Dynamic Femtocell Resource Allocation for Managing Inter-Tier Interference in Downlink of Heterogeneous Networks

Arsalan Saeed, Efstathios Katranaras, Mehrdad Dianati, and Muhammad Ali Imran
Institute for Communication Systems (ICS), home of the 5G Innovation Centre
University of Surrey, United Kingdom
Email: {arsalan.saeed, efstathios.katranaras, m.dianati, m.imran}@surrey.ac.uk

Abstract

This paper investigates the downlink resource allocation problem in Orthogonal Frequency Division Multiple Access (OFDMA) Heterogeneous Networks (HetNets) consisting of macrocells and femtocells sharing the same frequency band. The focus is to devise optimised policies for femtocells' access to the shared spectrum, in terms of femtocell transmissions, in order to maximise femto users sum data rate while ensuring that certain level of quality of service (QoS) for the macro-cell users in the vicinity of femtocells is provided. The optimal solution to this problem is obtained by employing the well-known Dual Lagrangian method and the optimal femtocell transmit power and resource allocation solution is derived in detail. However, the optimal solution introduces high computational complexity and may not be feasible to apply in real-time systems. To this end, we propose a heuristic solution to the problem. The algorithms to implement both optimal and efficient suboptimal schemes in a practical system are also given in detail while their complexity is compared. Simulation results show that our proposed dynamic resource allocation scheme a) ensures the macro users QoS requirements compared to the Reuse-1 scheme, where femtocells are allowed to transmit at full power and bandwidth; b) can maintain femto user data rates at high levels, compared to the Orthogonal Frequency Reuse scheme, where the network bandwidth resources are partially divided amongst macro and femtocells; and c) provides performance close to the optimal solution, while introducing much lower complexity.

Index Terms

Heterogeneous Networks, Femtocells, Inter-cell Interference, Resource Allocation, Binary Integer Linear Programming, Dual Lagrangian Problem.

I. Introduction

Heterogeneous Networks (HetNets) comprising macro cells and densely deployed small cells are considered as a promising solution for future 5G networks [1]. It is indicated in [2] that dense deployment of Femto Access Points (FAPs)[1] can provide higher spectral efficiency, as compared to WiFi offloading. However, mass deployment of small cells overlayed within the area of larger cells raises challenges regarding their joint operation. FAPs can usually operate in two modes: Open Subscriber Group (OSG) and Closed Subscriber Group (CSG). OSG FAPs are deployed and owned by the network operator and operate as open cells to serve macrocell users in HotSpots or near the edges of the cells. This type of FAPs are simple to manage and have demonstrated to improve access network capacity [3]. CSG FAPs are typically owned by the subscriber and are open only to a long term managed list of users. On the other hand, CSG FAPs are easy to manage if they are operated in a separate licence free band similar to Wi-Fi. However, these FAPs, serving indoor subscribers as part of the operators network, need to be operated in a licensed band. Since the licensed spectrum resources are expensive and scarce, operators prefer to deploy these FAPs under the so-called co-channel deployment, i.e. by spatially reusing the available spectrum. As a trade-off, this sharing of the frequency band amongst the macrocell and CSG FAPs increases Inter-cell Interference (ICI) within the network which, if left unmanaged, may significantly deteriorate overall network performance [4]. This highlights the need for introduction of efficient low-complexity radio resource management techniques which can be implemented in practical systems.

A. Related Work

ICI problem has been widely discussed in literature, with focus initially targeted at homogeneous[2] macrocell scenarios. The simplest downlink frequency allocation technique is to share the whole available frequency band amongst multiple transmission nodes. This so-called Reuse-1 technique has the highest spectrum usage but also results in severe ICI experienced amongst the neighbouring cells. To reduce ICI, Fractional Frequency Reuse (FFR) schemes were initially introduced [5]. However FFR schemes reduce the spectrum usage and are mostly preplanned in nature, prohibiting adaptive frequency allocation to system dynamics. More recently, Dynamic Fractional Frequency Reuse (D-FFR) techniques have been introduced. In [6], a central broker is considered which constantly updates users into groups, based on their signal strength. These groups are assigned sub-carriers which are further used to serve the users in each group. This scheme employs low spatial reuse, hence reducing the overall throughput of the network. In [7], a dynamic graph based FFR scheme is discussed where neighbouring macrocells are assigned orthogonal chunks of spectrum based on the load on each cell. This approach results in a greedy and low spatial reuse, especially when heavily loaded cells require high number of Resource Blocks (RBs). Another approach for dynamic FFR is discussed in [8], where each cell aims to minimise its transmit power on each RB. This leads to each cell utilising only the RBs with best channel quality (least interference) to serve its users. A similar approach is shown in [9], where neighbouring nodes notify each other about their RB usage, so that they

[1]We interchangeably use the terms femtocell node and FAP in this work.

[2]By homogeneous networks, we indicate the networks with same size and same access technology cells.

avoid assigning high transmission power in those RBs. A two step solution approach is proposed in this work: Dynamic Frequency Planning (DFP) takes places in the first step to distribute chunks of frequency bands to the participant sharing macrocells; at the next step, a resource allocation algorithm is proposed to take place within each macrocell. Furthermore, the authors in [10] and [11] apply the aforementioned concept of minimising transmission power and discuss the use of interference tolerance estimation [3] for performing resource allocation in homogeneous macrocell and femtocell deployments, respectively.

The aforementioned techniques, being only designed for homogeneous scenarios, cannot perfectly fit to networks with underlaid macrocells and overlaid densely deployed small cells; the reason is that the dominant interferers for a user in the homogeneous scenario are limited and usually not as strong as in the dense HetNet scenario. Thus, focusing on the HetNet scenario and on the femto-femto interference, [12] suggests that FAPs should serve their users on RBs with the least measured pilot signal strength from neighbouring FAPs (hence the least femto-femto interference). Similarly, [13] proposes a technique where FAPs assign the top best RBs to their users and adjust their transmit power subject to FAP users QoS constraints.

Although femto-femto interference is a notable aspect in HetNet scenario, the degradation of performance for macrocell served users due to interference caused from FAPs to macrocell users will be more critical than in case of FAP users; since there are fewer users served by FAPs as compared to macrocells, FAP served users are anyway allocated with more bandwidth resources. Thus, regarding the interference from FAPs to macrocell users, [14] presents a bandwidth partitioning amongst macrocells and FAPs, where FAPs are not allowed to transmit in the bandwidth assigned to a macrocell, hence, reducing the spatial reuse. In [15], authors elaborate on the presence of CSG FAPs further elevating the issue of ICI as compared to public FAPs and discuss the use of shared, separate and partially shared bandwidth for this case. Furthermore [16] suggests the use of higher level modulation and coding schemes for indoor femtocells as their users generally realise good signal strengths. In [17], a scheme is proposed which zones FAP served users for either link adoption or requirement of orthogonal sub-bands and a central entity assigns the users with separate subbands from a pool. Finally, in [18], a mathematical framework is presented to minimise the interference from FAPs to macrocells. FAPs are allowed to transmit on certain RBs based on the calculated distance between the FAPs and neighbouring macrocells. However, for enhanced performance FAP muting decisions should be more adaptive to the system dynamics and consider the presence of macrocell served users in the vicinity of FAPs. To the best of our knowledge there is no such analysis in literature based on the idea of interference tolerance estimation in heterogeneous networks where FAPs pose interference to macrocell users. Authors in [10], [11] have applied the concept of interference tolerance estimation but only in case of homogeneous networks. This is a notable shortcoming, as macro victim users trapped in the vicinity of CSG FAPs suffer from severe interference [15].

[3] Sum of interference signal that a user can tolerate from neighbouring interferers in order to achieve a signal strength level.

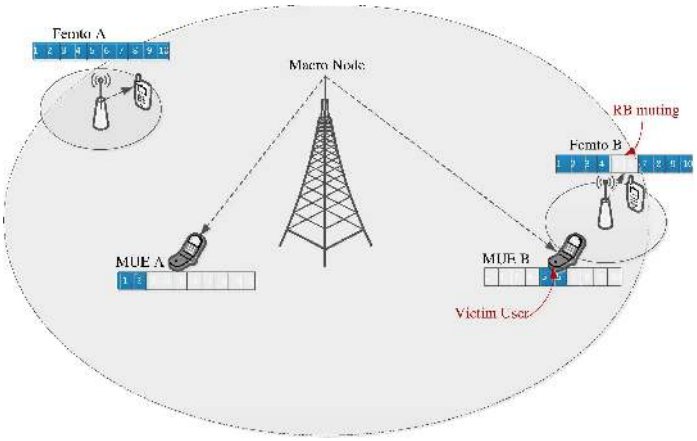

Fig. 1: Example of dynamic femtocell resource allocation for victim macro-user protection. Femtocell node A may use the full available resources while the transmission for femtocell node B is restricted in order to protect macrocell user B, which is in its vicinity at that specific time instance.

