Fundamentals of the Downlink Green Coverage and Energy Efficiency in Heterogeneous Networks

Chun-Hung Liu, Senior Member, IEEE and Kok Leong Fong

Abstract—This paper studies the proposed green (energy-efficient) coverage probability, link and network energy efficiencies in the downlink of a heterogeneous cellular network (HetNet) consisting of $K$ independent Poisson point processes (PPPs) of base stations (BSs). The important statistical properties of the universal (general) cell association functions are first studied and the cell load statistics for power-law cell association functions, which can characterize the accurate void cell probability of a BS in every tier, is also derived. A simple and feasible green channel-aware cell association (GCA) scheme is proposed and the green coverage probability is also proposed for any particular cell association scheme, such as the maximum received power association (MRPA) and nearest base station association (NBA) schemes. Then the link and network energy efficiencies are proposed to characterize the mean spectrum efficiency per unit power consumption for a BS and the mean area spectrum efficiency for a HetNet, respectively. All the tight bounds on the green coverage probability, link and network energy efficiencies for the GCA, MRPA and NBA schemes are found. They are theoretically shown to pose the fundamental maximum limits on the link and network energy efficiencies achieved by any other cell association schemes and such a fact is validated by numerical results as well.

Index Terms—Green communication, energy efficiency, heterogeneous networks, cell association, stochastic geometry.

I. INTRODUCTION

In recent years, the longstanding uniform topology of a cellular network has been gradually vanished and transformed into a complex and heterogeneous network (HetNet) consisting of different types of base stations (BSs), such as macro, micro, small cell (e.g., picocell and femtocell) BSs, etc. Such a HetNet is potentially able to support a huge amount of traffic flow from the explosive proliferations of smart wireless handsets since its network capacity seems not to reach a hard limit even when more and more BSs are deployed in a HetNet as long as BSs are well positioned without causing severe interference between them [2] [3]. Deploying more small cell BSs instead of macro BSs in a HetNet not only brings considerable network capacity but also consumes less power since small cell BSs have much less power consumption than macro BSs. However, some practical problems would arise in the dense HetNet, such as, operation and maintenance expenditure, traffic management and offloading, cell load balancing, energy consumption and management, etc., and how to leverage these problems to further improve the capacity and low power consumption benefits is an important issue that needs to be exhaustively investigated.

A. Prior Work and Motivation

HetNets with a dense small cell topology definitely help the entire information and communication technology (ICT) industry control the CO$_2$ emission, which is estimated to account for at least 6% of the global emission in 2020 [4], [5]. Effective energy-saving methods are certainly worth exploring in a dense HetNet since power control and operation at small cell BSs are much easier than macro BSs [6]. The majority of current works on energy-saving techniques for HetNets focus on the BS-level management such as sleep/active mode control, cell activation, scheduling, etc. (see some typical works in [7]–[14]). The main idea in these related prior works is to save power by switching BSs to the sleep/dormant mode if they satisfy certain conditions, such as no users, carrying low traffic and meeting the predesignated constraints on power consumption, etc.

From the green communications point of view, offloading traffic from macro BSs to small cell BSs should improve the (bits-per-joule) energy efficiency of BSs while serving users. However, how to do traffic allocation/offloading between different types of BSs to accomplish a certain level of energy efficiency is hardly investigated in the literatures. Traffic flow behaviors in a HetNet are significantly affected by the cell (user) association strategies. If there exists an energy-efficient cell association means to make all the traffic flow into BSs with high energy efficiency, saving huge amount of energy can be certainly anticipated. Most prior works on cell association and traffic offloading focused on the study of the energy-free interactions between cell association, coverage and throughput. For example, reference [15] proposed a biased received power association and studied the average throughput of a downlink channel in a HetNet, but no energy efficiency analysis is involved in this work. Although the optimal BS intensity problem for achieving high throughput and energy efficiency is studied in [16], the impact on the energy efficiency due to cell association is not addressed at all.

The energy-saving approach for a cellular network by considering dynamic BS operation and cell association together was studied in [17], [18]. For example, however, these works only considered the traditional one-tier macro BS model so that its proposed scheme cannot be applied in a multi-tier HetNet. In references [19]–[22], the joint resource allocation, cell activation and traffic offloading problems that originate dynamic traffic variations in a HetNet were investigated and they are approached from the perspective of energy-efficient resource allocation, not from the perspective of making users associate a BS having high energy efficiency. As a result,
in general their proposed schemes on cell association and traffic offloading cannot ensure to achieve high link and/or network energy efficiency. Although reference [23] proposed a crucial concept that user association schemes should consider the energy cost in the backhaul communications in order to attain a certain level of energy efficiency, it still does not consider the link energy efficiency of BSs in its proposed context-aware user association scheme. This motivates the core question delved in this paper—Can cell association enhance or even optimize the energy efficiency of a BS? Such a question is scarcely studied in prior works on green communications.

B. Contributions

To investigate the fundamental interplay between cell association schemes and link energy efficiency (i.e., mean spectrum efficiency per unit power consumption at a BS), in this paper we consider a HetNet consisting of $K$ different types of BSs and each particular type of BSs, referred as a tier, are assumed to form a Poisson point process (PPP). We first study the statistical properties of a universal cell association function which is so general that it can cover all pathloss-based cell association schemes in a $K$-tier HetNet and these properties are the basis of deriving all theoretical results in the paper. We then derive the cell load statistics for the power-law cell association function that importantly shows the load balancing status of BSs in every tier and the probability of void BSs in each tier that do not have any tagged users may not be negligible in a dense HetNet, yet the void BS issue is usually overlooked in almost all prior works on the modeling and analysis in HetNets. The green cell association (GCA) scheme is proposed with a simple and feasible green cell association function that is derived based on the (random) link energy efficiency defined as the spectrum efficiency per unit power consumption at a BS. The GCA function characterizes the ratio of the random received signal power to the power consumption of a BS in the active operating mode.

To assess the link energy efficiency of cell association schemes, we propose the concept of green coverage probability that can indicate how likely cell association schemes can attain the desired link energy efficiency. The tight bounds on the green coverage probability for the scheme GCA are found, which shows the proposed GCA scheme achieves the fundamental maximum limit on the green coverage probability that cannot be surmounted by any other deterministic and/or non-green cell association schemes, such as maximum received power association (MRPA) and nearest BS association (NBA). This maximum limit on the green coverage probability is only achievable whenever there does not exist random channel impairments such as fading and shadowing or the GCA scheme can completely exploit the channel power gain variations due to those channel impairments. In other words, there exists an achievability gap in the green coverage probability since cell association schemes in general are unable to complete a new cell association within the short time scale of channel fading variations.

The very tight bounds on the link energy efficiencies for the GCA, MRPA and NBA schemes are all found. Especially, the fundamental maximum limit on the link energy efficiency achieved by GCA is presented in a very neat expression since it is obtained by a special integral technique first devised in this paper. Since the link energy efficiency does not thoroughly reflect the impact from the energy consumption of the void BSs in the dormant operating mode, the network energy efficiency, defined as the mean area spectrum efficiency per unit network power consumption, is proposed to remedy this shortcoming and it is a more realistic and accurate networkwise energy efficiency metric than other energy efficiency metrics defined in the current literatures. All theoretical analyses and numerical results validate that the proposed GCA scheme outperforms non-green cell association schemes such as MRPA and NBA in terms of green coverage probability, link and network energy efficiencies. In addition, GCA also maintains cell load balancing as almost good as NBA that is the best in cell load balancing among all cell association schemes.

C. Paper Organization

The rest of this paper is organized as follows. In Section II, we first introduce the network model of a heterogeneous cellular network consisting of $K$ different types of base stations and the power consumption model of the base stations. Afterwards, we introduce some fundamental results regarding the universal cell association function and cell load statistics. Section III presents the analytical results of the green cell association and coverage probability. In Section IV the fundamental limits on the energy efficiency of the HetNet with different cell association schemes are investigated. Finally, Section V concludes our main findings and observations in this paper.

II. NETWORK MODEL AND PRELIMINARIES

A. Multi-tier Random Network Model

Consider a large-scale heterogeneous cellular network (HetNet) on plane $\mathbb{R}^2$ consisting of $K$ different types of base stations (e.g., macrocells, microcells, picocells, etc.) and each specific type of base stations (BSs) is referred as a tier, which is named a $K$-tier HetNet. In this network, all BSs in each tier are assumed to form an independent homogeneous Poisson point processes (PPPs). Specifically, the BSs in the $k$th tier can be described as a marked PPP of intensity $\lambda_k$ and their set $\Phi_k$ can be written as
\[ \Phi_k \triangleq \{(B_{kj}, V_{kj}, H_{kj}, \Psi_{kj}) : B_{kj} \in \mathbb{R}^2, V_{kj} \in \{0, 1\}, H_{kj} \Psi_{kj}, Y_{kj} \in \mathbb{R}_+\}, \]
where $k \in K \triangleq \{1, 2, \ldots, K\}$, $B_{kj}$ denotes the $j$th BS in the $k$th tier and its location, $V_{kj}$ is a Bernoulli random variable indicating if $B_{kj}$ is void cell or not, (i.e., if it is associated by any user or not. $V_{kj} = 1$ indicates $B_{kj}$ is not void and zero otherwise.), $\Psi_{kj} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is called the universal cell association (UCA) function used by $B_{kj}$ and it is bijective, and $H_{kj}$ is the channel (power) gain from BS $B_{kj}$ to its serving user and all $H_{kj}$’s are i.i.d. random variables for same $k$ with all $j \in \mathbb{N}_+$ and independent for all $k \in K$, $Y_{kj}$ stands for the power consumption of BS $B_{kj}$. Note that all $\Psi_{kj}$’s are i.i.d. for same $k$ with all $j \in \mathbb{N}_+$ as well and independent for all $k \in K$ if they contain random variables. Without loss

1This “universal” cell association function can cover different cell association strategies and it could be random or deterministic depending on the nature of cell association schemes.
of generality, we assume that the first tier in the network is composed of all macro BSs and other $K-1$ tiers are composed of small cell BSs, such as picocells and femtocells. Also, our following analysis will be based on a typical user located in the origin since the statistics of all location-dependent random parameters seen by this typical user is the same as that seen by all other users in the network due to Slivnyak’s theorem [24]. Also, all main variables, symbols, and functions used in this paper are summarized in Table I.

