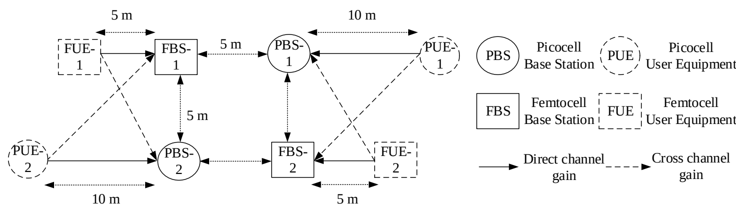

Figure 4. A heterogeneous network of 2 PBS and 2 FBS and their corresponding UEs.

that maximizes the augmented Lagrangian given the previous dual variables $\lambda_{ei}^k(t-1)$ and auxiliary variables $\theta_{ei}^k(t-1)$. Then it updates its dual variable $\lambda_{ei}^k(t)$ and auxiliary variable $\theta_{ei}^k(t)$ based on its local estimate $\beta_i^k(t)$ and its neighbor j's local estimate $\beta_j^k(t)$. This iteration of updating local estimates, dual variables, and auxiliary variables is repeated P times. Next, it is shown that this procedure will indeed converge.

Theorem 3: If DP is feasible[12], then the ADMM algorithm in Table IV converges to the optimal value $\mathbf{W}^G_{\text{distributed}}$ with a rate of convergence $\mathcal{O}(\frac{1}{P})$.

Step 4. Determining the cycle length and transmission times: At the end of Step 3, all the UEs have a consensus about the optimal fractions of time allocated to each MIS, namely $\boldsymbol{\beta}_i^* = \boldsymbol{\gamma}^* = (\gamma_1^*, ..., \gamma_H^*)$, $\forall i \in \{1, .., N\}$. The MISs transmit in the order of their indices (i.e., $\{1, .., H\}$) in cycles. In each cycle of transmission, MIS $I_k^{'}$ transmits for $\left\lceil \frac{\gamma_k^*}{\min_{i \in 1,...,N} \gamma_i^*} \times 10^d \right\rceil$ slots, where we multiply by 10^d such that the rounding error is reduced or eliminated in case that $\frac{\gamma_k^*}{\min_{i \in 1,...,N} \gamma_i^*}$ is not an integer.

B. A Motivation Example

Consider a network of 2 picocell base stations (PBS) and 2 femtocell base stations (FBS), each serving one UE. The network topology is shown in Fig. 2. We assume a path loss model for channel gains, with path loss exponent 4. The maximum transmit power of each UE is 80 mW, and the noise power at each SBS is 1.6×10^{-3} mW. UEs in different tiers have different minimum throughput requirements: FUE (femtocell UE) 1 and FUE 2 in the femtocells require a minimum throughput 0.4 bits/s/Hz, and PUE (picocell UE) 1 and PUE 2 in the picocells require 0.2 bits/s/Hz. The interference graph is constructed according to a distance based threshold rule similar to [28]. Specifically, an edge exists between two UE-BS pairs if the distance between

[12]DP is feasible, if the feasible region resulting from the constraints in DP is non-empty.

any pair of SBSs is less than a threshold, which is set to be 1.2m here. There are two MISs. MIS 1 consists of FUE 1 and FUE 2, and MIS 2 consists of PUE 1 and PUE 2. We consider two performance criteria: the max-min fairness and the sum throughput. We will compare with the following state-or-the-art policies:

1. Distributed Constant Power Control Policies [7], [8], [10]**:** In these policies, all the UEs choose *constant* power levels determined by distributed algorithms utilizing information (e.g., power levels used by neighbors) made available through local/global message exchange.

2. Optimal Centralized Constant Power Policies: In these policies, all the UEs choose *constant* power levels determined by a central controller utilizing global information.

3. Distributed MIS STDMA-1 [6] and **STDMA**-2 [4]**:** These policies construct a subset of the MISs of the interference graph in a distributed manner and propose fixed schedules of the MISs. Different works adopt different schedules, and we differentiate them by referring to them as MIS STDMA-1 [6] and STDMA-2 [4].

4. Distributed Joint Power Control and Spatial Reuse [22] [23]**:** These policies choose one UE from each cell to form a subset, and schedule these subsets of UEs based on their channel gains to maximize the sum throughput. The policies are named power matched scheduling (PMS).

In Table 1, we compare the performance of our proposed policy with state-of-the-art policies for the same setup as in Fig. 1. We compute the optimal centralized constant power control policy by exhaustive search, which serves as the performance upper bound of the distributed constant power control policies [7], [8], [10] centralized constant power control policies [15]. In PMS policies [22] [23], UEs within the same cell are scheduled in a time-division multiple access (TDMA) fashion, and the active UEs in different cells transmit simultaneously. In this motivating example, there is one UE in each cell, which will be scheduled to transmit all the time. Therefore, the PMS policy reduces to a constant power control policy, and is worse than the optimal centralized constant power control policy. We can see that our proposed policy outperforms all constant power control policies and distributed PMS policies by at least 375% and 32.8%, in terms of max-min fairness and sum throughput, respectively. The significant performance improvement over the constant power control policies results from the elimination of the high interference among the users through scheduling MISs. Our proposed policy also outperforms distributed STDMA policies by 30%-40%. As we will see in Section VI, the performance gain is even higher (160%-700%) in realistic deployment scenarios. Finally, in this motivating example,

Table I

COMPARISONS IN TERMS OF MAX-MIN FAIRNESS & SUM THROUGHPUT CRITERION

Policies	Max-min throughput (bits/s/Hz)	Performance Gain %	Sum throughput (bits/s/Hz)	Performance Gain %
Distributed constant power control [7], [8], [10]	<0.28	>375 %	6.1	32.8 %
Distributed PMS [22], [23]	<0.28	>375%	6.1	32.8 %
Optimal centralized constant power control	0.28	375%	6.1	32.8 %
Distributed MIS STDMA-2/1 [4], [6]	0.96	38.5%	6.25	30.0 %
Proposed (Section-V)	1.33	-	8.12	-
Benchmark Problem (BP) (Section- VI)	1.33	-	8.12	-

the proposed policy achieves the optimal performance of the benchmark problem defined in Section VI, which is a close approximation of the original problem (CP).

C. Performance Guarantees for Large Networks and Properties of Interference Graphs

In this subsection, we provide performance guarantees for our proposed framework described in Section V-A. Specifically, we prove that the network performance $\mathbf{W}^{G}_{\text{distributed}}$ achieved by the proposed distributed algorithms has a constant competitive ratio with respect to the optimal value $\mathbf{W}_{\text{opt}}$ of the PDP. Moreover, we prove that the competitive ratio does not depend on the network size. Our result is strong, because the solution to PDP needs to be computed by a centralized controller with global information and with NP complexity, while our proposed framework allows the UEs to compute the policy fast in a distributed manner with local information and local message exchange.

Before characterizing the competitive ratio analytically, we define some auxiliary variables. Define the upper and lower bounds on the UEs' maximum transmit power levels and throughput requirements as, $0 < p_{lb}^{max} \leq p_i^{max} \leq p_{ub}^{max}, \forall i \in \{1, ..., N\}$ and, $0 < R_{lb}^{min} \leq R_i^{min} \leq R_{ub}^{min}, \forall i \in \{1, ..., N\}$ respectively. Let D_{ij} is the distance between UE i and SBS j. Define upper and lower bounds on the distance between any UE and its serving SBS and the noise power at the SBSs as, $0 < D^{lb} \leq D_{iT(i)} \leq D^{ub}, \forall i \in \{1, ..., N\}$ and, $\sigma_{lb}^2 \leq \sigma_j^2 \leq \sigma_{ub}^2, \forall j \in \{1, ..., K\}$ respectively. We assume that the channel gain is $g_{ij} = \frac{1}{(D_{ij})^{np}}$, where np is the path loss exponent.

Definition 1 (Weak Non-neighboring Interference): The interference graph G exhibits ζ Weak Non-neighboring Interference (ζ-WNI) if for each UE i the maximum interference from its non-neighbors is bounded, namely $\sum_{j\notin\mathcal{N}(i),j\neq i} g_{jT(i)}p_j^{max} \leq (2^{\zeta}-1)\sigma_{ub}^2$, $\forall i \in \{1,...,N\}$.

Define $\Delta^{max} = \frac{\log_2(1+\frac{p_{lb}^{max}}{(D^{ub})^{np}2^{\zeta}\sigma_{ub}^2})}{R_{ub}^{min}} - 1$. Then we have the following theorem for the network performance criterion, sum throughput[13].

Theorem 4: For any connected interference graph, if the maximum degree $\Delta \leq \Delta^{max}$ and it exhibits ζ-WNI then, our proposed framework of interference management described in Section V-A achieves a performance $\mathbf{W}_{\text{distributed}}^G \geq \Gamma \cdot \mathbf{W}_{\text{opt}}$ with a probability no smaller than $(1-\frac{1}{N^{c_1-1}})(1-\frac{1}{N^{c_2-1}})$. Moreover, the competitive ratio $\Gamma = \frac{R_{ub}^{min}}{\log_2(1+\frac{p_{ub}^{max}}{(D^{lb})^{np}\sigma_{lb}^2})}$ is independent of the network size.