B. Contributions and Overview of the Paper

In this paper, we investigate the dynamic resource allocation problem for OFDMA heterogeneous networks by considering femto to macro inter-tier interference. Our objective is to improve the overall throughput of FAP served users without deteriorating the macro users performance by dynamically adjusting FAP resource allocation. This is a valid problem especially for the case where macro users happen to be in the vicinity of one or more interfering FAPs. We consider a scenario where the FAPs and the macrocell node are allowed to reuse the entire available bandwidth; however, in order to protect the macrocell served users from femtocell interference, a jointly optimised resource allocation scheme prohibits FAPs from accessing certain RBs. The general concept of the proposed approach is illustrated in Fig. 1.

The aforementioned problem is concisely formulated in this paper as a femtocell users' sum rate maximisation problem subject to minimum macro user rate requirement constraints. These constraints are translated into a maximum level of interference that each macro user can tolerate from all neighbouring FAPs. To determine the optimal FAP transmit power and RB allocation, the above Mixed Integer Non-linear Programming problem is relaxed to a computationally tractable dual problem and the well known Dual Lagrangian approach [19] is translated into this specific case. However, the optimal solution introduces high computational complexity for implementation in real-time systems. To this end, we consider a relaxed version of the original problem, where FAPs are either transmitting or being muted on each RB, and propose a low-complexity heuristic scheme to solve it. The big advantages of our proposed scheme are that: a) the optimisation problem can be solved considering the instantaneous throughput obtained in practical systems instead of the theoretical Shannon link capacity, thus, more practical aspects of the communication channel (such as modulation and coding scheme used) can be taken into account and evaluated; and b) provides gains close to the optimal solution with reasonable low complexity for practical implementation, despite the fact that power allocation per RB is kept constant. Focusing on the practical application of such dynamic

approaches, we analyse in detail how the optimal and heuristic schemes can be implemented in a real-world system such a Long-term evolution (LTE) networks and compare their computational complexities. Using Monte-Carlo simulations, we demonstrate that our proposed dynamic resource allocation scheme: a) ensures the macro users QoS, compared to Reuse-1 scheme; b) maintains femto user data rates at high levels, compared to the Orthogonal Frequency Reuse scheme.

The rest of this paper is organised as follows. Section II presents the system model. The mathematical formulation of the problem as a Dual Lagrangian problem is given in section III and the optimal solution is derived. Section IV introduces the efficient suboptimal RB allocation scheme while Section V presents the algorithms of the proposed schemes and compares their computational complexity. Numerical results and obtained insights are discussed in Section VI. Finally, Section VII concludes the paper.

II. System Model

We consider a system of $M+1$ cells, comprising one macrocell (identified as cell 0) and M femtocells within the macrocell area. The set of femtocells is defined as $\mathcal{M} = \{1, \ldots, M\}$. We assume that there are K active users in the system. We consider that each user can have only one serving node, but each cell can support multiple users; thus, $K \triangleq |\mathcal{K}| = |\mathcal{K}_0 \cup \mathcal{K}_1 \cdots \cup \mathcal{K}_M|$, where $\mathcal{K}$ denotes the set of all users in the system and $\mathcal{K}_m$ denotes the set of users served by node in cell m.

Following the binary RB allocation nature of the OFDMA systems the total system bandwidth is divided in N RBs and each RB can be allocated to only one user in each cell. Macrocell node can allocate all the available RBs to its associated *macro-users* (MUE). Moreover, macrocell users are assumed to have minimum data rate requirements.

On the other hand, femtocell nodes reuse the same resources to serve their *femto-users* (FUE) based on a resource allocation policy. We consider a central entity residing at the macrocell node which is able to collect relevant information to make resource allocation decisions and guide femtocells on the resource allocation policy to be adopted. Such a deployment could be considered semi-distributed since for a multi macrocell system their could be a central entity present at each macrocell, guiding the underlying femtocells.

We define binary indicator variables $\phi_{k,m,n} \in \{0, 1\}$, where $\phi_{k,m,n} = 1$ when femtocell m serves its k^{th} assigned user in the n^{th} RB; otherwise, the RB allocation parameters take the zero value. Thus, we can define the vector containing all RB allocation parameters $\boldsymbol{\phi} = [\phi_{1,1,1} \ldots \phi_{K_M,M,N}]$, which characterizes the femtocells *RB allocation policy*. Moreover, transmit power of the m^{th} femtocell in the n^{th} RB is denoted by $p_{m,n} \leq P_{\max}$, where $P_{\max}$ is the maximum allowed transmission power of any femtocell. Vector $\mathbf{p} = [p_{1,1} \ldots p_{M,N}]$ characterizes the femtocells *power allocation policy*.

A. User SINR and Rate Modelling

The SINR of MUE or FUE users can be modelled as follows. Using index 0 as macrocell identification, the SINR of the u^{th} MUE at RB n can be given by:

$$\gamma_{u,0,n} = \frac{p_{0,n}\Gamma^{0}_{u,0,n}}{\sum_{m=1}^{M}\left(\sum_{k\in\mathcal{K}_m}\phi_{k,m,n}\right)p_{m,n}\Gamma^{m}_{u,0,n} + N_0 B}, \tag{1}$$

where $p_{0,n}$ denotes the transmit power of macrocell node at RB n, $\Gamma^{i}_{k,m,n}$ is the channel gain between base station at cell i and user k being served at cell m in RB n, N_0 is the noise power spectral density and B is the bandwidth of each RB.

Similarly, the SINR of FUE k in cell m at RB n can be given by:

$$\gamma_{k,m,n} = \frac{p_{m,n}\Gamma^{m}_{k,m,n}}{p_{0,n}\Gamma^{0}_{k,m,n} + \sum\limits_{\substack{i=1 \\ i\neq m}}^{M}\left(\sum\limits_{l\in\mathcal{K}_i}\phi_{l,i,n}\right)p_{i,n}\Gamma^{i}_{k,m,n} + N_0 B}. \tag{2}$$

The rate of each user (FUE or MUE) can be expressed by the Shannon-Hartley Theorem as follows:

$$R_{k,m,n} = B\ \log_2\left(1+\gamma_{k,m,n}\right). \tag{3}$$

It should be noted that although (3) is not a practically achievable rate, it is used as a performance indicator for comparison purposes.

B. Maximum Interference Allowance

In this sub-section we formulate the *maximum interference allowance*, which is defined as the maximum amount of interference (sum of interference from all neighbouring transmitters) that a user can tolerate for a given minimum data rate demand. The minimum data rate demand for a MUE can be translated into a minimum data rate demand at each RB, allocated to that specific MUE. Moreover, the minimum MUE demand data rate at RB n can be translated into a specific minimum required $\gamma^{\text{req}}_{u,0,n}$ SINR value [10]. Having identified the minimum SINR value and considering (1) we can find the maximum interference power $\Omega^{\max}_{u,n}$ that MUE u can tolerate in RB n from all femtocell nodes to obtain this rate threshold:

$$\Omega^{\max}_{n} = \frac{p_{0,n}\Gamma^{0}_{u,0,n}}{\gamma^{\text{req}}_{u,0,n}} - N_0 B. \tag{4}$$

If the potential channel gain from any femtocell m to the MUE is denoted as $\Gamma^{(m)}_{0,u,n}$, the total interference caused to it by all femtocells in each RB can be given by:

$$\Omega^{\text{sum}}_{n} = \sum_{m=1}^{M}\left(\sum_{k\in\mathcal{K}_m}\phi_{k,m,n}\right)p_{m,n}\Gamma^{m}_{0,u,n} = \sum_{m=1}^{M}\left(\sum_{k\in\mathcal{K}_m}\phi_{k,m,n}\right)\omega^{m}_{0,u,n}, \tag{5}$$

where $\omega^{m}_{0,u,n} \triangleq p_{m,n}\Gamma^{m}_{u,0,n}$ can be interpreted as the interference that is caused to user u in cell 0 (macrocell) on RB n from femtocell m.

III. Optimal Resource Allocation (ORA)

Our problem is defined as a maximisation of the sum rate of all active users in the femtocells, while: 1) the individual rate of any MUE is ensured to be greater than a minimum value and; 2) FAP transmit power as well as RB allocation constraints are satisfied.

The achievable sum rate of all active users in femtocells over the whole allocated system bandwidth is given by:

$$R = \sum_{n=1}^{N}\sum_{m=1}^{M}\left(\sum_{k\in\mathcal{K}_m}\phi_{k,m,n}R_{k,m,n}\right), \tag{6}$$

where $R_{k,m,n}$ denotes the achievable rate of k^{th} user served by femtocell m on RB n. From equation (6), considering also (2) and (3), it can be observed that the FUEs sum rate is a function of both femtocell RB and power allocation policy, i.e. $R = f(\boldsymbol{\phi}, \mathbf{p})$. In the following we formulate the respective sum rate optimisation problem and examine its solution.