In this paper, our study is based on the downlink transmission scenario and thus all cell association functions are a downlink-based design no matter which user association strategies are adopted2. Each user associates with a BS that has the maximum cell association function among others. By considering the typical user associating with BS $B^* \in \bigcup_{k=1}^{K} \Phi_k$, the UCA scheme, without loss of generality, can be written as

$$B^* = \arg \{ \Psi^*(\|B^*\|) \} = \arg \sup_{B_k \in \bigcup_{k=1}^{K} \Phi_k} \Psi_{k_b}(\|B_{k_b}\|),$$

(2)

where $\|X\|$ is the Euclidean distance from node X to the typical user and $\Psi^*(\cdot) \in \{ \Psi_{k_b}(\cdot), \forall k \in K, j \in \mathbb{N}^+ \}$ is the bijective UCA function used by the associated BS $B^*$. The cell association scheme in (2) can cover any distance-based cell association schemes. For instance, users will associate with their nearest BS, i.e., they adopt the (unbiased) nearest BS association (NBA) scheme with $\Psi_{k_b}(x) = x^{-\alpha}$ only accounting for the path loss of signals with $\alpha > 2$ denotes

Since the total power consumption of BSs is much higher than that of mobile handsets and the downlink traffic is usually much higher than the uplink traffic. Accordingly, in this paper we only consider the downlink network and cell association models because such a downlink formulation can significantly and effectively improve the energy efficiency of BSs in the HetNet.

### B. Preliminaries of Universal Cell Association

The following theorem shows the distribution of the value of the user association function corresponding to BS $B^*$, which is important since it can be used to characterize some key performance metrics, such as coverage probability, throughput, and energy efficiency, etc.

**Theorem 1.** Suppose the UCA function $\Psi_{k_b} : \mathbb{R}^+ \to \mathbb{R}^+$ is monotonic decreasing, random and independent for all $k \in K$ and $j \in \mathbb{N}^+$. If random variable $X \in \mathbb{R}^+$, is independent of all $\Psi_{k_b}$’s and the conditional second moment $\mathbb{E}[\Psi_{k_b}(\Psi_{k_b}^{-1}(X))^2 | \Psi_{k_b}]$ exists for all $k \in K$ where $\Psi_{k_b}^{-1}(\cdot)$ denotes the inverse function of $\Psi_{k_b}(\cdot)$, we have

$$\mathbb{E} \left[ F_{\Psi^*(\|B^*\|)}(X) \right] = \mathbb{E} \left\{ \exp \left( - \pi \sum_{k=1}^{K} \lambda_k \times \mathbb{E}_{\Psi_{k_b}} \left[ (\Psi_{k_b}^{-1}(X))^2 | \Psi_{k_b} \right] \right) \right\},$$

(4)

where $F_{\Psi^*}(\cdot)$ is the cumulative distribution function (CDF) of random variable Z. However, if random variable X and $\Psi_{k_b}$
are correlated, then we have the following

\[ E[F_{\Psi^r(||B||)}(X)] = E\left[\exp\left(-\pi \sum_{k=1}^{K} \lambda_k (\Psi_k^{-1}(X))^2\right)\right]. \quad (5) \]

**Proof:** See Appendix A. ■

Due to the generality of the UCA function \( \Psi_k(\cdot) \), the results in Theorem 1 can be applied in many different user association contexts. For instance, if the UCA function of BS \( B_k \) is the power-law (received-signal-power) model from \( B_k \) to the typical user, for example, the MRPA scheme with \( \Psi_k(||B_k||) = P_k H_k ||B_k||^{-\alpha} \), then \( \Psi_k^{-1}(y) = (P_k H_k)^{\frac{1}{\alpha}} \) and the complementary cumulative density function (CCDF) of the maximum received power of a user can be found based on (4) as

\[ P\left[\Psi^r(||B||) \geq \frac{1}{x^\alpha}\right] = \mathbb{P}\left[\sup_{B_k \in \bigcup_{k \in K} \Phi_k} \frac{P_k H_k}{||B_k||^\alpha} \geq \frac{1}{x^\alpha}\right] = 1 - \exp\left(-\pi x^2 \sum_{k=1}^{K} \lambda_k P_k^2 E[H_k^2]\right), \quad (6) \]

which indicates this CCDF is the same as the CDF of the path-loss-only received power from the nearest point of an equivalent homogeneous PPP of intensity \( \sum_{k \in K} \lambda_k P_k^2 E[H_k^2] \).

This fact can be used to calculate the distributions of some performance metrics (such as signal-to-interference ratio (SIR)) if all users associate with their strongest BS, i.e., MRPA is adopted. In addition, if \( \Psi_k(\cdot) \) is deterministic and only depends on the path loss, i.e., \( \Psi_k(x) = x^{-\alpha}, \) (6) further reduces to

\[ P\left[\Psi^r(||B||) \geq \frac{1}{x^\alpha}\right] = \mathbb{P}\left[\sup_{B_k \in \bigcup_{k \in K} \Phi_k} ||B_k||^{-\alpha} \geq \frac{1}{x^\alpha}\right] = 1 - \exp\left(-\pi x^2 \sum_{k=1}^{K} \lambda_k\right), \quad (7) \]

which can be used to find the distributions of some performance metrics pertaining to (unbiased) nearest BS association (NBA).

According to the results in Theorem 1, we can derive the probability of associating a BS from any tier and the CDF of the distance between the associated BS and the typical user for deterministic cell association functions as shown in the following theorem.

**Theorem 2.** If all UCA functions are random, the probability that the associated BS \( B^* \) in (2) from tier \( k \), i.e., \( \vartheta_k = P[B^* \in \Phi_k] \) is given by

\[ \vartheta_k = 2\pi \lambda_k \int_0^\infty y E_{\Psi_k} \left[\exp\left(-\pi \sum_{m=1}^{K} \lambda_m x^2\right)\right] dy, \quad (8) \]

where \( \Psi_k(\cdot) \) and \( \hat{\Psi}_k(\cdot) \) are i.i.d. for all \( k \in K \). On the other hand, if all cell association functions are deterministic, the CDF of the distance from BS \( B^* \) given in (3) to the typical user is given by

\[ F_{||B^*||}(x) = 1 - \exp\left(-\pi \sum_{m=1}^{K} \lambda_m (\Psi_m^{-1} \circ \hat{\Psi}_k(x))^2\right) \quad \text{for } k \in K. \]

in which \( \vartheta_k \) in (8) reduces to

\[ \vartheta_k = 2\pi \lambda_k \int_0^\infty y \exp\left(-\pi \sum_{m=1}^{K} \lambda_m (\Psi_m^{-1} \circ \hat{\Psi}_k(y))^2\right) dy. \quad (10) \]

**Proof:** See Appendix B. ■

With the result in (8), we can easily find the cell association probability for any cell association scheme. In the case of the aforementioned MRPA scheme with \( \Psi_k(x) = x^{-\alpha}, \) for example, we can find

\[ \vartheta_k = \lambda_k^\frac{2}{\alpha} E[H_k^\frac{2}{\alpha}] \left[\sum_{m=1}^{K} \lambda_m P_m^2 E[H_m^\frac{2}{\alpha}]\right], \quad (11) \]

For the NBA scheme with \( \Psi_k(x) = x^{-\alpha}, \) we have the tier-\( k \) cell association probability and the CDF of \( ||B^*|| \) as follows

\[ \vartheta_k = \frac{\lambda_k}{\sum_{m=1}^{K} \lambda_m} \quad \text{and } F_{||B^*||}(x) = 1 - e^{-\pi x^2 \sum_{k \in \lambda_k}} \]

which coincides with the result in [15]. These two user association examples indicate that Theorems 1 and 2 play a pivotal role in finding the statistics of the signal-to-interference ratio (SIR) and mean spectrum efficiency of a user for any cell association scheme as long as the user association functions of the scheme satisfy the constraints stated in Theorem 1.

**C. Cell Load Statistics for Power-Law Cell Association Functions**

In this subsection, we study the distribution of the number of users associated with a BS in a particular tier when the UCA functions in (2) have a power-law form, which is called the cell load statistics that refers to the random number of users carried by a cell. Assume all users in the network also form an independent homogeneous PPP of intensity \( \lambda_u \) denoted by \( \Phi_u \triangleq \{U_j \in \mathbb{R}^2 : j \in \mathbb{N}_+\} \) where \( U_j \) represents the \( j \)th user and its location. Let \( C_k \subset \mathbb{R}^2 \) denote the cell region of BS \( B_k \) where all users associating \( B_k \) are located and \( \nu(C_k) \) denote the Lebesgue measure of \( C_k \). Since the users is a PPP of intensity \( \lambda_u \), the probability that a tier-\( k \) BS is associated by \( n \) users can be expressed as

\[ p_{k,n} = \mathbb{P}\left[\Phi_u(C_k) = n\right] = \frac{1}{n!} E\left[\nu(C_k)^n e^{-\lambda_u \nu(C_k)}\right], \quad (13) \]

where \( \Phi_u(C_k) \) stands for the number of the users associating with BS \( B_k \) and assuming all \( \nu(C_k) \)'s are i.i.d. for same \( k \). Then the accurate expression of \( p_{k,n} \) is provided in the following lemma.

**Lemma 1.** If all users adopt the scheme in (2) to associate their BSs and \( \Psi_k(x) = \psi_k x^{-\alpha} \) with \( E[\psi_k^2] \leq E[\psi_k^\frac{2}{\alpha}] < \infty \)
for all $k \in K$, the probability that there are $n$ users associating with a tier-$k$ BS is given by

$$p_{k,n} = \frac{\Gamma(n + \hat{\zeta}_k)}{n! \Gamma(\hat{\zeta}_k)} \left( \frac{\lambda_u}{\hat{\zeta}_k \lambda_k} \right)^n \left( \frac{\hat{\zeta}_k \lambda_k}{\zeta_k \lambda_k + \lambda_u} \right)^{n + \hat{\zeta}_k}, \tag{14}$$

where $\hat{\zeta}_k = \sum_{m=1}^{K} \lambda_m m \{ \psi_k^m \} / \{ \psi_k^m \}$ and $\zeta_k = \frac{\lambda_u}{\hat{\zeta}_k \lambda_k} = 2 \mathbb{E} \left[ \psi_k^2 \right] \mathbb{E} \left[ \psi_k^2 - \psi_k^m \right]$. Then the mean number of the users associating with a tier-$k$ BS, call tier-$k$ cell load denoted by $L_k$, can be shown as

$$L_k = \mathbb{E} \{ \Phi_u(C_k) \} = \frac{\lambda_u \mathbb{E} \left[ \psi_k^2 \right]}{\sum_{m=1}^{K} \lambda_m \mathbb{E} \left[ \psi_k^m \right]} \tag{15}.$$

