Multi-Channel Analysis of Cell Range Expansion and Resource Partitioning in Two-Tier Heterogeneous Cellular Networks

Yamuna Dhungana and Chintha Tellambura, *Fellow, IEEE*

Abstract—Cellular heterogeneous networks (HetNets) can improve capacity by offloading users from congested macro cells to lightly-loaded small cells through biased association known as cell range expansion (CRE). However, the offloaded (range-expanded) users must be protected from macro interference through time/frequency resource partitioning. In this paper, we develop an analytical framework to evaluate the performance gain due to CRE further supported by resource partitioning in two-tier (macro-pico) networks with multi-channel downlinks, for example, those based on orthogonal frequency division multiple access (OFDMA). By exploiting the flexibility in subchannel allocation offered by OFDMA, frequency-domain resource partitioning is proposed in which the macro tier is muted on a fraction of total subchannels, which are allocated exclusively to range-expanded pico users. The load perceived by a base-station is a key factor in determining its interference contribution over the network and is directly affected by user offloading and resource partitioning. Thus, the analysis of such systems must incorporate cell load. While previous studies mostly rely on full-load assumption, in this paper, we properly characterize cell load as the function of user density, association bias and resource partitioning fraction. We then, evaluate the performance in terms of average user data rate over the entire network, and also investigate the optimal choice of association bias and resource partitioning fraction.

Index Terms—Cell load, cell range expansion (CRE), heterogeneous networks (HetNet), orthogonal frequency domain multiple access (OFDMA), Poisson point process (PPP), stochastic geometry.

I. Introduction

Co-channel deployment of non-conventional, low-power base-stations (BSs) such as picos and femtos within the areas covered by the existing macro cellular infrastructure is the most viable solution to meet the unrelenting growth in data traffic demand [1]–[3]. The resulting mixed network of macro and low-power BSs is referred to as heterogeneous network (HetNet) [1]–[3]. The increased capacity is achieved by offloading users from the congested macro tier to pico/femto tier to give them access to a larger fraction of radio resources and to reduce the macro cellular load so as to serve the remaining macro users with improved rates. The user offloading however, may be limited due to transmit power disparities of macro and pico/femto BSs, thereby limiting the capacity gain.

The macro offloading can be increased by biased association known as cell range expansion (CRE) [2], [3], in which a user is offloaded to a small cell if the received power from it is less than that from a macro cell by at most some amount known as association bias. Such offloaded users are referred

to as range-expanded users. With this technique, the number of offloaded users can be controlled with the bias value to obtain a balanced distribution of user loads across the tiers. However, the load balancing offered by CRE comes at the cost of severe downlink co-channel interference to range-expanded users from the macro tier, which must be mitigated by interference coordination techniques.

The interference coordination in such scenarios can be implemented by resource partitioning [2]–[4], in which a certain fraction of time or frequency resources is provided exclusively to small cells by muting the macro-tier transmissions in these resources. The range-expanded users are then served by the small cells in these resources, thus isolating them from the macro tier interference. Simulation results [5]–[8] demonstrate that CRE with resource partitioning highly enhances the otherwise limited performance gains from the deployment of small cells. These interdependent techniques however, must be jointly tuned for optimal system performance. To this end, the optimal association bias at the given resource partitioning fraction was investigated for sum capacity and related performance metrics in [9] through a semi-analytical approach. Analytical approaches to determine the optimal combination of bias value and resource partitioning fraction were presented in [10] and [11] based on per user spectral efficiency and downlink rate distribution, respectively. However, both assumed fully loaded network, i.e., all the BSs are simultaneously active all the time.

The full-load assumption is not applicable for small cells, unless they are deployed in hot-spots and very large biasing towards them is introduced. On the other hand, with very large biasing, macro cells may no longer be fully loaded. Thus, full-load is not a reasonable assumption to study biasing. The analytical results in [12] show that biasing has detrimental impact on the average rate of the overall network in fully loaded condition. The motivation behind biasing is to improve the network rate through load balancing, i.e., relieving the heavily loaded macro cells and better utilizing the resources of lightly loaded small cells. But if the macro and small cells are assumed to be always fully loaded, then biasing makes no sense. Thus, the full-load assumption cannot reflect the benefits of biasing. Meanwhile, the interference from a BS is a direct function of its load. For example, the BSs that receive more load have higher probability of being active at a given time instant, thus contributing more interference to the network. As the load perceived by a BS is significantly affected by the number of users offloaded to/from and the fraction of resources allocated, the interference to a given user strongly

depends on association bias and resource partitioning fraction. Such effects cannot be captured if the full-load assumption is used. Thus, an analytical framework for the performance evaluation of cellular HetNets with biased association and resource partitioning, while appropriately modeling the cell load, becomes essential and this paper aims to fulfill this need.

We focus on the downlink performance analysis of two-tier (macro and pico) HetNets. The spatial distributions of the network nodes in such networks have recently been modeled using Stochastic Geometry, with each tier modeled by a point process [13]–[16]. The locations of macro and pico BSs are thus modeled as independent Poisson point processes (PPPs) in this paper. The downlink analysis of cellular HetNets usually assumes a time-shared single channel per cell [10]–[12], [15], [16]. Thus, only time-domain approach to resource partitioning has been mainly analyzed [10], [11], [17], [18], in which the macro tier is periodically muted on certain fraction of subframes, known as almost blank subframes (ABSFs). In contrast, we consider a multi-channel downlink, for example the one based on orthogonal frequency division multiple access (OFDMA), in which multiple users are simultaneously served in orthogonal subchannels. In LTE networks, multiple access in the downlink is established by OFDMA. Due to the flexibility in subchannel allocation offered by OFDMA, we propose frequency-domain resource partitioning in which the macro tier is restricted from using a fraction of total subchannels so that they are allocated exclusively to range-expanded pico users.

The main contributions of this paper are summarized as follows:

1) Based on the proposed multi-channel model, we first define the load perceived by a BS as a direct function of the number of associated users and the number of available subchannels. Such characterization effectively captures the effect of user density, association bias and resource partitioning fraction on cell load.
2) Next, we evaluate the performance of the proposed system in terms of average user data rate that can be attained over the entire network, while incorporating the cell load into the analysis.
3) We comprehensively analyze the average rate performance under different bias values and resource partitioning fractions and finally investigate their optimal combination.
4) We numerically demonstrate that if the bias value and resource partitioning fraction are carefully selected, the rate performance can be highly improved in comparison to the CRE only system (i.e., no resource partitioning).
5) We show that the optimal combination of association bias and resource partitioning fraction is strongly dependent on the network load.

We hasten to add that although the stochastic geometry approach to multi-channel downlink analysis of cellular Het-Nets has been considered before, for example [19], [20], the problem of CRE with resource partitioning in multi-channel environment, while successfully capturing their impact on cell load is addressed here **for the first time.**

The rest of the paper is organized as follows. The network model, user association policy and resource partitioning scheme are described in Section II. Section III utilizes the user association probability and cell load derived in Section II to derive the average user data rate over the entire network. The special case of no resource partitioning is also analyzed in section III. Validation of the analytical results through Monte-Carlo simulation is done in Section IV, along with the extensive numerical analysis to assess the impact of biasing and resource partitioning on user data rate. Finally, Section V concludes the paper.

Notations: Throughout the paper, $\mathbb{R}^2$ denotes the two-dimensional (2-D) Euclidean space and $||x-y||$, the Euclidean distance between two points $x, y \in \mathbb{R}^2$. $\mathbb{P}(\cdot)$, $\mathbb{E}(\cdot)$ and $\text{Exp}[1]$ denote the probability measure, expectation operator, and unit exponential distribution, respectively. A homogeneous PPP of density $\lambda > 0$ on $\mathbb{R}^2$ is denoted by a set of points $\boldsymbol{\Phi} = \{x_1, x_2, \ldots\}$, where $x_i \in \mathbb{R}^2$ is the location of the i-th point. According to the definition of PPP [21], the numbers of points in disjoint areas $A_1, A_2, \ldots A_k$ are independent Poisson random variables (RVs) with mean $\lambda A_1, \lambda A_2, \ldots, \lambda A_k$, respectively.

II. SYSTEM MODEL

A. Network and Channel model

We consider an OFDMA based two-tier downlink cellular HetNet consisting of macro and pico BSs, which are assumed to be spatially distributed in $\mathbb{R}^2$ plane as independent homogeneous PPPs $\boldsymbol{\Phi}_m$ of density λ_m and $\boldsymbol{\Phi}_p$ of density λ_p, respectively. For the macro tier, the PPP model provides tight lower bounds for performance measures, as tight as the upper bound results provided by the popular grid model, when compared to the actual 4G network [22]. However, the analytical tractability of the PPP model is a key benefit over the grid model. Its adoption for the pico tier is justified because the randomness in pico BS locations is expected. Similarly, user locations are modeled as an independent PPP $\boldsymbol{\Phi}_u$ with density λ_u. The two network tiers share the same spectrum which is evenly divided into $L > 1$ subchannels[1]. We consider flat transmit power spectrum on the downlink[2] and thus, the power per subchannel is kept constant at P_m and P_p for macro and pico tiers, respectively. If $P_{\max}^m$ and $P_{\max}^p$ are the maximum allowable transmit powers of macro and pico BSs, respectively, then $P_m = P_{\max}^m/L$ and $P_p = P_{\max}^p/L$.

Independent Rayleigh multipath fading with power-law path loss is assumed between any BS-user pair. The channel power gains from the macro BS located at $x_m \in \boldsymbol{\Phi}_m$ and pico BS located at $x_p \in \boldsymbol{\Phi}_p$ to a typical user located, without loss of generality, at the origin, are thus given by $h_{x_m}||x_m||^{-\alpha_m}$ and $h_{x_p}||x_p||^{-\alpha_p}$, respectively, where $h_{x_m} \sim \text{Exp}[1]$ and $h_{x_p} \sim \text{Exp}[1]$ are the corresponding fading powers, and α_m and α_p are the path-loss exponents of macro and pico tier, respectively.

---

[1]A subchannel may refer to one or multiple resource blocks (RBs) in LTE systems.