A. Problem Formulation and Solution Approach

The general sum rate optimisation problem comprising the objective function and the imposed constraints can be formulated as follows:

$$\max_{\mathbf{p},\boldsymbol{\phi}} R \tag{7}$$

subject to:

$$\phi_{k,m,n} \in \{0,1\}, \forall k \in \mathcal{K} \setminus \mathcal{K}_0, m \in \mathcal{M}, n; \tag{7a}$$

$$\sum_{k \in \mathcal{K}_m} \phi_{k,m,n} \in \{0,1\}, \forall m \in \mathcal{M}, n; \tag{7b}$$

$$\Omega_n^{\text{sum}} \leq \Omega_n^{\max}, \quad \forall n; \tag{7c}$$

$$\sum_{n=1}^{N} \left(\sum_{k \in \mathcal{K}_m} \phi_{k,m,n} \right) p_{m,n} \leq P_{\max}, \forall m \in \mathcal{M}; \tag{7d}$$

$$p_{m,n} \geq 0, \forall m \in \mathcal{M}, n. \tag{7e}$$

Constraint (7b) indicates that RBs are exclusively allocated to one user served by each cell pair to avoid intra-cell interference; constraint (7c) denotes the total maximum interference that a MUE served by macrocell on RB n can tolerate from all femtocells in the macro area in order to satisfy its minimum rate needs; finally, constraints (7d)-(7e) stand for the maximum and minimum transmission power constraints at each femtocell node.

The optimisation problem in (7) contains both continuous[4] ($\mathbf{p}$) and binary ($\boldsymbol{\phi}$) decision variables and it is categorised in general as a mixed integer nonlinear programming problem (MINP) since the objective function (R) is nonlinear in $\mathbf{p}$ considering equations (2) and (3). Finding the optimal solution to these non-convex problems requires computationally complex exhaustive search, rendering its implementation in practical systems impossible and becomes even harder when QoS constraints are added on top (as is the case here with the minimum MUE rate constraints). However, to make the problem tractable, we relax the resource allocation integer constraints to take any real value between 0 and 1. This *time-sharing condition* essentially considers the time sharing of each subcarrier in practice and it is proved in [20] that the duality gap of any optimisation problem satisfying the time sharing condition is negligible as the number of subcarriers becomes sufficiently large. Therefore, our relaxed optimisation problem of (7) can be solved optimally by using the dual method [19], [20].

[4]Considering that femtocells allocate power to RBs according to some predefined power levels, vector $\mathbf{p}$ can instead contain integer variables. This of course renders the optimisation problem even harder to solve.

B. Dual Method for Optimal Joint Power and RB Allocation

The dual method applied in our case will comprise the following steps [19]: a) translating the original optimisation problem into its Lagrangian dual, associating QoS and power constraints with dual variables; b) decomposing the dual problem into independently solvable subproblems by removing the coupling between RBs via Lagrangian relaxation; c) further decomposing the subproblems through a two phase second level primal decomposition where power and RB allocation optimisation is performed sequentially and; d) using the subgradient method to iteratively update the dual variables in parallel until they (and essentially the original problem) converge into the optimal values. In the following, the various steps of the dual method are presented in detail.

1) Dual Problem: The Lagrangian function of the problem in (7) can be given by:

$$\begin{aligned}
&L(\boldsymbol{\phi}, \mathbf{p}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \\
&= \sum_{n=1}^{N} \sum_{m \in \mathcal{M}} \sum_{k \in \mathcal{K}_m} \phi_{k,m,n} R_{k,m,n} + \sum_{n} \lambda_n \Big(\sum_{m=1}^{M} \Omega_n^{\max} - \big(\sum_{k \in \mathcal{K}_m} \phi_{k,m,n} \big) p_{m,n} \Gamma_{0,u,n}^{m} \Big) \\
&\quad + \sum_{m} \mu_m \Big(P_{\max} - \sum_{n} \big(\sum_{k \in \mathcal{K}_m} \phi_{k,m,n} \big) p_{m,n} \Big) \\
&= \sum_{n} \left[\sum_{m} \sum_{k \in \mathcal{K}_m} \phi_{k,m,n} R_{k,m,n} - \lambda_n \Big(\sum_{m=1}^{M} \big(\sum_{k \in \mathcal{K}_m} \phi_{k,m,n} \big) p_{m,n} \Gamma_{0,u,n}^{m} \Big) - \sum_{m} \mu_m \big(\sum_{k \in \mathcal{K}_m} \phi_{k,m,n} \big) p_{m,n} \right] \\
&\quad + \sum_{n} \lambda_n \Omega_n^{\max} + \sum_{m} \mu_m P_{\max},
\end{aligned} \tag{8}$$

where $\boldsymbol{\lambda} = [\lambda_1, ...\lambda_N]$ and $\boldsymbol{\mu} = [\mu_1, ...\mu_M]$ are the dual variable vectors associated with the individual interference constraints on MUEs and the femtocells transmit power constraint, respectively. The Lagrangian dual function can be given as:

$$g(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \begin{cases} \max\limits_{\boldsymbol{\phi}, \mathbf{p}} \quad L(\boldsymbol{\phi}, \mathbf{p}, \boldsymbol{\lambda}, \boldsymbol{\mu}), \\ s.t. \\ 0 \leq \phi_{k,m,n} \leq 1, \forall k \in \mathcal{K} \setminus \mathcal{K}_0, m \in \mathcal{M}, n; \\ \sum\limits_{k \in \mathcal{K}_m} \phi_{k,m,n} \leq 1, \forall m \in \mathcal{M}, n; \\ p_{m,n} \geq 0, \forall m \in \mathcal{M}, n. \end{cases} \tag{9}$$

Hence, the dual optimisation problem is formulated as:

$$\min_{\boldsymbol{\lambda}, \boldsymbol{\mu} \geq 0} \quad g(\boldsymbol{\lambda}, \boldsymbol{\mu}). \tag{10}$$

2) Decomposition: The coupling between RBs can be removed by Lagrangian relaxation and equation (9) can be decomposed into N subproblems at each RB with each subproblem given as:

$$\max_{\boldsymbol{\phi},\mathbf{p}} L_n\left(\boldsymbol{\phi}_n, \mathbf{p}_n\right) \quad = \quad \begin{cases} \max\limits_{\boldsymbol{\phi},\mathbf{p}} \sum\limits_{m} \sum\limits_{k\in\mathcal{K}_m} \phi_{k,m,n} R_{k,m,n} \\ -\lambda_n \left(\sum\limits_{m=1}^{M} \left(\sum\limits_{k\in\mathcal{K}_m} \phi_{k,m,n} \right) p_{m,n} \Gamma^m_{0,u,n} \right) \\ -\sum\limits_{m} \mu_m \left(\sum\limits_{k\in\mathcal{K}_m} \phi_{k,m,n} \right) p_{m,n} \end{cases} \tag{11}$$

$$s.t.$$

$$0 \leq \phi_{k,m,n} \leq 1, \forall k \in \mathcal{K} \setminus \mathcal{K}_0, m \in \mathcal{M}, n;$$

$$\sum_{k\in\mathcal{K}_m} \phi_{k,m,n} \leq 1, \forall m \in \mathcal{M}, n;$$

$$p_{m,n} \geq 0, \forall m \in \mathcal{M}, n,$$

where $\boldsymbol{\phi}_n \triangleq [\phi_{1,1,n} \dots \phi_{k,m,n}]$ and $\mathbf{p}_n \triangleq [p_{1,n} \dots p_{m,n}]$. This dual problem can be further decomposed through a second level primal decomposition and solved in two phases: optimal power allocation and optimal RB allocation.

3) Optimal Power Allocation for a Given RB Allocation: Let for RB n, $\phi_{k,m,n} = 1$. Then, optimal power allocation over this RB can be determined by the following problem:

$$\begin{aligned} \max_{\mathbf{p}_{m,n}} \quad & L_n, \forall m \\ s.t. \quad & p_{m,n} \geq 0. \end{aligned} \tag{12}$$

In the following, without loss of generality, we consider the scenario where femto-femto interference is negligible compared to macro-femto interference to simplify mathematical analysis. This assumption is generally valid in scenarios with femtocells overlaid by a macrocell and users are also provisioned to be served by the macrocell when femtocell coverage is weak. In that case the rate of each FUE (and subsequently R) becomes linear in $\mathbf{p}$:

$$R_{k,m,n} = B \ \log_2\left(1 + \frac{p_{m,n}\Gamma^m_{k,m,n}}{p_{0,n}\Gamma^0_{k,m,n} + N_0 B}\right) \triangleq B \ \log_2\left(1 + p_{m,n} \ \alpha_{k,m,n}\right). \tag{13}$$

Thus, we substitute the rate equation (13) in (12) and differentiate L with respect to $p_{m,n}$, getting:

$$\frac{\partial L}{\partial p_{m,n}} = \frac{\alpha_{k,m,n}}{ln(2)(1 + p_{m,n}\,\alpha_{k,m,n})} - \lambda_n \Gamma^m_{0,u,n} - \mu_m. \tag{14}$$

Furthermore, applying the Karush-Kuhn-Tucker (KKT) condition [19], the optimal power allocation can be obtained by setting (14) equal to zero as follows:

$$p^*_{m,n} = \left[\frac{1}{ln(2)(\lambda_n \Gamma^m_{0,u,n} + \mu_m)} - \frac{1}{\alpha_{k,m,n}} \right]^+, \tag{15}$$

where $[x]^+ = max[x, 0]$. This process is explained in detail in Appendix A.