**Proof:** See Appendix C.

The probability in (14) not only indicates how the number of users is distributed over different tiers but also reveals whether traffic load is balanced between these $K$ tiers or not, and it highly depends on $p_{k}$ and $\lambda_k$. For example, the first moment of the number of users associating a tier-$k$ BS given in (15) indicates the cell loads between tiers are essentially unbalanced unless all $\lambda_k$’s are the same. In other words, all cell loads will get more balanced if $\lambda_k \mathbb{E} \left[ \psi_k^2 \right] \approx \lambda_2 \mathbb{E} \left[ \psi_2^2 \right] \approx \cdots \approx \lambda_K \mathbb{E} \left[ \psi_K^2 \right]$. The parameter $\hat{\zeta}_k$ has an interesting physical meaning, that is, it can be interpreted as the equivalent intensity of the network provided that all BSs are composed of the tier-$k$ BSs, and thus $\lambda_k / \hat{\zeta}_k$ is exactly the probability that a user associates with a tier-$k$ BS. In addition, (14) also importantly shows that it is significantly affected by user association schemes and leads to the void cell probability of a tier-$k$ BS given by

$$\mathbb{P} \{ V_k = 0 \} = p_{k,0} = \left( 1 + \frac{\lambda_u}{\hat{\zeta}_k \lambda_k} \right)^{-\hat{\zeta}_k}. \tag{16}$$

Therefore, the void cell probability for each BS in any tier indeed exists and cannot be negligible if the cell load of a BS is small. To validate this void cell issue, a simulation example for a 3-tier HetNet with three cell schemes – green cell association (GCA) with $\Psi_{k,j}(x) = P_k H_k x^{-\alpha} / \gamma_k^\alpha$, MRPA with $\Psi_{k,j}(x) = P_k H_k x^{-\alpha}$ and NBA with $\Psi_{k,j}(x) = x^{-\alpha}$. The network parameters for simulation are $\lambda_1 = 10^{-6}$ (BS/m$^2$), $\lambda_2 : \lambda_3 = 1 : 10 : 50$, $P_1 = 40$W, $P_2 = 1$W, $P_3 = 0.5$W, $\gamma_1^\alpha = 118.7$W, $\gamma_2^\alpha = 6.8$, $\gamma_3^\alpha = 4.8$W, $\alpha = 4$, $\omega_2 = 2.66$, $\omega_3 = 4.0$ and $\omega_3 = 7.5$, exponential (Rayleigh) fading gain with unit mean and log-normal shadowing gain with zero mean and 3dB variance.

### III. Green Cell Association and Coverage Probability

Almost all current schemes pertaining to cell association, traffic offloading and cell range expansion are merely developed based on the strength of the received signal power between a user and a BS. The basic spirit of these signal-power-based cell association schemes is to make users associate with a BS providing satisfactory coverage, rate and load balance, however, these schemes fail to characterize the energy usage of BSs and they may attain the cell association objectives of high throughput, coverage and load balance performance at the expense of high power consumption. In a HetNet where densely deployed small cell BSs may induce a large amount of energy consumption, maintaining high data rate and low power consumption is important in order to fulfill the green (energy-efficient) communication goal for such a network.

The energy efficiency of a BS can be defined as how much the mean spectrum efficiency can be achieved at the cost of the unit power consumption of a BS. Since the downlink throughput of a user depends on the SIR of the user, cell association schemes certainly affect the link throughput as well as the power consumption of a BS. To characterize the fundamental relationship between link spectrum efficiency and power consumption at a BS, the green cell association function is proposed as follows.

**Definition 1** (Green Cell Association Function). For the $j$th BS in the $k$th tier of the interference-limited HetNet with $K$ tiers, its green cell association (GCA) function is defined as

$$\Psi_{k,j}(\|B_k\|) \triangleq \log_2 \left( 1 + \frac{P_k H_k}{I_k \|B_k\|^{\alpha}} \right) / \gamma_k^\alpha. \tag{17}$$

![Fig. 1. The void cell probabilities for a 3-tier HetNet with three cell association schemes – GCA with $\Psi_{k,j}(x) = P_k H_k x^{-\alpha} / \gamma_k^\alpha$, MRPA with $\Psi_{k,j}(x) = P_k H_k x^{-\alpha}$ and NBA with $\Psi_{k,j}(x) = x^{-\alpha}$. The network parameters for simulation are $\lambda_1 = 10^{-6}$ (BS/m$^2$), $\lambda_2 : \lambda_3 = 1 : 10 : 50$, $P_1 = 40$W, $P_2 = 1$W, $P_3 = 0.5$W, $\gamma_1^\alpha = 118.7$W, $\gamma_2^\alpha = 6.8$, $\gamma_3^\alpha = 4.8$W, $\alpha = 4$, $\omega_2 = 2.66$, $\omega_3 = 4.0$ and $\omega_3 = 7.5$, exponential (Rayleigh) fading gain with unit mean and log-normal shadowing gain with zero mean and 3dB variance.](image)
where $I_k = \sum_{m=1}^{K} k \not\in B_k \cdot P_m \cdot H_m \cdot V_m ||B_m||^{-\alpha}$ denotes the interference evaluated at the origin $\nu_k$ in addition to the signal power from BS $B_k$. $P_B H_B / \|B_k\|^\alpha$ is the SIR from BS $B_k$ to the typical user and $\Psi_k^\alpha$ defined in (13) is the power consumption of the BS in the active operating mode.

Obviously, this GCA function not only characterizes the spectrum efficiency per unit power consumption, but also leverages the SIR and the power consumption at a BS. Thus, the key to improving the energy efficiency of the network is to make a user associate a BS having the largest GCA function for the user. In the following, we will propose a more feasible GCA function for easy implementation that is able to make users associate a BS with high energy efficiency and low computational complexity.

A. Green Cell Association

According to the GCA function defined in (17), we propose the following GCA scheme:

$$B^* = \arg \left\{ \Psi^*(\|B^*\|) \right\} \triangleq \arg \sup_{B_k \in \Phi_k} \Psi_k(|B_k|), \quad (18)$$

where $\Psi^*(\cdot)$ is the GCA function corresponding to $B^*$. Such a GCA scheme is somewhat complicated in computation and have a difficulty in implementation if the interference cannot be accurately estimated. Accordingly, the following lemma renders an equivalent low-complexity GCA scheme.

**Lemma 2.** The green cell association scheme in (18) is equivalently rewritten as the following scheme

$$B^* = \arg \left\{ \Psi^*(\|B^*\|) \right\} = \arg \sup_{B_k \in \Phi_k} \frac{P_k H_k}{\|B_k\|^\alpha}$$

$$\quad \quad \quad \quad = \arg \sup_{B_k \in \Phi_k} \Psi_k(|B_k|), \quad (19)$$

in which the GCA function of BS $B_k$ becomes $\Psi_k(x) \triangleq P_k H_k(1 - x^{-\alpha}) / \|B_k\|^\alpha$.

**Proof:** See Appendix D.

The GCA scheme in Lemma 2 is essentially a “biased” MRPA scheme in which the cell association function needs to characterize the ratio of the received signal power to the power consumption of an active BS in order to achieve high energy efficiency. In fact, the important implication of the GCA scheme in Lemma 2 is that users tend to associate with a BS that has a high transmit power fraction of the power consumption, which intuitively makes sense because having high transmit power (i.e., high SIR and spectrum efficiency) and low hardware power consumption is the key to attaining high energy efficiency. For tractability and simplicity in the following analysis, we will use the GCA scheme in (19) instead of that in (18). The probability in (14) for the GCA scheme, for instance, can be easily found by substituting $\Psi_k(x) = \frac{P_k H_k}{\|B_k\|^\alpha}$ into (14), and $\lambda_k$, i.e.,

$$\zeta_k = \frac{\mu}{2} \mathbb{E} \left[ H_k^\frac{\alpha}{2} \right] \mathbb{E} \left[ H_k^\frac{\alpha}{2} \right]$$

(20)

4Note that all $I_k$’s are correlated for all $k \in K$ and $j \in \mathbb{N}_k$ since they are generated by the same $K$ sets of BSs.

and

$$\tilde{\lambda}_k = \frac{\sum_{m=1}^{K} \lambda_m (P_m / T_m^\alpha) \mathbb{E} \left[ H_m^\frac{\alpha}{2} \right] \mathbb{E} \left[ H_m^\frac{\alpha}{2} \right]}{(P_k / T_k^\alpha) \mathbb{E} \left[ H_k^\frac{\alpha}{2} \right] \mathbb{E} \left[ H_k^\frac{\alpha}{2} \right]} = \frac{\lambda_k}{\tilde{\nu}_k}, \quad (21)$$

which also immediately gives the probability $\nu_k$ of associating a tier-$k$ BS and can be used to find the tier-$k$ cell load in (15) for the GCA scheme.

B. Green Coverage Probability

We have realized that user association schemes indeed significantly influence the energy efficiency of a BS. To quantitatively assess the fundamental interplay between energy efficiency and user association, the green coverage probability of an associated BS, based on the energy efficiency concept, is defined as follows.

**Definition 2** (Green Coverage Probability). We define the green coverage probability as the performance metric of indicating how likely a BS can maintain the desired energy efficiency under any particular cell association scheme. As a result, the green coverage probability for cell association scheme “a”, denoted by $\rho_a$, can be mathematically expressed as

$$\rho_a(\eta) \triangleq \mathbb{P} \sum_{k=1}^{K} \log_2 \left( 1 + \frac{P_k H_k}{I^* k} \right) \frac{I(B^* \in \Phi_k)}{T_m^\alpha} \geq \eta,$$

where subscript $a \in \{ge, mr, nb, \ldots\}$, $\eta > 0$ is the threshold of the desired energy efficiency, $I^* \triangleq \sum_{m=1}^{K} \lambda_m \|B_m\|^{-\alpha}$ denotes the interference received by the typical user, and $1(E)$ is the indicator function that is equal to one if event $E$ is true and zero otherwise.

According to the definition of the green coverage probability, a BS is said in the outage of energy efficiency if it cannot reach its desired energy efficiency while serving users. For the GCA scheme in (19), its green coverage probability can be shown in the following theorem.

**Theorem 3.** For the GCA scheme in (19), the tight lower bound on the green coverage probability can be found and shown in (25) where $a \succeq b$ (a $\preceq b$) means $b$ is the tight lower (upper) bound on $a$.

$$p_{m,a}(\rho) \succeq \frac{\mu}{2} \mathbb{E} \left[ H_k^\frac{\alpha}{2} \right] \mathbb{E} \left[ H_k^\frac{\alpha}{2} \right]$$

$\rho_{m,a}$ is given in (25) with $\zeta_k$ given in (20), and $\lambda_k$ given in (21) and $\mathcal{L}^{-1}(\mathcal{G}(s))(\tau)$ is the inverse Laplace transform operator of function $G(s)$.