[2]This assumption is consistent with LTE downlink power allocation [23].

B. User association scheme

The user association scheme is based on biased received power, i.e., each user is associated with the BS offering the maximum biased received power [11], [12]. Fading effect is ignored in the association metric to avoid the ping-pong handover effect [24]. If B is the association bias introduced for pico CRE, a typical user at the origin is associated with the nearest macro BS only if $P_m R_m^{-\alpha_m} \geq P_p B R_p^{-\alpha_p}$, where $R_m = \min\limits_{x_m \in \Phi_m} \|x_m\|$ and $R_p = \min\limits_{x_m \in \Phi_p} \|x_p\|$ are the distances from the origin to the nearest macro and pico BSs, respectively. It is otherwise associated with the nearest pico BS. When associated with the nearest pico BS, it is registered in its user list as *unbiased user* if $P_p R_p^{-\alpha_p} \geq P_m R_m^{-\alpha_m}$ and as *range-expanded user* if $P_p R_p^{-\alpha_p} \leq P_m R_m^{-\alpha_m} < P_p B R_p^{-\alpha_p}$. The nomenclatures for pico users (*unbiased* and *range-expanded*) have been adopted from [11].

For the given user association scheme, if we randomly pick a user, it may turn out be a macro user, an unbiased pico user, or a range-expanded pico user with certain probabilities. The following lemma expresses these probabilities.

Lemma 1. *Let U_m, U_o, and U_e denote the probabilities that a randomly chosen user is a macro user, unbiased pico user and range-expanded pico user, respectively. Then,*

$$U_m = 2\pi\lambda_m \int_0^\infty r e^{-\pi\lambda_m r^2} \exp\Big(-\pi\lambda_p \Big(\frac{BP_p}{P_m}\Big)^{\frac{2}{\alpha_p}} r^{\frac{2\alpha_m}{\alpha_p}}\Big) dr \tag{1}$$

$$U_o = 2\pi\lambda_p \int_0^\infty r e^{-\pi\lambda_p r^2} \exp\Big(-\pi\lambda_m \Big(\frac{P_m}{P_p}\Big)^{\frac{2}{\alpha_m}} r^{\frac{2\alpha_p}{\alpha_m}}\Big) dr \tag{2}$$

$$U_e = 2\pi\lambda_p \int_0^\infty r e^{-\pi\lambda_p r^2} \Big\{ \exp\Big(-\pi\lambda_m \Big(\frac{P_m}{BP_p}\Big)^{\frac{2}{\alpha_m}} r^{\frac{2\alpha_p}{\alpha_m}}\Big)$$
$$- \exp\Big(-\pi\lambda_m \Big(\frac{P_m}{P_p}\Big)^{\frac{2}{\alpha_m}} r^{\frac{2\alpha_p}{\alpha_m}}\Big) \Big\} dr. \tag{3}$$

Proof: Since the analysis conducted on a typical user located at the origin is valid for any randomly chosen user according to Slivnyak's theorem [21], U_m can be derived as

$$U_m = \mathbb{P}(P_m R_m^{-\alpha_m} \geq P_p B R_p^{-\alpha_p})$$
$$= \mathbb{E}_{R_m}\Big[\mathbb{P}\big(R_p \geq (P_p/P_m B)^{1/\alpha_p} R_m^{\alpha_m/\alpha_p}\big)\Big],$$

which can be solved by using the probability distributions of R_m and R_p. We know that $\mathbb{P}(R_l > r)$, $l \in \{m, p\}$, is the probability that no points of Φ_l lie within a circle of radius r, centered at the origin. Since Φ_l is a PPP with density λ_l, we have, $\bar{F}_{R_l}(r) = \mathbb{P}(R_l > r) = \exp(-\pi\lambda_l r^2)$. The probability density function (PDF) of R_l can then be obtained as $f_{R_l}(r) = -d\bar{F}_{R_l}(r)/dr = 2\pi\lambda_l r \exp(-\pi\lambda_l r^2)$. Similarly, by using these distributions, U_o and U_e can be obtained as $U_o = \mathbb{P}(P_p R_p^{-\alpha_p} \geq P_m R_m^{-\alpha_m})$ and $U_e = \mathbb{P}(P_p R_p^{-\alpha_p} \leq P_m R_m^{-\alpha_m} < P_p B R_p^{-\alpha_p})$. ∎

For the special case of equal path-loss exponents, i.e., $\alpha_m = \alpha_p = \alpha$, the integrals in (1)-(3) can be reduced to following simple closed-form expressions by using $\int_0^\infty r \exp(-\beta r^2) dr = 1/(2\beta)$.

$$U_m = \frac{\lambda_m P_m^{2/\alpha}}{\lambda_m P_m^{2/\alpha} + \lambda_p (BP_p)^{2/\alpha}}, \ U_o = \frac{\lambda_p P_p^{2/\alpha}}{\lambda_m P_m^{2/\alpha} + \lambda_p P_p^{2/\alpha}},$$

$$U_e = \frac{\lambda_p (BP_p)^{2/\alpha}}{\lambda_m P_m^{2/\alpha} + \lambda_p (BP_p)^{2/\alpha}} - \frac{\lambda_p P_p^{2/\alpha}}{\lambda_m P_m^{2/\alpha} + \lambda_p (P_p)^{2/\alpha}}. \tag{4}$$

The probabilities (1)-(4) are also derived in [11] and are given here for the sake of completeness.

As per the given user association scheme, the set of total users in the network, Φ_u can be divided into three subsets:
1) Φ_u^m, the set of **macro users**,
2) Φ_u^o, the set of **unbiased pico users**, and
3) Φ_u^e, the set of **range-expanded pico users**,

such that $\Phi_u = \Phi_u^m \cup \Phi_u^o \cup \Phi_u^e$. Since each user in Φ_u can belong to exactly one of these three sets, they are disjoint. The probabilities U_m, U_o, and U_e can also be interpreted as the average fraction of users belonging to the sets Φ_u^m, Φ_u^o and Φ_u^e, respectively. For each user-set, we are interested in the number of users associated with a typical BS to characterize the typical cell load of each tier and the average share of radio resources received by a typical user. The actual locations of users with respect to each other in each Φ_u^l, $l \in \{m, o, e\}$ are less important to us. Thus, Φ_u^m, Φ_u^o and Φ_u^e can be equivalently modeled as independent PPPs with densities $U_m\lambda_u$, $U_o\lambda_u$ and $U_e\lambda_u$, respectively. In other words, they can be modeled as thinned versions of the original process Φ_u with retention probabilities U_m, U_o and U_e, respectively, independent of the locations of the users.

Each user in Φ_u^m is always associated with the nearest macro BS and each user in $\Phi_u^o \cup \Phi_u^e$ is always associated with the nearest pico BS. The network can thus be viewed as a superposition of two independent Voronoi tessellations of the macro and pico tier, respectively. The Voronoi cells of each tessellation are disjoint and their sizes are independent and identically distributed (i.i.d.) random variables (RVs) [25]. Hence, together with the independent scattering property of the PPP, which states that the number of points of a PPP in disjoint sets are independent RVs [21], [26], the number of macro users in different macro Voronoi cells are i.i.d RVs, and so are the number unbiased pico users and the number of range-expanded pico users in different pico Voronoi cells. Given in the following lemma are their probability mass functions (PMFs), which are essential for calculating the typical cell load of each tier and the average share of radio resources received by a typical user, later in this paper.

Lemma 2. *Let N_m be the number of users associated with a randomly chosen macro BS, and N_o and N_e be the number of unbiased and range-expanded users of a randomly chosen pico BS. Their PMFs are given by*

$$\mathbb{P}(N_l = n) = \frac{3.5^{3.5}\Gamma(3.5 + n)(U_l\lambda_u/\lambda_{\zeta(l)})^n}{\Gamma(3.5)n!(U_l\lambda_u/\lambda_{\zeta(l)} + 3.5)^{n+3.5}}, \ n \geq 0,$$
$$\forall l \in \{m, o, e\}, \tag{5}$$

where $\lambda_{\zeta(m)} = \lambda_m$ and $\lambda_{\zeta(o)} = \lambda_{\zeta(e)} = \lambda_p$.

Proof: Since Φ_u^m is a PPP of density $U_m\lambda_u$, the number of macro users in a typical macro Voronoi cell of given area A

is Poisson distributed with mean $U_m \lambda_u A$. The unconditional PMF of N_m in (5) is then obtained by averaging over the distribution of Voronoi cell area A approximated by the Gamma distribution, $f_A(a) = (3.5\lambda_m)^{3.5} a^{2.5} \exp(-3.5\lambda_m a)/\Gamma(3.5)$ [25]. The PMFs of N_o and N_e can be similarly obtained. ∎

C. Resource Allocation and Partitioning

For any range-expanded pico user in Φ_u^e, the average received power from its nearest macro BS is greater than that from the serving pico BS. The range-expanded users thus need to be protected from high-power macro interference. If each macro BS leaves a set of L_r subchannels unutilized out of the total L subchannels, each pico BS can serve its range-expanded users in these macro-interference free subchannels. Each macro BS thus allocates $L - L_r = L_c$ number of subchannels to serve its users. The unbiased users in each pico cell are also served in the same set of L_c subchannels since L_r macro-interference free subchannels are reserved exclusively for its range-expanded users.

Let the set of L_r macro-interference free subchannels reserved exclusively for range-expanded users be denoted by S_r and the set of L_c common subchannels shared by the macro and pico tiers be denoted by S_c. The subchannels in each BS are allocated to individual users according to one subchannel per user and they are uniformly and independently selected from the available set. However, if the number of users associated with a BS is greater than the number of subchannels available, the resources are time-shared equally among the users. This is basically frequency- and time-domain round-robin scheduling, which gives equal share of resources to all the users. This simple scheduling algorithm leads to analytical tractability and provides important insights on system parameters. Sophisticated scheduling algorithms like max-rate and proportional fair schedulars, which add significant complexities to the analysis will be considered in future work. The current analysis serves as a lower bound on the performance of these sophisticated algorithms.