4) Optimal RB allocation: By eliminating the power variable in equation (12) and substituting into equation (8), the dual function can be alternatively expressed as:

$$g(\boldsymbol{\lambda},\boldsymbol{\mu}) = \begin{cases} \max\limits_{\boldsymbol{\phi}} \sum\limits_{n}\sum\limits_{m}\sum\limits_{k\in\mathcal{K}_m} \phi_{k,m,n} H_{k,m,n}(\boldsymbol{\lambda},\boldsymbol{\mu}) \\ +\sum\limits_{n}\lambda_n \Omega_n^{\max} + \sum\limits_{m}\mu_m P_{\max}, \\ s.t. \\ 0 \le \phi_{k,m,n} \le 1, \forall k \in \mathcal{K}\setminus\mathcal{K}_0, m \in \mathcal{M}, n; \\ \sum\limits_{k\in\mathcal{K}_m} \phi_{k,m,n} \le 1, \forall m \in \mathcal{M}, n \end{cases}, \quad (16)$$

where the function $H_{k,m,n}(\lambda,\mu)$ is given by:

$$H_{k,m,n} = \log_2(1 + p^*_{n,m}\alpha_{k,m,n}) - \lambda_n p^*_{n,m}\Gamma^m_{0,u,n} - \mu_m p^*_{n,m}. \quad (17)$$

Here, $H_{k,m,n}$ can be regarded as the potential profit or loss from femtocell m transmitting to its k^{th} user on RB n. Intuitively we can define the first term of the expression as the maximum achieved rate of a user if its serving femtocell transmits on RB n, the second term as the interference penalty and the third term as the power constraint price. Thus, the optimal RB allocation will be obtained according to the following criterion:

$$\phi^*_{k,m,n} = \begin{cases} 1, k^* = \arg\max\limits_{k\in\mathcal{K}_m} H_{k,m,n} \text{ and } H_{k^*,m,n} > 0 \\ 0, \text{otherwise}, \quad \forall m \in \mathcal{M}, n. \end{cases} \quad (18)$$

5) Variable Update: As the dual function in equation (9) is convex by definition, the subgradient method is used to minimise $g(\boldsymbol{\lambda},\boldsymbol{\mu})$ [19]. Thus, dual variable vectors $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ are updated in parallel using the appropriate subgradients of $g(\boldsymbol{\lambda},\boldsymbol{\mu})$ at each iteration (see Appendix B):

$$\lambda_n(i+1) = \left[\lambda_n(i) + \psi(i)\left(\sum_{n}\sum_{m=1}^{M}\Big(\sum_{k\in\mathcal{K}_m}\phi^*_{k,m,n}\Big)p^*_{m,n}\Gamma^m_{0,u,n} - \Omega_n^{\max}\right)\right], \quad (19)$$

$$\mu_m(i+1) = \left[\mu_m(i) + \kappa(i)\left(\sum_{n}\Big(\sum_{k\in\mathcal{K}_m}\phi^*_{k,m,n}\Big)p^*_{m,n} - P_{\max}\right)\right]. \quad (20)$$

where, $\psi(i)$ and $\kappa(i)$ are the diminishing step sizes and i denotes the iteration index. If the step sizes are selected according to the diminishing step size policy [20], the subgradient method converges to the optimal dual variables, thus, the optimal joint power and RB allocation can be computed algorithmically.

The process of the decomposed dual problem is shown in Fig. 2. It can be summarised that the dual problem decomposition includes the following steps: the power allocation values are calculated using equation (15), which are then replaced into equation (17) to determine the profit matrix H; further equation (18) is used to determine the optimal pair of power and transmitting femtocell for each RB.

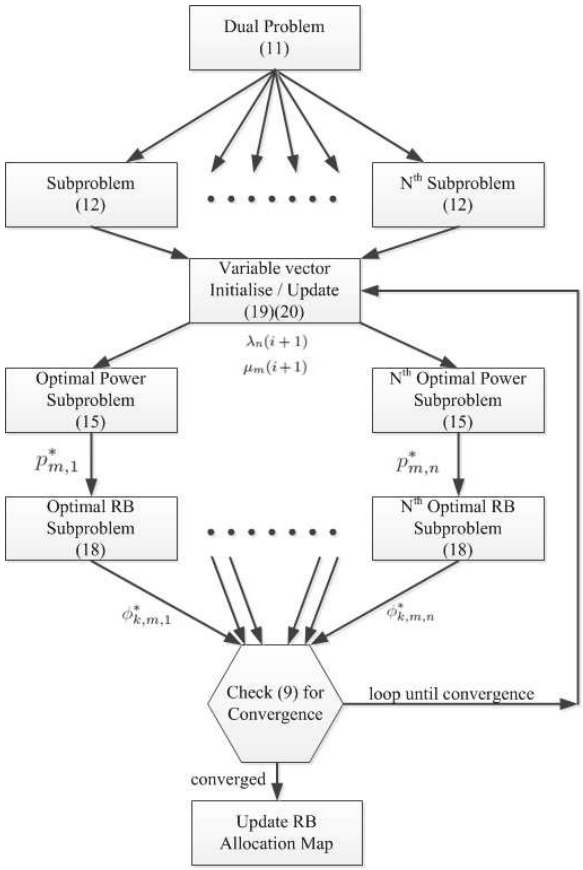

Fig. 2: Dual problem decomposition flow diagram.

IV. Efficient Suboptimal RB Allocation (ESRA)

The computational complexity of the ORA scheme will still be high for implementation in a real system when the number of femtocells and users per cell grows large. To address the complexity issues of ORA scheme, we further propose a heuristic efficient suboptimal RB allocation (ESRA) scheme. As will be shown in later sections, the proposed ESRA scheme significantly reduces the computational complexity with minimal degradation in the performance compared to ORA scheme. Computational complexity of ORA and ESRA schemes are further discussed in Section V.

In order to simplify the problem in (7), we focus only on the RB allocation. To this end, we assume maximum transmit power at femtocell nodes and equal power allocation across RBs, i.e. $p_{m,n} = \frac{P_{\max}}{N}$ for any femtocell m. In that case, the sum rate maximisation problem is transformed into a pure binary linear optimisation problem (BLP) which is formulated as follows:

$$\max_{\boldsymbol{\phi}} \sum_{n=1}^{N} \sum_{m=1}^{M} \left(\sum_{k \in \mathcal{K}_m} \phi_{k,m,n} R_{k,m,n} \right), \tag{21}$$

subject to:

$$\phi_{k,m,n} \in \{0,1\}, \forall m,n,k; \tag{21a}$$

$$\sum_{k \in \mathcal{K}_m} \phi_{k,m,n} \in \{0,1\}, \forall m,n; \tag{21b}$$

$$\sum_{m=1}^{M} \left(\sum_{k \in \mathcal{K}_m} \phi_{k,m,n} \right) p_{m,n} \Gamma^{m}_{0,u,n} \leq \Omega^{\max}_{u,n}, \forall u \in \mathcal{K}_m, n. \tag{21c}$$

The key benefit of the efficient sub-optimal RB allocation scheme is expected to come from the significant reduction of the optimisation problem search space by considering only RB allocation. This reduces the complexity and convergence time of the problem; hence, it can be easily solved for multiple or even every Transmission Time Interval (TTI) in LTE networks. The computational complexity comparison for ORA and ESRA schemes is presented in Fig. 3. It can be seen that as we increase the number of FAPs in the network, the computational complexity increase for both the scheme. However for ESRA, this increase is relatively negligible as compare to ORA scheme.

An additional significant benefit offered by the efficient ESRA scheme, apart from the reduced complexity, is that the optimisation problem in (21) can be solved considering the instantaneous throughput obtained in practical OFDMA systems instead of the theoretical Shannon link capacity of (3). In general, the instantaneous throughput of any user k served by femtocell m on RB n in OFDMA systems can be given as [10]:

$$\hat{R}_{k,m,n} = \text{BR}\left(r\right) \cdot \left[1 - \text{BLER}\left(r, \gamma_{k,m,n}\right)\right], \tag{22}$$

where BR is the theoretical bit rate for any MCS r when there are no errors which is depended on the network configuration, i.e. for N^{n}_{SC} number of data sub-carriers per RB, N^{n}_{SY} number of symbols per RB, RB's duration T^{n}_{RB} and e_r efficiency (in bits per symbol) of MCS r allocated to the user of interest, the BR for MCS r is given by:

$$\text{BR}\left(r\right) = \frac{N^{n}_{\text{SC}} N^{n}_{\text{SY}}}{T^{n}_{\text{RB}}} \cdot e_r \ . \tag{23}$$

Moreover, BLER denotes the block error rate suffered by this user on RB n which is a function of the realised SINR and the MCS used.

