A Tractable Framework for the Analysis of General Multi-Tier Heterogeneous Cellular Networks

Serkan Ak, Student Member, IEEE, Hazer Inaltekin, Member, IEEE, H. Vincent Poor, Fellow, IEEE

Abstract

This paper investigates the downlink performance of $K$-tier heterogeneous cellular networks (HCNs) under general settings. First, Gaussian approximation bounds for the standardized aggregate wireless interference (AWI) in dense $K$-tier HCNs are obtained for when base stations (BSs) in each tier are distributed over the plane according to a spatial and general Poisson point process. The Kolmogorov-Smirnov (KS) distance is used to measure deviations of the distribution of the standardized AWI from the standard normal distribution. An explicit and analytical expression bounding the KS distance between these two distributions is obtained as a function of a broad range of network parameters such as per-tier transmission power levels, per-tier BS intensity, BS locations, general fading statistics, and general bounded path-loss models. Bounds achieve a good statistical match between the standardized AWI distribution and its normal approximation even for moderately dense HCNs. Second, various spatial performance metrics of interest such as outage capacity, ergodic capacity and area spectral efficiency in the downlink of $K$-tier HCNs for general signal propagation models are investigated by making use of the derived distribution approximation results. Considering two specific BS association policies, it is shown that the derived performance bounds track the actual performance metrics reasonably well for a wide range of BS intensities, with the gap among them becoming negligibly small for denser HCN deployments.

Index Terms

S. Ak is with the Wireless Networking and Communications Group (WNCG), The University of Texas at Austin, Austin, TX 78701, USA (e-mail: serkanak@utexas.edu). H. Inaltekin and H. V. Poor are with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA (e-mail: {hinaltek, poor}@princeton.edu).

This research was supported in part by the European Union Research Executive Agency Marie Curie FP7-Reintegration-Grants under Grant PCIG10-GA-2011-303713, in part by the Scientific and Technological Research Council of Turkey (TUBITAK) under Grant 115E162, and in part by the U.S. National Science Foundation under Grant ECCS-1343210. This work was presented in part at the IEEE International Symposium on Information Theory (ISIT), in Barcelona, Spain, July 2016 [1], [2].
I. INTRODUCTION

A. Background and Motivation

Fifth generation (5G) wireless networks are conceived as highly heterogeneous consisting of multiple-tiers of network elements with much denser deployments of base stations (BSs) and more advanced communication protocols to deal with excessive data demand from mobile users [3]–[8]. Modeling and analyzing performance of such multi-tier heterogeneous cellular networks (HCNs) using spatial point processes have recently gained increasing popularity [9]–[13]. In particular, it is shown in [9] that using a Poisson point process (PPP) model even for macro cell BS locations provides us with an approximation as good as the one provided by the conventional grid based model [14] for the actual network performance. With the introduction of more irregularly deployed network elements such as femto BSs and distributed antennas, it is expected that the PPP model will more accurately track the actual HCN performance.

A major design issue for such PPP based HCN models is characterization and mitigation of aggregate wireless interference (AWI), e.g., see [15] and [16]. Although the statistical characterization of AWI is possible for the special case of Rayleigh fading and the classical unbounded path-loss model to compute various performance metrics such as outage probability in closed form [10]–[12], computation of the exact AWI distribution is a very challenging task that usually does not result in closed form expressions for general signal propagation models [9], [17]. The main objective of the current paper is to generalize those previously known results and techniques to study the dense HCN performance under more general and heterogeneous communication scenarios that take into account the general bounded path-loss models, general fading distributions and spatial distributions of BSs including non-homogeneous PPPs. This is achieved by leveraging our methodological analysis approximating the standardized AWI distribution as a normal distribution.

B. Main Contributions

A thorough mathematical analysis of the AWI distribution under general network settings is often a prohibitively hard task. To overcome this difficulty, we provide a simple and effective methodological approach for examining the statistical structure of AWI in the downlink of a dense $K$-tier HCN, wherein the network tiers are differentiated from each other in terms of transmission power levels, spatial BS distributions and RF signal propagation characteristics. The proposed approach provides us
with a principled framework to analyze the downlink performance of spatially dense $K$-tier HCNs, where BS locations in each tier follow a PPP with possibly non-homogeneous spatial BS intensity. The signal power attenuation due to path-loss is modeled through a general bounded and power-law decaying path-loss function, which can vary from one tier to another. Fading and shadowing are also accounted for in the employed signal propagation model without assuming any specific distribution functions for these other random wireless channel dynamics. Similar to [17], we only consider the statistics of the AWI power in this paper. Amplitude statistics of the AWI, strongly related to the type of employed signaling scheme, are out of scope of the current paper. Our specific contributions are further explained below.