**Proof:** See Appendix E.

The tight lower bound in (25) is derived based on the fact stated in the proof of Theorem 3 that the distribution of nonvoid cells still behaves very similarly to PPPs even though they

5Throughout this paper, subscripts gc, mr, and nb denote the GCA, MRPA, and NBA schemes, respectively.

6Actually, the green coverage probability is another form of the traditional SIR-based coverage probability. Thus, the following analytical results of the green coverage probability can be transformed to the traditional coverage probability by replacing $2^\nu T^\nu - 1$ with the desired SIR threshold.
theoretically do not form PPPs any more due to the presence of the location correlations between the non-void BSs. Thus, as the user intensity goes to infinity (i.e., the void cell probability goes to zero), the interference increases so that $\rho_{gc}$ reduces and eventually converges to its lower bound without void cell probabilities, that is,

$$\lim_{\lambda_u \to \infty} \rho_{gc}(\eta) = \sum_{k=1}^{K} \phi_k \int_{0}^{1} L^{-1} \left\{ \left[ 1 + \sum_{m=1}^{K} \phi_m \times h \left( \frac{\lambda_m \gamma_k \eta^{\alpha+1}}{\gamma_k + \eta^{\alpha}}, \frac{2}{\alpha} \right) \right]^{-1} \right\} (\tau) d\tau, \quad (24)$$

which is the lowest limit $\rho_{gc}$ can achieve and does not depend on the intensities of users and BSs. Also, this lowest limit is the green coverage probability if void cells are overlooked in the interference model and this implies the green coverage probability without considering the void cell impact is actually underestimated. On the other hand, we should have $\lim_{\lambda_u \to \infty} \rho_{gc} = 1$ since the void cell probabilities go to unity and that leads to an infinitely large SIR. Deploying more BSs indeed benefits the green coverage probability, but it suffers the problem of diminishing marginal returns on the green coverage gain. In general, the tight lower bound on $\rho_{gc}$ in (23) cannot be obtained in closed-form since the inverse Laplace transform is usually intractable. Nonetheless, we can resort to numerical Laplace transform methods to find the exact tight lower bound, or by applying $h(x, \frac{2}{\alpha}) \geq x^{\frac{2}{\alpha}} \frac{\Gamma(1 - \frac{2}{\alpha})}{\Gamma(\frac{1}{\alpha})} - 1$ for $x > 0$ in (23), a tight upper bound on $\rho_{gc}$ can be obtained as shown in (25), which has a closed form for $\alpha = 4$ (see the Laplace transform table in (27)) and it also has the following closed-form result when $\lambda_u \to \infty$:

$$\lim_{\lambda_u \to \infty} \rho_{gc}(\eta) \leq \frac{\alpha \sin(2\pi/\alpha) \sum_{k=1}^{K} \phi_k \left( \frac{\gamma_k}{\gamma_k + \eta^{\alpha}} - 1 \right)^{\frac{2}{\alpha}}}{2\pi \sum_{m=1}^{K} \phi_m \left( \frac{\gamma_m}{\gamma_m + \eta^{\alpha}} \right)^{\frac{2}{\alpha}}}. \quad (26)$$

It is worth pointing out that the green coverage probability in (23) is the fundamental maximum limit of the green coverage probability in that the GCA scheme in (19) captures all random channel variations to exploit the “multi-BS diversity gain” while performing cell association, and this limit cannot be surpassed by any other non-green cell association schemes. This limit is only achievable whenever there are no random channel impairments such as fading and shadowing, in other words, it may not be practically achievable provided that channel power gains experience small–scaling fast fading since users may not be able to finish another new cell association action within a very short channel coherence time. As a result, if the channel power variations cannot be exploited by green cell association, the green coverage probability in (23) will reduce, as shown in the following theorem.

**Theorem 4.** Suppose all channels in the network undergo independent Rayleigh fading and shadowing and the channel gain from BS $B_k$ to the typical user can be specifically expressed as $H_{kj} = H_{kj}^f Q_{kj}$, where fading gain $H_{kj}^f$ is an exponential random variable with unit mean for all $k$, and $Q_{kj}$ is the random channel gain due to shadowing. If the GCA scheme only can exploit the channel shadowing gain, the GCA function in (19) for BS $B_k$ becomes $\Psi_{kj}(x) = P_k Q_{kj} x^{-\alpha}/\gamma_k^\alpha$ and the tight lower bound on the green coverage probability can be found in closed-form as

$$\rho_{gc} \geq \sum_{k=1}^{K} \phi_k \left( 1 + \sum_{m=1}^{K} \alpha \phi_m \right)^{-1} \left( \frac{\lambda_m \gamma_k \eta^{\alpha+1}}{\gamma_k + \eta^{\alpha}}, \frac{2}{\alpha} \right) \eta_m \left( \frac{\gamma_m}{\gamma_m + \eta^{\alpha}} - 1 \right)^{-1}, \quad (27)$$

where $\phi_k = \lambda_k \left( \frac{P_k}{\gamma_k} \right)^{\frac{2}{\alpha}} \mathbb{E} \left[ Q_{kj}^\frac{2}{\alpha} \right] / \mathbb{E} \left[ Q_{kj}^\frac{2}{\alpha} \right]$, $\lambda_m \phi_m$ is already given in (16) with $\lambda_k = \frac{\lambda_k \gamma_k}{\gamma_k}$ and $\zeta_k = \frac{\gamma_k}{\gamma_k} \mathbb{E} \left[ Q_{kj}^\frac{2}{\alpha} \right]$, and $\ell(\cdot, \cdot)$ is defined as

$$\ell(x, z) = x^\frac{2}{\alpha} \left( \frac{\pi z}{\sin(\pi z)} \right) - \int_0^{x-z} \frac{dt}{1 + t^\frac{2}{\alpha}} \int_0^{x-z} \frac{dt}{1 + t^\frac{2}{\alpha}} = \frac{2}{\alpha} \frac{dt}{1 + t^\frac{2}{\alpha}} - 1. \quad (28)$$

**Proof:** See Appendix [F].

Note that $\rho_{gc}$ in (27) should be smaller than that in (23) because it does not exploit the multi-BS diversity induced by Rayleigh fading via cell association schemes. In other words, $\rho_{gc}$ in (27) is much more achievable than that in (23) if channel fading is unavoidable in practice so that there exists a fundamental gap between them, thereby we can further deduce that the green coverage probability will reduce more if channel shadowing variations cannot be exploited while performing cell association. Also, as the user intensity goes to infinity $\rho_{gc}$ in (27) will also decrease and converge to the lowest limit given by

$$\lim_{\lambda_u \to \infty} \rho_{gc} = \sum_{k=1}^{K} \phi_k \left( 1 + \sum_{m=1}^{K} \alpha \phi_m \right)^{-1} \left( \frac{\gamma_k}{\gamma_k + \eta^{\alpha}} - 1 \right)^{-1} \left( \frac{\gamma_m}{\gamma_m + \eta^{\alpha}} - 1 \right)^{-1}, \quad (29)$$

which can be viewed as the “more achievable” lowest limit of the green coverage probability.

In addition to the green coverage probability for the GCA scheme found in above, the green coverage probabilities for other cell association schemes can also be found based on the similar approaches to deriving the green coverage probability of the GCA scheme. For example, the tight lower and upper bounds on the green coverage probability of the MRPA scheme with $\Psi_{kj}(x) = P_k H_{kj} x^{-\alpha}$ can be easily found by slightly modifying the results in (23) and (25), respectively, and they are given in (30) where $\phi_k = \lambda_k P_k \mathbb{E} \left[ H_{kj}^f \right] / \mathbb{E} \left[ H_{kj}^f \right]$ and their asymptotic results for $\lambda_u \to \infty$ are given in (31). Intuitively, $\rho_{mr}$ is definitely smaller than $\rho_{gc}$.
since $\rho_{gc}$ is the maximum green coverage probability that can be achieved by cell association. This fact can also be inferred by comparing (20) with (23), the integral and $\vartheta_k$ in (23) both decrease as $P_{k}^{on}$ (or $\Psi_k$) increases, but the integral in (20) decreases and $\vartheta_k$ remains unchanged as $P_{k}^{on}$ increases. This reveals that $\rho_{gc}$ is more optimal than $\rho_{mn}$ since $\rho_{gc}$ gets more contribution from the BSs with less hardware power consumption, which means $\rho_{gc}$ is “greener” than $\rho_{mn}$ from the energy-efficiency perspective. Another example is the NBA scheme with $\Psi_k(x) = x^{-\alpha}$ and the tight lower and upper bounds on the green coverage probability of the NBA scheme, based on Proposition 2 in [14] and the proof of Theorem 3 can be derived as shown in (32) where $p_g = p_1^{0} = \cdots = p_{K}^{0}$, $s_k = s_k^{(2\Psi_k^{\tau} - 1)}$, and $\vartheta_k = \lambda_k / \sum_{m=1}^{K} \lambda_m$ and we also have $\lim_{\lambda_u \to \infty} \rho_{hn}$ as shown in (33). Note that (32) and (33) are valid for any channel models and can be further simplified to a neater expression if all channels undergo Rayleigh fading. Since NBA does not exploit any channel variations at all, its green coverage probability must be smaller than that of GCA and might be also smaller than that of MRPA if the power consumption of BSs does not significantly degrade the link energy efficiency. As a result, we can conclude that $\rho_{gc} (\eta) > \max \{ \rho_{hn} (\eta), \rho_{mn} (\eta) \}$ for a given $\eta$. This fact will be demonstrated in the following subsection of numerical results.