Note that the analytical expression of competitive ratio, $\Gamma = \frac{R_{ub}^{min}}{\log_2(1+\frac{p_{ub}^{max}}{(D^{lb})^{np}\sigma_{lb}^2})}$, does not depend on the size of the network. Our results are derived under the conditions that the interference graph has a maximum degree bounded by Δ^{max}, and that the interference from non-neighbors is bounded (i.e. $\zeta-$WNI). These conditions do not restrict the size of the network, next example illustrates this. In addition, our results hold for any interference graph that satisfy the conditions in Theorem 4, regardless of how the graph is constructed.

Example: Consider a layout of SBSs in a $K \times K$ square grid, i.e. K^2 SBSs with a distance of 5m between the nearest SBSs. Assume that each UE is located vertically below its SBS at a distance of 1 m. Fix the parameters $p_i^{max} = 100$ mW, $\sigma_i^2 = 3$ mW, $R_i^{min} = 0.1$bits/s/Hz, $\forall i \in \{1,..,K^2\}$, $np = 4$. We construct the interference graph based on the distance rule [28], namely there is an edge between two pairs if the distance between their SBSs exceeds 6m, which gives us the maximum degree $\Delta = 4$. We can also verify that the interference graphs under any number K^2 of SBSs exhibit ζ-WNI with $\zeta = 0.15$ and $\Delta < \Delta^{max}$, where $\Delta^{max} = 48$. Given $\Delta = 4$ and $\zeta = 0.15$, from Theorem 4, we get the performance guarantee of 0.17 for any network size K^2. Note that the number 0.17 is a performance guarantee, and that the actual performance is much higher compared to the performance guarantee as well as those achieved by state-of-the-art policies (see Section VI).

Both Theorem 1 and 4 required the maximum degree of the interference graph to be bounded by a given constant. Here, we show that constraint on the degree is natural and is a must to ensure feasibility, i.e. to satisfy the minimum throughput requirements of every UE. Specifically, we

[13]We can extend this result for weighted sum throughput, with weights $w_i = \Theta(\frac{1}{N})$, it is not done to avoid complex notations.

prove that if the maximum degree exceeds some threshold, then no policy based on scheduling MISs in Π^{MIS} (a large space of policies, see Section III) is feasible. Let the construction of interference graph be based on a distance based threshold rule similar to [28]. An edge exists between two UE-SBS pairs if and only if, the distance between two SBS is no greater than D^{th}. We define the threshold of the maximum degree as Δ^* (See the Appendix for the expression).

Theorem 5: If the maximum degree of the interference graph $\Delta \geq \Delta^*$, then any policy based on scheduling MISs in Π^{MIS} fails to satisfy the minimum throughput requirements of the UEs.

The intuition behind Theorem 5 is that, if the degree of the interference graph is large then there must be a large number of UE-SBS pairs which interfere with each other strongly (mutually connected) which makes it impossible to allocate each UE enough transmission time to satisfy its minimum throughput requirement.

D. Self-Adjusting Mechanism for Dynamic Entry/Exit of UEs

We now describe how the proposed framework can adjust to dynamic entry/exit by the UEs in the network without recomputing all the four steps. We allow the UEs to enter and exit, but number of SBSs is fixed. We only allow let one UE enter or leave the network in any time slot.

1. *UE leaves the network:* Suppose a UE i which was transmitting to SBS $T(i)$ leaves the network. If the UE i was transmitting in a set of colors C_i, then as soon as it leaves, these colors can be potentially used by some neighbors, $\mathcal{N}(i)$. The SBS $T(i)$ which was serving the UE i can have other UEs which are still in the network and transmitting to it. Then for each color $c' \in C_i$ it first searches among the UEs which it serves that are not already transmitting in c' and who also do not have a neighboring UE-SBS pair which is already transmitting in c'. Let the set of such UEs be $UE^{c'}_{i,left}$. SBS $T(i)$ allocates color c' to the UE whose index is arg $\max_{j \in UE^{c'}_{i,left}} R^{c'}_j$. In case $UE^{c'}_{i,left}$ is empty then that color, $c^{'}$ is left unused.

2. *UE enters the network:* Suppose a UE i registered with SBS $T(i)$ enters the network.

i). Given the minimum throughput requirement of the UE i the SBS $T(i)$ first creates a list of UEs, $UE_{i,enter}$, which consists of the UEs it is serving and who are transmitting at more than their minimum throughput requirement.

ii). SBS $T(i)$ creates the list of colors, $C_{i,enter}$ in which UEs in $UE_{i,enter}$ are transmitting, it also consists of the colors that are not being used by any UE served by $T(i)$. Next, it creates valid colors list i.e. $C^{valid}_{i,enter}$ from $C_{i,enter}$, where a color $c \in C^{valid}_{i,enter}$ if $c \in C_{i,enter}$ and if none of the neighbors of i in $\mathcal{N}(i)$ that are not in $UE_{i,enter}$ are already using that color.

iii). Next, the SBS $T(i)$ has to allocate some portions from the fractions of time allocated to the colors in $C_{i,enter}^{valid}$, such that UE-i can transmit and its minimum throughput requirement is satisfied to the best possible extent. The allocation is done as follows, let $C_{i,enter}^{valid} = \{c_1^{'}, ..., c_s^{'}\}$. Proceeding sequentially, for each color $c_i^{'}$, SBS $T(i)$ selects the maximum possible portion to satisfy the minimum throughput requirement of UE-i, such that the minimum throughput requirements of UEs in $UE_{i,enter}$, who are using this color, $c_i^{'}$ are not violated.

iv). If the requirement of UE-i is not satisfied then, SBS $T(i)$ requests the neighboring UE-SBSs (apart from the UEs that are served by $T(i)$) to announce the set of colors which are either not being used or in which their corresponding UEs are operating at more than the minimum throughput requirement. From the set of colors that are received, the SBS-$T(i)$ chooses each color from the list if it is not being used by any other neighboring UE apart from the ones who sent the announcement. The resulting list of colors is $Cl_{i,enter}^{valid} = \{c_1^{'}, ..., c_l^{'}\}$.

v). Proceeding sequentially with the colors in $Cl_{i,enter}^{valid}$, for each color, SBS-$T(i)$ requests a portion from the fraction of time allocated to that color, to the neighboring UE-SBSs allocated that color, such that the throughput requirement of UE-i is satisfied. The neighboring UE-SBSs either allow the requested portion or send the portion which is acceptable to them, i.e. their throughput requirements are not violated. SBS-$T(i)$ allocates the minimum acceptable portion to UE-i and proceeds to the next color in the list if the throughput requirements are not satisfied.

E. Extensions

In our model, UEs operate in the same frequency band. However, our methodology can be extended to scenarios where UEs operate in different frequency channels (frequency reuse) and transmit at the same time. In this case, the problem is to find the optimal frequency allocation with the same objective function and constraints as in PDP. To solve this problem, the first two steps of the framework remain the same. In Step 3, the UEs compute distributedly the optimal fractions of *bandwidth* to be allocated to each MIS. This step is equivalent to computing the optimal fraction of time allocated to each MIS as in our current formulation. In Step 4, the UEs compute the number of frequency channels allocated to each MIS based on the bandwidth allocation.

Note that we do not implement beamforming, although beamforming can be used in conjunction with our policy. If the UEs transmitting to the same SBS cooperate to do beamforming, we can delete the edge between them in the interference graph, and use the new interference graph in the scenario with beamforming.

VI. ILLUSTRATIVE RESULTS

In this section, we evaluate our proposed policy under a variety of scenarios with different levels of interference, large numbers of UEs, different performance criteria, time-varying channel conditions, and dynamic entry and exit of UEs.

We compare our policy with the optimal centralized constant power control policy, the distributed MIS STDMA-1 [6] and STDMA-2 [4], distributed PMS [22] [23], in terms of sum throughput and max-min fairness. We do not separately compare with distributed/centralized constant power control policies in [7], [8], [10] [15], because their performance is upper bounded by the optimal centralized power control. Since it is difficult to compute the solution to the NP-hard PDP, we define a benchmark problem, where we restrict our search to policies in which a UE either transmits at its maximum power level or does not transmit.The space of such policies can be writtenas $\Pi_{BC} = \{\boldsymbol{\pi} = (\pi_1, ..., \pi_N) : \pi_i : \mathbb{Z}_+ \rightarrow \{0, p_i^{max}\} \, \forall i \in \{1, .., N\}\}$. The policy space Π_{BC} is a subset of all policies Π and is a superset of MIS based policies Π^{MIS}. In other words, the benchmark problem has the same objective and constraints as PDP; the only difference is the policy space to search . Hence, the benchmark problem is a close approximation of the PDP. Note that the benchmark problem is also NP-hard (see the appendix).