1) An Analytical Approximation for the Distribution of AWI in Multi-Tier HCNs: Measuring the distance between the standardized downlink AWI and normal distributions by means of the Kolmogorov-Smirnov (KS) distance, we obtain an analytical expression for deviations between them, which paves the way for illuminating statistical behavior of the AWI and in turn for designing an efficient $K$-tier HCN. Briefly, the stated distance consists of two parts: (i) a scaling coefficient and (ii) a multiplicative positive function $c(x)$ with $x \in \mathbb{R}$ being the point at which we want to estimate the value of the standardized AWI distribution.

The scaling coefficient has two important properties. First, it is related to the skewness characteristic of the individual interference terms constituting the standardized AWI, which mainly depends on various network parameters at each tier such as transmission powers, BS distribution and signal propagation characteristics. The second property is its monotonically decaying nature to zero with denser deployments of BSs per tier. The function $c(x)$ is uniformly bounded by a small constant and approaches zero for large absolute values of $x$ at a rate $|x|^{-3}$, which makes the derived bounds on the tails of the standardized downlink AWI distribution tight even for sparsely deployed HCNs.

2) Bounds on Various HCN Performance Metrics Under General Settings: We derive analytical expressions for upper and lower bounds on performance metrics of interest in multi-tier HCNs under general network settings and general PPPs. In particular, we focus on the downlink data rates in a dense $K$-tier HCN under two different BS association policies. We obtain tight performance bounds on the downlink outage capacity, ergodic capacity and area spectral efficiency (ASE) in HCNs under both association policies for general signal propagation models. The derived bounds approximate the actual HCN performance metrics accurately for a wide range of BS intensities in each tier. The gap between bounds becomes negligibly small as the BSs are more densely deployed.

Based on the derived analytical expressions, we show that the outage probability is an increasing function of each BS intensity. Our simulations corroborate these analytical results indicating that an
increase in the network intensity induces degradations in outage capacity, ergodic capacity and ASE, each of which scales at least as fast as $\Theta \left(\|\lambda\|^{-1}_2\right)$, where $\lambda = [\lambda_1,\ldots,\lambda_K]^T$ is the BS intensity vector. Since our general bounded path loss model can be extended to multi-slope models, it is also suitable for emerging millimeter wave (mmWave) communications in dense HCN settings [18], [19]. Finally, the proposed approach can be extended to other HCN control policies and performance metrics, with the potential of shedding light on HCN performance and design beyond specific selections of the path-loss model and the fading distribution.

C. Paper Organization and Notation

The rest of the paper is organized as follows. In Section II, we present a survey of relevant previous work by comparing and contrasting existing results with those we obtain in the current paper. In Section III, we introduce the system model. In Section IV, we present our methodological approach for approximating the AWI distribution in HCNs as a Gaussian distribution and illustrate numerical examples of these analytical findings. In Section V, we introduce BS association policies for performance analysis and derive our bounds tracking the performance metrics, especially outage characteristics, of mobile users under these association policies. We illustrate simulation results on these performance bounds in Section VI too. Finally, Section VI concludes the paper.

We use boldface letters, upper-case letters and calligraphic letters to denote vector quantities, random variables and sets, respectively. $| \cdot |$ notation is used to measure the magnitudes of scalar quantities, whereas $\| \cdot \|_2$ notation is used to measure the Euclidean norms of vector quantities. $\mathbb{I}_{\{\cdot\}}$ is used to denote the indicator function. Expected value and variance of a random variable $X$ are denoted by $\mathbb{E}\left[X\right]$ and $\text{Var}\left(X\right)$, respectively. As is standard, when we write $f(t) = O\left(g(t)\right)$, $f(t) = \Omega\left(g(t)\right)$ and $f(t) = o\left(g(t)\right)$ as $t \to t_0$ for two positive functions $f(t)$ and $g(t)$, we mean $\lim_{t \to t_0} \frac{f(t)}{g(t)} < \infty$, $\lim_{t \to t_0} \frac{f(t)}{g(t)} > 0$ and $\lim_{t \to t_0} \frac{f(t)}{g(t)} = 0$, respectively. $f(t)$ is said to be $\Theta\left(g(t)\right)$ as $t \to t_0$ if $f(t) = O\left(g(t)\right)$ and $f(t) = \Omega\left(g(t)\right)$ as $t \to t_0$.