### C. Numerical Results

In this subsection, we first provide some simulation results in a 3-tier HetNet consisting of macrocell, picocell and femtocell BSs to numerically verify the green coverage probabilities derived in Section III-B for the GCA, MRPA and NBA schemes. Assume all channels undergo Rayleigh fading and log-normal shadowing and all network parameters for

$$
\rho_{gc}(\eta) \approx \sum_{k=1}^{K} \vartheta_k \int_{0}^{1} L^{-1} \left\{ \frac{1}{\Gamma(1 - \frac{2}{\alpha})} \left( s_k^{(2\Psi_k^{\tau} - 1)} - 1 \right) \right\} \left( 1 - p_{m, \beta} \right) \vartheta_m \left( \frac{1}{\tau} \right) d\tau,
$$

(25)

$$
\rho_{mn}(\eta) \approx \sum_{k=1}^{K} \vartheta_k \int_{0}^{1} L^{-1} \left\{ \left( 1 + h \left( s_k^{(2\Psi_k^{\tau} - 1)} \right) \right)^{-1} \right\} \left( 1 - p_{m, \beta} \right) \vartheta_m \left( \frac{1}{\tau} \right) d\tau,
$$

(30)

$$
\rho_{hn}(\eta) \approx \sum_{k=1}^{K} \vartheta_k \int_{0}^{1} L^{-1} \left\{ \left( 1 + h \left( s_k^{(2\Psi_k^{\tau} - 1)} \right) \right)^{-1} \right\} \left( 1 - p_{m, \beta} \right) \vartheta_m \left( \frac{1}{\tau} \right) d\tau
$$

(32)

$$
\lim_{\lambda_u \to \infty} \rho_{mn}(\eta) \approx \frac{\alpha \sin(2\pi/\alpha)}{2\pi} \sum_{k=1}^{K} \vartheta_k \left( s_k^{(2\Psi_k^{\tau} - 1)} - 1 \right)^{-\frac{1}{\alpha}}
$$

(31)

**TABLE II**

| Parameter \ BS (Tier #) | Macro (1) | Picocell (2) | Femtocell (3) |
|--------------------------|-----------|--------------|---------------|
| Transmit Power $P_k^{on}$ (W) | 40 | 1.0 | 0.05 |
| Active Power $P_k^{on}$ (W) | 118.7 | 6.8 | 4.8 |
| Dormant Power $P_k^{off}$ (W) | 93 | 4.3 | 2.9 |
| Power Scaling Factor $\omega_k$ | 6.02 | 4.0 | 7.5 |
| Intensity $\lambda_k$ (BSs/m²) | $1 \times 10^{-6}$ | $10\lambda_1$ | $50\lambda_1$ |
| Threshold $\eta$ (bps/joule) | 0.1 | | |
| Gain $H_k = H_k^{1/2} Q_{k_j}$ | $H_k^{1/2} \sim \text{Exp}(1,1)$, $Q_{k_j} \sim \ln N(0, 3\text{dB})$ | | |
| Pathloss Exponent $\alpha$ | 4 | | |

![Fig. 2. Simulation results of the green coverage probabilities of the GCA scheme that captures different channel gains.](image-url)
The link energy efficiency can directly display how much energy efficiency a BS in the active operating mode can achieve on average. Nonetheless, it cannot offer a good evaluation for the energy efficiency of the entire network since there could be some void BSs that still consume some power and we will very likely have to endure a data transmission without the anticipated energy efficiency if the GCA scheme does not contain channel state information (see how random fading and shadowing channel gains severely degrade the green coverage probability in Fig. 2). The simulation results of the green coverage probabilities for the GCA, MRPA and NBA schemes are shown in Fig. 3. Unsurprisingly, GCA outperforms the other two schemes and MRPA outperforms NBA. The tightness of all the derived lower bounds on the green coverage probabilities of the three schemes in (23), (30) and (32) is also confirmed.

Another good aspect of evaluating the performance of a cell association scheme is to observe whether the scheme is able to balance the cell loads between the BSs at different tiers or not. Having load balancing between BSs is important since it can make users acquire more resource and better service. For any cell association scheme, its theoretical tier-\(k\) cell load is given by (15) as long as its tier-\(k\) cell association function is known. The simulation results of the cell load of each tier for the GCA, MRPA and NBA schemes are illustrated in Fig. 4. According to Fig. 4 GCA and NBA both have a very similar performance in cell load balancing and this suggests GCA can balance the cell loads very well since NBA treats different BSs equally and makes the cell loads in different tiers equal, i.e., it completely balances the cell loads. MRPA is intrinsically unable to balance the cell loads well because it makes users favor the BSs having large transmit power.

IV. ENERGY EFFICIENCY ANALYSIS

The green coverage probabilities for different cell association schemes have been studied in Section III and they are the basis of calculating the energy efficiency of communication links. The link energy efficiency can directly display how much energy efficiency a BS in the active operating mode can achieve on average. Nonetheless, it cannot offer a good evaluation for the energy efficiency of the entire network since there could be some void BSs that still consume some power as shown in [3] even though they are in the dormant operating mode. In this section, first we investigate how to analyze the link energy efficiency and then propose a new metric of the network energy efficiency to realistically characterize the mean area spectrum efficiency at the cost of the power consumption per unit area in the network.

A. Analysis of the Link Energy Efficiency

According to the green coverage probability in (22), the link energy efficiency of cell association scheme a can be

$$
\lim_{\lambda_{A} \to \infty} \rho_{nb} = \sum_{k=1}^{K} \beta_{k} \int_{0}^{\infty} L^{-1} \left\{ E_{H_{k}} \left[ \left( 1 + \sum_{m=1}^{K} E_{H_{m}} \left[ h \left( s_{k} H_{m} P_{m} \right) \right] \right) \vartheta_{m} \right]^{-1} \right\} (\tau) d\tau.
$$

(33)
Theorem 5. Consider all users adopt the GCA scheme in (18) to associate their BS in the network. The tight lower bound on the link energy efficiency (bits/Hz/Joule) of the GCA scheme is shown as
\[
\delta_{\text{gc}} \geq \frac{1}{\ln 2} \sum_{k=1}^{K} \frac{\vartheta_k}{Y_k} C_{\text{gc},k} \tag{36}
\]
where \(C_{\text{gc},k}\) is the lower bound on \(C_{\text{gc},k}\) and it is given by
\[
C_{\text{gc},k} = \int_{0}^{\infty} \frac{(1 - e^{-s})ds}{s \left(1 + \sum_{m=1}^{K} \vartheta_m (1 - p_{m,0}) \varrho \left[\frac{\alpha Y_m}{Y_k} - \frac{2}{\alpha}\right] \right)} \tag{37}
\]
in which the exact results of \(\vartheta_k\) and \(h(\cdot, \cdot)\) can be referred to Theorem 4.

Proof: See Appendix G.

The link energy efficiency in (36) achieved by the GCA scheme, like the case of \(\rho_{\text{gc}}\), is the fundamental maximum limit on the link energy efficiency in the network and it cannot be surmounted by any other non-green cell association schemes. Applying \(h(x, \frac{2}{\alpha}) \geq x^2 \Gamma(1 - \frac{2}{\alpha}) - 1\) in (36) gives the tight upper bound on \(\delta_{\text{gc}}\) as shown in (38). As all the BS intensities go to infinity, the link energy efficiency goes to infinity as well because the void probabilities go to unity and the SIR goes to infinity as well, that is, \(\lim_{\lambda_\text{u} \to \infty} \delta_{\text{gc}} = \infty\), which means deploying more BSs can increase the link energy efficiency for a given user intensity. Furthermore, as the user intensity goes to infinity \(\delta_{\text{gc}}\) and its tight bound will converge to their lower limits given in (39) which do not depend on BS intensities and this means deploying more BSs only improves the link energy efficiency whenever the user intensity is extremely larger than the BS intensities. This lowest limit also represents the link energy efficiency for the context where void cells are overlooked in the interference model, as the similar case explained in the green coverage probability. Be aware that \(\delta_{\text{gc}}\) in (36) can only be achieved whenever channels do not suffer any random channel impairments such as fading and/or shadowing, whereas it is fairly difficult to be achieved in a fading environment since the GCA scheme hardly exploits the fast fading channel variations in general, like the green coverage probability in (23). The more achievable link energy efficiency, given in the following theorem, is obtained by making the GCA scheme exploit the shadowing variations in the channels.

Theorem 6. Consider the same fading channel gain model and the same GCA function specified in Theorem 7. The tight lower bound on \(\delta_{\text{gc}}\) is given by
\[
\delta_{\text{gc}} \geq \frac{1}{\ln 2} \sum_{k=1}^{K} \vartheta_k \int_{0}^{\infty} \left[(1 + s) \times \left(1 + \frac{K}{\ln 2} \sum_{m=1}^{\infty} \ell \left(\frac{\alpha Y_m}{Y_k}, \frac{2}{\alpha}\right) \varrho_m \right)^{-1} ds\right] \tag{40}
\]
where \(\vartheta_k\) and \(p_{m,0}\) are defined in Theorem 4. If the user intensity goes to infinity, then all void cell probabilities vanish and it follows that
\[
\lim_{\lambda_\text{u} \to \infty} \delta_{\text{gc}} = \frac{1}{\ln 2} \sum_{k=1}^{K} \vartheta_k \int_{0}^{\infty} \left[(1 + s) \times \left(1 + \frac{K}{\ln 2} \sum_{m=1}^{\infty} \ell \left(\frac{\alpha Y_m}{Y_k}, \frac{2}{\alpha}\right) \varrho_m \right)^{-1} ds\right] \tag{41}
\]
Proof: Since the tight lower bound on the green coverage probability has already been found in (27), we can integrate it with respect to \(\eta\) to find the tight lower bound on \(\delta_{\text{gc}}\) as shown in the following
\[
\delta_{\text{gc}} = \int_{0}^{\infty} \rho_{\text{gc}}(\eta) d\eta \geq \frac{1}{\ln 2} \sum_{k=1}^{K} \vartheta_k \int_{0}^{\infty} \left[(1 + s) \times \left(1 + \frac{K}{\ln 2} \sum_{m=1}^{\infty} \ell \left(\frac{\alpha Y_m}{Y_k}, \frac{2}{\alpha}\right) \varrho_m \right)^{-1} ds\right] d\eta.
\]
Then replacing variable \(\eta\) in the integral with \(\log_2(1 + s)/Y_k\) leads to (40) and letting \(p_{m,0}\) go to zero yields (41).