A. Performance under time-varying channel conditions

Consider a 3x3 square grid of 9 SBSs with the minimum distance between any two SBSs being $d = 4.7$m. Each SBS serves one UE, who has a maximum power of 1000 mW and a minimum throughput requirement of 0.45 bits/s/Hz. The UEs and the SBSs are in two parallel horizontal hyperplanes, and each SBS is vertically above its UE with a distance of $\sqrt{10}m$. Then the distance from UE i to another SBS j is $D_{ij} = \sqrt{10 + (D_{ij}^{BS})^2}$, where D_{ij}^{BS} is the distance between SBSs i and j. The channel gain from UE i to SBS j is a product of path loss and Rayleigh fading $f_{ij} \sim Rayleigh(\beta)$, namely $g_{ij} = \frac{1}{(D_{ij})^2} f_{ij}$. The density function of $Rayleigh(\beta)$ is $v(z) = \frac{z}{\beta^2} e^{-\frac{z^2}{2\beta^2}}$ for $z \geq 0$, and $v(z) = 0$ for $z < 0$. The SBSs identify neighbors using a distance based rule with the threshold distance as in Section V-C with $D^{th} = 7$m. Note that different thresholds lead to different interference graphs, and hence different performance, which will be discussed next. Although, we use a distance based threshold rule, our framework

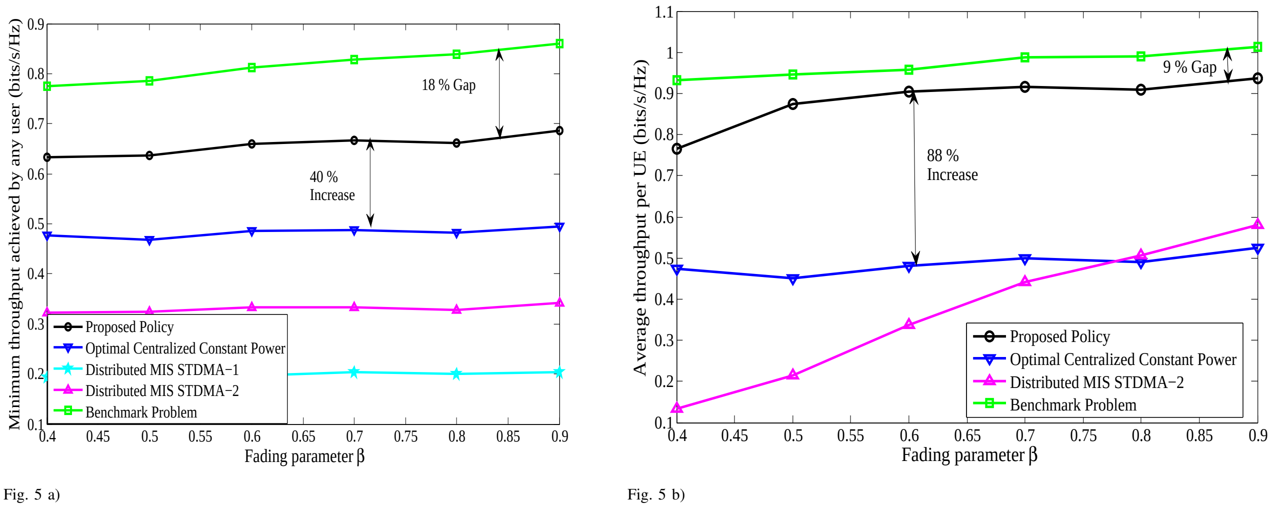

Figure 5. Comparison of the proposed policy with state of the art under different interference strength and time-varying channel conditions

is general and does not rely on a particular rule. The resulting interference graph for this setting is graph 3 shown in Fig. 7 a).

At the beginning, the UE-SBS pairs generate the set of MISs (Step 2 of the design framework in Section V), and compute the optimal fractions of time allocated to each MIS (Step 3). In our simulation, we assume a block fading model [38] and the fading changes every 100 time slots independently. To reduce complexity, the UEs do not recompute the interference graph and the MISs, but will recompute the optimal fractions of time under the new channel gains every 100 time slots. In Fig. 5, we compare the performance of the proposed policy with state of the art policies under different variances β of Rayleigh fading. We do not plot the performance of distributed PMS for this scenario since it is upper bounded by optimal centralized constant power control (because there is one UE per cell). We do not plot the distributed MIS STDMA -1 either, when the performance criterion is average throughput per UE (i.e., $\frac{\text{sum throughput}}{N}$), because it cannot satisfy the minimum throughput constraints. From Fig. 5, we can see that in terms of both average throughput and max-min fairness, our proposed policy achieves large performance gain (up to 88%) over existing policies, and achieves performance close to the benchmark (as close as 9%).

Selecting the Optimal Interference Graph : For different values of d, there can be five possible interference graphs, which are shown in Fig. 7 a). In Fig. 6 a) we show that as the grid size

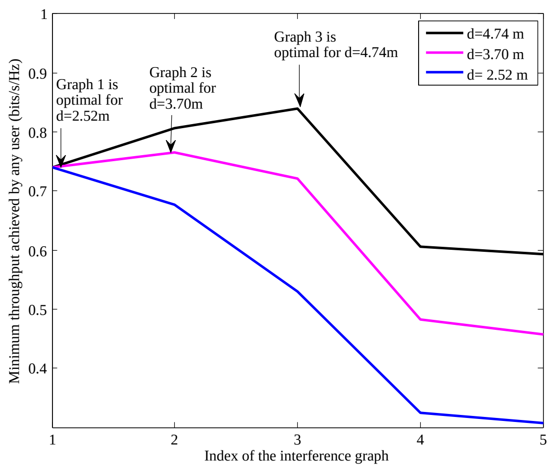

Fig. 6 a)

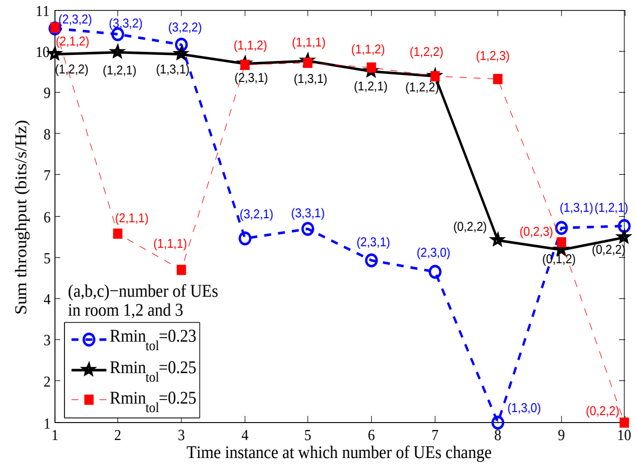

Fig. 6 b)

Figure 6. a) Comparison of max-min fairness under different grid sizes, b) Sample paths of sum throughput under dynamic entry/exit of UEs in the network

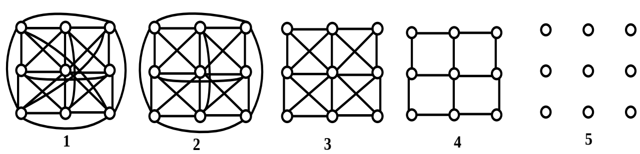

Fig. 7 a)

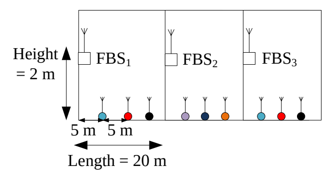

Fig. 7 b)

Figure 7. a) Different interference graphs for the 3 x 3 BS grid, b) Illustration of setup with 3 rooms.

d decreases ($d = 4.7m, d = 3.7m$ and $d = 2.5m$), the levels of interference from the adjacent UEs increases, and as a result, the interference graph with higher degrees perform better (as d decreases, the optimal graph changes from graph 3 to graph 1) .

B. Performance scaling in large networks

Consider the uplink of a femtocell network in a building with 12 rooms adjacent to each other. Fig. 7 b) illustrates 3 of the 12 rooms with 5 UEs in each room. For simplicity, we consider a 2-dimensional geometry. Each room has a length of 20 meters. In each room, there are P uniformly spaced UEs, and one SBS installed on the left wall of the room at a height of 2m. The distance from the left wall to the first UE, as well as the distance between two adjacent UEs in a room, is $\frac{20}{(1+P)}$ meters. Based on the path loss model in [39], the channel gain from each SBS i to a UE j is $\frac{1}{(D_{ij})^2 \Delta^{n_{ij}}}$, where $\Delta = 10^{0.25}$ is the coefficient representing the loss from the wall, and n_{ij} is the number of walls between UE i and SBS j. Each UE has a maximum transmit power level of 50 mW, a minimum throughput requirement of $R_i^{min} = 0.025$ bits/s/Hz, and a

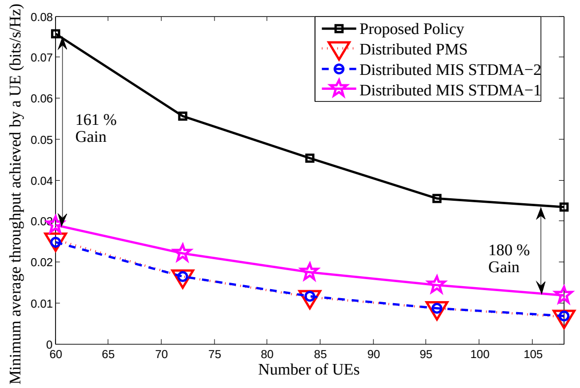

Fig. 8 a)

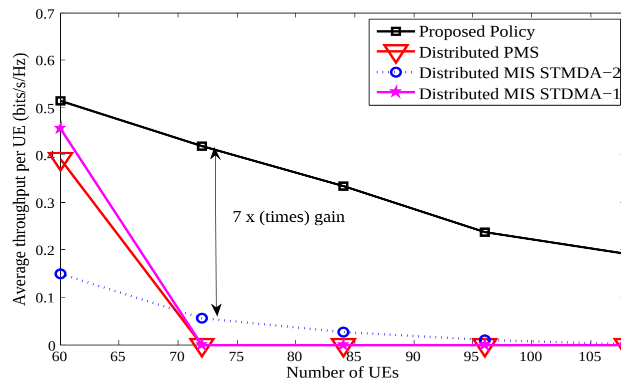

Fig. 8 b)

Figure 8. Comparison of max-min fairness and average throughput per UE against state of the art for large networks

noise power level of 10^{-11}mW at its receiver. Here, we consider that the UEs use a distance based threshold rule as in Section V-B with $D^{th} = 30$ m. This results in interference graphs which connects all the UE-SBS pairs within the room and in the adjacent rooms. We vary the number P of UEs in each room from 5 to 9 and compare the performance in Fig. 8. Note that the optimal centralized constant power policy cannot satisfy the feasibility conditions for any number of UEs in each room. Hence, only the performance of distributed MIS STDMA-1,2 and distributed PMS is shown in Fig. 8. We can see that under both criteria, the performance gain of our proposed policy is significant (from 160% to 700%). Note that since the number of UEs is large, it is impossible to solve the benchmark problem (which is NP-hard) is not possible.