Similarly the instantaneous throughput of any MUE u served on RB n can be given as:

$$\hat{R}_{u,0,n} = \text{BR}\left(r\right) \cdot \left[1 - \text{BLER}\left(r, \gamma_{u,0,n}\right)\right]. \tag{24}$$

As discussed in the previous section, a minimum overall data rate demand for a MUE can be translated into a minimum data rate demand at each RB. Moreover, according to equation (24), the minimum MUE demand data rate at RB n can be translated into a specific MCS ($r_{\min}$) that has to be used and a minimum required $\gamma^{\text{req}}_{u,0,n}$ SINR value. Having identified the minimum SINR value and considering equation (1) we can find the maximum interference power $\Omega^{\max}_{n}$ that MUE being served on RB n can tolerate from all femto base stations to obtain this rate threshold.

V. Algorithms and Implementation

In this section, a high level description is provided on how the investigated optimal and suboptimal resource allocation schemes can be implemented in LTE heterogeneous networks comprising macrocells and femtocells. The following arguments explain how the key functions and elements of LTE architecture can be used for this reason.

A: UEs report their Channel Quality Indicator (CQI) and demand rate to their serving cells on frequent basis which determines the user channel gain on that specific RB. Based on these reports received from MUEs, equations (3) and (4) can be used to estimate the maximum interference, $\Omega_n^{\max}$, that a MUE can tolerate on a certain RB. Note that this estimation will also decide the MCS and Transport Block (TB) size of the future transmissions from the serving node to that UE.

B: UEs also report to their serving cell, the neighbouring cell's Reference Signal Received Quality (RSRQ) along with the Physical Cell ID (PCI) of the neighbouring cell. These reports are generally used for A2, A3 and A4 measurements based handovers. In our case, the respective MUE reports can be used to estimate the top neighbouring interfering femtocells; then, this information can be used to estimate the total interference caused to it by all femtocells in each RB, $\Omega_n^{\max}$, and formulates the optimisation constraint (7c).

C: Moreover, the addition of X2 logical interface in LTE provides the means for cells to communicate. Amongst the macrocell and the neighbouring femtocells, X2 can act as an interface to guide the neighbouring femtocells to restrict their transmissions. Thus, X2 interface can be used to input each femtocell utility (i.e. expected rate of FUEs in the femtocell based on equation (3)) at each RB to the central entity at the macrocell. The input from all femtocells, formulates our objective function in (6) (i.e. expected sum rate of all FUEs in the system).

D: Finally, the optimisation process of either the optimal problem in (7) or the suboptimal problem in (21) is performed at the central entity. The optimisation function returns $\phi_{k,m,n}$ and $p_{m,n}$ for the optimal case and only $\phi_{k,m,n}$ for the suboptimal case. These parameters are passed to femtocells over the X2 interface and act as a restriction matrix for each femtocell. Furthermore, in order to avoid introducing unnecessary control overheads into the network, restriction matrix can only be forwarded subject to change in the optimisation parameters, $\phi_{k,m,n}$ and $p_{m,n}$. In that case, femtocells continue to use the last updated restriction matrix until a new update is passed by the central entity.

Our proposed solutions can be considered for semi-distributed implementation case of a practical multi macrocell system i.e. a central entity could be placed at each macrocell which guides the under laying femtocells in a distributed manner. Furthermore, spectrum sensing techniques such as wideband CQI sensing [21] could be employed for systems with limited coordination possibilities. The following tables provide a summarising pseudocode for the processes required at each scheme.

The time complexity of the optimal exhaustive search in our case shall be $O(2^{(KMN)})$, which is exponential in nature. However, for the ORA scheme, the complexity is mainly dependent on solving the dual problem. The number of computations required to solve the RB allocation is $K(M+1)$ and N number of allocations are required to solve for all RBs. The complexity for each complete iteration is $O(NK(M+1))$. The total complexity of the subgradient method is polynomial in the number of dual variable, and is $O(N+M)$. Therefore, the overall complexity of the

Algorithm 1 ORA Scheme

Calculate : $\Omega_{u,n}^{\text{sum}}, R_{k,m,n}, \Omega_{u,n}^{\max}$ using eq (4), (5) and (13)
Initialize $\lambda_n \quad and \quad \mu_m$
while $g(\lambda, \mu)$ is not converged in eq (9), **do**
  Calculate : $P_{m,n}^*$ in eq (15)
  Calculate : $H_{k,m,n}$ using eq (17)
  Update : $\phi_{k,m,n}$ using eq (18)
    If $H_{k,m,n} > 0$ **then** update $\phi_{k,m,n} = 1,\ 0\ otherwise$
  Calculate : $g(\lambda, \mu)$ using eq (16)
  Update : λ_n and μ_m using eq (19) and (20)
end
Notify neighbouring femtocells with $P_{m,n}^*$ and $\phi_{k,m,n}$

Algorithm 2 ESRA Scheme

Calculate : $\Omega_{u,n}^{\text{sum}}, \hat{R}_{k,m,n}, \Omega_{u,n}^{\max}$ using eq (4), (5) and (22)
$[\phi_{k,m,n}, \hat{R}_{k,m,n}] = bintprog(\hat{R}_{k,m,n}, \Omega_{u,n}^{\max}, \Omega_{u,n}^{\text{sum}})$
Notify neighbouring femtocells with $\phi_{k,m,n}$

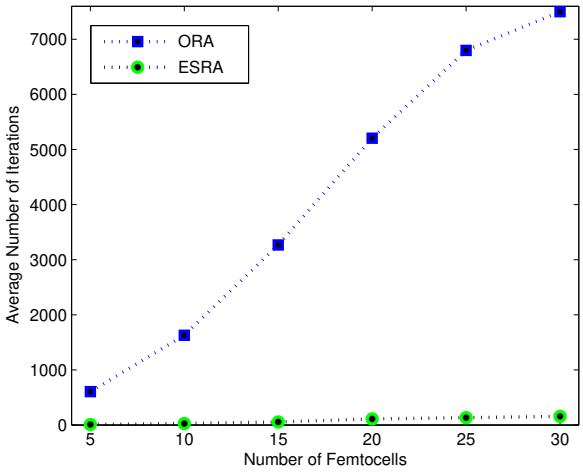

Fig. 3: Computational complexity comparison between ORA and ESRA scheme.

ORA scheme is $O((N+M)^2(NMK))$. The ESRA scheme is solved by binary linear integer programming. There are several linear programming relaxations applied to such algorithms, which make them very effective in practice but it is difficult to prove theoretical complexity bounds on the performance of such algorithms. A comparison in terms of number of iterations between the ORA and ESRA scheme is presented in Fig. 3, emphasizing on the lower complexity, therefore, higher the practicality of the ESRA scheme.

VI. Simulation Results

In this section, we evaluate the performance of the dynamic resource allocation schemes in the context of a real-world cellular network scenario. We simulate a single LTE macrocell with a fixed number of users attached to it and several femtocells, within the operational area of the macrocell. Macrocell serves the MUEs with a persistent scheduling (resource allocation within the macrocell remains fixed for multiple frames). On the other hand each femtocell has a single user attached to it, being served with the potential to use all the available RBs. Further details of the simulation parameters are given in Table I.

TABLE I: LTE-Based Scenario - Simulation Parameters.

Parameter	**Macro**	**Femto**
Number of nodes	1	5
Carrier frequency	2.1 GHz	
Bandwidth	10 MHz	
Node transmit power	43 dBm	23 dBm
Path loss model	$128.1 + 37.6 \log_{10}(d[\text{Km}])$	
Number of UEs	5	1 UE per FAP
Noise Figure at UE	9 dB	
Thermal noise density	-174 dBm/Hz	
Cell Radius	800m	50m

In order to evaluate the average performance of the ORA and ESRA scheme, we first consider a large number of system snapshots with uniform distribution of randomly deployed MUE and FAP nodes within the macrocell area at each snapshot. We also compare the performance of the proposed dynamic resource allocation schemes with the two benchmark cases: a) Reuse-1, where macrocell and FAPs transmit on all the RBs, and b) Orthogonal Reuse, where 50% of the RBs are reserved for macrocell and the remaining 50% RBs are shared amongst FAPs. At the second step we validate and compare in more detail, the operation of ORA and ESRA schemes considering deterministically placed nodes.

A. Randomly Placed Nodes

To evaluate and compare the overall performance of the proposed schemes we find the achieved MUE and FUE rates for a large number of uniform random MUE and FAP node placement scenarios. Results are averaged for 10^3 independent system snapshots.