Without considering the fading effect in the GCA scheme, \(\delta_{\text{gc}}\) in (40) must be smaller than that in (36), like the aforementioned discussion about the green coverage probability, the fundamental gap also exists in the link energy efficiency. As for the MRPA scheme, its tight bounds on link energy efficiency that can be easily inferred from (36) and (38) are
\[
\delta_{\text{mr}} \left\{ \begin{array}{ll}
\geq \frac{1}{\ln 2} \int_{0}^{\infty} \left[(1 - e^{-s})ds\right] \sum_{k=1}^{K} \vartheta_k \int_{0}^{\infty} \left[(1 + s) \times \left(1 + \frac{K}{\ln 2} \sum_{m=1}^{\infty} \ell \left(\frac{\alpha Y_m}{Y_k}, \frac{2}{\alpha}\right) \varrho_m \right)^{-1} ds\right] d\eta,
\end{array} \right.
\]
where \(\vartheta_k\) is the same as that in (30) and
\[
\lim_{\lambda_\text{u} \to \infty} \delta_{\text{mr}} = \frac{1}{\ln 2} \int_{0}^{\infty} \left[(1 - e^{-s})ds\right] \sum_{k=1}^{K} \vartheta_k \int_{0}^{\infty} \left[(1 + s) \times \left(1 + \frac{K}{\ln 2} \sum_{m=1}^{\infty} \ell \left(\frac{\alpha Y_m}{Y_k}, \frac{2}{\alpha}\right) \varrho_m \right)^{-1} ds\right] d\eta.
\]
For the NBA scheme, its tight bounds on the link energy efficiency can be found by (32) and they are shown in (44) in which \(\vartheta_k\) is the same as that in (22) so that we have \(\lim_{\lambda_\text{u} \to \infty} \delta_{\text{nb}}\) shown in (45). Similarly, \(\delta_{\text{gc}}\) also must be greater than \(\delta_{\text{nb}}\), but the ordering between \(\delta_{\text{mr}}\) and \(\delta_{\text{nb}}\) cannot be absolutely discerned. Thus, we also can conclude \(\delta_{\text{gc}} > \max\{\delta_{\text{mr}}, \delta_{\text{nb}}\}\), which will be validated in the section of numerical results.
\[
\delta_{ge} \lesssim \frac{1}{\ln 2} \sum_{k=1}^{K} \int_{1^{+}}^{\infty} \frac{\partial_k (\Upsilon_k^{on})^{\frac{1}{2}} - (1 - e^{-s})}{s \left( \frac{s}{2} \sum_{m=1}^{K} (1 - p_{m,0}) (\Upsilon_m^{on})^{\frac{1}{2}} \partial_m + \sum_{m=1}^{K} p_{m,0} \Upsilon_m \right)} \, ds.
\]  
(38)

\[
\lim_{\lambda_u \rightarrow \infty} \delta_{ge} = \left\{ \begin{array}{ll}
\frac{1}{\ln 2} \sum_{k=1}^{K} \frac{\partial_k \Upsilon_k}{\Upsilon_k^{on}} \int_{1^{+}}^{\infty} \left( 1 + \sum_{m=1}^{K} \mathcal{E}_m \left[ h \left( s \frac{H_m}{H_k}, \frac{2}{\alpha} \right) \right] \partial_m \right)^{-1} \frac{(1 - e^{-s})}{s \left( 1 + \frac{2}{\alpha} \right) \mathcal{E}_m} ds & , \quad (44)
\end{array} \right.
\]  

\[
\lim_{\lambda_u \rightarrow \infty} \delta_{nb} = \left\{ \begin{array}{ll}
= \frac{1}{\ln 2} \sum_{k=1}^{K} \frac{\partial_k \Upsilon_k}{\Upsilon_k^{on}} \int_{1^{+}}^{\infty} \left( 1 + \sum_{m=1}^{K} \mathcal{E}_m \left[ h \left( s \frac{H_m}{H_k}, \frac{2}{\alpha} \right) \right] \partial_m \right)^{-1} \frac{(1 - e^{-s})}{s \left( 1 + \frac{2}{\alpha} \right) \mathcal{E}_m} ds & , \quad (45)
\end{array} \right.
\]

B. Analysis of the Network Energy Efficiency

The link energy efficiency has a drawback, that is, it fails to characterize the power consumed by those void BSSs and cannot provide an accurate overall index of how the entire HetNet efficiently uses its own energy. Hence, we propose the following network energy efficiency to remedy this drawback.

**Definition 3 (Network Energy Efficiency).** The network energy efficiency for cell association scheme “a” is defined as the ratio of the mean area spectrum efficiency of the network to the total power consumption of the network per unit area and its expression can be written as [1]

\[
\Delta_a \triangleq \frac{\sum_{k=1}^{K} \lambda_k (1 - p_{k,0}) C_{a,k}}{\sum_{m=1}^{K} \lambda_m \mathcal{E}[\Gamma_m]} = \frac{\sum_{k=1}^{K} \lambda_k (1 - p_{k,0}) C_{a,k}}{\sum_{m=1}^{K} \lambda_m \left[ (1 - p_{m,0}) \Upsilon_m^{on} + p_{m,0} \Upsilon_m \right]}, \quad (46)
\]

where \( \Upsilon_m \) is defined in (3) and \( C_{a,k} \) is the mean spectrum efficiency of an associated BS in tier \( k \) and already defined in [35].

The salient feature of this network energy efficiency, compared with the network-wise energy efficiency defined in the prior works (such as [11] and [29]), is to characterize the void cell issue that impacts network capacity and energy consumption of BSSs so that this network energy efficiency is a more accurate metric of evaluating how much area spectrum efficiency is achieved at the cost of the unit energy consumed in the network per unit area.

\footnote{Note that the network energy efficiency defined in this paper is essentially based on the concept of “the link throughput intensity”, which may be different from the network energy efficiency defined based on the entire network throughput and power consumption.}

The network energy efficiencies of the GCA, MRPA and NBA schemes can be directly inferred from the link energy efficiencies \( \delta_{ge} \), \( \delta_{nb} \) and \( \delta_{jb} \), respectively. For example, according to the proof of Theorem 5, the tight lower bound on \( C_{ge,k} \) is obtained, i.e., \( C_{ge,k} \). and substituting it into (46) readily yields the network energy efficiency of the GCA scheme given by

\[
\Delta_{ge} \lesssim \frac{\sum_{k=1}^{K} \lambda_k (1 - p_{k,0}) C_{ge,k}}{\mathcal{E}[\Gamma_m] (2) \sum_{m=1}^{K} \lambda_m \left[ (1 - p_{m,0}) \Upsilon_m^{on} + p_{m,0} \Upsilon_m \right]}, \quad (47)
\]

and note that \( p_{k,0} \) and \( \delta_{ge} \) are the same as those given in Theorem 3. As the user intensity goes to infinity, \( \Delta_{ge} \) converges to the following lowest limit

\[
\lim_{\lambda_u \rightarrow \infty} \Delta_{ge} = \frac{\sum_{k=1}^{K} \lambda_k C_{ge,k}}{\mathcal{E}[\Gamma_m] (2) \sum_{m=1}^{K} \lambda_m \Upsilon_m}, \quad (48)
\]

where \( C_{ge,k} \triangleq \lim_{\lambda_u \rightarrow \infty} C_{ge,k} \), and \( \Delta_{ge} \) goes to zero as all BS intensities go to infinity, i.e., \( \lim_{\lambda_u \rightarrow \infty} \Delta_{ge} = 0 \), because the total power consumed in the network goes to infinity whereas the mean area spectrum efficiency of the network converges to a constant. Due to the limited space, the tight bounds on the network energy efficiencies of the MRPA and NBA schemes are not specified here, however, we can expect that they have the same convergence properties as those in the case of the GCA scheme.

C. Numerical Results

In this subsection, two numerical simulations are provided to validate the previous derivations of the link and network energy efficiencies and our findings. All simulation assumptions are the same as those in Section III-C and the network parameters for simulation are also the same as those listed in
Network energy efficiencies that are quite different from those join in the network. Fig. 6 shows the simulation results of the GCA, MRPA, and NBA schemes. All link energy efficiencies reduce and eventually converge to the maximum limit on the link energy efficiency, as expected. Table II. The simulation results of the link energy efficiencies are shown in Fig. 5 and GCA is also the best among the three schemes and saves a lot of power (see the void probabilities shown in Fig. 1). Since GCA also has a higher link energy efficiency that MRPA and NBA, its network efficiency is significantly higher than those of MRPA and NBA due to low network power consumption. When $\lambda_3/\lambda_3$ increases, all network energy efficiencies increase and approach to constants since all mean area spectrum efficiencies increase faster than all network power consumptions of different cell association schemes that increase up to a constant.

V. CONCLUSION

In this paper, we investigate the fundamental limits on the energy efficiency in a HetNet from the cell association perspective. We propose the green cell association scheme that can leverage the received (random) signal power and the power consumption of an active BS and also propose the green coverage probability to evaluate how likely a cell association scheme can achieve the desired link energy efficiency: the more green coverage probability, the better link energy efficiency. The network energy efficiency is also proposed to realistically assess the area spectrum efficiency per unit network power consumption contributed by all active and dormant BSs. Our important findings are first to derive all tight bounds on the green coverage probability and link energy efficiency of the GCA scheme and then show that they are the fundamental maximum limits on those achieved by other cell association schemes, such as MRPA and NBA. Namely, the proposed GCA scheme outperforms other schemes in terms of the link and network energy efficiencies.

APPENDIX

PROOFS OF LEMMAS AND THEOREMS

A. Proof of Theorem 7

First consider random variable $X$ are independent of all $\Psi_{kj}$’s. In this case, $E[F_{\Psi^*}(\|B^*\|)(X)]$ is equal to $P\left[\sup_{B_{kj} \in \cup_{k=1}^K \Phi_k} |B_{kj}| \leq X \right]$ given by

$$
= E_X \left[ \prod_{k=1}^K \mathbb{E}_{\Psi_k} \left( \prod_{B_{kj} \in \Phi_k} P \left[ \Psi_k(\|B_{kj}\|) \leq X \mid X, \Phi_k \right] \right) \right] 
\equiv (a) E_X \left[ \prod_{k=1}^K \exp \left( -2\pi \lambda_k \int_0^\infty P[\Psi_k(r) \geq X \mid r] dr \right) \right] 
\equiv (b) E_X \left[ \exp \left( -\pi \sum_{k=1}^K \lambda_k \int_0^\infty P \left[ y \leq (\Psi_k^{-1}(X))^2 \mid X \right] dy \right) \right] 
\equiv (c) E_X \left[ \exp \left( -\pi \sum_{k=1}^K \lambda_k E_{\Psi_k} \left[ (\Psi_k^{-1}(X))^2 \mid X \right] \right) \right],
$$

where $(a)$ follows from the probability generating functional (PGFL) of $K$ independent homogeneous PPPs [24], $(b)$ is due to the fact that we let $y = r^2$ and $\Psi_k(r)$ is a monotonic decreasing and invertible function, and $(c)$ is obtained since $\int_0^\infty P \left[ y \leq (\Psi_k^{-1}(X))^2 \mid X \right] dy = E_{\Psi_k} \left[ (\Psi_k^{-1}(X))^2 \mid X \right]$. Thus, the result in (4) is obtained. Now
consider $X$ is correlated with $\Psi_k$, and thus we have

$$E[F_{\Psi^+ (\|B^*\|)}(X)] = \mathbb{P} \left( \sup_{{B_k} \in \cup_{k=1}^K \Phi_k} \Psi_k(\|B_k\|) \leq X \right)$$

$$= \prod_{k=1}^K E_{\Psi_k, X} \left\{ \prod_{B_{k \in} \Phi_k} \mathbb{P} \left( \Psi_k(\|B_k\|) \leq X | \Psi_k, X \right) \right\}$$

$$= \mathbb{E} \left\{ \exp \left( -2\pi \sum_{k=1}^K \lambda_k \int_0^\infty \mathbb{P}(\Psi_k(x) \geq X | \Psi_k, X) \right) \right\}$$

$$= \mathbb{E} \left\{ e^{-\pi \sum_{k=1}^K \lambda_k \int_0^\infty \Psi_k(\|x\|)(\Psi_k(\|x\|))^2 | \Psi_k, X) \right\},$$

which equals to (5) since $\mathbb{P}(y \leq (\Psi_k^{-1}(X))^2 | \Psi_k, X) = 1$ for $y \in [0, (\Psi_k^{-1}(X))^2]$.  