II. RELATED WORK

The early work in the literature focusing on the design and analysis of wireless networks by means of stochastic geometry based models includes [20]–[24]. These papers considered traditional single-tier macro cell deployments and obtained various approximations on the distribution of AWI using characteristic functions [20], LePage series [21], Edgeworth expansion [22], geometrical considerations [23] and skewed stable distributions [24]. In [20], a closed form expression was also obtained for the AWI distribution under the assumption of no fading and unbounded power-law decaying path-loss.
function when the path-loss exponent is 4. More recently, generalized shot-gun models are considered for one-, two- and three-dimensional wireless networks to derive semi-analytical expressions for the downlink coverage probability for arbitrary fading and general path-loss models in [25], optimum downlink coverage for Poisson cellular networks subject to transmit power, BS density and transmit power density constraints is derived in [26], and Berry-Esseen types of bounds were obtained in [27]–[29], but again by considering only single-tier wireless networks.

The current paper differs from the above previous work in several important aspects. In particular, this paper extends the previous known results approximating the AWI distribution for macro cell deployments to more heterogeneous and complex wireless communication environments when compared to [20]–[29]. Functional dependencies among different tiers to approximate the AWI distribution in the downlink of a K-tier HCN are clearly identified. While the authors in [24] showed that the AWI distribution can be modeled as a skewed stable distribution with unity skew parameter under the unbounded path-loss model, our approximation technique uses general bounded path-loss models and demonstrates that the skewness of the standardized AWI distribution decreases to zero as the K-tier HCN in question gets denser. We discover that the behavior of AWI changes from being heavy-tailed to an exponentially decaying light-tailed one for bounded power-law decaying path-loss models.

This paper differs from [25]–[28] by focusing on a multi-tier network scenario with distinct and general network parameters in each tier, which enables us to discover further insights into the behavior of AWI in dense HCNs via Lemmas 1 and 2. The authors in [25] focuses only on single-tier networks, and their equivalence results are between one-, two- and three-dimensional Poisson wireless networks. As a result, the results presented in [25] are not applicable to our case to first reduce the multi-tier network to a single-tier network, and then use the Gaussian approximation results presented in our previous work [28], [29]. In comparison with the preliminary results presented in [29], the association policies (APs) studied for K-tier HCNs in Section V are much richer in terms of the network parameters that they include, which provides additional insights into various capacity metrics for dense HCNs.

The related work also includes the papers that use stochastic geometry to model and analyze HCNs such as [30]–[36]. To start with, the papers [30] and [31] calculate the statistics of signal-to-interference-plus-noise ratio (SINR) in the downlink of HCNs by utilizing moment generating functions (MGF). In [32], the authors investigated a gamma distribution approximation for the distribution of AWI clogging a fixed-size cell with a guard zone and a dominant interferer. In [33], the author derived the downlink SINR distribution for K-tier HCNs by assuming the classical unbounded path-loss model, Rayleigh faded wireless links and the nearest base-station (BS) association rule. In [34], they considered vector broadcast channels operating according to opportunistic beamforming in a K-tier HCN setting,
and obtained tight approximations for beam outage probabilities and ergodic aggregate data rates for Rayleigh faded propagation environments with homogenous PPP distributions for the locations of access points.

In [35], the authors extended the previous results in [25] for four different association policies in HCNs (i.e., max-SINR, nearest-BS, maximum received instantaneous power and maximum biased received power association models) for an unbounded path-loss model with specific functional form (having varying path-loss exponents from tier-to-tier) and with arbitrary fading when the BS locations in each tier are given by homogeneous PPPs. They obtained semi-analytical expressions for the downlink performance of HCNs involving complex-valued integrals. In [36], the authors used the factorial moments for the SINR process to show that a generic multi-tier HCN can be transformed into a stochastically equivalent single-tier network when the BSs in each tier are distributed over the plane according to homogenous PPPs and the path-loss model is given by the unbounded inverse power-law function.

The current paper differs from [30] and [31] by utilizing general and tier-dependent bounded path-loss models as well as using general fading distributions. In comparison with wireless network performance results obtained for homogeneous PPPs, we show that incorporating the skewness of the standardized AWI via the Berry-Esseen theorem makes the Gaussian approximation analytically more useful and numerically more tractable to accurately approximate the AWI distribution even for non-homogenous PPPs and general fading models. When compared with the results reported in [32]–[34], our network set-up is much richer, allowing non-homogeneous PPPs for BS locations and general signal propagation models including fading and shadowing.