Fig. 4 and 5 show the cumulative distribution function (CDF) of the achievable MUE data rates for a MUE demand of 0.5Mbps and 1.2Mbps respectively. It can be observed that Reuse-1 scheme results into a MUE outage (i.e. when the MUE achieved rate is below the demand rate) of 20% and 50% respectively. Orthogonal Reuse also results into a 10% MUE outage but only at higher MUE demand rate (1.2Mbps). On the other hand, ORA and ESRA schemes successfully eliminate MUE outage. Moreover, comparing ORA and ESRA schemes performance, we observe that ORA scheme manages to keep the MUE achieved data rates close to the MUE demand; nearly 55% and 90% MUE achieved data rates for 0.5Mbps and 1.2Mbps MUE demand respectively. While in ESRA,

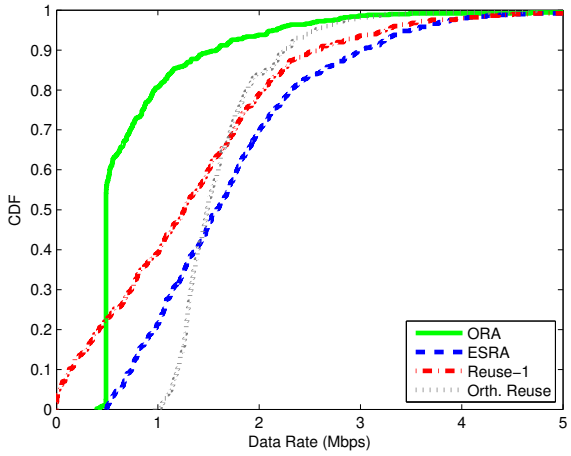

Fig. 4: CDF of MUE data rates for MUE Demand of 0.5Mbps.

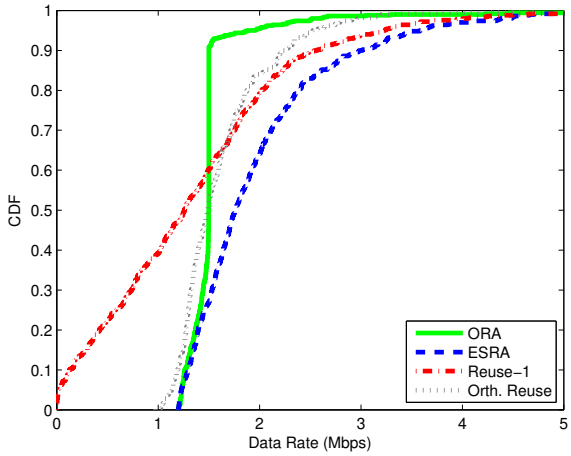

Fig. 5: CDF of MUE data rates for MUE Demand of 1.2Mbps.

MUE rates are always above the demand. The ORA approach is beneficial to facilitate FAPs to maximize RB usage to serve their users. This behaviour is clearly depicted in Fig. 6 and 7, where the CDF of FUE data rates is presented for the same MUE demand rates. ORA scheme has a similar performance to that of Reuse-1, whereas ESRA scheme slightly lags behind. It can also be observed from these latter plots that ORA and ESRA schemes outperform Orthogonal Reuse scheme at higher percentiles in terms of achieved FUE data rates. This is due to the fact that since FFR scheme is static in nature, the reserved bandwidth for MUEs may not be fully utilised when the MUE demand is low. Specific numerical values supporting the aforementioned observations are presented in Table II, where the 50^{th} and 95^{th} percentile average MUE and FUE rates are given for the various MUE demand rates.

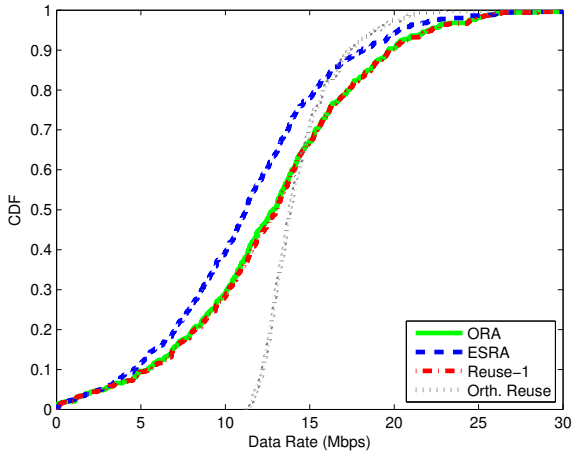

Fig. 6: CDF of FUE data rates for MUE Demand of 0.5Mbps.

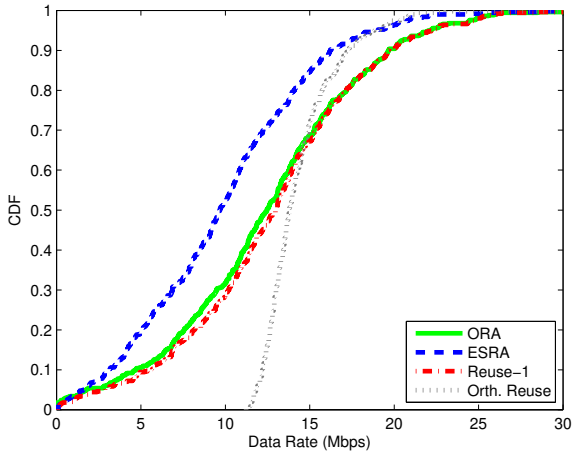

Fig. 7: CDF of FUE data rates for MUE Demand of 1.2Mbps.

B. Fixed Node Locations

In order to present an in-depth working of the ORA and ESRA scheme we simulate a deterministic case with fix node locations. The purpose of such analysis is to demonstrate that ORA scheme has higher liberty in terms of optimising the transmit power level as well as the RB allocation, on the other hand ESRA scheme reduces the complexity by only optimising the RB allocation. The placement of the nodes in the static scenario is illustrated in Fig. 8, where all the femtocell nodes are placed close to the MUEs except for one, i.e. FAP-3.

In this scenario all the femtocell nodes are placed close to the MUEs, except for one, i.e. FAP-3. For clearer presentation, we consider only 10 RBs in total for this case and assume that each MUE is assigned two RBs in a numeric order, i.e. MUE-x is assigned RB-(2x-1) and RB-(2x).

To this end, Fig. 9 (a) and (b) depict the achieved MUE data rates, for a MUE demand of 0.2Mbps: for ORA

TABLE II: Performance Comparison of various resource allocation schemes.

		Schemes			
		ORA	ESRA	Reuse-1	Orth. Reuse
MUEs below demand rate [%]					
MUE Demand (Mbps)	*0.2*	0.0	0.0	10.8	0.0
	0.5	0.0	0.0	22.2	0.0
	0.8	0.0	0.0	29.4	0.0
	1.2	0.0	0.0	47.2	0.0
MUE rate 50th percentile [Mbps]					
MUE Demand (Mbps)	*0.2*	0.48	1.54	1.39	1.57
	0.5	0.52	1.58	1.26	1.49
	0.8	0.89	1.77	1.26	1.49
	1.2	1.40	1.97	1.26	1.48
FUE rate 50th percentile [Mbps]					
MUE Demand (Mbps)	*0.2*	12.74	11.98	12.76	13.89
	0.5	12.82	11.28	13.06	13.81
	0.8	12.61	10.57	13.02	13.81
	1.2	12.46	9.68	13.00	13.79
FUE rate 80th percentile [Mbps]					
MUE Demand (Mbps)	*0.2*	17.45	16.25	17.45	15.91
	0.5	17.27	15.30	17.27	15.80
	0.8	17.27	14.63	17.27	15.80
	1.2	17.17	14.08	17.17	15.73

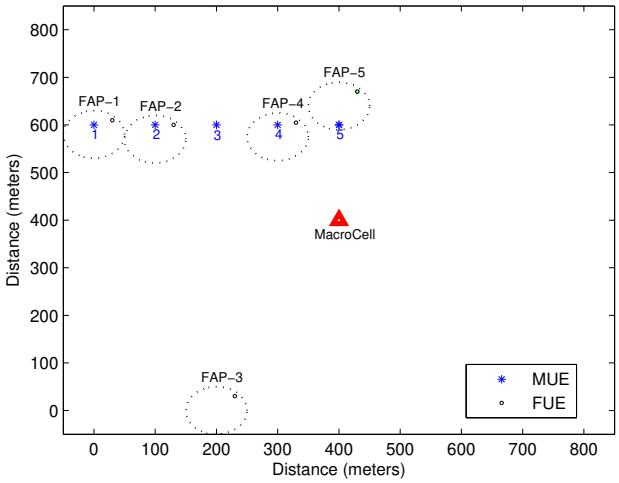

Fig. 8: Node locations in static scenario. FAP-3 is the only femtocell node not close to a victim user.

and ESRA scheme, respectively. We observe clearer now that in case of ORA, majority of the MUE's achieved rate does not exceed the demand. However, in case of ESRA, MUE's achieved rate is not as close to the demand.