C. Self-adjusting mechanism for dynamic entry/exit of the UEs

The self-adjusting mechanism proposed in Section V-D is aimed to provide incoming UEs with the maximum possible throughput without affecting the incumbent UEs, and to reuse the time slots left vacant by exiting UEs efficiently. Consider the same setup as in Section VI-B with 3 rooms and a maximum of $P = 3$ UEs in each room. Each UE has a maximum transmit power of 1000 mW and a minimum throughput requirement of 0.25 bits/s/Hz.

We assume that at a given time only one UE either enters or leaves the network. In Fig. 6 b) we show different sample paths of the sum throughput under different entry and exit processes. In the legends (i.e., $Rmin_{tol}$), we show the minimum throughput achieved at any point in the sample path. We repeated the same procedure 100 times. We can see that the self-adjusting mechanism works well by guaranteeing a worst-case minimum throughput requirement of 0.23

bits/s/Hz, which is just 0.02 bits/s/Hz below the original requirement more than 80% of the time.

VII. Conclusion

We proposed a design framework for distributed interference management in large-scale, heterogeneous networks, which are composed of different types of cells (e.g. femtocell, picocell), have different number of UEs in each cell, and have UEs with different minimum throughput requirements and channel conditions. Our framework allows each UE to have only local knowledge about the network and communicate only with its interfering neighbors. There are two key steps in our framework. First, we propose a novel distributed algorithm for the UEs to generate a set of MISs that span all the UEs. The distributed algorithm for generating MISs requires $\mathcal{O}(\log N)$ steps (which is much faster than state-of-the-art) before it converges to the set of MISs with a high probability. Second, we reformulate the problem of determining the optimal fractions of time allocated to the MISs in a novel manner such that the optimal solution can be determined by a distributed algorithm based on ADMM. Importantly, we prove that under wide range of conditions, the proposed policy can achieve a constant competitive ratio with respect to the policy design problem which is NP-hard. Moreover, we show that our framework can adjust to UEs entering or leaving the network. Our simulation results show that the proposed policy can achieve large performance gains (up to 85%).

Appendix

Discussion on max-min fairness: We now discuss as to how the proposed framework can be extended to incorporate inseparable function like max-min fairness. The coupled problem with max-min fairness objective is restated below:

$$\textbf{Coupled Problem (CP)} \quad \max_{\boldsymbol{\alpha}} \quad \min_{i\in\{1,..,N\}} W_i\Big(\sum_{k=1}^{H} \alpha_k R_i^k\Big)$$

$$\text{subject to} \quad \sum_{k=1}^{H} \alpha_k R_i^k \geq R_i^{min},\ \forall i \in \{1,...N\}$$

$$\sum_{k=1}^{H} \alpha_k = 1,\ \alpha_k \geq 0,\ \forall k \in \{1,...,H\}$$

Transforming the above problem into an equivalent problem with auxiliary variable t is given as

Table II

GENERATING MISs IN A DISTRIBUTED MANNER, ALGORITHM FOR UE i

Phase 1- Initialization: $\text{Tx}^i_{\text{tent}} = \phi$, $\text{Tx}^i_{\text{final}} = \phi$, tentative and final choice of UE i, $\text{Rx}^{\mathcal{N}(i)}_{\text{tent}} = \phi$,$\text{Rx}^{\mathcal{N}(i)}_{\text{final}} = \phi$ tentative and final choice made by the neighbors, $C^0_i = \{1, ..., H\} \cap \{1, .., d_i + 1\}$ the current list of subset of available colors, $C_i = \phi$, list of colors used by i, $\text{F}^i_{\text{colored}} = \phi$, $C1^0_i = \{1, ..., H\}$,the current list of all available colors

for $n = 0$ to $\lceil c_1 \log_{\frac{4}{3}} N \rceil$
$\text{Tx}^i_{\text{tent}} = \phi$, $\text{Tx}^i_{\text{final}} = \phi$
if($\text{F}^i_{\text{colored}} = \phi$)
$\text{Tx}^i_{\text{tent}} = \text{rand}\{C^n_i\}$, rand represents randomly selecting a color and informing the neighbors about it.
$\text{Rx}^{\mathcal{N}(i)}_{\text{tent}} = \{\text{Tx}^k_{\text{tent}}, \forall k \in \mathcal{N}(i)\}$
If($\text{Tx}^i_{\text{tent}} \neq \text{Rx}^{\mathcal{N}(i)}_{\text{tent}}(j)$, $\forall j \in \mathcal{N}(i)$), here UE-$i$ checks if there is a conflict with any of the neighbor's choice
$\text{Tx}^i_{\text{final}} = \text{Tx}^i_{\text{tent}}$, $C_i = \{\text{Tx}^i_{\text{final}}\}$,if there is no conflict then UE-i transmits its final color choice to the neighbors,
else
$\text{Tx}^i_{\text{final}} = \phi$
end
end
$\text{Rx}^{\mathcal{N}(i)}_{\text{final}} = \{\text{Tx}^k_{\text{final}}, \forall k \in \mathcal{N}(i)\}$
$C^{n+1}_i = C^n_i \cap \{\text{Rx}^{\mathcal{N}(i)}_{\text{final}} \cup \text{Tx}^i_{\text{final}}\}^c$, $C1^{n+1}_i = C1^n_i \cap \{\text{Rx}^{\mathcal{N}(i)}_{\text{final}} \cup \text{Tx}^i_{\text{final}}\}^c$
if($\text{Tx}^i_{\text{final}} \neq \phi$)
$\text{F}^i_{\text{colored}} = 1$
end
end

Table III

ADMM UPDATE ALGORITHM FOR UE i

Initialization: arbitrary $\boldsymbol{\beta}_i(0) \in \mathcal{B}_i$, $\theta^k_{ei}(0)$ such that $\boldsymbol{\theta}^k \in \boldsymbol{\Theta}^k$, and $\lambda^k_{ei}(0) = 0$, $\forall k \in \{1, ..., H\}$,$\forall e$ such that $i \in e$

For $t = 0$ to $P - 1$
$\boldsymbol{\beta}_i(t+1) = \arg\min_{\boldsymbol{\beta}_i \in \mathcal{B}_i} -\sum_{i=1}^N W_i(\boldsymbol{\beta}_i^T \boldsymbol{R}_i) + \sum_{k=1}^H \sum_{e\in E} \sum_{q \in e} \left[\lambda^k_{eq}\left(D^k_{eq}\beta^k_q - \theta^k_{eq}\right) + \frac{y}{2}\left(D^k_{eq}\beta^k_q - \theta^k_{eq}\right)^2\right]$
$\boldsymbol{\beta}_i(t+1)$ is transmitted to all of its neighbors in $\mathcal{N}(i)$.
$\lambda^k_{ei}(t)$ is transmitted to its neighbor connected with edge e, $\forall k \in \{1, ..., H\}$ and $\forall e$ such that $i \in e$
Update $\forall k \in \{1, ..., H\}$ and $\forall e$ such that $i \in e$
$\lambda^k_{ei}(t+1) = \frac{1}{2}(\lambda^k_{ei}(t) + \lambda^k_{ej}(t)) - \frac{y}{2}(D^k_{ei}\beta^k_i(t+1) + D^k_{ej}\beta^k_j(t+1))$, where j is the other endpoint of e.
$\theta^k_{ei}(t+1) = \frac{1}{y}(\lambda^k_{ei}(t+1) - \lambda^k_{e,i}(t)) + D^k_{ei}\beta^k_i(t+1)$
end

$$\begin{aligned}
\max_{\boldsymbol{\alpha},t} \quad & t \\
\text{subject to} \quad & W_i(\sum_{k=1}^{H} \alpha_k R_i^k) \geq t, \; \forall i \in \{1,...,N\} \\
& \sum_{k=1}^{H} \alpha_k R_i^k \geq R_i^{min}, \; \forall i \in \{1,...N\} \\
& \sum_{k=1}^{H} \alpha_k = 1, \; \alpha_k \geq 0, \; \forall k \in \{1,...,H\}
\end{aligned}$$

To decouple the above problem, we introduce local variables for each UE i given as,$\{\beta_i^1,...,\beta_i^{H+1}\}$. Now we state a problem which we claim is equivalent to CP,(the proof to this claim is very similar to the proof of Theorem 2 and we will highlight this fact in the proof clearly).