This paper differs from [35] in two important directions. Firstly, our network model is not restricted to homogenous PPPs and the unbounded path-loss model. Secondly, our Gaussian approximation technique does not necessitate an inversion of complex-valued characteristic functions. More specifically, our model generalizes the results in [35] to non-homogenous PPPs with arbitrary mean-measure and to bounded path-loss models with arbitrary functional forms by only requiring computation of integrals with respect to the generalized fading distributions on the real line to obtain tight performance bounds on the downlink of $K$-tier HCNs. The approach presented in [36] for mapping a $K$-tier HCN to a stochastically equivalent single-tier planar wireless network is promising, but the key propagation invariance property disappears for general bounded path-loss models. Hence, such stochastic equivalence results are not directly applicable to the $K$-tier HCN setup with general bounded path-loss models and non-homogenous PPPs studied in this paper. More generally, we observe that the SINR invariance property does not hold in our case, which reinforces the similar observations obtained for multi-slope
III. System Model

In this section, we will introduce the details of the studied downlink model in a $K$-tier cellular topology, the details of the spatial processes determining BS locations, the signal propagation characteristics and the association policy under which the network performance of a $K$-tier HCN is determined.

A. The Downlink Model in a $K$-Tier Cellular Topology

We consider an overlay $K$-tier HCN in which the BSs in all tiers are fully-loaded (i.e., no empty queues) and have access to the same communication resources both in time and frequency. The BSs in different tiers are differentiated mainly on the basis of their transmission powers, with $P_k > 0$ being the transmission power of a tier-$k$ BS for $k = 1, \ldots, K$. As is standard in stochastic geometric modeling, it is assumed that BSs are distributed over the plane according to a general PPP with differing spatial density among the tiers. Further, the signal propagation characteristics (including both large-scale path-loss and small-scale fading) also vary from one tier to another. The details of BS location processes and signal propagation are elaborated below.

We place a test user at an arbitrary point $x^{(o)} = (x_1^{(o)}, x_2^{(o)}) \in \mathbb{R}^2$ and consider signals coming from all the BSs in all tiers as the downlink AWI experienced by this test user. Since we focus on the downlink analysis, we assume that the uplink and downlink do not share any common communication resources. Therefore, the uplink interference can be ignored for the analysis of downlink AWI. This setting is general enough to illuminate the effects of various network parameters such as transmission powers and BS intensity in each tier on the distribution of the AWI seen by the test user.

B. BS Location Processes

The BS locations in tier-$k$, $k = 1, \ldots, K$, independently form a spatial planar PPP $\Phi_k$. $\Lambda_k$ represents the mean measure (alternatively called the intensity measure or spatial density) of the $k$th tier BSs. We do not assume any specific functional form for $\Lambda_k$ and hence do not restrict our attention only to homogeneous PPPs. For each (Borel) subset $S$ of $\mathbb{R}^2$, $\Lambda_k (S)$ gives us the average number of BSs lying in $S$. We will assume that $\Lambda_k$ is locally finite i.e., $\Lambda_k (S) < \infty$ for all bounded subsets $S$ of $\mathbb{R}^2$, and $\Lambda_k (\mathbb{R}^2) = \infty$, i.e., there is an infinite population of tier-$k$ BSs scattered all around in $\mathbb{R}^2$. For the whole HCN, the aggregate BS location process, which is the superposition of all individual position processes, is denoted by $\Phi = \bigcup_{k=1}^K \Phi_k$. 
For mathematical convenience, we also express $\Phi_k$ as a discrete sum of Dirac measures as $\Phi_k(S) = \sum_{j \geq 1} \delta_{X_j^{(k)}}(S)$, where $\delta_{X_j^{(k)}}(S) = 1$ if $X_j^{(k)} \in S \subseteq \mathbb{R}^2$, and zero otherwise. The level of AWI at $x^{(o)}$ from tier-$k$ BSs depends critically on the distances between the points of $\Phi_k$ and $x^{(o)}$. It is well-known from the theory of Poisson processes that the transformed process $\sum_{j \geq 1} \delta_{T(X_j^{(k)})}$ is still Poisson (on the positive real line) with mean measure given by $\Lambda_k \circ T^{-1}$, where $T(x) = \|x - x^{(o)}\|_2 = \sqrt{(x_1 - x_1^{(o)})^2 + (x_2 - x_2^{(o)})^2}$ and $T^{-1}(S) = \{x \in \mathbb{R}^2 : T(x) \in S\}$ for all $S \subseteq \mathbb{R}$ [37]. We will assume that $\Lambda_k \circ T^{-1}$ has a density in the form $\Lambda_k \circ T^{-1}(S) = \lambda_k \int_S \mu_k(t)dt$. Here, $\lambda_k$ is a modeling parameter pertaining to the $k$th tier, which can be interpreted as the BS intensity parameter, that will enable us to control the average number of tier-$k$ BSs whose distances from $x^{(o)}$ belong to $S$ and interfere with the signal reception at the test user.