We assume that each BS has saturated downlink transmission queue for each associated user and thus, each user always has data to receive from its serving BS. While some BSs may have more users than the available subchannels, some may have less. Thus, depending upon the number of users associated with it, a BS, may or may not be active on all of its available subchannels. In the following lemma, we derive the probability that a typical BS of each tier is active on a given subchannel.

Lemma 3. *Let p_m be the probability that a randomly chosen macro BS is active on a given subchannel from the set S_c. Similarly, let p_o and p_e be the probabilities that a randomly chosen pico BS is active on a given subchannel from the set S_c and S_r, respectively. Then,*

$$p_l = 1 - \frac{3.5^{3.5}}{\Gamma(3.5)} \frac{1}{L_{\kappa(l)}} \sum_{n=0}^{L_{\kappa(l)}-1} \left[\frac{(L_{\kappa(l)} - n)\Gamma(3.5 + n)}{n!} \right.$$
$$\left. \times \frac{(U_l \lambda_u / \lambda_{\zeta(l)})^n}{(U_l \lambda_u / \lambda_{\zeta(l)} + 3.5)^{3.5+n}} \right], \forall l \in \{m, o, e\} \quad (6)$$

where $\lambda_{\zeta(m)} = \lambda_m$, $\lambda_{\zeta(o)} = \lambda_{\zeta(e)} = \lambda_p$, $L_{\kappa(m)} = L_{\kappa(o)} = L_c$ and $L_{\kappa(e)} = L_r$.

Proof: If the number of users associated with a typical macro cell is less than L_c (i.e. $N_m < L_c$), the probability that a subchannel of S_c is used in the cell is N_m/L_c. However, if $N_m \geq L_c$, all the subchannels of S_c are used in the cell with probability 1. Thus, p_m can be expressed as

$$p_m = \sum_{n=0}^{L_c-1} \frac{n}{L_c} \mathbb{P}(N_m = n) + \sum_{n=L_c}^{\infty} \mathbb{P}(N_m = n)$$
$$= 1 - \sum_{n=0}^{L_c-1} \left(1 - \frac{n}{L_c}\right) \mathbb{P}(N_m = n) \quad (7)$$

The final expression for p_m is then obtained by substituting the PMF of N_m in the above equation. The probabilities p_o and p_e can be similarly obtained. ∎

Remark. As we explained earlier, the number of users associated with different macro cells are independent. Thus, the probability that any randomly chosen macro BS is active on a given subchannel from the set S_c, which is p_m, is independent of any other macro BS. Similarly, the probabilities p_o and p_e of any arbitrary pico BS are independent of any other pico BS.

We refer to p_m as the load of a typical macro cell. It can also be interpreted as the probability that a typical macro BS is contributing to network interference because a typical user being served on a subchannel receives interference from only those BSs which are active on that particular subchannel. While the authors in [27] used mean statistic of the number of users associated with a typical BS to approximate this probability by $\min\left(\frac{\text{Average number of users}}{\text{Total number of frequency resource blocks available}}, 1\right)$, we derive the exact probability in this paper. Similarly, p_o and p_e are referred to as the loads of a typical pico cell in two different groups of frequency resources S_c and S_r, respectively. We can observe that user density, association bias and the degree of resource partitioning directly affect the cell load.

Having derived the cell loads, the signal-to-interference-and-noise ratio (SINR) of a typical user u located at the origin when it belongs to Φ_u^l, denoted by SINR_l, for each $l \in \{m, o, e\}$ can be computed as

$$\text{SINR}_m = \frac{P_m h_{b_m} D_m^{-\alpha_m}}{P_m \sum_{x_m \in \Psi_m \backslash b_m} h_{x_m} ||x_m||^{-\alpha_m} + P_p \sum_{x_o \in \Psi_o} h_{x_o} ||x_o||^{-\alpha_p} + \sigma^2},$$
$$(8)$$

$$\text{SINR}_o = \frac{P_p h_{b_o} D_o^{-\alpha_p}}{P_m \sum_{x_m \in \Psi_m} h_{x_m} ||x_m||^{-\alpha_m} + P_p \sum_{x_o \in \Psi_o \backslash b_o} h_{x_o} ||x_o||^{-\alpha_p} + \sigma^2},$$
$$(9)$$

$$\text{SINR}_e = \frac{P_p h_{b_e} D_e^{-\alpha_p}}{P_p \sum_{x_e \in \Psi_e \backslash b_e} h_{x_e} ||x_e||^{-\alpha_p} + \sigma^2},$$
$$(10)$$

where b_l is the serving BS at a distance D_l from the user u when it belongs to Φ_u^l and σ^2 is the noise power. If the

user u is being served on a subchannel from the set S_c (i.e. when it is either a macro user or an unbiased-pico user), $\mathbf{\Psi}_m$ and $\mathbf{\Psi}_o$ are the sets of macro and pico BSs respectively, that are active on that particular subchannel. The sets $\mathbf{\Psi}_m$ and $\mathbf{\Psi}_o$ are independent thinnings of the the original PPPs $\mathbf{\Phi}_m$ and $\mathbf{\Phi}_p$, respectively, with retention probabilities p_m and p_o, respectively. Hence, they are independent PPPs with densities $p_m\lambda_m$ and $p_o\lambda_p$, respectively. Similarly, $\mathbf{\Psi}_e$ is the set of active pico BSs on the subchannel the user u being served from the set S_r (i.e. when u is a range-expanded pico user) and is also a PPP with density $p_e\lambda_p$.

III. AVERAGE USER DATA RATE

The motivation behind CRE supported by resource partitioning in cellular HetNets is to provide high data rate to users through load balancing. Thus, the performance metric chosen in this paper is the average user data rate that can be attained over the entire network.

Theorem 1. *The average data rate per unit bandwidth of a typical user u is given by*

$$\bar{R} = \sum_{l\in\{m,o,e\}} U_l\, T_l\, \mathbb{E}\left[\log_2(1 + \textit{SINR}_l)\right], \qquad (11)$$

where $U_l = \mathbb{P}(u \in \mathbf{\Phi}_u^l)$, which is given by Lemma 1, and T_l is the average time share of frequency resources received by the user u when it belongs to $\mathbf{\Phi}_u^l$.

Proof. With adaptive transmission scheme so that the Shannon bound can be achieved and treating interference as noise [22], the data rate per unit bandwidth of a typical user u, conditioned on u belonging to $\mathbf{\Phi}_u^l$, is given by

$$R_l = t_l\log_2(1 + \text{SINR}_l), \ \forall l \in \{m,o,e\}, \qquad (12)$$

where t_l is the fraction of time the user u is served on a subchannel. Let N_l' be the number of other users in the cell to which the user u belongs. If the total number of users is no greater than $L_{\kappa(l)}$ (i.e. $N_l' + 1 \le L_{\kappa(l)}$), where $L_{\kappa(m)} = L_{\kappa(o)} = L_c$ and $L_{\kappa(e)} = L_r$, the user u can exclusively occupy a subchannel without time sharing and thus, $t_l = 1$. Otherwise, the subchannels are time-shared equally among the total users and thus, $t_l = L_{\kappa(l)}/(N_l' + 1)$.

The average data rate per unit bandwidth of the user $u \in \mathbf{\Phi}_u^l$ is $\bar{R}_l = \mathbb{E}[R_l] = \mathbb{E}[t_l\log_2(1 + \text{SINR}_l)]$. Since the number of users associated with a BS determines its probability of being active on a certain subchannel (Lemma 3), the total interference received by the user u, and thus its SINR depend on the number of users associated with the cells other than the serving cell in the network. t_l, however is the function of number of users in the serving cell. As we discussed in II-B, the number of users in different cells are independent, which implies that t_l and SINR_l are independent. Thus, $\mathbb{E}[t_l\log_2(1 + \text{SINR}_l)] = T_l\,\mathbb{E}[\log_2(1 + \text{SINR}_l)]$, where $T_l = \mathbb{E}[t_l]$. According to the law of total expectation, the overall data rate of a typical user u is then given by $\bar{R} = \sum_{l\in\{m,o,e\}} U_l\,\bar{R}_l$. By using Lemma 4 and 5, which derive $T_l = \mathbb{E}[t_l]$ and the average spectral efficiency $\mathbb{E}\left[\log_2(1 + \text{SINR}_l)\right]$, respectively, the final expression for the average user data rate is obtained. $\square$

In the following Lemma, we derive T_l, which is required to compute the average user data rate in (11).

Lemma 4. *The average time-share of frequency resources, T_l received by a typical user u when $u \in \mathbf{\Phi}_u^l$ is given by*

$$T_l = \frac{L_{\kappa(l)}\lambda_{\zeta(l)}}{U_l\lambda_u}\left(1 - \left(1 + 3.5^{-1}U_l\lambda_u/\lambda_{\zeta(l)}\right)^{-3.5}\right)$$
$$- \frac{3.5^{3.5}}{\Gamma(3.5)}\sum_{n=1}^{L_{\kappa(l)}} \frac{\Gamma(3.5+n)\left(U_l\lambda_u/\lambda_{\zeta(l)}\right)^{n-1}\left(L_{\kappa(l)} - n\right)}{n!\left(U_l\lambda_u/\lambda_{\zeta(l)} + 3.5\right)^{3.5+n}},$$
$$\forall l \in \{m,o,e\}, \quad (13)$$

where $\lambda_{\zeta(m)} = \lambda_m$, $\lambda_{\zeta(o)} = \lambda_{\zeta(e)} = \lambda_p$, $L_{\kappa(m)} = L_{\kappa(o)} = L_c$ and $L_{\kappa(e)} = L_r$.

Proof. The proof is given in Appendix A. $\square$

To finally compute the average user data rate in (11), we now derive the average link spectral efficiency $\mathbb{E}\left[\log_2(1 + \text{SINR}_l)\right]$.