$$\begin{aligned}
\textbf{P1} \qquad \max_{\boldsymbol{\beta}} \quad & \sum_{i=1}^{N} \beta_i^{H+1} \\
\text{subject to} \quad & W_i(\sum_{k=1}^{H} \beta_i^k R_i^k) \geq \beta_i^{H+1}, \; \forall i \in \{1,...,N\} \\
& \sum_{k=1}^{H} \beta_i^k R_i^k \geq R_i^{min}, \; \forall i \in \{1,...N\} \\
& \sum_{k=1}^{H} \beta_i^k = 1, \; \beta_i^k \geq 0, \; \forall k \in \{1,...,H\}, \forall i \in \{1,...,N\} \\
& \beta_i^k = \beta_j^k, \forall j \in \mathcal{N}(i), \forall k \in \{1,...,H+1\}
\end{aligned}$$

Here, $\boldsymbol{\beta} = (\boldsymbol{\beta_1,..,\beta_N})$, with $\boldsymbol{\beta_i} = (\beta_i^1,...,\beta_i^{H+1}), \forall i \in \{1,...,N\}$. Now, given the two problems CP and the problem P1 are equivalent, we focus on solving P1. P1 can be changed to a problem similar to DP. To do that we introduce some additional variables similar to the ones introduced for DP. If UE i and l are connected by an edge (i,l) then for each set $I_k^{'}$ define $\theta_{(i,l)i}^k = \beta_i^k$ and $\theta_{(i,l)l}^k = -\beta_l^k$, note that these auxiliary variables are introduced to formulate the problem into the ADMM framework [37]. Define a polyhedron for each i, $\mathcal{T}_i^{'} = \{(\boldsymbol{\beta 1})_{\boldsymbol{i}} | \text{s.t. } \mathbf{1^t}(\boldsymbol{\beta_i^{''}}) = 1, (\boldsymbol{\beta 1})_{\boldsymbol{i}} \geq \mathbf{0}, \boldsymbol{R_i^{'}}(\boldsymbol{\beta_i^{''}}) \geq R_i^{min}, W_i(\boldsymbol{R_i^{'}}(\boldsymbol{\beta_i^{''}})) - \beta_i^{H+1} \geq 0\}$, here $\boldsymbol{\beta_i^{''}} = (\beta_i^1,...,\beta_i^H)$ and $\boldsymbol{R_i} = (R_i^1,...,R_i^H)$ and $()^{'}$corresponds to the transpose. Let $\boldsymbol{\beta} = (\boldsymbol{\beta_1,...,\beta_N}) \in \boldsymbol{\mathcal{T}^{'}}$, where $\boldsymbol{\mathcal{T}^{'}} = \prod_{i=1}^{N} \mathcal{T}_i^{'}$ and $\prod$ corresponds to the Cartesian product of the sets. Also, let $\boldsymbol{\beta^k} = (\beta_1^k, ..., \beta_N^k), \forall k \in \{1, .., H\}$. Define another polyhedron $\Theta_{(i,l)}^k = \{(\theta_{(i,l)i}^k, \theta_{(i,l)l}^k) : \theta_{(i,l)i}^k + \theta_{(i,l)l}^k = 0, -1 \leq \theta_{(i,l)s}^k \leq 1, \forall s \in \{i, l\}\}$, $\boldsymbol{\Theta^k} = \prod_{(i,l)\in E} \Theta_{(i,l)}^k$ here $E = (e_1, ..e_M)$ is the set of all the M edges in the interference graph. A vector $\boldsymbol{\theta^k} \in \boldsymbol{\Theta^k}$ is written as $\boldsymbol{\theta^k} = (\theta_{e_1,z(e_1)}^k, \theta_{e_1,t(e_1)}^k, .., \theta_{e_M,z(e_M)}^k, \theta_{e_M,t(e_M)}^k)$, here $z(e_i)$, $t(e_i)$ correspond to the vertices in the edge, e_i. Similarly define, $\boldsymbol{\theta} = (\boldsymbol{\theta^1}, ..., \boldsymbol{\theta^{H+1}}) \in \boldsymbol{\Theta'}$, where $\boldsymbol{\Theta'} = \prod_{k=1}^{H+1} \boldsymbol{\Theta^k}$. The reformulated problem is stated as follows:

$$\textbf{DP1} \qquad \min_{\boldsymbol{\beta}\in\boldsymbol{\mathcal{T}'}, \boldsymbol{\theta}\in\boldsymbol{\Theta'}} -\sum_{i=1}^N \boldsymbol{W_i}(\boldsymbol{R_i}'\boldsymbol{\beta}_i)$$

$$\text{subject to } \boldsymbol{D^k\beta^k - \theta^k = 0}, \forall k \in \{1, .., H+1\}$$

Then, DP1 can be solved using the ADMM procedure similar to the one described for DP.

Discussion on Benchmark Problem's complexity: Benchmark Problem is restated here for convenience:

$$\textbf{Benchmark Problem (BP)} \quad \max_{\boldsymbol{\pi}\in\boldsymbol{\Pi_{BC}}} W(R_1(\boldsymbol{\pi}), ..., R_N(\boldsymbol{\pi}))$$

$$\text{subject to. } R_i(\boldsymbol{\pi}) \geq R_i^{min}, \forall i \in \{1, ..., N\}$$

Let the power set of $\mathcal{U}$ be $S_\mathcal{U}$, where $S_\mathcal{U}$ consists of 2^N subsets of UEs. Let $S_\mathcal{U}(j)$ denote the j^{th} element of $S_\mathcal{U}$. Define a set of power profiles,$\mathcal{P}^{S_\mathcal{U}}$, where the $\mathcal{P}^{S_\mathcal{U}}(j)$ corresponds to the j^{th} element in the set and it corresponds to the power profile when the UEs in set $S_\mathcal{U}(j)$ transmit at their maximum power levels and the rest of the UEs do not transmit. Note that for $\boldsymbol{\pi} \in \boldsymbol{\Pi_{BC}}$, $\boldsymbol{\pi(t)}$ corresponds to a power profile in $\mathcal{P}^{S_\mathcal{U}}$. Therefore, the average throughput achieved by UE i, $R_i(\boldsymbol{\pi})$, where $\boldsymbol{\pi} \in \boldsymbol{\Pi_{BC}}$, can also be expressed as $R_i(\boldsymbol{\pi}) = \sum_{j=1}^{2^N} \alpha_j r_i(\mathcal{P}^{S_\mathcal{U}}(j))$, with $\alpha_j \geq 0, \forall j \in \{1, .., 2^N\}$and $\sum_{j=1}^{2^N} \alpha_j = 1$. Here the fraction α_j associated with each profile $\mathcal{P}^{S_\mathcal{U}}(j)$ corresponds to the fraction of transmission time associated with that power profile.

Consider the following problem:

$$
\begin{aligned}
\textbf{BP1} \quad & \max_{\mathbf{y},\boldsymbol{\alpha}} W(y_1, ..., y_N) \\
& \text{subject to. } y_i \geq R_i^{min}, \forall i \in \{1, ..., N\} \\
& y_i = \sum_{i=1}^{2^N} \alpha_i r_i(\mathcal{P}^{S_{\mathcal{U}}}(j)), \forall i \in \{1, ..., N\} \\
& \alpha_j \geq 0, \forall j \in \{1, .., 2^N\}, \sum_{j=1}^{2^N} \alpha_j = 1
\end{aligned}
$$

Next, in order to show that the above problem is NP-hard we will show intuitively why is it so, but the detailed proof follows from proof of Theorem 1 in [40]. Consider $W(y_1, .., y_N) = \sum_{i=1}^{N} y_i$,to be a linear function, $R_i^{min} = 0$, $\forall i \in \{1, ..., N\}$ and the cross channel gains amongst some users who do not share an edge in the interference graph to be 0 and the cross channel gains amongst the interfering neighbors to be ∞. This implies that in any optimal solution will correspond to the transmission by a MIS of the interference graph. This can be justified as follows. Consider an optimal solution in which two neighboring UEs are transmitting, making one of the UEs not transmit will definitely increase the sum throughput contradicting the optimality. Specifically, this problem reduces to finding the maximum weighted maximum indpendenet set which is NP hard. Here the weight of each MIS corresponds to $\sum_{i=1}^{N} r_i(\boldsymbol{p}^{\boldsymbol{I}_j})$.

Proof of Theorem 1: The success probability of Phase 1 is high, $(1 - \frac{1}{N^{c_1 - 1}})$ (lower bound), (see [35] for detail), here we analyze Phase 2.

We first show that, if the list of remaining colors given as, $C1_i^n$ is empty at $n \geq \lceil c_1 \log_{\frac{4}{3}} N \rceil + \lceil c_2 \log_x N \rceil + 2$ and if this holds $\forall i \in \{1, ..., N\}$ then the Phase 2 has converged to a set of H MISs which span all the UEs. Let us assume otherwise, i.e. $C1_i^n$ is empty $\forall i \in \{1, ..., N\}$ however, the set corresponding to some color $h \in \{1, ..., H\}$, $I_h^{'}$ is not a MIS. $I_h^{'}$ has to be an IS. Assume otherwise, i.e. $I_h^{'}$ is not an IS, which implies that there must exist a pair of UEs, i and j, which are neighbors and are a part of $I_h^{'}$. If this is true then both then both acquired the color h either in the same time slot or in different time slots, in Phase 1 or 2. In case the color is acquired in different time slots, then after the first time slot when either of the UEs in the pair acquires the color it will transmit the final color choice, h to the neighbors (see Table II and III) who in turn delete that color. However, if the color is deleted by the neighbor then it cannot acquire it in the future thus, ruling out the case that the colors were acquired in two different time slots. If the color was acquired by the UEs in the same time slot, then it implies that despite the conflict in tentative choice the UEs acquire the color which is not possible (see Table II and III). This shows that $I_h^{'}$ is an IS.