C. Signal Propagation Model

We model the large scale signal attenuation for tier-$k$, $k = 1, \ldots, K$, by a bounded monotone non-increasing path-loss function $G_k : [0, \infty) \mapsto [0, \infty)$. $G_k$ asymptotically decays to zero at least as fast as $t^{-\alpha_k}$ for some path-loss exponent $\alpha_k > 2$. To ensure the finiteness of AWI at the test user, we require the relationship $\mu_k(t) = O(t^{\alpha_k - 1 - \epsilon})$ as $t \to \infty$ to hold for some $\epsilon > 0$.

The fading (power) coefficient for the wireless link between a BS located at point $X \in \Phi$ and the test user is denoted by $H_X$. The fading coefficients $\{H_X\}_{X \in \Phi}$ form a collection of independent random variables (also independent of $\Phi$), with those belonging to the same tier, say tier-$k$, having a common probability distribution with density $q_k(h), h \geq 0$. The first, second and third order moments of fading coefficients are assumed to be finite, and are denoted by $m^{(k)}_1, m^{(k)}_2$ and $m^{(k)}_3$, respectively, for tier-$k$. We note that this signal propagation model is general enough that $H_X$’s could also be thought to incorporate shadow fading effects due to blocking of signals by large obstacles existing in the communication environment, although we do not model such random factors explicitly and separately in this paper.

D. Association Policy, Interference Power and Performance Measures

Association policy is a key mechanism that determines the outage and rate performance experienced by the test user as it regulates the useful signal power as well as the interference power at the test

\footnote{For simplicity, we only assign a single fading coefficient to each BS. In reality, it is expected that the channels between a BS and all potential receivers (intended or unintended) experience different (and possibly independent) fading processes. Our simplified notation does not cause any ambiguity here since we focus on the outage and rate performance of the test user at a given arbitrary position in $\mathbb{R}^2$ in the remainder of the paper.}
user. Hence, we first formally define it to facilitate the upcoming discussion.

Definition 1: An association policy \( \mathcal{A} : \Omega \times \mathbb{R}_+^\infty \times \mathbb{R}_+^K \times \mathbb{R}_+^K \to \mathbb{R}^2 \) is a mapping that takes a BS configuration \( \varphi \in \Omega \) (i.e., a countable point measure), fading coefficients \( \{h_x\}_{x \in \varphi} \), transmission power levels \( \{P_k\}_{k=1}^K \) and biasing factors \( \{\beta_k\}_{k=1}^K \) as an input and determines the BS location to which the test user is associated as an output.

For the HCN model explained above, the output of \( \mathcal{A} \) is a random point \( \mathbf{X}^* = (X^*_1, X^*_2) \in \Phi \) since the BS locations and fading coefficients are random elements. Biasing coefficients are important design parameters to offload data from bigger cells to the smaller ones. Two other important random quantities related to \( \mathbf{X}^* \) are the tier index \( \mathcal{A}^* \) to which \( \mathbf{X}^* \) belongs and the distance between \( \mathbf{X}^* \) and the test user \( x^{(o)} = (x^{(o)}_1, x^{(o)}_2) \in \mathbb{R}^2 \), which is denoted by \( R^* = \|\mathbf{X}^* - \mathbf{x}^{(o)}\|_2 = \sqrt{(X^*_1 - x^{(o)}_1)^2 + (X^*_2 - x^{(o)}_2)^2} \). Using these definitions along with considering all the signal impairments due to fading and path-loss, the total interference power at the test user is written as

\[
I_\lambda = \sum_{\mathbf{X} \in \Phi \setminus \{\mathbf{x}^{(o)}\}} P_\mathbf{X} H_\mathbf{X} G_\mathbf{X} (T(\mathbf{X})),
\]

where \( \lambda = [\lambda_1, \ldots, \lambda_K]^T \), and it is understood that \( P_\mathbf{X} = P_k \) and \( G_\mathbf{X} = G_k \) if \( \mathbf{X} \in \Phi_k \). This parametrization of AWI is chosen to emphasize the dependence of its distribution on the BS intensity parameter \( \lambda_k \) of each tier.