Lemma 5. *The average link spectral efficiency $\mathbb{E}\left[\log_2(1 + SINR_l)\right]$ of the user u when it belongs to $\mathbf{\Phi}_u^l$, denoted by C_l is given by*

$$C_l = \frac{1}{\ln 2}\int_0^\infty \frac{\bar{F}_l(t)}{1+t}dt, \ \forall l \in \{m,o,e\}, \qquad (14)$$

where $\bar{F}_l(t) = \mathbb{P}(SINR_l \ge t)$ is the conditional complementary cumulative distribution function (CCDF) of the SINR of the user u when it belongs to $\mathbf{\Phi}_u^l$, and is given by (15)-(17) for each $l \in \{m,o,e\}$.

If the noise is ignored (i.e., the network is interference limited) and the path-loss exponents are assumed equal ($\alpha_m = \alpha_p = \alpha$), then $\bar{F}_m(t)$, $\bar{F}_o(t)$ and $\bar{F}_e(t)$ can be simplified to (18)-(20).

Proof. The proof is given in Appendix B. $\square$

These conditional SINR distributions can also be interpreted as conditional coverage probability as $\mathbb{P}(\text{SINR}_l > \tau)$ is the probability that a randomly chosen user can achieve the target SINR τ under the condition that the user belongs to $\mathbf{\Phi}_u^l$.

Unlike the SINR distribution in [11], in this paper, it is also dependent on user density and the degree of resource partitioning, apart from association bias. In the full-load model, the impact of biasing on SINR distribution is only due to the associated users' geometry (i.e., with biased association, only very good geometry users are served by macro cells while pico users now also include worse geometry users). In this paper, biasing also affects the interference power through cell load. Thus, the impact of biasing on SINR distribution is better captured than [11]. The SINR distribution of unbiased pico users is also dependent on association bias unlike the one in [11] as the interference from macro tier to unbiased pico users depends on biasing. These claims will be verified through numerical results in Figure 2.

The special case of no resource partitioning will be addressed next so that the average data rate with resource partitioning can be compared against the no resource partitioning case, subsequently quantifying the gain.

$$\bar{F}_m(t) = \frac{2\pi\lambda_m}{U_m} \int_0^\infty r \exp\left\{-tP_m^{-1}\sigma^2 r^{\alpha_m} - \pi\lambda_m r^2\left(1 + \frac{2p_m}{(\alpha_m-2)}\frac{1}{(1+1/t)}\,_2F_1\left[1,1,2-\frac{2}{\alpha_m},\frac{1}{1+1/t}\right]\right)\right.$$
$$\left. -\pi\lambda_p\left(P_pP_m^{-1}\right)^{2/\alpha_p} r^{2\alpha_m/\alpha_p}\left(B^{2/\alpha_p} + \frac{2p_o}{(\alpha_p-2)}\frac{B^{2/\alpha_p}}{(1+B/t)}\,_2F_1\left[1,1,2-\frac{2}{\alpha_p},\frac{1}{1+B/t}\right]\right)\right\} dr, \tag{15}$$

$$\bar{F}_o(t) = \frac{2\pi\lambda_p}{U_o} \int_0^\infty r \exp\left\{-tP_p^{-1}\sigma^2 r^{\alpha_p} - \pi\lambda_p r^2\left(1 + \frac{2p_o}{(\alpha_p-2)}\frac{1}{(1+1/t)}\,_2F_1\left[1,1,2-\frac{2}{\alpha_p},\frac{1}{1+1/t}\right]\right)\right.$$
$$\left. -\pi\lambda_m\left(P_mP_p^{-1}\right)^{2/\alpha_m} r^{2\alpha_p/\alpha_m}\left(1 + \frac{2p_m}{(\alpha_m-2)}\frac{1}{(1+1/t)}\,_2F_1\left[1,1,2-\frac{2}{\alpha_m},\frac{1}{1+1/t}\right]\right)\right\} dr, \tag{16}$$

$$\bar{F}_e(t) = \frac{2\pi\lambda_p}{U_e} \int_0^\infty r \exp\left\{-tP_p^{-1}\sigma^2 r^{\alpha_p} - \pi\lambda_p r^2\left(1 + \frac{2p_e}{(\alpha_p-2)}\frac{1}{(1+1/t)}\,_2F_1\left[1,1,2-\frac{2}{\alpha_p},\frac{1}{1+1/t}\right]\right)\right.$$
$$\left. -\pi\lambda_m\left(P_mP_p^{-1}\right)^{2/\alpha_m} r^{2\alpha_p/\alpha_m}\right\}\left\{\exp\left(-\pi\lambda_m\left(P_mP_p^{-1}\right)^{2/\alpha_m} r^{2\alpha_p/\alpha_m}(B^{-2/\alpha_m}-1)\right) - 1\right\} dr. \tag{17}$$

$$\bar{F}_m(t) = U_m^{-1}\left\{1 + \frac{2p_m}{(\alpha-2)}\frac{1}{(1+1/t)}\,_2F_1\left[1,1,2-\frac{2}{\alpha},\frac{1}{1+1/t}\right] + \frac{\lambda_p}{\lambda_m}\left(\frac{P_p}{P_m}B\right)^{2/\alpha}\right.$$
$$\left. \times\left(1 + \frac{2p_o}{(\alpha-2)}\frac{1}{(1+B/t)}\,_2F_1\left[1,1,2-\frac{2}{\alpha},\frac{1}{1+B/t}\right]\right)\right\}^{-1}, \tag{18}$$

$$\bar{F}_o(t) = U_o^{-1}\left\{1 + \frac{\lambda_m}{\lambda_p}\left(\frac{P_m}{P_p}\right)^{2/\alpha} + \frac{2}{(\alpha-2)}\frac{1}{(1+1/t)}\right.$$
$$\left. \times\,_2F_1\left[1,1,2-\frac{2}{\alpha},\frac{1}{1+1/t}\right]\left(p_o + \frac{\lambda_m p_m}{\lambda_p}\left(\frac{P_m}{P_p}\right)^{2/\alpha}\right)\right\}^{-1}, \tag{19}$$

$$\bar{F}_e(t) = U_e^{-1}\left[\left\{1 + \frac{2p_e}{(\alpha-2)}\frac{1}{(1+1/t)}\,_2F_1\left[1,1,2-\frac{2}{\alpha_p},\frac{1}{1+1/t}\right] + \frac{\lambda_m}{\lambda_p}\left(\frac{P_m}{P_pB}\right)^{2/\alpha}\right\}^{-1}\right.$$
$$\left. -\left\{1 + \frac{2p_e}{(\alpha-2)}\frac{1}{(1+1/t)}\,_2F_1\left[1,1,2-\frac{2}{\alpha_p},\frac{1}{1+1/t}\right] + \frac{\lambda_m}{\lambda_p}\left(\frac{P_m}{P_p}\right)^{2/\alpha}\right\}^{-1}\right]. \tag{20}$$

A. Special Case: No resource partitioning

If no resource partitioning is applied, the pico users need not be categorized as unbiased and range-expanded users because they are served from the same pool of total L subchannels in each pico cell. Let U_p denotes the probability that a randomly chosen user is a pico user. Then,

$$U_p = 1 - U_m,$$

where U_m is the probability that a randomly chosen user is a macro user, and is derived in Lemma 1. The total users in the network can thus be divided into two sets: Φ_u^m, the set of macro users, and Φ_u^p, the set of pico users. Φ_u^m and Φ_u^p are independent PPPs of densities $U_m\lambda_u$ and $U_p\lambda_u$, respectively.

With no resource partitioning, each macro BS can access all the available subchannels in the system to serve its users (i.e., a total of L subchannels). Meanwhile, in each pico cell, as explained earlier, the users (either unbiased or range-expanded users) are served from the same pool of L subchannels. The cell loads of a typical macro BS and a typical pico BS, denoted by p_m and p_p, respectively, are thus given by

$$p_j = 1 - \frac{3.5^{3.5}}{\Gamma(3.5)}\frac{1}{L}\sum_{n=0}^{L-1}\left[\frac{(L-n)\Gamma(3.5+n)}{n!}\right.$$
$$\left. \times\frac{(U_j\lambda_u/\lambda_j)^n}{(U_j\lambda_u/\lambda_j + 3.5)^{3.5+n}}\right], \forall j \in \{m,p\}. \tag{21}$$

In any cell, either a macro or a pico cell, if the total associated users are less than L, all the users can exclusively occupy a subchannel, otherwise the subchannels have to be time shared among the users. The average time shares received by a typical macro user and a typical pico user, denoted by T_m and T_p, respectively, are given by

$$T_j = \frac{L\lambda_j}{U_j\lambda_u}\left(1 - \left(1 + 3.5^{-1}U_j\lambda_u/\lambda_j\right)^{-3.5}\right)$$
$$-\frac{3.5^{3.5}}{\Gamma(3.5)}\sum_{n=1}^{L}\frac{\Gamma(3.5+n)(U_j\lambda_u/\lambda_j)^{n-1}(L-n)}{n!(U_j\lambda_u/\lambda_j+3.5)^{3.5+n}},$$
$$\forall j \in \{m,p\}. \tag{22}$$

The average user data rate per unit bandwidth for the case with no resource partitioning can finally be expressed as

$$\bar{R} = U_m T_m C_m + U_p T_p C_p, \qquad (23)$$

where $C_j = \mathbb{E}\left[\log_2(1+\text{SINR}_m)\right]$ and $C_p = \mathbb{E}\left[\log_2(1+\text{SINR}_p)\right]$ are the spectral efficiencies of a typical user u located at the origin when it belongs to $\mathbf{\Phi}_u^m$ and $\mathbf{\Phi}_u^p$, respectively. The corresponding SINRs are given by

$$\text{SINR}_m = \frac{P_m h_{b_m} D_m^{-\alpha_m}}{P_m \sum\limits_{x_m \in \mathbf{\Psi}_m \backslash b_m} h_{x_m}||x_m||^{-\alpha_m} + P_p \sum\limits_{x_p \in \mathbf{\Psi}_p} h_{x_p}||x_p||^{-\alpha_p} + \sigma^2}, \qquad (24)$$

$$\text{SINR}_p = \frac{P_p h_{b_p} D_p^{-\alpha_p}}{P_m \sum\limits_{x_m \in \mathbf{\Psi}_m} h_{x_m}||x_m||^{-\alpha_m} + P_p \sum\limits_{x_p \in \mathbf{\Psi}_p \backslash b_p} h_{x_p}||x_p||^{-\alpha_p} + \sigma^2}, \qquad (25)$$

where b_j is the serving BS at a distance D_j from u when $u \in \mathbf{\Phi}_u^j$, $j \in \{m,p\}$. $\mathbf{\Psi}_m$ and $\mathbf{\Psi}_p$ are the sets of macro and pico BSs respectively, that are active on the subchannel the user u is being served and they are PPPs with densities $p_m \lambda_m$ and $p_p \lambda_p$, respectively. As in Lemma 5, the spectral efficiencies can be derived as

$$C_j = \frac{1}{\ln 2} \int_0^\infty \frac{\bar{F}_j(t)}{1+t} dt, \ \forall j \in \{m,p\}, \qquad (26)$$

where, $\bar{F}_m(t)$ and $\bar{F}_p(t)$ are given by (27)-(28).