### B. Proof of Theorem 2

First consider all UCA functions are random. According to [4], the probability that $B^*$ belongs to $\Phi_k$ can be written as

$$\mathbb{P}(B^* \in \Phi_k) = \mathbb{E}_{\Phi_k} \left[ \sup_{{B_k} \in \Phi_k} \Psi_k(\|B_k\|) \right]$$

$$= \mathbb{E}_{\Phi_k} \left[ \sup_{{B_m} \in \cup_{m=K+1}^K \Phi_m} \Psi_m(\|B_m\|) \right]$$

$$= \mathbb{E}_{\Psi_m} \left[ \exp \left( -\pi \sum_{k=1}^K \lambda_k \mathbb{E}_{\Psi_m} \left[ (\Psi_m^{-1}(Z_k))^2 | Z_k \right] \right) \right],$$

where $Z_k = \sup_{{B_k} \in \Phi_k} \Psi_k(\|B_k\|) \triangleq \Psi_k(\|B_k^*\|)$ for $B_k^* \triangleq \arg \sup_{{B_k} \in \Phi_k} \Psi_k(\|B_k\|)$. Applying Theorem 1 to the network with one tier, the CDF of $Z_k$ can be found as $F_{Z_k}(z) = e^{-\pi \lambda Z_k \Psi_k^{-1}(z)^2}$ and thus the pdf of $Z_k$ is given by

$$f_{Z_k}(z) = 2\pi \lambda Z_k \Psi_k^{-1}(z) \frac{d\Psi_k^{-1}(z)}{dz}$$

$$\exp \left( -\pi \lambda Z_k \Psi_k^{-1}(z)^2 \right)$$

and note that $\frac{d\Psi_k^{-1}(z)}{dz} < 0$. Therefore, it follows that

$$\mathbb{E}_{\Psi_m} \left[ \exp \left( -\pi \sum_{k=1}^K \lambda_k \mathbb{E}_{\Psi_m} \left[ (\Psi_m^{-1}(Z_k))^2 \right] \right) \right]$$

$$= -2\pi \lambda Z_k \Psi_k^{-1}(z) \frac{d\Psi_k^{-1}(z)}{dz} \int_0^\infty \mathbb{E}_{\Psi_m} \left[ (\Psi_m^{-1}(z))^2 \right]$$

which equals to (5). Now if all the UCA functions are deterministic, then using the total probability law the CDF of $\|B_0\|$ can be written as

$$F_{\|B_0\|}(x) = \sum_{m=K}^\infty \mathbb{P}(\|B_0\| \leq x | B_0 \in \Phi_m) \mathbb{P}(B_0 \in \Phi_m)$$

$$= \sum_{m=K}^\infty \mathbb{P}(\|B_0\| \geq \Psi_m(\|x\|) | B_0 \in \Phi_m) \mathbb{P}(B_0 \in \Phi_m) \mathbb{E}_{\Psi_m} \left[ \exp \left( -\pi \sum_{k=1}^K \lambda_k \left[ (\Psi_m^{-1}(\Psi_k(\|x\|)))^2 \right] \right) \right]$$

$$= 1 - \sum_{k=1}^K \exp \left( -\pi \sum_{m=K}^\infty \lambda_m \left[ (\Psi_m^{-1}(\Psi_k(\|x\|)))^2 \right] \right)$$

(50)

where (a) is due to the fact that $\Psi_k(\cdot)$ is monotonic decreasing and (b) follows from the result in (5). Substituting (49) into (50) results in (5).
which exactly equals to (14). The tier-\(k\) cell load can be found as
\[
\mathbb{E}[\Phi_u(C_k)] = \sum_{n=1}^{\infty} n p_{k,n} = \int_0^{\infty} (\lambda_u \mathbb{P}[B^* \in \Phi_k]) x f_x(C_k)(x) \, dx
\]
\[
= \frac{\lambda_u \lambda_k}{\sum_{m=1}^K \lambda_m \mathbb{E}[\psi_k] \frac{\alpha}{\pi_m}} \times 1 / \lambda_k
\]
which is the result in (15) and the proof is complete.

D. Proof of Lemma 2

According to the energy efficiency function defined in (17), the GCA scheme in (18) can be equivalently expressed as follows
\[
B^* = \arg \sup_{B_j \in \cup_{k=1}^K \Phi_k} \left( 1 + \frac{P_k H_{k,j}}{I_k \|B_k\|^\alpha} \right)^{1/\alpha}
\]
\[
= \arg \sup_{B_j \in \cup_{k=1}^K \Phi_k} \left[ 1 + \frac{P_k H_{k,j} / \mathbb{Y}^\alpha_k}{I_k \|B_k\|^\alpha} \right] + O \left( \frac{P_k H_{k,j} / \mathbb{Y}^\alpha_k}{I_k \|B_k\|^\alpha} \right),
\]
where (a) follows from the binomial expansion of \((1 + X)^r = \sum_{j=0}^{\infty} \binom{r}{j} X^j\). \(\mathcal{O}(X)\) denotes the linear sum of the higher order terms of \(X\). Since all terms in \(\mathcal{O}(\cdot)\) consists of the first term \(P_k H_{k,j} / \mathbb{Y}^\alpha_k\), removing them and constant 1 does not affect the result of finding \(B^*\) since \(\|B^*\|\) that maximizes the first term \(P_k H_{k,j} / \mathbb{Y}^\alpha_k\) also maximizes the higher order terms in \(\mathcal{O}(\cdot)\). Hence, we have
\[
B^* = \arg \sup_{B_j \in \cup_{k=1}^K \Phi_k} \left( \frac{P_k H_{k,j}}{\mathbb{Y}^\alpha_k I_k \|B_k\|^\alpha} \right) .
\]
In addition, note that \(I_k\) is the interference at the typical user for a given point set \(\cup_{k=1}^K \Phi_k\) and \(B_j\). All \(I_k\)’s are identical in that they are completely correlated and has the same distribution based on the Slivnyak theorem. Therefore, removing \(I_k\) from \(\frac{P_k H_{k,j}}{\mathbb{Y}^\alpha_k I_k \|B_k\|^\alpha}\) does not affect the result of finding \(B^*\) as well, which yields the result in (19).

E. Proof of Theorem 3

Consider users use the GCA scheme in (19) to associate with their serving BSs. According to Theorem 1 and \(\Psi^{-1}(x) = (P_k H_{k,j} / \mathbb{Y}^\alpha_k)^{1/2} x^{1/2}\), we have the following:
\[
-1 \mathbb{F}_{\Psi^*(\|B^*\|)}(x) = \mathbb{P}[\Psi^*(\|B^*\|) \geq x]
\]
\[
= 1 - \mathbb{P} \left[ \sup_{B_j \in \cup_{k=1}^K \Phi_k} \frac{P_k H_{k,j}}{\mathbb{Y}^\alpha_k I_k \|B_k\|^\alpha} \leq x \right]
\]
\[
= 1 - \exp \left[ -\pi x^{\alpha/2} \sum_{k=1}^K \lambda_k \left( \frac{P_k}{\mathbb{Y}^\alpha_k} \right)^2 \mathbb{E}[H_k^\alpha] \right]
\]
\[
= 1 - \exp \left[ -\pi y^{\alpha/2} \sum_{k=1}^K \hat{\lambda}_k \right],
\]
where \(\|B^*\| \triangleq (\Psi(\|B^*\|))^{-1/2}, \hat{\lambda}_k = \lambda_k \left( \frac{P_k}{\mathbb{Y}^\alpha_k} \right)^2 \mathbb{E}[H_k^\alpha] \) and \(y \triangleq x^{-\alpha/2}\). Thus, \(B^* \in \bigcup_{k=1}^K \hat{\Phi}_k\) can be viewed as the point in \(\bigcup_{k=1}^K \hat{\Phi}_k\) nearest to the typical user where \(\hat{\Phi}_k\) is a homogeneous PPP of intensity \(\hat{\lambda}_k\) and all \(\hat{\Phi}_k\)’s are independent. Moreover, for \(B^* \in \Phi_k\) we can have the following
\[
\frac{H_k P_k}{\|B^*\|^\alpha} = \frac{\mathbb{Y}^\alpha_k \Psi^*(\|B^*\|)}{\|B^*\|^\alpha} = \frac{\sum_{B_m,j \in \cup_{k=1}^K \Phi_k \cup B^* \cup B_m,j} \mathbb{Y}^\alpha_k V_m,j \Psi_m,j(\|B_m,j\|)}{\|B^*\|^\alpha},
\]
where \(\hat{\Phi} \triangleq \bigcup_{k=1}^K \hat{\Phi}_k\). The result (d) follows from the identity that \(\mathbb{P}[Z \leq z] = \mathcal{L}^{-1} \mathbb{E}[e^{-sZ}]\) (z). Using the algebraic technique in the proof of Proposition 2 in (14) and letting \(s_m = \mathbb{Y}^{\alpha/2}(\mathbb{Y}^{-\alpha/2} - 1) \mathbb{Y}^{\alpha/2}\), the Laplace transform of \(\|B^*\|^\alpha\) \(\hat{T}^*\) can be found as follows
\[
\mathbb{E} \left[ \exp \left\{ -s (\frac{\mathbb{Y}^\alpha_k}{\mathbb{Y}^\alpha_k} - 1) \|B^*\|^\alpha \hat{T}^* \right\} \right] = \mathbb{E} \left[ \exp \left\{ -\pi \|B^*\|^\alpha \sum_{m=1}^K \frac{\hat{\lambda}_m (1 - p_{m,b}) \mathbb{Y}^{\alpha/2} \left( \int_1^\infty (1 - e^{-s_m x^{-\alpha/2}}) \, dx \right)}{s_m \mathbb{Y}^{\alpha/2}} \right\} \right],
\]
where (b) follows by assuming all \(V_m,j\)’s are independent and the non-void BS are independent PPPs that give rise to slightly larger interference so that this lower bound is fairly tight [14], [26], (c) follows from the probability generating functional of \(K\) independent homogeneous PPPs [24], [29], and the integral can be simplified as follows
\[
\int_1^\infty (1 - e^{-s_m x^{\alpha/2}}) \, dx = \int_0^{s_m \mathbb{Y}^{\alpha/2}} (1 - e^{-u}) \, du = \frac{s_m \mathbb{Y}^{\alpha/2}}{\alpha} \left( \Gamma \left( 1 - \frac{\alpha}{2} \right) \right) + \frac{2}{\alpha} \Gamma \left( -\frac{\alpha}{2}, s_m \mathbb{Y}^{\alpha/2} \right) - 1 = h \left( s_m, \frac{\alpha}{2} \right),
\]
where (d) follows from assuming \(\mathbb{Y}\) is an exponential random variable with unit mean and \(\Gamma(z, x) = \int_x^{\infty} t^{z-1} e^{-t} \, dt\) is the
lower incomplete gamma function. Then we can have
\[ E \left[ e^{-\frac{i\pi \mathbf{Y}_k}{\hat{a}} \| \hat{B}^* \|_{2,\alpha}} \right] \geq \left( 1 + \sum_{m=1}^{K} (1 - p_{m,0}) \times h \left( \frac{m}{2 \alpha}, \theta_m \right) \right)^{-1} \] (54)
since \( f_{\hat{B},\|} (x) = 2\pi (\sum_{k=1}^{K} \lambda_k) e^{-\pi x^2} \sum_{k=1}^{K} \lambda_k \) and substituting (54) into (53) yields the following result
\[ \rho_{ge}(\eta) \geq \sum_{k=1}^{K} \hat{\varphi}_k \mathbb{E} \left[ \exp \left( - \pi \| \hat{B}^* \|_{2,\alpha} \right) \right] \geq \sum_{k=1}^{K} \mathbb{E} \left[ \exp \left( - \pi \| \hat{B}^* \|_{2,\alpha} \right) \right] \] which equals to (23) due to \( \int_{0}^{\infty} g(\tau) d\tau = \mathcal{L}^{-1} \{ G(s)/s \} (t) \).