Since $I_h^{'}$ is not maximal then $\exists$ at least one UE-$j \notin I_k^{'}$ which can be added to this set without violating independence. From the assumption, we have $C1_j^n = \phi$ which implies that the color h was deleted at some stage from the original list of all the colors either in Phase 1 or 2. The deletion of h was a result of that color being acquired finally by at least one of the neighbors $k \in \mathcal{N}(j)$ since $j \notin I_k^{'}$. In that case, j cannot acquire h as it will violate the independence property.

Next, we show that indeed the list of all colors available $C1_i^n$ is empty at the end of Phase 2 with a high probability. Let U^n correspond to the number of UEs which have a non-empty list at the beginning of time slot n and, let $Tn(U^n)$ correspond to the total time needed before all the UEs have an empty list. The probability that a UE at time slot n with a non-empty list will have an empty list in next time slot is always greater than $c^H(1-c)^{H^2}$. This can be explained as, if the UE chooses all the colors in the list assuming (worst case H number of colors remain) and all the neighbors (worst case H neighbors) do not choose any color, then all the colors in the UE's list will be deleted. From this, we get $E(U^{n+1}) \leq (1 - c^H(1-c)^{H^2})U^n = \frac{1}{x}U^n$ and $Tn(U^n) = 1 + Tn(U^{n+1})$. Assuming that the Phase 2 will start with N UEs whose list are non-empty (worst case) and from [41] we get $P(Tn(N) \geq \lceil c_2 \log_x N \rceil) \leq \frac{1}{N^{c_2 - 1}}$. This gives the lower bound on success probability of Phase 2 and thereby the result in the Theorem. (Q.E.D)

Proof of Theorem 2: The two problems which are introduced to transit from CP to DP are,

$$\textbf{Global Primal Problem (GPP)} \qquad \max_{\{\beta_i^k\}_{i,k}} \sum_{k=1}^{H} W_i\Big(\sum_{i=1}^{N} \beta_i^k R_i^k\Big)$$

$$\text{subject to} \quad \sum_{k=1}^{H} \beta_i^k R_i^k \geq R_i^{min}, \sum_{k=1}^{H} \beta_i^k = 1,\ \forall i \in \{1, ..., N\}$$

$$\beta_i^k = \beta_l^k,\ \forall i \neq l, \forall k \in \{1, ..., H\}, \qquad \beta_i^k \geq 0,\ \forall i \in \{1, ..., N\}, \forall k \in \{1, ...H\}$$

The second problem, Local Primal Problem (LPP) is the same as GPP except we choose a subset of the constraints from the above problem. Basically, instead of an equality constraint between the UE's estimate and every other UE in the network, we only keep the equality constraints between the UE and its neighbors, i.e. $\beta_i^k = \beta_l^k,\ \forall k \in \{1, ..., H\}, \forall l \in \mathcal{N}(i)$. This is formally stated below:

Local Primal Problem (LPP) $\max_{\{\beta_i^k\}_{i,k}} \sum_{k=1}^{H} W_i(\sum_{i=1}^{N} \beta_i^k R_i^k)$

subject to $\sum_{k=1}^{H} \beta_i^k R_i^k \geq R_i^{min}, \sum_{k=1}^{H} \beta_i^k = 1,\ \forall i \in \{1,...,N\}$

$\beta_i^k = \beta_l^k,\ \forall l \notin \mathcal{N}(i), \forall k \in \{1,...,H\},$ $\beta_i^k \geq 0,\ \forall i \in \{1,...,N\}, \forall k \in \{1,...H\}$

To show that problems CP and GPP are equivalent, we need to show that from $\boldsymbol{\beta}^* = (\boldsymbol{\beta_1^*, .., \beta_N^*})$, an optimal argument of GPP, we can obtain an optimal argument of CP, i.e. $\boldsymbol{\alpha}^*$ and vice versa. Since $\boldsymbol{\beta}^*$ is the optimal value (assuming feasibility) we know that $\boldsymbol{\beta}_i^* = \boldsymbol{\beta}_j^*$ (component-wise) holds $\forall i, j \in \{1, ..., N\}$.

a).Let $\boldsymbol{\alpha}' = \boldsymbol{\beta_i^*}$. $\boldsymbol{\alpha}'$ satisfies the constraints in CP. The objective of CP at $\boldsymbol{\alpha}'$ attains the optimal value of GPP. We need to establish that $\boldsymbol{\alpha}'$ is indeed the optimal argument of CP. Assume that $\boldsymbol{\alpha}'$is not the optimal value, then there exists another $\boldsymbol{\alpha}^*$ which is indeed the optimal. Next, using $\boldsymbol{\alpha}^*$, we can obtain another $\boldsymbol{\beta}'$as follows, $\boldsymbol{\beta_1' = \alpha^*}$and $\boldsymbol{\beta_i' = \beta_1'}$, $\forall i \in \{1, ..., N\}$. The objective of GPP at $\boldsymbol{\beta}'$ should be higher than $\boldsymbol{\beta}^*$ which contradicts $\boldsymbol{\beta}^*$ being the optimal argument. Note that if either of CP or GPP is infeasible then the other problem can be shown to be infeasible as well. On the same lines we can show that from an $\boldsymbol{\alpha}^*$ we can obtain $\boldsymbol{\beta}^*$ as well.

b). Let $\boldsymbol{\alpha}^*$ be the optimal solution to CP, and define $\boldsymbol{\beta}''$a solution to GPP as follows. Let $\boldsymbol{\beta_1'' = \alpha^*}$ and $\boldsymbol{\beta_i''} = \boldsymbol{\beta_j''}$, $\forall j \neq i$ and since $\boldsymbol{\alpha}^*$ satisfies the constraints of CP, i.e. it is feasible, implies that $\boldsymbol{\beta}''$ as well satisfies constraints of GPP. We want to show that $\boldsymbol{\beta}''$is the optimal value as well, assume that it is not and there exists an argument $\boldsymbol{\beta}^*$ for which the objective takes a higher value. If this is the case then, from $\boldsymbol{\beta}^*$ we can construct a $\boldsymbol{\alpha}'$as in part a). which, if $\boldsymbol{\beta}^*$ takes a higher value than $\boldsymbol{\beta}''$, takes a higher value than $\boldsymbol{\alpha}^*$ thus, contradicting optimality.

To show that GPP and LPP are equivalent, we use the following fact, since LPP consists of a subset of the constraints then the solution of LPP is an upper bound of the solution to GPP. We need to show that the gap between the solution of LPP and GPP is always 0. Note that for an optimal solution of LPP, $\boldsymbol{\gamma}^* = (\boldsymbol{\gamma_1^*, .., \gamma_N^*})$ we know that $\boldsymbol{\gamma}_i^* = \boldsymbol{\gamma}_j^*\ \forall j \in \mathcal{N}(i)$ (component-wise). If we can show that $\boldsymbol{\gamma}_i^* = \boldsymbol{\gamma}_j^*\ \forall j \in \{1, ..., N\}$ then LPP and GPP will be equivalent, since it will also satisfy all the constraints of GPP. Assume that this does not hold then $\exists\, i,\, j$ such that $\boldsymbol{\gamma}_i^* \neq \boldsymbol{\gamma}_j^*$. Since, the interference graph is connected $\exists$ a path $i \to j = \{i_1, ..., i_s\}$ which implies, $\boldsymbol{\gamma}_i^* = \boldsymbol{\gamma}_{i_1}^* ... = \boldsymbol{\gamma}_j^*$. This leads to a contradiction, thereby establishing the claim.

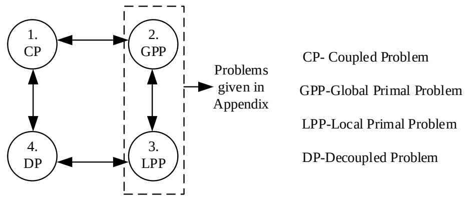

Figure 9. Problems used to transit from the Coupled Problem (CP) to Decoupled Problem (DP).

Lastly, to show that DP is equivalent LPP. Given γ^*, define $\boldsymbol{\kappa} = \boldsymbol{\gamma}^*$ and a $\boldsymbol{\theta} = (\boldsymbol{\theta^1}, ..., \boldsymbol{\theta^H})$ to satisfy $\boldsymbol{D^k \kappa^k} - \boldsymbol{\theta^k} = \mathbf{0}$, $\forall k \in \{1, .., H\}$, where $\boldsymbol{\kappa^k} = (\gamma_1^{*,k}, .., \gamma_N^{*,k})$. It can be shown using the same approach as we did for GPP and CP that $(\boldsymbol{\kappa}, \boldsymbol{\theta})$ is indeed optimal argument for DP. Assume that $(\boldsymbol{\kappa}, \boldsymbol{\theta})$ is not the optimal solution then we know that there exists $(\boldsymbol{\kappa}^*, \boldsymbol{\theta^*})$ for which the objective in DP takes a higher value. If this is the case, let us define $\boldsymbol{\gamma}' = \boldsymbol{\kappa}^*$, here $\boldsymbol{\gamma}'$ satisfies the constraints in LPP. Also, since the objective in DP at $(\boldsymbol{\kappa}^*, \boldsymbol{\theta^*})$ takes a higher value than that at $(\boldsymbol{\kappa}, \boldsymbol{\theta})$, this yields that the objective in LPP at $\boldsymbol{\gamma}'$should take a higher value than that at $\boldsymbol{\gamma}^*$, which contradicts optimality of $\boldsymbol{\gamma}^*$. On the same lines, it can be easily shown that from $(\boldsymbol{\kappa}^*, \boldsymbol{\theta^*})$ we can construct the optimal solution $\boldsymbol{\gamma}^*$ of the LPP. This, will establish equivalence between LPP and DP. Hence, all the four problems are equivalent. This is shown in Fig. 10.