When the network is interference limited (i.e., the noise is ignored) and $\alpha_m = \alpha_p = \alpha$, the simplified expressions for $\bar{F}_m(t)$ and $\bar{F}_p(t)$ can be obtained as (29)-(30)

Remark: The results for the case of neither biasing nor resource partitioning can be found by substituting $B = 1$ in the above results.

IV. SIMULATION AND NUMERICAL RESULTS

In this section, we present numerical analysis and validation of our analytical results. Unless otherwise stated, we choose $L = 20$, $\lambda_m = 1\text{BS/km}^2$, $P_{\max}^m = 46$ dBm, $P_{\max}^p = 30$ dBm, $\sigma^2 = -109$ dBm, $\alpha_m = 3.5$ and $\alpha_p = 4$. The degree of resource partitioning is expressed by the fraction $\mu = L_r/L$.

The average user data rate (11) is validated in Figure 1 via Monte Carlo simulations on a square window of $20 \times 20\,\text{km}^2$. We can observe that the analytical results match the simulation results quite well. The small gaps are mainly due to the approximation for cell areas distribution.

Before numerically analyzing the average user data rate, we first analyze the conditional coverage probabilities (15)-(17) so that the data rate trends can be better understood. We can observe in Figure 2 that the coverage probability of unbiased pico user increases with the increase in bias. This is because, when more macro users are offloaded to pico tier, the cell load p_m of macro tier decreases and so does the interference from macro tier. Similarly, for macro user and range-expanded pico user, apart from users' geometry, the variation in the coverage probability with bias is mainly due to the change in p_m and p_e. The coverage probability of each user type decreases as user density increases because more BSs
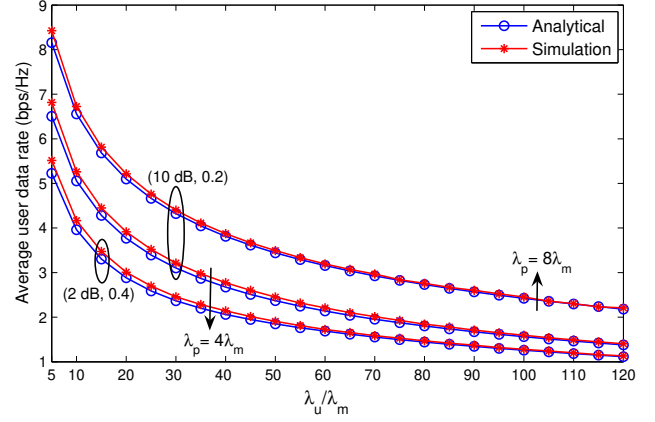

Fig. 1. Validation of the analytical result for average user data rate (11) via Monte Carlo simulations for different values of user density λ_u, pico cell density λ_p, and association bias and resource partitioning fraction (B, μ).
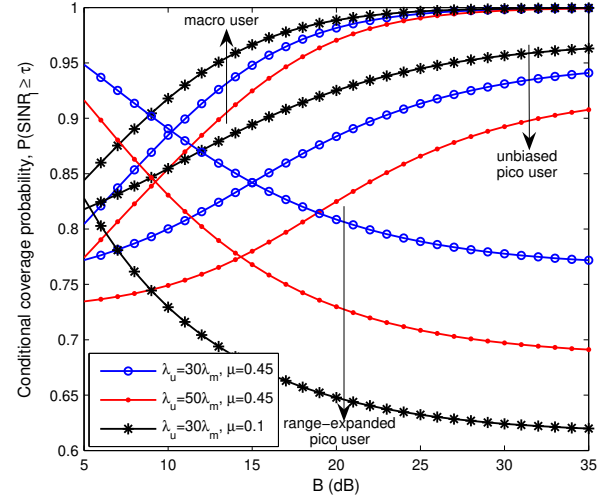

Fig. 2. Impact of association bias B, user density λ_u and resource partitioning fraction μ on conditional coverage probabilities of macro user and both types of pico user (unbiased and range-expanded): $\lambda_p = 4\lambda_m$, $\tau = 0.5$.

become active to serve the increased number of users which directly affect the interference. We can observe that resource partitioning fraction μ directly impact the coverage probability. As μ decreases, more subchannels become available for macro users and unbiased pico user, which decreases p_m and p_o, consequently increasing the coverage probabilities of macro and unbiased pico users due to decrease in interference. The coverage probability of range-expanded user on the other hand decreases due to increase in p_e.

In Figure 3, the average user data rate of biased association with and without resource partitioning is compared against that of unbiased association in different load conditions. In our model, the network load is directly proportional to user density. The user data rate decreases with increasing load due to increase in interference and decrease in the users' share of resources. In biased association without resource partitioning, the SINR of offloaded users degrades. However, in a lightly loaded network, they are offloaded to BSs offering the higher

$$\bar{F}_m(t) = \frac{2\pi\lambda_m}{U_m}\int_0^\infty r\exp\left\{-tP_m^{-1}\sigma^2 r^{\alpha_m} - \pi\lambda_m r^2\left(1 + \frac{2p_m}{(\alpha_m-2)}\frac{1}{(1+1/t)}\,_2F_1\left[1,1,2-\frac{2}{\alpha_m},\frac{1}{1+1/t}\right]\right)\right.$$
$$\left.-\pi\lambda_p\left(P_pP_m^{-1}\right)^{2/\alpha_p}r^{2\alpha_m/\alpha_p}\left(B^{2/\alpha_p} + \frac{2p_p}{(\alpha_p-2)}\frac{B^{2/\alpha_p}}{(1+B/t)}\,_2F_1\left[1,1,2-\frac{2}{\alpha_p},\frac{1}{1+B/t}\right]\right)\right\}dr, \tag{27}$$

$$\bar{F}_p(t) = \frac{2\pi\lambda_p}{U_p}\int_0^\infty r\exp\left\{-tP_p^{-1}\sigma^2 r^{\alpha_p} - \pi\lambda_p r^2\left(1 + \frac{2p_p}{(\alpha_p-2)}\frac{1}{(1+1/t)}\,_2F_1\left[1,1,2-\frac{2}{\alpha_p},\frac{1}{1+1/t}\right]\right)\right.$$
$$\left.-\pi\lambda_m\left(P_mP_p^{-1}\right)^{2/\alpha_m}r^{2\alpha_p/\alpha_m}\left(B^{-2/\alpha_m} + \frac{2p_m}{(\alpha_m-2)}\frac{B^{-2/\alpha_m}}{(1+1/(Bt))}\,_2F_1\left[1,1,2-\frac{2}{\alpha_m},\frac{1}{1+1/(Bt)}\right]\right)\right\}dr. \tag{28}$$

$$\bar{F}_m(t) = U_m^{-1}\left\{1 + \frac{2p_m}{(\alpha-2)}\frac{1}{(1+1/t)}\,_2F_1\left[1,1,2-\frac{2}{\alpha},\frac{1}{1+1/t}\right] + \frac{\lambda_p}{\lambda_m}\left(\frac{P_p}{P_m}B\right)^{2/\alpha}\right.$$
$$\left.\times\left(1 + \frac{2p_p}{(\alpha-2)}\frac{1}{(1+B/t)}\,_2F_1\left[1,1,2-\frac{2}{\alpha},\frac{1}{1+B/t}\right]\right)\right\}^{-1}, \tag{29}$$

$$\bar{F}_p(t) = U_p^{-1}\left\{1 + \frac{2p_p}{(\alpha-2)}\frac{1}{(1+1/t)}\,_2F_1\left[1,1,2-\frac{2}{\alpha},\frac{1}{1+1/t}\right] + \frac{\lambda_m}{\lambda_p}\left(\frac{P_m}{P_pB}\right)^{2/\alpha}\right.$$
$$\left.\times\left(1 + \frac{2p_m}{(\alpha-2)}\frac{1}{(1+1/(Bt))}\,_2F_1\left[1,1,2-\frac{2}{\alpha},\frac{1}{1+1/(Bt)}\right]\right)\right\}^{-1}. \tag{30}$$
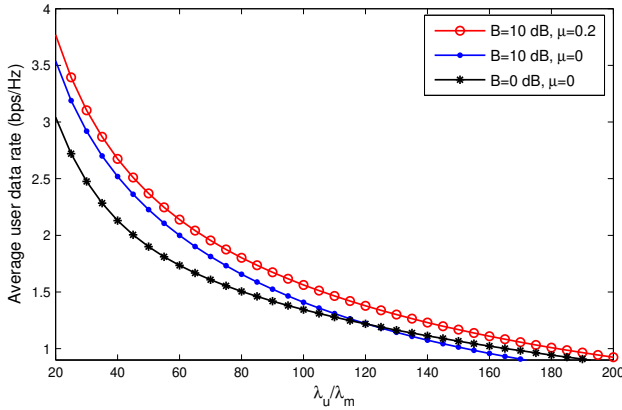

Fig. 3. Effect of CRE with and without resource partitioning on user data rate, as user density is varied: $\lambda_p = 4\lambda_m$

share of resources and thus the user data rate improves compared to unbiased association. But when the network is heavily loaded, the decrease in SINR dominates and unbiased association outperforms biasing. The SINR degradation of offloaded users can be compensated by resources partitioning. The resource partitioning fraction of 0.2, for example, is shown to outperform no resource partitioning scenario in terms of average user data rate in Figure 3 in any load condition.