SINR is the main performance determinant for the HCN model in question. Given an association policy \( \mathcal{A} \), the SINR level experienced by the test user is equal to

\[
\text{SINR}_\mathcal{A} = \frac{P_{\mathcal{A}^*} H_{\mathbf{X}^*} G_{\mathbf{X}^*} (T(\mathbf{X}^*))}{N_0 + \frac{1}{P_G} I_\lambda},
\]

where \( N_0 \) is the constant background noise power and \( P_G \geq 1 \) is a processing gain constant that signifies the interference reduction capability, if possible, of the test user. We also let \( \text{SNR}_k = \frac{P_k}{N_0} \) to denote the signal-to-noise ratio (SNR) for tier-\( k \). Next, we define the main performance metrics used to measure the HCN outage and rate performance.

Definition 2: For a target bit rate \( \tau \), \( \tau \)-outage probability is equal to

\[
\Pr(\tau\text{-outage}) = \Pr \{ \log(1 + \text{SINR}_\mathcal{A}) < \tau \}.
\]

Similarly, for a target outage probability \( \gamma \), the outage capacity achieved by the test user under the association policy \( \mathcal{A} \) is equal to

\[
C_o(\gamma) = \sup \{ \tau \geq 0 : \Pr(\tau\text{-outage}) \leq \gamma \},
\]

which is the maximum data rate supported with outage probability not exceeding \( \gamma \).
Unlike the outage capacity in which (the instantaneous) SINR$_A$ is assumed to be constant for the duration of channel coherence time, ergodic capacity is the average of (the instantaneous) capacity that can be achieved by averaging over a large number of coherence time intervals leading to Definition 3 below.

**Definition 3:** The ergodic capacity achieved by the test user under the association policy $A$ is equal to

$$C_{\text{erg}} = \mathbb{E}[\log (1 + \text{SINR}_A)].$$

In the next section, we will first present our methodological approach establishing the bounds on the KS distance to measure the proximity of the AWI distribution to the normal distribution. These approximation results will be leveraged in Section V to obtain the outage and rate performance of HCNs.

### IV. Gaussian Approximation for the AWI Distribution

In this section, we will establish the Gaussian approximation bounds under different spatial distribution assumptions for the standardized AWI distribution in the downlink of an HCN. Then, we will present numerical examples validating our theoretical work.

#### A. Analytical Results

In this part, we present our theorems providing explicit upper and lower bounds on the AWI distribution, which the test user experiences at a specific location in a $K$-tier HCN. These bounds will clearly show the functional dependence between the downlink AWI distribution and a broad range of network parameters such as transmission power levels, BS distribution over the plane and signal propagation characteristics in each tier. We will also specialize these approximation results to the commonly used homogeneous PPPs at the end of this part.

**Theorem 1:** For all $x \in \mathbb{R}$,

$$\left| \Pr \left\{ \frac{I_A - \mathbb{E}[I_A]}{\sqrt{\text{Var}[I_A]}} \leq x \right\} - \Psi(x) \right| \leq \Xi \cdot c(x),$$

where $\Xi = \sum_{k=1}^{K} \frac{\lambda_k P_k \mu_k(t)}{\lambda_k P_k \mu_k(t)} \int_{0}^{\infty} \frac{G^2_k(t) \mu_k(t) dt}{\sqrt{2\pi}}$, $c(x) = \min\left(0.4785, \frac{31.935}{1+|x|^2}\right)$ and $\Psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$, which is the standard normal cumulative distribution function (CDF).

---

3We do not assume any interference protection or cancellation to derive the results in this section. In Section V, we show that the results in this section can be modified in a straightforward manner to incorporate the effects of association policies on the $K$-tier HCN performance.
Proof: Please see Appendix B.

Measuring the KS distance, Theorem 1 provides us with an explicit expression for the deviations between the standardized AWI and normal distributions. Several important remarks about this result are in order. First, the standardized AWI can take negative values due to the centering operation, which makes the deviations from the normal distribution bounded for negative values of $x$, as desired in Theorem 1. Second, since a bounded path-loss model is used in each tier, the decay rate of the tails of the AWI distribution also depends on the fading distribution parameters in each tier as indicated by our bound in Theorem 1. A similar phenomenon was also observed in [17], [39]. We refer interested readers to [17], [39] for a thorough comparison between bounded and unbound path-loss models.