Moreover, Fig. 10 shows the resource (RB and power) allocation map of the FAPs for MUE demand rate of 0.2Mbps. We can see that ORA scheme mutes FAP-1 and FAP-2 in first four RBs on which the nearby MUEs are being served, however transmits with lower power in the RBs where MUE-3 is being served. On the other hand

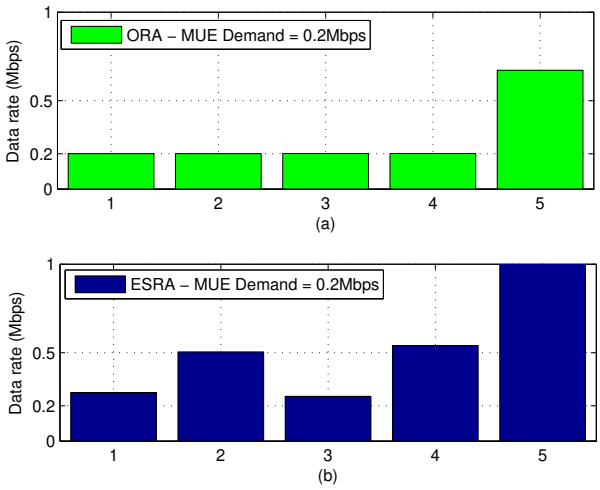

Fig. 9: MUE achieved data rate for MUE Demand of 0.2Mbps. The X-axis shows the MUE index. (a) ORA scheme (b) ESRA scheme.

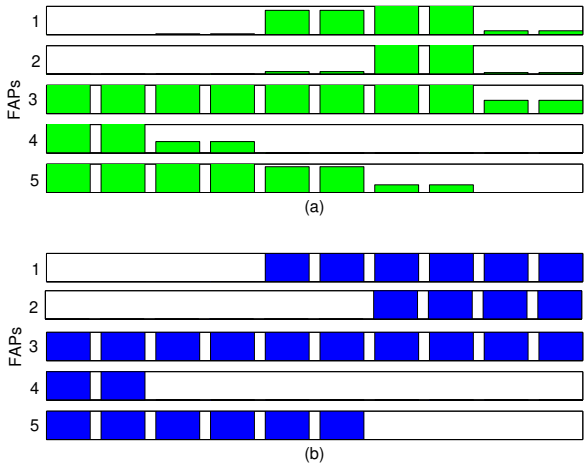

Fig. 10: FAP RA Allocation Map for MUE Demand of 0.2Mbps. The Y-axis for the bar graphs indicates the transmit power of FAPs (ranging from 0-20mW). The X-axis indicated the RB index (RB-1 to RB-10, from left to right). (a) ORA scheme (b) ESRA scheme.

a complete muting for those RBs takes place in case of ESRA scheme. Similarly to protect MUE-4 and MUE-5, ESRA completely mutes transmissions of FAP-4 and FAP-5, in their serving RBs. However, the ORA scheme still transmits in some of the RBs with lower transmit power. Such a behaviour is observed since the ORA scheme has the liberty to optimise not only the RB allocation as well as the transmit power of each femtocell. It is noted that the optimal scheme is more effective in such cases as it does not necessarily completely mutes the femtocells in such *critical RBs* but in fact reduces transmit power as much as needed. However, this advantageous behaviour of

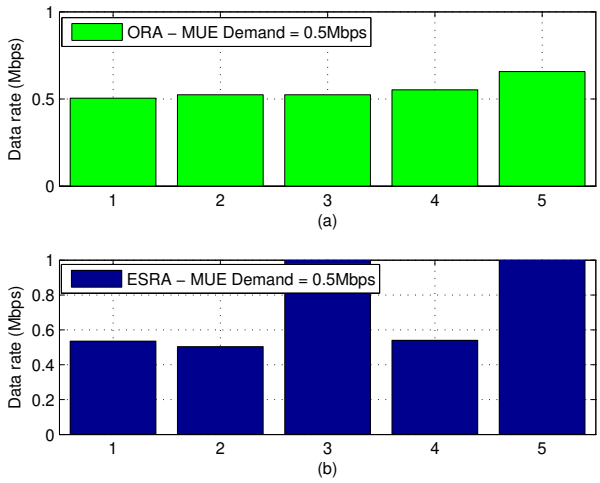

Fig. 11: MUE achieved data rate for MUE Demand of 0.5Mbps. The X-axis shows the MUE index. (a) ORA scheme (b) ESRA scheme.

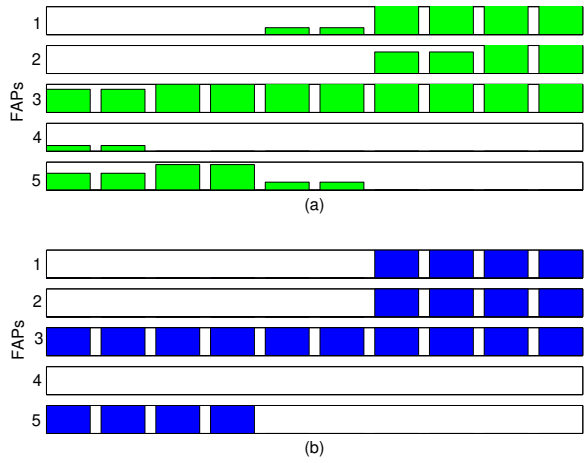

Fig. 12: FAP RA Allocation Map for MUE Demand of 0.5Mbps. The Y-axis for the bar graphs indicates the transmit power of FAPs (ranging from 0-20mW). The X-axis indicated the RB index (RB-1 to RB-10, from left to right). (a) ORA scheme (b) ESRA scheme.

ORA scheme comes at the cost of extra computational complexity as explained in section IV. We can observe a similar trend of MUE achieved data rates and resource allocation map in Fig. 11 and 12, where the MUE demand rate is 0.5Mbps. Furthermore, focusing on FAP-3 which is away from the MUEs, we observe that it is allowed to transmit on all the RBs with high power, even for the higher MUE demand case.

VII. Conclusion

In this paper, we tackled the inter-tier interference issue which deteriorates the performance of mobile macrocell-served users in a LTE HetNet environment comprising macrocells and femtocells sharing the same frequency band. We propose dynamic resource allocation at femtocells to maximise their sum data rate while at the same time the interference faced by the macrocell-served users is kept below a tolerance threshold, estimated based on their minimum rate requirement. We analysed the optimal solution to this problem and also proposed a more practical scheme which considers femtocell RB muting and significantly reduces computational complexity. Focusing on the practical application of these dynamic approaches, we furthermore design algorithms to implement them in a real-world system such as Long-Term Evolution (LTE) networks. Our simulation results compare the dynamic resource allocation schemes with the conventional Reuse-1 and Orthogonal Reuse scheme, and demonstrate that macro users QoS requirements can be ensured while keeping the femto users data rates at similar high levels.

Appendix A
Optimal Power for a Given RB Allocation

For the sake of simplicity of understanding, we suppress the notations in equation (11) and write L as:
$L = \phi R - \lambda\phi p\Gamma - \mu\phi p$, now replacing $\phi = 1$ and R with equation (3), we get $L = \log_2(1 + p\,\alpha) - \lambda\, p\,\Gamma - \mu\, p$.

Let, $y = \log_2(1 + p\,\alpha)$,
$\frac{\partial L}{\partial p} = \frac{\partial y}{\partial p} - \lambda\Gamma - \mu$.
Now let, $x = (1 + p\,\alpha)$; $y = \log_2(x)$,
$\because \frac{\partial y}{\partial p} = \frac{\partial y}{\partial x}\frac{\partial x}{\partial p} = \frac{1}{ln(2)x}\alpha = \frac{\alpha}{ln(2)(1+p\alpha)}$.
$\therefore \frac{\partial L}{\partial p} = \frac{\alpha}{ln(2)(1+p\alpha)} - \lambda\Gamma - \mu$.

Applying the KKT condition, we equate $\frac{\partial L}{\partial p} = 0$;
$\frac{\alpha}{ln(2)(1+p\alpha)} - \lambda\Gamma - \mu = 0$. It then follows,
$\implies \frac{\alpha}{ln(2)(1+p\alpha)} = \mu + \lambda\Gamma$,
$\implies \frac{\alpha}{(1+p\alpha)} = ln(2)(\mu + \lambda\Gamma)$,
$\implies (1 + p\alpha) = \frac{\alpha}{ln(2)(\mu+\lambda\Gamma)}$,
$\implies p\alpha = \frac{\alpha}{ln(2)(\mu+\lambda\Gamma)} - 1$, and we solve for p as:
$p = \frac{1}{ln(2)(\mu+\lambda\Gamma)} - \frac{1}{\alpha}$.