Since resource partitioning costs macro tier its available resources, the resource partitioning fraction μ must be coordinated within the network for optimal user data rate. The optimal pair (B, μ) for the given network parameters is investigated in Figure 4. It is found to be strongly dependent on user density (i.e., network load). For $\lambda_u = 30\lambda_m$, the optimal

pair is $(29\text{dB}, 0.45)$, while for $\lambda_u = 100\lambda_m$, the optimal pair is $(30\text{dB}, 0.1)$.

With resource partitioning, for the given value of μ, the average link spectral efficiency C_m of a typical macro user increases with the increase in bias because more macro users with low SINR (users far from the serving macro BS) are offloaded to pico cells, and the interference from other macro cells also decreases due to the decrease in macro cell load. Note that the interference from pico tier is invariant to biasing because the pico cell load in the resource group S_c is independent of bias. Meanwhile, the share of radio resources received by macro users also increases. Thus, the contribution of macro users towards the average data rate increases with the increase in bias. But after a certain association bias, it eventually decreases due to the decreasing fraction of macro users. Similarly, the contribution from range-expanded pico users initially increases with increase in bias due to the increasing fraction of range-expanded users, but eventually decreases after a certain bias due to the decrease in average link spectral efficiency C_e and the increase in number of users sharing the resources. The decrease in C_e is due to the fact that more users with low SINR (users far from the prospective pico BS) are associated with the pico cell, and in the meantime, the interference from other pico cells increases due to increase in pico cell load in the resource group S_r. On the other hand, the fraction of unbiased pico users is invariant to biasing and hence, so is the share of radio resources received by unbiased users. However, with increasing bias, the contribution to average data rate from these unbiased users significantly increases because of the decreasing macro-tier interference as a result of the decrease in macro cell load. Overall, the average data rate initially increases with the increase in bias,
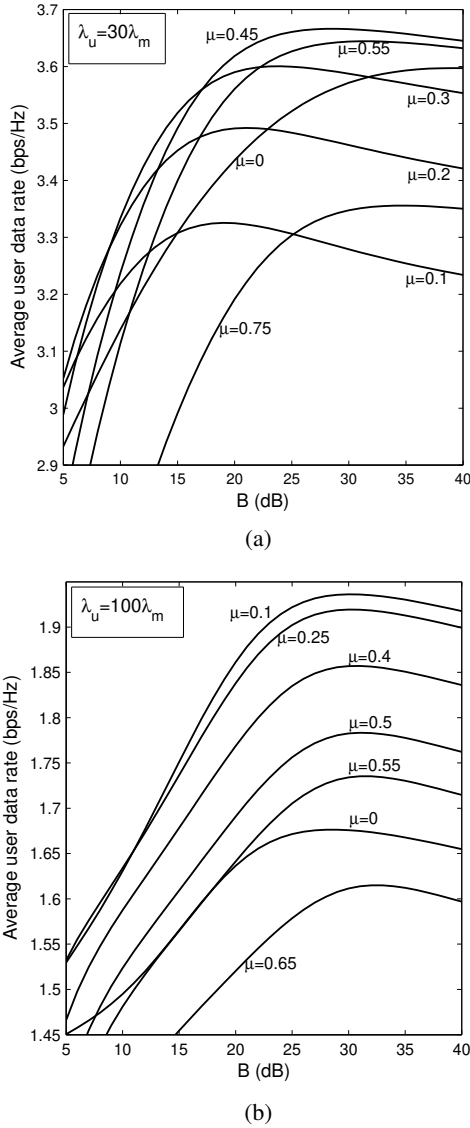
Fig. 4. Average user data rate vs. association bias B for different values of resource partitioning fraction μ in lightly loaded (a) and heavily loaded (b) network conditions: $\lambda_p = 4\lambda_m$, $\alpha_m = \alpha_p = 4$, $\sigma^2 = 0$.

but decreases beyond a certain association bias and hence, the optimal bias exist.

It can be observed from Figure 4 that for the case with resource partitioning, the optimal bias increases with increasing μ because more resources are reserved for the offloaded users. For $\lambda_u = 30\lambda_m$ (lightly loaded network), the optimal bias lies in the range of 19 dB and above as μ increases from 0.1, whereas, it lies in the range of 30 dB and above for $\lambda_u = 100\lambda_m$ (heavily loaded network).

With no resource partitioning (i.e., $\mu = 0$), the contribution from macro users to average user data rate varies with bias in the same way as in the case of resource partitioning. The contribution from pico users however, has a different variation as both the unbiased and range-expanded users are now served from the same pool of L subchannels. The average link spectral efficiency C_p of a randomly selected pico user (either unbiased or range-expanded) initially drops as more users with

poor SINR (users with higher average received power from the nearest macro BS as compared to the nearest pico BS) are associated with the pico cell. But, beyond a certain bias, the decreasing macro-tier interference causes C_p to improve. If the increasing fraction of pico users dominates the initial drop in C_p and the decreasing share of radio resources, the contribution from pico users to average data rate increases with increasing bias. Otherwise, it may drop initially, but eventually increases as C_p improves. When the number of pico users sharing the resources become large, the contribution towards the average user rate subsequently decreases. The initial drop in the contribution from pico users towards average user data rate is the reason for the initially low rate of increase in the average user data rate for $\lambda_u = 100\lambda_m$ in Figure 4.

The optimal bias values for $\lambda_u = 30\lambda_m$ and $\lambda_u = 100\lambda_m$ are found to be 39 dB and 29 dB, respectively, in Figure 4 with $\mu = 0$. Thus, with no resource partitioning, the optimal bias decreases with increasing user density as large bias values will make the pico cells overly congested with poor SINR users in a heavily loaded network.

The variation of average user data rate with resource partitioning fraction μ for the given bias value is plotted in Figure 5. As previously explained while analyzing conditional coverage probabilities in Figure 2, the cell loads p_m and p_o increase with increasing μ. The average spectral efficiency of a typical macro user thus decreases with increasing μ due to the increasing interference from macro and pico tier as a result of the increasing cell load, and so does the average spectral efficiency of a typical unbiased pico user. This, together with the decrease in the average share of radio resources received by users, causes the average data rate of both macro and unbiased pico users to decrease with increasing μ. On the other hand, as more subchannels are available for range-expanded pico users with increasing μ, their average data rate increases. The net result is the initial increase in the average data rate with increasing μ and the subsequent decrease beyond a certain value of μ. With full-load assumption, the spectral efficiency would be independent of μ and data rate would vary only due to the change in the users' share of radio resources.

As observed in Figure 5, in a lightly loaded network ($\lambda_u = 30\lambda_m$), the optimal resource partitioning fraction shifts towards higher values as association bias increases, whereas in a heavily loaded network ($\lambda_u = 100\lambda_m$), it shifts towards lower values. This shows that in a heavily loaded network, when a large number of macro users are offloaded to pico cells, allocating more resources to serve these offloaded users will rather highly degrade the rate of remaining macro users and unbiased pico users instead of improving the average data rate. Thus, lower resource partitioning fraction is desirable.

In Figure 6, we analyze how small cell density affects the optimal choices of association bias and resource partitioning fraction. It is clearly visible from the figure that irrespective of the user density and association bias, the average user data rate always increases with the increase in pico cell density as the number of users served by each BS decreases and thus users get access to larger fraction of resources. With no resource partitioning, whether the network is lightly or heavily loaded, the optimal bias is found to increase as pico
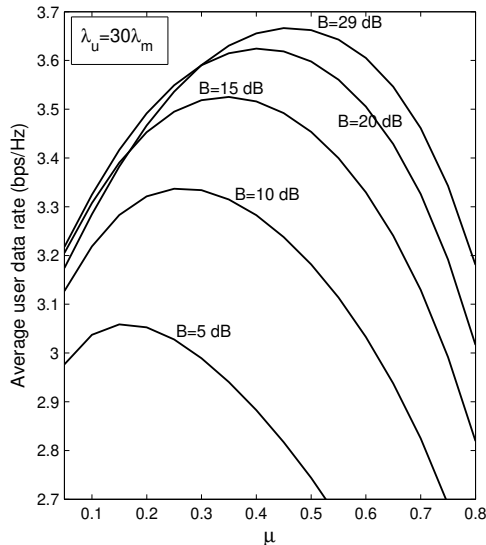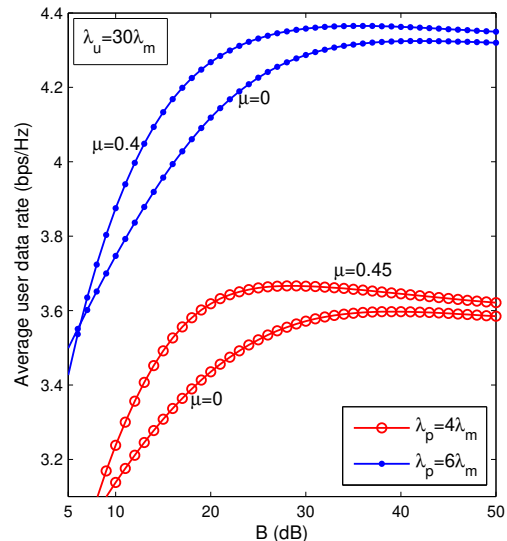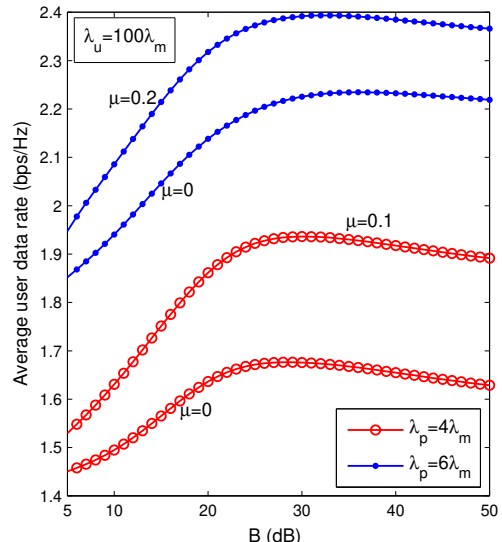
(a)
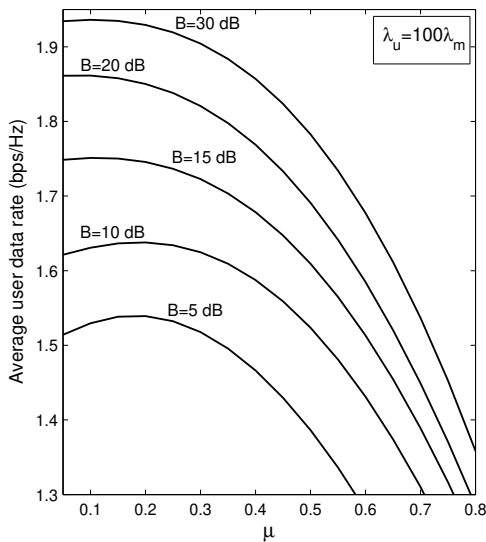
(b)

Fig. 5. Average user data rate vs. resource partitioning fraction μ for different values of association bias B in lightly loaded (a) and heavily loaded (b) network conditions: $\lambda_p = 4\lambda_m$, $\alpha_m = \alpha_p = 4$, $\sigma^2 = 0$.