Third, the scaling coefficient $\Xi$ appearing in Theorem 1 is linked to the main network parameters such as transmission power levels, distribution of BSs over the plane and signal propagation characteristics. Starting with the BS intensity parameters $\lambda_k$, $k = 1, \ldots, K$, we observe that the rate of growth of the expression appearing in the denominator of $\Xi$ is half an order larger than that of the expression appearing in the numerator of $\Xi$ as a function of $\lambda_k$. This observation implies that the derived Gaussian approximation becomes tighter for denser deployments of HCNs. A formal statement of this result is given in Lemma 1.

Fourth, the same functional form for $c(x)$ was also obtained in papers [28], [29]. The reason is the fact that the same form of the Berry-Esseen theorem from [38] is applicable to both single-tier and multi-tier networks, but not the stochastic equivalence between a single-tier network and multi-tier HCNs [35], [36]. Fifth, an approach alternative to our Gaussian approximation method in Theorem 1 would be to use characteristic functions and/or Laplace transforms in an inversion formula to derive performance metrics for $K$-tier HCNs. However, this is not tractable for the HCN setup with non-homogeneous PPPs, general path-loss models and general fading distributions considered in this paper.

For example, an important feature of our Gaussian approximation result in Theorem 1 is that it only depends on the fading distributions through their second and third moments. On the other hand, fading distributions appear in a more convoluted manner in the method of characteristic functions and/or Laplace transforms that necessitates the computation of integrals with respect to both the mean measure of the underlying BS processes and the per-tier fading distributions. Finally, the skewness of the standardized AWI is related to its third moment, whose absolute value is upper bounded by $\Xi$ in Theorem 1 and hence a decreasing function of each BS intensity.

Lemma 1: The scaling coefficient $\Xi$ appearing in the Gaussian approximation result in Theorem 1 is bounded above by $\Xi \leq \frac{\delta}{\sqrt{||\lambda||_2}}$ for some finite positive constant $\delta$. 
Proof: Let $a_k = P^3 k m_H^{(k)} \int_0^\infty G^3_k(t) \mu_k(t) dt$ and $b_k = P^2 k m_{H^2}^{(k)} \int_0^\infty G^2_k(t) \mu_k(t) dt$. Then,

$$\Xi = \frac{\sum_{k=1}^K a_k \lambda_k}{\left( \sum_{k=1}^K \lambda_k b_k \right)^{\frac{3}{2}}} \leq \frac{\| \lambda \|_2}{\| a \|_2} \left( \sum_{k=1}^K \lambda_k b_k \right)^{\frac{3}{2}},$$

due to the Cauchy-Schwarz inequality. Further, we can lower-bound the sum in the denominator above as

$$\left( \sum_{k=1}^K \lambda_k b_k \right)^{\frac{3}{2}} \geq \left( \min_{1 \leq k \leq K} b_k \sum_{k=1}^K |\lambda_k| \right)^{\frac{3}{2}} \geq \epsilon \left( \| \lambda \|_2 \right)^{\frac{3}{2}},$$

where the last inequality follows from the equivalence of all the norms in finite dimensional vector spaces. Combining these two inequalities, we conclude the proof.

In the case of ultra-dense HCNs, the asymptotic behavior of the scaling coefficient $\Xi$ (and the related skewness behavior) is formally described by Lemma 1, i.e., $\Xi = O \left( \frac{1}{\sqrt{\| \lambda \|_2}} \right)$ as $\| \lambda \|_2$ goes to infinity. Further, the effect of an association policy on $\delta$ in Lemma 1 can be seen more clearly by considering specific association policies (i.e., see Lemma 3 or Lemma 4). In particular, the association policies considered in this paper modify the lower limits of the integrals in the numerator and denominator of $\Xi$. The numerical value of the ratio of these integrals are usually small numbers a little greater than one, i.e., see [28] for one specific example. Moreover, the existence of the multiplying function $c(x)$ makes our bounds further sharpened, and hence the effect of a specific association policies on our Gaussian approximation stays limited.

Following a similar approach above, we can also see that changing transmission powers is not as effective as changing BS intensity parameters to improve the Gaussian approximation bound in Theorem 1. This is expected since the power levels are assumed to be deterministic (i.e., no power control is exercised) and therefore they do not really add to the randomness coming from the underlying spatial BS distribution over the plane and the path-loss plus fading characteristics modulating transmitted signals. Another important observation we have in regards to the combined effect of the selection of transmission powers per tier and the moments of fading processes in each tier on the Gaussian approximation result in Theorem 1 is that our approximation bounds benefit from the fading distributions with restricted dynamic ranges and the alignment of received AWI powers due to fading and path-loss components. This observation is made rigorous through the following lemma.