(Q.E.D)

Proof of Theorem 3: According to [37], the ADMM algorithm converges with rate $O(1/P)$ if the DP is feasible and if the feasible set is compact. Since $\mathcal{B}_i$ and Θ^k are all closed and bounded polyhedrons, the feasible set is compact. (Q.E.D)

Proof of Theorem 4: Here, we need to show three things,

i). if $\Delta \leq \Delta^{max}$ then the distributed policy yields a feasible solution,

ii). the size of any MIS is $\geq \frac{N}{\Delta+1}$, thereby using this to show that the distributed policy, if feasible will yield a network performance of at least $\frac{N}{\Delta+1}\log_2(1+\frac{p_{lb}^{max}}{(D^{ub})^{np}2^{\zeta}\sigma^2})$ and

iii). the upper bound on the network performance, sum throughput here is $N\log_2(1+\frac{p_{ub}^{max}}{(D^{lb})^{np}\sigma^2})$.

i). In the Phase 1 of the algorithm the maximum number of colors used is $\Delta+1$, since each UE selects colors from subset of $\{1, ..., H\} \cap \{1, ..., d_i+1\}$. The first $\Delta+1$ output MISs, $\{I_1', ..., I_{\Delta+1}'\}$ span all the UEs in the network. If the fraction of time assigned to each of these $\Delta + 1$ MISs is, $\alpha_k' = \frac{R_{ub}^{min}}{\log_2(1+\frac{p_{lb}^{max}}{(D^{ub})^{np}2^{\zeta}\sigma_{ub}^2})}$, $\forall k \in \{1, .., \Delta+1\}$ then such an assignment satisfies the constraint that sum of fractions assigned to all the colors cannot be more than 1, i.e. since $\Delta \leq \Delta^{max} \implies$

$(\Delta+1)\frac{R_{ub}^{min}}{\log_2(1+\frac{p_{lb}^{max}}{(D^{ub})^{np}2^{\zeta}\sigma^2})} \leq 1$. Using the fact that network exhibits $\zeta-$WNI we can write the minimum instantaneous throughput that can be obtained by UE-i as, $\log_2(1+\frac{p_i^{max}}{(D_{iT(i)})^{np}2^{\zeta}\sigma_{ub}^2})$, and minimum instantaneous throughput of any UE as, $\log_2(1+\frac{p_{lb}^{max}}{(D^{ub})^{np}2^{\zeta}\sigma_{ub}^2})$. Thus, given the fractions assigned to the MISs, $\alpha_k^{'} = \frac{R_{ub}^{min}}{\log_2(1+\frac{p_{lb}^{max}}{(D^{ub})^{np}2^{\zeta}\sigma_{ub}^2})}$, $\forall k \in \{1,..,\Delta+1\}$, which span all the UEs. each UE i's throughput requirement is satisfied, , $\frac{R_{ub}^{min}}{\log_2(1+\frac{p_{lb}^{max}}{(D^{ub})^{np}2^{\zeta}\sigma_{ub}^2})}\log_2(1+\frac{p_i^{max}}{(D_{iT(i)})^{np}2^{\zeta}\sigma_{ub}^2}) \geq R_{ub}^{min}$.

ii). Assume that $\exists$ an MIS whose size is $S < \frac{N}{\Delta+1}$. Each UE in the MIS can exclude a maximum of Δ UEs from being included in the MIS. This implies that $S(\Delta+1)$, represents the total number of UEs excluded and the UEs in the MIS which put together should exceed N. Since this is not the case here, the contradiction implies that $S \geq \frac{N}{\Delta+1}$. This combined with minimum instantaneous throughput of any UE, we get the lower bound $\frac{N}{\Delta+1}\log_2(1+\frac{p_{lb}^{max}}{(D^{ub})^{np}2^{\zeta}\sigma_{ub}^2})$,for our policy.

iii). The upper bound on the optimal network performance is obtained by summing maximum instantaneous throughput of any UE $\log_2(1+\frac{p_{ub}^{max}}{(D^{lb})^{np}\sigma_{lb}^2})$ for all UEs, $N\log_2(1+\frac{p_{ub}^{max}}{(D^{lb})^{np}\sigma_{lb}^2})$. Computing the ratio of the lower bound of proposed scheme $\frac{N}{\Delta+1}\log_2(1+\frac{p_{lb}^{max}}{(D^{ub})^{np}2^{\zeta}\sigma^2})$ and $N\log_2(1+\frac{p_{ub}^{max}}{(D^{lb})^{np}\sigma^2})$, we get $\frac{\log_2(1+\frac{p_{lb}^{max}}{(D^{ub})^{np}2^{\zeta}\sigma^2})}{(\Delta+1)\log_2(1+\frac{p_{ub}^{max}}{(D^{lb})^{np}\sigma^2})}$ which is no less than, $\Gamma = \frac{R_{ub}^{min}}{\log_2(1+\frac{p_{ub}^{max}}{(D^{lb})^{np}\sigma^2})}$ since $\Delta \leq \Delta^{max}$. (Q.E.D)

Proof of Theorem 5: Let $\Delta^* = 6\eta$ with $\eta = \lceil\frac{\log_2(1+\frac{1}{(D^{lb})^{np}\sigma_{lb}^2}p_{ub}^{max})}{R_{lb}^{min}}\rceil$. We assume that the interference graph is constructed using a distance threshold rule (Subsection V-B). Note that each UE's minimum throughput requirement is at least R_{lb}^{min}, this combined with maximum instantaneous throughput of any UE $\log_2(1+\frac{p_{ub}^{max}}{(D^{lb})^{np}\sigma_{lb}^2})$ yields that each UE needs at least $\frac{R_{lb}^{min}}{\log_2(1+\frac{p_{ub}^{max}}{(D^{lb})^{np}\sigma_{lb}^2})}$ fraction of time slots. First, we need to show that if there exists a clique (a subset of vertices in the graph which are mutually connected) in the interference graph of size, X greater than η then the minimum throughput constraints cannot be satisfied. Assume that there does exist such a clique, then any MIS based scheduling policy will allocate separate time slots to each UE in the clique. This is true because no two UEs in the clique will belong to the same MIS. This implies that $X\frac{R_{lb}^{min}}{\log_2(1+\frac{p_{ub}^{max}}{(D^{lb})^{np}\sigma_{lb}^2})}$ is the total fraction separate time slots needed which has to be less than 1. But as $X \geq \eta$, this leads to infeasibility. Next, if $\Delta \geq \Delta^*$, we claim that we will have at least one clique in the graph satisfying this condition. Then $\exists$ UE-i with a degree $d_i \geq 6\eta$, this implies that within a radius of D^{th} around SBS-$T(i)$ $\exists$ 6η SBSs. Also, this circle around SBS-$T(i)$ can be partitioned into 6 sectors subtending $\frac{\pi}{3}$ at the center.The distance between any two points located in the sector is $\leq D^{th}$, which we justify next. Hence, all the points in a sector are mutually connected, thus forming a clique.

Let the 2-D polar coordinates of two points i, j in a sector be $(r_i, 0)$ and (r_j, θ), where $0 \leq r_i \leq D^{th}, 0 \leq r_j \leq D^{th}$ and $0 \leq \theta \leq \frac{\pi}{3}$. Hence, the square of the distance between the two points is expressed as $f(r_i, r_j, \theta) = r_i^2 + r_j^2 - 2r_i r_j cos\theta$ and our claim is that the maximum value $f(r_i, r_j, \theta)$, in the set of constraints above is no greater than $(D^{th})^2$. We formally state this as an optimization problem below:

$$\max_{r_i, r_j, \theta} \quad f(r_i, r_j, \theta)$$

$$0 \leq r_i \leq D^{th}, 0 \leq r_j \leq D^{th}$$

$$0 \leq \theta \leq \tfrac{\pi}{3}$$

Since, both r_i, r_j are non-negative, this implies that in the above optimization problem, $\theta = \frac{\pi}{3}$ has to be satisfied in the optimal argument. Substituting $\theta = \frac{\pi}{3}$ in $f(r_i, r_j, \theta)$ we get, $f(r_i, r_j, \frac{\pi}{3}) = r_i^2 + r_j^2 - r_i r_j$. Next, we show that $r_i^2 + r_j^2 - r_i r_j \leq (D^{th})^2$ for $0 \leq r_i \leq D^{th}, 0 \leq r_j \leq D^{th}$. Fix a $0 \leq r_j \leq D^{th}$,then $r_i^2 + r_j^2 - r_i r_j$ takes its maximum value at $r_i = D^{th}$, which gives $(D^{th})^2 + r_j^2 - D^{th} r_j$. Since $0 \leq r_j \leq D^{th}$, this yields $(D^{th})^2 + r_j^2 - D^{th} r_j \leq (D^{th})^2$ which establishes the claim.

If we have a total of 6η SBSs in the circle then at least one sector has to have more than η SBSs (Pigeonhole principle), which implies that a clique of size $X \geq \eta$ will exist. (Q.E.D)

References

[1] E. Hossain, L. B. Le, and D. Niyato, "Self-organizing small cell networks," in *Radio Resource Management in Multi-Tier Cellular Wireless Networks*. Hoboken, NJ, USA: John Wiley Sons, Inc., 2013.