(a)

(b)

Fig. 6. Effect of pico cell density λ_p on the optimal choices of association bias B and resource partitioning fraction μ in lightly loaded (a) and heavily loaded (b) network conditions: $\alpha_m = \alpha_p = 4$, $\sigma^2 = 0$.

cell density increases. The optimal bias increased from 39 dB to 42 dB for $\lambda_u = 30\lambda_m$ and 29 dB to 36 dB for $\lambda_u = 100\lambda_m$ when λ_p changed from $4\lambda_m$ to $6\lambda_m$. With resource partitioning, the optimal pair (B, μ) changed from $(29\text{dB}, 0.45)$ to $(36\text{dB}, 0.4)$ for $\lambda_u = 30\lambda_m$ and $(30\text{dB}, 0.1)$ to $(33\text{dB}, 0.2)$ for $\lambda_u = 100\lambda_m$ as λ_p increased from $4\lambda_m$ to $6\lambda_m$. When λ_p further increased to $8\lambda_m$, the optimal pairs for $\lambda_u = 30\lambda_m$ and $\lambda_u = 100\lambda_m$ were found to be $(38\text{dB}, 0.3)$ and $(34\text{dB}, 0.25)$, respectively (not shown in the figure). Thus, as in the case with no resource partitioning, the optimal bias increases with increasing pico cell density in both lightly and heavily loaded network conditions. The optimal μ, however decreases with increasing pico cell density in a lightly loaded network.

V. CONCLUSION

We developed an analytical framework to evaluate the downlink performance of cellular HetNets with CRE and resource partitioning in a multi-channel environment, while taking cell load into account. The incorporated cell load model effectively captures the impact of user offloading and resource partitioning on network interference. The performance is evaluated in terms of average user data rate that can be achieved over the entire network. We observed that if CRE is supported by resource partitioning, the average user data rate can be highly boosted. However, the bias value B and resource partitioning fraction μ must be carefully tuned. With the optimal pair (B, μ), the gain can be as high as 115%. Our analysis showed that the optimal pair must be updated in accordance with the changing network load.

APPENDIX

A. Proof of Lemma 4

We have,

$$t_l = \begin{cases} 1 & \text{if } N'_l + 1 \le L_{\kappa(l)} \\ \frac{L_{\kappa(l)}}{(N'_l+1)} & \text{otherwise.} \end{cases} \quad (31)$$

Thus, $T_l = \mathbb{E}[t_l]$ can be derived as

$$\begin{aligned} T_l &= \sum_{n=0}^{L_{\kappa(l)}-1} \mathbb{P}(N'_l = n) + \sum_{n=L_{\kappa(l)}}^{\infty} \frac{L_{\kappa(l)}}{n+1} \mathbb{P}(N'_l = n) \\ &= \sum_{n=1}^{\infty} \frac{L_{\kappa(l)}}{n} \mathbb{P}(N'_l = n-1) \\ &\quad - \sum_{n=1}^{L_{\kappa(l)}} \left(\frac{L_{\kappa(l)}}{n} - 1 \right) \mathbb{P}(N'_l = n-1). \end{aligned} \quad (32)$$

We know, the probability that a typical user belongs to a given cell is directly proportional to the area of the cell. Thus, the conditional PDF of the area of a Voronoi cell given that a typical user belongs to it is given by $f_{A'}(a) = caf_A(a)$, where $f_A(a)$ is the unconditional PDF and c is a constant such that $\int_0^{\infty} f_{A'}(a)da = 1$. The PMF of N'_l can then be similarly derived as in Lemma 2 as

$$\mathbb{P}(N'_l = n) = \frac{3.5^{3.5}\Gamma(4.5+n)(U_l\lambda_u/\lambda_{\zeta(l)})^n}{\Gamma(3.5)n!(U_l\lambda_u/\lambda_{\zeta(l)} + 3.5)^{n+4.5}}, \, n \ge 0,$$
$$\forall l \in \{m, o, e\}, \quad (33)$$

where $\lambda_{\zeta(m)} = \lambda_m$ and $\lambda_{\zeta(o)} = \lambda_{\zeta(e)} = \lambda_p$. The final expression for t_l in (13) is obtained by substituting the PMF of N'_l in (32), where the first term is further simplified by using $\mathbb{P}(N'_l = n-1)/n = \lambda_{\zeta(l)}/(U_l\lambda_u)\mathbb{P}(N_l = n)$, followed by $\sum_{n=0}^{\infty} \mathbb{P}(N_l = n) = 1$.

B. Proof of Lemma 5

The average link spectral efficiency C_l of the user u when $u \in \Phi_u^l$, can be expressed as

$$\begin{aligned} C_l &= \frac{1}{\ln 2} \int_0^{\infty} \ln(1+t) f_l(t) dt \\ &= -\frac{1}{\ln 2} \int_0^{\infty} \ln(1+t) d\bar{F}_l(t), \end{aligned}$$

where $f_l(t)$ is the conditional PDF of the SINR of the user u, given that $u \in \Phi_u^l$. The second equality is obtained by expressing $f_l(t)$ in terms of CCDF $\bar{F}_l(t) = \mathbb{P}(\text{SINR}_l > t)$ as $f_l(t) = -\frac{d\bar{F}_l(t)}{dt}$. Equation (14) is then obtained by using integration by parts.

By using the SINR expression (8), the SINR distribution of the typical user u conditioned on $u \in \Phi_u^m$ can be derived as follows:

$$\bar{F}_m(t) = \mathbb{P}\left(\frac{P_m h_{b_m} D_m^{-\alpha_m}}{I_{b_m,m} + I_{b_m,o} + \sigma^2} > t \right), \quad (34)$$

where $I_{b_m,m} = P_m \sum_{x_m \in \Psi_m \backslash b_m} h_{x_m} ||x_m||^{-\alpha_m}$ and $I_{b_m,o} = P_p \sum_{x_o \in \Psi_o} h_{x_o} ||x_o||^{-\alpha_p}$. By utilizing the fact that $h_{b_m} \sim$

Exp(1) and the independence between $I_{b_m,m}$ and $I_{b_m,o}$, $\bar{F}_m(t)$ can be further expressed as

$$\begin{aligned} \bar{F}_m(t) = \int_0^{\infty} &\exp\left(-t\sigma^2 P_m^{-1} r^{\alpha_m}\right) \mathcal{L}_{I_{b_m,m}}\left(tP_m^{-1}r^{\alpha_m}\right) \\ &\times \mathcal{L}_{I_{b_m,o}}\left(tP_m^{-1}r^{\alpha_m}\right) f_{D_m}(r)dr, \quad (35) \end{aligned}$$

where $\mathcal{L}_{I_{b_m,m}}(\cdot)$ and $\mathcal{L}_{I_{b_m,o}}(\cdot)$ are the Laplace transforms of $I_{b_m,m}$ and $I_{b_m,o}$, respectively, and $f_{D_m}(r)$ is the PDF of the distance D_m between the user u and the serving BS b_m. The cumulative distribution function (CDF) of D_m, $F_{D_m}(r) = \mathbb{P}(D_m \le r)$ can be expressed as

$$\begin{aligned} F_{D_m}(r) &= \mathbb{P}(R_m \le r | u \in \Phi_u^m) = \frac{\mathbb{P}(R_m \le r, u \in \Phi_u^m)}{\mathbb{P}(u \in \Phi_u^m)} \\ &= \frac{1}{U_m} \int_0^r \mathbb{P}\left(R_p \ge \left(\frac{BP_p}{P_m}\right)^{1/\alpha_p} R_m^{\alpha_m/\alpha_p} \Big| R_m = y \right) \\ &\quad \times f_{R_m}(y)\, dy. \quad (36) \end{aligned}$$

After using the distributions of R_p and R_m, which are derived in the proof of Lemma 1, the required PDF $f_{D_m}(r)$ can be obtained as

$$\begin{aligned} f_{D_m}(r) &= \frac{dF_{D_m}(r)}{dr} \\ &= \frac{2\pi\lambda_m}{U_m} r e^{-\pi\lambda_m r^2} \exp\left(-\pi\lambda_p \left(\frac{BP_p}{P_m}\right)^{\frac{2}{\alpha_p}} r^{\frac{2\alpha_m}{\alpha_p}} \right). \end{aligned} \quad (37)$$