F. Proof of Theorem 4
According to the proof of Theorem 3 we also can know the SIR in (22) has the following equivalence in distribution:
\[ \frac{H_k^T \Psi_k'}{\| B^* \|_{t,\alpha}} = \frac{H_k^T \Psi_k' \cdot \text{SINR}(\| B^* \|)}{\| B^* \|_{t,\alpha}} \]
due to (23). From the proof of Proposition 2 in [14]. 

Now consider all \( \hat{\varphi}_k \)'s are \( K \) independent PPPs and the Laplace transform of \( \hat{\varphi}_k \|_{\hat{B}^*} \|_{t,\alpha} \| \) with parameter \( (2\pi \mathbf{Y}_k)/\mathbf{Y}_k \) can be found by using the proof of Proposition 2 in [14]. Accordingly, we have
\[ \rho_{ge}(\eta) \geq \sum_{k=1}^{K} \hat{\varphi}_k \mathbb{E} \left[ \exp \left( - \pi \| \hat{B}^* \|_{2,\alpha} \right) \right] \geq \sum_{k=1}^{K} \mathbb{E} \left[ \exp \left( - \pi \| \hat{B}^* \|_{2,\alpha} \right) \right] \] and then averaging this tight lower bound over \( \| \hat{B}^* \|_{2} \) with \( f_{\| \hat{B}^* \|} (x) = 2\pi (\sum_{k=1}^{K} \lambda_k) e^{-\pi x^2} (\sum_{k=1}^{K} \lambda_k) \) results in (27).

G. Proof of Theorem 5
Although \( \delta_{ge} \) can be found by integrating \( \rho_{ge}(\eta) \) in (23) with respect to \( \eta \) over [0, \infty], this method is unable to yield a much neat result. Here we find \( \delta_{ge} \) by calculating the mean spectrum efficiency of the associated BS in a more tractable way shown in the following
\[ C_{ge, k} = \frac{1}{2 \ln 2} \mathbb{E} \left[ \ln \left( 1 + \frac{P_k H_k}{\| B^* \|_{2,\alpha}} \right) \right] = \frac{1}{2 \ln 2} \int_{0}^{\infty} \left[ \left( \frac{1}{P_k H_k} + y \right)^{-1} \right] dy. \] Also, we know
\[ \mathbb{E} \left[ \left( \frac{1}{P_k H_k} + y \right)^{-1} \right] = \int_{0}^{\infty} e^{-sy} \mathbb{E} \left[ \left( \frac{1}{P_k H_k} + y \right)^{-1} \right] ds \]
and then substituting this result into (34) yields (36).

REFERENCES
[1] C.-H. Liu, “On the energy efficiency limit of dense heterogeneous cellular networks,” in Proc. IEEE Global Commun. Conf., Dec. 2016, pp. 1–7.
[2] A. Ghosh, N. Mangalvedhe et al., “Heterogeneous cellular networks: From theory to practice,” IEEE Commun. Mag., pp. 54–64, June 2012.
[3] J. G. Andrews, “Seven ways that HetNets are a cellular paradigm shift,” IEEE Commun. Mag., pp. 136–144, Mar. 2013.
[4] A. Fehske, G. Fettweis, J. Malmovin, and G. Biczok, “The global footprint of mobile communications: The ecological and economic perspective,” IEEE Commun. Mag., vol. 49, no. 8, pp. 55–62, Aug. 2011.
[5] Z. Hasan, H. Boosanimmeh, and V. K. Bhargava, “Green cellular networks: A survey, some research issues and challenges,” IEEE Communications Surveys and Tutorials, vol. 13, no. 4, pp. 524–540, Nov. 2011.
[6] I. Asghar, F. Boccardi, and L. Ho, “SLEEP mode techniques for small cell deployments,” IEEE Commun. Mag., vol. 49, no. 8, pp. 72–79, Aug. 2011.
[7] W. Wang and G. Shen, “Energy efficiency of heterogeneous cellular network,” in Proc. IEEE Vehicular Technology Conference, Sep. 2010, pp. 1–5.
[8] S.-R. Cho and W. Choi, “Energy-efficient repulsive cell activation for heterogeneous cellular networks,” IEEE J. Sel. Areas Commun., vol. 31, no. 5, pp. 870–882, May 2013.
[9] Y. S. Soh, T. Q. S. Quek, M. Kontouris, and H. Shin, “Energy efficient heterogeneous cellular networks,” IEEE J. Sel. Areas Commun., vol. 31, no. 5, pp. 840–850, May 2013.
[10] P. H. Jinlin Peng, Hao Tang and K. Xue, “Stochastic geometry analysis of energy efficiency in heterogeneous network with sleep control,” IEEE Wireless Communications Letters, vol. 2, no. 6, pp. 615–618, Dec. 2013.
[11] E. Mugume and D. K. C. So, “Sleep mode mechanisms in dense small cell networks,” in Proc. IEEE Int. Conf. on Commun., Jun. 2015.
[12] C.-T. Peng, L.-C. Wang, and C.-H. Liu, “Optimal base station deployment for small cell networks with energy-efficient power control,” in Proc. IEEE Int. Conf. on Commun., 2015, pp. 1863–1868.
[13] C.-H. Liu and L.-C. Wang, “On the optimality of green cell load and throughput in small cell networks with weighted channel-aware cell association,” in Proc. IEEE Global Commun. Conf., Dec. 2015, pp. 1–6.
[14] ——, “Optimal cell load and throughput in green small cell networks with generalized cell association,” IEEE J. Sel. Areas Commun., vol. 34, no. 5, pp. 1058–1072, May 2016.
[15] H.-S. Jo, Y. J. Sang, X. Ping, 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.
[16] C. Li, J. Zhang, and K. B. Letaief, “Throughput and energy efficiency analysis of small cell networks with multi-antenna base stations,” IEEE Trans. Wireless Commun., vol. 13, no. 5, pp. 2505–2517, May 2014.
[17] Z. Niu, Y. Wu, J. Gong, and Z. Yang, “Cell zooming for cost-efficient green cellular networks,” IEEE Commun. Mag., vol. 48, no. 11, pp. 74–79, Nov. 2010.
[18] K. Son, H. Kim, Y. Yi, and B. Krishnamachari, “Base station operation and user association mechanisms for energy-delay tradeoffs in green cellular networks,” IEEE J. Sel. Areas Commun., vol. 29, no. 8, pp. 1525–1536, Aug. 2011.
[19] D. Fooladivanda and C. Rosenberg, “Joint resource allocation and user association for heterogeneous wireless cellular networks,” IEEE Trans. Wireless Commun., vol. 12, no. 1, pp. 248–257, Jan. 2012.
[20] S. Singh, H. S. Dhillon, and J. G. Andrews, “Offloading in heterogeneous networks: modeling, analysis, and design insights,” vol. 12, no. 5, pp. 2484–2497, May 2013.
[21] S. Singh and J. G. Andrews, “Joint resource partitioning and offloading in heterogeneous cellular networks,” IEEE Trans. Wireless Commun., vol. 13, no. 2, pp. 888–901, Feb. 2014.
[22] B. Zhuang, D. Guo, and M. L. Honig, “Energy-efficient cell activation, user association, and spectrum allocation in heterogeneous networks,” IEEE J. Sel. Areas Commun., vol. 34, no. 4, pp. 823–831, Apr. 2016.
[23] A. Mesodiakaki, F. Adelantado, L. Alonso, and C. Verikoukis, “Energy-efficient user association in cognitive heterogeneous networks,” IEEE Commun. Mag., vol. 52, no. 7, pp. 22–29, Jul. 2014.
[24] D. Stoyan, W. Kendall, and J. Mecke, Stochastic Geometry and Its Applications, 2nd ed. New York: John Wiley and Sons, Inc., 1996.
[25] G. Auer, V. Giannini, C. Desset et al., “How much energy is needed to run a wireless network?” IEEE Commun. Mag., vol. 18, no. 5, pp. 40–49, Oct. 2011.
[26] C.-H. Liu and L.-C. Wang, “Random cell association and void probability in poisson-distributed cellular networks,” in Proc. IEEE Int. Conf. on Commun., Jun. 2015, pp. 2816–2821.
[27] M. Abramowitz and I. A. Stegun, Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables, 9th ed. Dover Publications, 1972.
[28] L. Xiang, X. Ge, C.-X. Wang, F. Y. Li, and F. Reichert, “Energy efficiency evaluation of cellular networks based on spatial distributions of traffic load and power consumption,” IEEE Trans. Wireless Commun., vol. 12, no. 3, pp. 961–973, Mar. 2013.
[29] M. Haenggi and R. K. Ganti, “Interference in large wireless networks,” Foundations and Trends in Networking, vol. 3, no. 2, pp. 127–248, 2009.
[30] J.-S. Ferenc and Z. Néda, “On the size distribution of Poisson Voronoi cells,” Physica A: Statistical Mechanics and its Applications, vol. 385, no. 2, pp. 518–526, 2007.