[2] A. Ghosh, N. Mangalvedhe, R. Ratasuk, B. Mondal, M. Cudak, E. Visotsky, T. A. Thomas, J. G. Andrews, P. Xia, H. S. Jo *et al.*, "Heterogeneous cellular networks: From theory to practice," *Communications Magazine, IEEE*, vol. 50, no. 6, pp. 54–64, 2012.

[3] J. G. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. C. Reed, "Femtocells: Past, present, and future," *Selected Areas in Communications, IEEE Journal on*, vol. 30, no. 3, pp. 497–508, 2012.

[4] R. Ramaswami and K. K. Parhi, "Distributed scheduling of broadcasts in a radio network," in *INFOCOM'89. Proceedings of the Eighth Annual Joint Conference of the IEEE Computer and Communications Societies. Technology: Emerging or Converging, IEEE*. IEEE, 1989, pp. 497–504.

[5] I. Cidon and M. Sidi, "Distributed assignment algorithms for multihop packet radio networks," *Computers, IEEE Transactions on*, vol. 38, no. 10, pp. 1353–1361, 1989.

[6] A. Ephremides and T. V. Truong, “Scheduling broadcasts in multihop radio networks,” *Communications, IEEE Transactions on*, vol. 38, no. 4, pp. 456–460, 1990.

[7] J. Huang, R. A. Berry, and M. L. Honig, “Distributed interference compensation for wireless networks,” *Selected Areas in Communications, IEEE Journal on*, vol. 24, no. 5, pp. 1074–1084, 2006.

[8] C. U. Saraydar, N. B. Mandayam, and D. Goodman, “Pricing and power control in a multicell wireless data network,” *Selected Areas in Communications, IEEE Journal on*, vol. 19, no. 10, pp. 1883–1892, 2001.

[9] ——, “Efficient power control via pricing in wireless data networks,” *Communications, IEEE Transactions on*, vol. 50, no. 2, pp. 291–303, 2002.

[10] E. J. Hong, S. Y. Yun, and D.-H. Cho, “Decentralized power control scheme in femtocell networks: A game theoretic approach,” in *Personal, Indoor and Mobile Radio Communications, 2009 IEEE 20th International Symposium on*, Sept 2009, pp. 415–419.

[11] L. Giupponi and C. Ibars, “Distributed interference control in ofdma-based femtocells,” in *Personal Indoor and Mobile Radio Communications (PIMRC), 2010 IEEE 21st International Symposium on*. IEEE, 2010, pp. 1201–1206.

[12] A. Ladanyi, D. Lopez-Perez, A. Juttner, X. Chu, and J. Zhang, “Distributed resource allocation for femtocell interference coordination via power minimisation,” in *GLOBECOM Workshops (GC Wkshps), 2011 IEEE*, Dec 2011, pp. 744–749.

[13] H.-S. Jo, C. Mun, J. Moon, and J.-G. Yook, “Interference mitigation using uplink power control for two-tier femtocell networks,” *Wireless Communications, IEEE Transactions on*, vol. 8, no. 10, pp. 4906–4910, 2009.

[14] M. Bennis, S. M. Perlaza, P. Blasco, Z. Han, and H. V. Poor, “Self-organization in small cell networks: A reinforcement learning approach,” *Wireless Communications, IEEE Transactions on*, vol. 12, no. 7, pp. 3202–3212, 2013.

[15] M. Chiang, C. W. Tan, D. P. Palomar, D. O’Neill, and D. Julian, “Power control by geometric programming,” *Wireless Communications, IEEE Transactions on*, vol. 6, no. 7, pp. 2640–2651, 2007.

[16] A. Gjendemsj, D. Gesbert, G. E. Oien, and S. G. Kiani, “Binary power control for sum rate maximization over multiple interfering links,” *Wireless Communications, IEEE Transactions on*, vol. 7, no. 8, pp. 3164–3173, 2008.

[17] S. Stanczak, M. Wiczanowski, and H. Boche, *Fundamentals of resource allocation in wireless networks: theory and algorithms*. Springer, 2009, vol. 3.

[18] J. Ling, D. Chizhik, and R. Valenzuela, “On resource allocation in dense femto-deployments,” in *Microwaves, Communications, Antennas and Electronics Systems, 2009. COMCAS 2009. IEEE International Conference on*. IEEE, 2009, pp. 1–6.

[19] G. J. Foschini and Z. Miljanic, “A simple distributed autonomous power control algorithm and its convergence,” *Vehicular Technology, IEEE Transactions on*, vol. 42, no. 4, pp. 641–646, 1993.

[20] S. G. Kiani and D. Gesbert, “Maximizing the capacity of large wireless networks: optimal and distributed solutions,” in *Information Theory, 2006 IEEE International Symposium on*. IEEE, 2006, pp. 2501–2505.

[21] S. G. Kiani, G. E. Oien, and D. Gesbert, “Maximizing multicell capacity using distributed power allocation and scheduling,” in *Wireless Communications and Networking Conference, 2007. WCNC 2007. IEEE*. IEEE, 2007, pp. 1690–1694.

[22] S. G. Kiani and D. Gesbert, “Optimal and distributed scheduling for multicell capacity maximization,” *Wireless Communications, IEEE Transactions on*, vol. 7, no. 1, pp. 288–297, 2008.

[23] D. Gesbert, S. G. Kiani, A. Gjendemsjo *et al.*, “Adaptation, coordination, and distributed resource allocation in interference-limited wireless networks,” *Proceedings of the IEEE*, vol. 95, no. 12, pp. 2393–2409, 2007.

[24] C.-C. Chen and D.-S. Lee, “A joint design of distributed qos scheduling and power control for wireless networks.” in *INFOCOM*, 2006.

[25] V. Chandrasekhar, J. G. Andrews, T. Muharemovict, Z. Shen, and A. Gatherer, “Power control in two-tier femtocell networks,” *Wireless Communications, IEEE Transactions on*, vol. 8, no. 8, pp. 4316–4328, 2009.

[26] R. Etkin, A. Parekh, and D. Tse, “Spectrum sharing for unlicensed bands,” *Selected Areas in Communications, IEEE Journal on*, vol. 25, no. 3, pp. 517–528, 2007.

[27] Y. Wu, B. Wang, K. R. Liu, and T. C. Clancy, “Repeated open spectrum sharing game with cheat-proof strategies,” *Wireless Communications, IEEE Transactions on*, vol. 8, no. 4, pp. 1922–1933, 2009.

[28] K. Jain, J. Padhye, V. N. Padmanabhan, and L. Qiu, “Impact of interference on multi-hop wireless network performance,” *Wireless networks*, vol. 11, no. 4, pp. 471–487, 2005.

[29] H. Li, X. Xu, D. Hu, X. Qu, X. Tao, and P. Zhang, “Graph method based clustering strategy for femtocell interference management and spectrum efficiency improvement,” in *Wireless Communications Networking and Mobile Computing (WiCOM), 2010 6th International Conference on*. IEEE, 2010, pp. 1–5.

[30] S. Uygungelen, G. Auer, and Z. Bharucha, “Graph-based dynamic frequency reuse in femtocell networks,” in *Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd*. IEEE, 2011, pp. 1–6.

[31] P. Lee, T. Lee, J. Jeong, and J. Shin, “Interference management in lte femtocell systems using fractional frequency reuse,” in *Advanced Communication Technology (ICACT), 2010 The 12th International Conference on*, vol. 2. IEEE, 2010, pp. 1047–1051.

[32] C. W. Tan, S. Friedland, and S. H. Low, “Spectrum management in multiuser cognitive wireless networks: Optimality and algorithm,” *Selected Areas in Communications, IEEE Journal on*, vol. 29, no. 2, pp. 421–430, 2011.

[33] K. Han, S. Woo, D. Kang, and S. Choi, “Automatic neighboring bs list generation scheme for femtocell network,” in *Ubiquitous and Future Networks (ICUFN), 2010 Second International Conference on*. IEEE, 2010, pp. 251–255.

[34] D. Lopez-Perez, I. Guvenc, G. De La Roche, M. Kountouris, T. Q. Quek, and J. Zhang, “Enhanced intercell interference coordination challenges in heterogeneous networks,” *Wireless Communications, IEEE*, vol. 18, no. 3, pp. 22–30, 2011.

[35] Ö. Johansson, “Simple distributed $\Delta + 1$ coloring of graphs,” *Information Processing Letters*, vol. 70, no. 5, pp. 229–232, 1999.

[36] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and distributed computation: numerical methods*. Prentice-Hall, Inc., 1989.

[37] E. Wei and A. Ozdaglar, “On the o (1/k) convergence of asynchronous distributed alternating direction method of multipliers,” *arXiv preprint arXiv:1307.8254*, 2013.

[38] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

[39] S. Y. Seidel and T. S. Rappaport, “914 mhz path loss prediction models for indoor wireless communications in multifloored buildings,” *Antennas and Propagation, IEEE Transactions on*, vol. 40, no. 2, pp. 207–217, 1992.

[40] Z.-Q. Luo and S. Zhang, “Dynamic spectrum management: Complexity and duality,” *Selected Topics in Signal Processing, IEEE Journal of*, vol. 2, no. 1, pp. 57–73, 2008.

[41] R. M. Karp, “Probabilistic recurrence relations,” *Journal of the ACM (JACM)*, vol. 41, no. 6, pp. 1136–1150, 1994.