Appendix B
Subgradients of Dual Function

Considering the Lagrangian dual function g, in equation (9) at two different points $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ and $(\boldsymbol{\lambda}', \boldsymbol{\mu}')$ in the dual variable multidimensional space, where $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, ...\lambda_n, ...\lambda_N)$, $\boldsymbol{\lambda}' = (\lambda_1, \lambda_2, ...\lambda'_n, ...\lambda_N)$, $\boldsymbol{\mu} = (\mu_1, \mu_2, ...\mu_m, ...\mu_M)$ and $\boldsymbol{\mu}' = (\mu_1, \mu_2, ...\mu'_m, ...\mu_M)$, we have:

$$g(\boldsymbol{\lambda},\boldsymbol{\mu}) \quad = \quad \begin{cases} \max_{\boldsymbol{\phi},\mathbf{p}} \quad L(\boldsymbol{\phi},\mathbf{p},\boldsymbol{\lambda},\boldsymbol{\mu}), \\ s.t. \\ 0 \le \phi_{k,m,n} \le 1, \forall k \in \mathcal{K} \setminus \mathcal{K}_0, m \in \mathcal{M}, n; \\ \sum\limits_{k \in \mathcal{K}_m} \phi_{k,m,n} \le 1, \forall m \in \mathcal{M}, n; \\ p_{m,n} \ge 0, \forall m \in \mathcal{M}, n. \end{cases} \tag{25}$$

$$g(\boldsymbol{\lambda}',\boldsymbol{\mu}') \quad = \quad \begin{cases} \max_{\phi,p} \quad L(\boldsymbol{\phi},\mathbf{p},\boldsymbol{\lambda}',\boldsymbol{\mu}'), \\ s.t. \\ 0 \le \phi_{k,m,n} \le 1, \forall k \in \mathcal{K} \setminus \mathcal{K}_0, m \in \mathcal{M}, n; \\ \sum\limits_{k \in \mathcal{K}_m} \phi_{k,m,n} \le 1, \forall m \in \mathcal{M}, n; \\ p_{m,n} \ge 0, \forall m \in \mathcal{M}, n. \end{cases} \tag{26}$$

Substituting the values of $\boldsymbol{\phi}$ and $\mathbf{p}$ with the optimal values, we get the subgradient of g at λ as:

$$\begin{aligned} &[g(\boldsymbol{\lambda}',\boldsymbol{\mu}') - g(\boldsymbol{\lambda},\boldsymbol{\mu})] \\ &= \max_{\boldsymbol{\phi},\mathbf{p}} \quad L(\boldsymbol{\phi},\mathbf{p},\boldsymbol{\lambda}',\boldsymbol{\mu}') - \max_{\boldsymbol{\phi},\mathbf{p}} \quad L(\boldsymbol{\phi},\mathbf{p},\boldsymbol{\lambda},\boldsymbol{\mu}) \ge L(\boldsymbol{\phi}^*,\mathbf{p}^*,\boldsymbol{\lambda}',\boldsymbol{\mu}') - L(\boldsymbol{\phi}^*,\mathbf{p}^*,\boldsymbol{\lambda},\boldsymbol{\mu}), \\ &= (\lambda_n' - \lambda_n) \sum_n (\sum_{m=1}^{M} (\sum_{k \in \mathcal{K}_m} \phi_{k,m,n}^*) p_{m,n}^* \Gamma_{0,u,n}^{m*} - \Omega_n^{max}) + (\mu_m' - \mu_m)(P_{\max} - \sum_n (\sum_{k \in \mathcal{K}_m} \phi_{k,m,n}^*) p_{m,n}^*). \end{aligned} \tag{27}$$

The inequality in equation (27) exists because of the definition of dual function and Lagrange in equation (9) and (10).

$$g(\boldsymbol{\lambda}',\boldsymbol{\mu}') \ge g(\boldsymbol{\lambda},\boldsymbol{\mu}) + (\lambda_n' - \lambda_n) \sum_n (\sum_{m=1}^{M} (\sum_{k \in \mathcal{K}_m} \phi_{k,m,n}^*) p_{m,n}^* \Gamma_{0,u,n}^{m*} - \Omega_n^{max}) - (\mu_m' - \mu_m)(P_{\max} - \sum_n (\sum_{k \in \mathcal{K}_m} \phi_{k,m,n}^*) p_{m,n}^*) \tag{28}$$

Hence, the subgradients of $g(\boldsymbol{\lambda},\boldsymbol{\mu})$ at the point λ_n are,

$$\begin{aligned} \Delta\lambda_n &= \sum_n (\sum_{m=1}^{M} (\sum_{k \in \mathcal{K}_m} \phi_{k,m,n}^*) p_{m,n}^* \Gamma_{0,u,n}^{m*} - \Omega_n^{max}), \\ \Delta\mu_m &= P_{\max} - \sum_n (\sum_{k \in \mathcal{K}_m} \phi_{k,m,n}^*) p_{m,n}^*. \end{aligned} \tag{29}$$

REFERENCES

[1] V. Chandrasekhar, J. Andrews, and A. Gatherer, "Femtocell Networks: a Survey," *IEEE Communications Magazine*, vol. 46, no. 9, pp. 59–67, 2008.

[2] I. Hwang, B. Song, and S. Soliman, "A Holistic View on Hyper-dense Heterogeneous and Small Cell Networks," *IEEE Communications Magazine*, vol. 51, no. 6, pp. 20–27, June 2013.

[3] S. Navaratnarajah, A. Saeed, M. Dianati, and M. Imran, "Energy Efficiency in Heterogeneous Wireless Access Networks," *IEEE Wireless Communications Magazine*, vol. 20, no. 5, pp. 37–43, October 2013.

[4] A. Saeed, A. Akbari, M. Dianati, and M. A. Imran, "Energy Efficiency Analysis for LTE Macro-Femto HetNets," in *Proceedings of 19th European Wireless Conference (EW)*, April 2013.

[5] S.-E. Elayoubi and B. Fourestie, "On Frequency Allocation in 3G LTE Systems," in *Proceedings of IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sept 2006.

[6] S. Ali and V. C. M. Leung, "Dynamic Frequency Allocation in Fractional Frequency Reused OFDMA Networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 8, pp. 4286–4295, August 2009.

[7] R. Y. Chang, Z. Tao, J. Zhang, and C.-C. J. Kuo, "A Graph Approach to Dynamic Fractional Frequency Reuse (FFR) in Multi-cell OFDMA Networks," in *Proceedings of IEEE International Conference on Communications*, 2009, pp. 3993–3998.

[8] A. Stolyar and H. Viswanathan, "Self-Organizing Dynamic Fractional Frequency Reuse in OFDMA Systems," in *Proceedings of IEEE 27th Conference on Computer Communications (INFOCOM)*, April 2008.

[9] D. Lopez-Perez, C. Xiaoli, and Z. Jie, "Dynamic Downlink Frequency and Power Allocation in OFDMA Cellular Networks," *IEEE Transactions on Communications*, vol. 60, no. 10, pp. 2904–2914, 2012.

[10] D. Lopez-Perez, X. Chu, A. V. Vasilakos, and H. Claussen, "Power Minimization Based Resource Allocation for Interference Mitigation in OFDMA Femtocell Networks," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 2, pp. 333–344, February 2014.

[11] D. Lopez-Perez, X. Chu, A. Vasilakos, and H. Claussen, "On Distributed and Coordinated Resource Allocation for Interference Mitigation in Self-Organizing LTE Networks," *IEEE/ACM Transactions on Networking*, vol. 21, no. 4, pp. 1145–1158, Aug 2013.

[12] J. Ling, D. Chizhik, and R. Valenzuela, "On Resource Allocation in Dense Femto-Deployments," in *Proceedings of IEEE International Conference on Microwaves, Communications, Antennas and Electronics Systems*, Nov 2009.

[13] F. Bernardo, R. Agusti, J. Cordero, and C. Crespo, "Self-Optimization of Spectrum Assignment and Transmission Power in OFDMA Femtocells," in *Proceedings of Sixth Advanced International Conference on Telecommunications (AICT)*, May 2010.

[14] V. Chandrasekhar and J. Andrews, "Spectrum Allocation in Tiered Cellular Networks," *IEEE Transactions on Communications*, vol. 57, no. 10, pp. 3059–3068, October 2009.

[15] J. D. Hobby and H. Claussen, "Deployment Options for Femtocells and Their Impact on Existing Macrocellular Networks," *Bell Labs Technical Journal*, vol. 13, no. 4, pp. 145–160, 2009.

[16] H. Claussen, "Co-Channel Operation of Macro- and Femtocells in a Hierarchical Cell Structure," *International Journal of Wireless Information Networks*, vol. 15, no. 3-4, pp. 137–147, 2008.

[17] M. Arslan, J. Yoon, K. Sundaresan, S. Krishnamurthy, and S. Banerjee, "A Resource Management System for Interference Mitigation in Enterprise OFDMA Femtocells," *IEEE/ACM Transactions on Networking*, vol. 21, no. 5, pp. 1447–1460, Oct 2013.

[18] X. Chu, Y. Wu, D. Lopez-Perez, and X. Tao, "On Providing Downlink Services in Collocated Spectrum-Sharing Macro and Femto Networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 12, pp. 4306–4315, December 2011.

[19] M. S. Alam, J. W. Mark, and X. S. Shen, "Relay Selection and Resource Allocation for Multi-User Cooperative OFDMA Networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 5, pp. 2193–2205, May 2013.

[20] W. Yu and R. Lui, "Dual Methods for Nonconvex Spectrum Optimization of Multicarrier Systems," *IEEE Transactions on Communications*, vol. 54, no. 7, pp. 1310–1322, July 2006.

[21] W. Yi, Z. Dongmei, J. Hai, and W. Ye, "A Novel Spectrum Arrangement Scheme for Femto Cell Deployment in LTE Macro Cells," in *Personal, Indoor and Mobile Radio Communications, 2009 IEEE 20th International Symposium on*, 2009.