The Laplace transform $\mathcal{L}_{I_{b_m,l}}(s) = \mathbb{E}[\exp(-sI_{b_m,l})], \, \forall l \in \{m, o\}$ at $s = t/P_m r^{\alpha_m}$ can be derived as

$$\begin{aligned} \mathcal{L}_{I_{b_m,l}}\left(\frac{t}{P_m}r^{\alpha_m}\right) = \mathbb{E}_{\Psi_l} &\Bigg[\prod_{x_l \in \Psi_l \backslash b_m} \mathbb{E}_{h_{x_l}} \Big[\exp\left(-tP_m^{-1}P_{\zeta(l)}\right. \\ &\times \left. r^{\alpha_m} h_{x_l} ||x_l||^{-\alpha_{\zeta(l)}}\right) \Big] \Bigg], \, \forall l \in \{m, o\}, \end{aligned} \quad (38)$$

where $P_{\zeta(m)} = P_m$, $P_{\zeta(o)} = P_p$, $\alpha_{\zeta(m)} = \alpha_m$ and $\alpha_{\zeta(o)} = \alpha_p$. By using the probability generating functional of PPP Ψ_l with density $p_l \lambda_{\zeta(l)}$, followed by the expectation with respect to exponential RV h_{x_l}, we have

$$\begin{aligned} &\mathcal{L}_{I_{b_m,l}}\left(\frac{t}{P_m}r^{\alpha_m}\right) = \\ &\exp\left(-\int_{\eta_l}^{\infty} \left(1 - \frac{1}{1 + tP_m^{-1}P_{\zeta(l)}r^{\alpha_m}y^{-\alpha_{\zeta(l)}}} \right) 2\pi p_l \lambda_{\zeta(l)} y\, dy \right), \\ &\forall l \in \{m, o\}, \quad (39) \end{aligned}$$

where $\lambda_{\zeta(m)} = \lambda_m$ and $\lambda_{\zeta(o)} = \lambda_p$; η_m and η_o are the distances from the user u to the closest interferer in the macro and pico tier, respectively, given that the user u is served by the macro BS b_m at a distance $D_m = r$. Thus, $\eta_m = r$ and $\eta_o = (BP_p/P_m)^{1/\alpha_p} r^{\alpha_m/\alpha_p}$. Now, with the change in variables $(tP_m^{-1}P_{\zeta(l)}r^{\alpha_m})^{-2/\alpha_{\zeta(l)}}y^2 = u$ and further simplifi-

cation, we get

$$\mathcal{L}_{I_{b_m,l}}\left(\frac{t}{P_m}r^{\alpha_m}\right) = \exp\left(-\pi p_l \lambda_{\zeta(l)} \left(tP_m^{-1}P_{\zeta(l)}r^{\alpha_m}\right)^{2/\alpha_{\zeta(l)}} \right.$$
$$\left. \times \int_{\nu_l}^{\infty} \frac{1}{1+u^{\alpha_{\zeta(l)}/2}}\,du\right),$$
$$\forall l \in \{m,o\}, \qquad (40)$$

where $\nu_m = t^{-2/\alpha_m}$ and $\nu_p = (t/B)^{-2/\alpha_p}$. The integral in the above equation can be solved in terms of Gauss Hypergeometric function as [16, eqn. (24)]

$$\int_{\nu_l}^{\infty} \frac{1}{1+u^{\alpha_{\zeta(l)}/2}} = \frac{2}{(\alpha_{\zeta(l)}-2)} \frac{\nu_l}{\left(1+\nu_l^{\alpha_{\zeta(l)}/2}\right)}$$
$$\times {}_2F_1\left[1,1,2-\frac{2}{\alpha_{\zeta(l)}},\frac{1}{1+\nu_l^{\alpha_{\zeta(l)}/2}}\right]$$
$$(41)$$

The final expression for $\bar{F}_m(t)$ in (15) is obtained by substituting (37) and (40) into (35). The conditional SINR distribution for unbiased and range-expanded pico user, (16)-(17) can be similarly derived.

The simplified expressions for $\bar{F}_m(t)$, $\bar{F}_o(t)$ and $\bar{F}_e(t)$ in (18)-(20) can be obtained by substituting $\sigma^2 = 0$, $\alpha_m = \alpha_p = \alpha$ in (15)-(17), and then solving the integrals as $\int_0^{\infty} r\exp(-\beta r^2)dr = 1/(2\beta)$.

REFERENCES

[1] Qualcomm Incorporated, "LTE Advanced: Heterogeneous networks," white paper, Jan. 2011.

[2] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A survey on 3GPP heterogeneous networks," *IEEE Wireless Commun. Mag.*, vol. 18, no. 3, pp. 10–21, Jun. 2011.

[3] D. Lopez-Perez, I. Guvenc, G. de la Roche, M. Kountouris, T. Quek, and J. Zhang, "Enhanced intercell interference coordination challenges in heterogeneous networks," *IEEE Wireless Commun. Mag.*, vol. 18, no. 3, pp. 22 – 30, Jun. 2011.

[4] R. Madan, J. Borran, A. Sampath, N. Bhushan, A. Khandekar, and T. Ji, "Cell association and interference coordination in heterogeneous LTE-A cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1479–1489, Dec. 2010.

[5] K. Balachandran, J. Kang, K. Karakayali, and K. Rege, "Cell selection with downlink resource partitioning in heterogeneous networks," in *IEEE Int. Conf. on Commun. Workshops (ICC'11)*, Kyoto, Japan, Jun. 2011, pp. 1–6.

[6] M. Vajapeyam, A. Damnjanovic, J. Montojo, T. Ji, Y. Wei, and D. Malladi, "Downlink FTP performance of heterogeneous networks for LTE-Advanced," in *IEEE Int. Conf. on Commun. Workshops (ICC'11)*, Kyoto, Japan, Jun. 2011, pp. 1–5.

[7] S. Strzyz, K. Pedersen, J. Lachowski, and F. Frederiksen, "Performance optimization of pico node deployment in LTE macro cells," in *Proc. Future Network and Mobile Summit*, Warsaw, Poland, Jun. 2011, pp. 1–9.

[8] C.-H. Huang and C.-Y. Liao, "An interference management scheme for heterogeneous network with cell range extension," in *Proc. Asia-Pacific Network Operations and Management Symposium (APNOMS)*, Taipei, Taiwan, Sep. 2011, pp. 1–5.

[9] I. Guvenc, "Capacity and fairness analysis of heterogeneous networks with range expansion and interference coordination," *IEEE Commun. Letters*, vol. 15, no. 10, pp. 1084 – 1087, Oct. 2011.

[10] S. Mukherjee and I. Guvenc, "Effects of range expansion and interference coordination on capacity and fairness in heterogeneous networks," in *Proc. IEEE Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 2011, pp. 1855 – 1859.

[11] S. Singh and J. Andrews, "Joint resource partitioning and offloading in heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. PP, no. 99, pp. 1 – 14, Dec. 2013.

[12] H.-S. Jo, Y. J. Sang, P. Xia, and J. G. Andrews, "Heterogeneous cellular networks with flexible cell association: A comprehensive downlink SINR analysis." *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3484–3495, Oct. 2012.

[13] A. Ghosh, N. Mangalvedhe, R. Ratasuk, B. Mondal, M. Cudak, E. Visotsky, T. A. Thomas, J. G. Andrews, P. Xia, H.-S. Jo, H. S. Dhillon, and T. D. Novlan, "Heterogeneous cellular networks: From theory to practice," *IEEE Commun. Mag.*, vol. 50, no. 6, pp. 54–64, 2012.

[14] H. ElSawy, E. Hossain, and M. Haenggi, "Stochastic geometry for modeling, analysis, and design of multi-tier and cognitive cellular wireless networks: A survey," *IEEE Commun. Surveys and Tutorials*, vol. 15, no. 3, pp. 996–1019, Jul. 2013.

[15] H. S. Dhillon, R. K. Ganti, F. Baccelli, and J. G. Andrews, "Modeling and analysis of k-tier downlink heterogeneous cellular networks," *IEEE J. Sel. Areas in Commun.*, vol. 30, no. 3, pp. 550–560, Apr. 2012.

[16] S. Mukherjee, "Distribution of downlink SINR in heterogeneous cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 54–64, Apr. 2012.

[17] M. Cierny, H. Wang, R. Wichman, Z. Ding, and C. Wijting, "On number of almost blank subframes in heterogeneous cellular networks," *IEEE Trans. on Wireless Commun.*, vol. 12, no. 10, pp. 5061–5073, Oct. 2013.

[18] Y. Jin and L. Qiu, "Joint user association and interference coordination in heterogeneous cellular networks," *IEEE Commun. Letters*, vol. 17, no. 12, pp. 2296–2299, Dec. 2013.

[19] H. ElSawy and E. Hossain, "Two-tier hetnets with cognitive femtocells: Downlink performance modeling and analysis in a multichannel environment," *IEEE Trans. on Mobile Computing*, vol. 13, no. 3, pp. 649–663, Mar. 2014.

[20] T. D. Novlan, R. K. Ganti, A. Ghosh, and J. G. Andrews, "Analytical evaluation of fractional frequency reuse for heterogeneous cellular networks," *IEEE Trans. on Commun.*, vol. 60, no. 7, pp. 2029–2039, 2012.

[21] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic Geometry and its Applications*, 3rd ed. John Wiley and Sons, 2013.

[22] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 3122–3134, Nov. 2011.

[23] N. L.-E. L. C. T. D. Colombi, B. Thors, "Measurements of downlink power level distributions in lte networks," in *Proc. IEEE Electromagnetics in Advanced Applications (ICEAA)*, Portland, Oregon, Sep. 2013, pp. 98–101.

[24] L. Barolli, F. Xhafa, A. Durresi, and A. Koyama, "A fuzzy-based handover system for avoiding ping-pong effect in wireless cellular networks." in *Proc. IEEE Int. Conf. on Parallel Processing-Workshops (ICPP-W)*, Portland, Oregon, Sep. 2008, pp. 135–142.

[25] J.-S. Ferenc and Z. Neda, "On the size distribution of Poisson-Voronoi cells," *Physica A-Statistical Mechanics And Its Applications*, vol. 385, no. 2, pp. 518–526, 2007.

[26] J. Kingman, *Poisson Processes*. Oxford University Press Inc., 1993.

[27] X. Lin, J. G. Andrews, and A. Ghosh, "Modeling, analysis and design for carrier aggregation in heterogeneous cellular networks," *IEEE Transactions on Communications*, vol. 61, no. 9, pp. 4002–4015, 2013.