Lemma 2: Let $a_k = \lambda_k \int_0^\infty G^3_k(t) \mu_k(t) dt$, $b_k = \lambda_k \int_0^\infty G^2_k(t) \mu_k(t) dt$ and $c_k = P^2 k m_{H^2}^{(k)}$. Then, the scaling coefficient $\Xi$ appearing in the Gaussian approximation result in Theorem 1 is bounded below by

$$\Xi \geq \left( \frac{1}{\| c \|_2 \| b \|_2} \right)^{\frac{3}{2}} \sum_{k=1}^K a_k c_k^{\frac{3}{2}},$$
with equality achieved if fading processes in all tiers are deterministic and the vectors \( b = [b_1, \ldots, b_K]^\top \) and \( c = [c_1, \ldots, c_K]^\top \) are parallel.

**Proof:** Using \( a_k, b_k \) and \( c_k \) introduced above, we can write a lower bound for \( \Xi \) as

\[
\Xi = \frac{\sum_{k=1}^K a_k P_k^3 m_{H^3}^{(k)}}{\left( \sum_{k=1}^K b_k c_k \right)^\frac{3}{2}} = \frac{\sum_{k=1}^K a_k P_k^3 m_{H^3}^{(k)}}{\left( \| c \|_2 \right)^\frac{3}{2} \left( \sum_{k=1}^K b_k \| c \|_2 \right)^\frac{3}{2}} \geq \left( \frac{1}{\| c \|_2 b_2} \right)^\frac{3}{2} \sum_{k=1}^K a_k P_k^3 m_{H^3}^{(k)}.
\]

Using Jensen’s inequality, we also have \( m_{H^3}^{(k)} \geq \left( m_{H^2}^{(k)} \right)^{\frac{3}{2}} \). Using this lower bound on \( m_{H^3}^{(k)} \) in the above expression, we finally have \( \Xi \geq \left( \frac{1}{\| c \|_2 b_2} \right)^\frac{3}{2} \sum_{k=1}^K a_k c_k^3 \).

In addition to the above fundamental properties of the scaling coefficient \( \Xi \), it is also worthwhile to mention that the Gaussian approximation bound derived in Theorem 1 is a combination of two different types of Berry-Esseen bounds embedded in the function \( c(x) \). One of these bounds is a uniform bound that helps us to estimate the standardized AWI distribution uniformly as

\[
\Pr \left\{ \frac{I_\lambda - E[I_\lambda]}{\sqrt{\text{Var}[I_\lambda]}} \leq x \right\} - \Psi(x) \leq \Xi \cdot 0.4785
\]

for all \( x \in \mathbb{R} \). On the other hand, the other one is a non-uniform bound that helps us to estimate the tails of the standardized AWI distribution as

\[
\Pr \left\{ \frac{I_\lambda - E[I_\lambda]}{\sqrt{\text{Var}[I_\lambda]}} \leq x \right\} - \Psi(x) \leq \Xi \cdot \frac{31.935}{1 + |x|^3}
\]

and decays to zero as a third order inverse power law.

Up to now, we considered general PPPs for the distribution of BSs in each tier. One simplifying assumption in the literature is to assume that PPPs determining the locations of BSs are homogeneous. In this case, \( \mu_k(t) \) for all tiers is given by \( \mu_k(t) = 2\pi t 1_{\{t \geq 0\}} \), where \( 1_{\{\cdot\}} \) is the indicator function.

Using this expression for \( \mu_k(t) \) in Theorem 1 we obtain the following approximation result for the distribution of AWI when all BSs are homogeneously distributed over the plane according to a PPP with differing BS intensity parameters \( \lambda_k \) from tier to tier.

**Theorem 2:** Assume that \( \Phi_k \) is a homogeneous PPP with a mean measure given \( \Lambda_k(S) = \lambda_k \cdot \text{area}(S) \). Then, for all \( x \in \mathbb{R} \),

\[
\Pr \left\{ \frac{I_\lambda - E[I_\lambda]}{\sqrt{\text{Var}[I_\lambda]}} \leq x \right\} - \Psi(x) \leq \Xi \cdot c(x),
\]

where \( \Xi = \frac{1}{\sqrt{2\pi}} \sum_{k=1}^K \frac{\lambda_k P_k^3 m_{H^3}^{(k)} \int_0^\infty G_k^2(t)tdt}{\left( \sum_{k=1}^K \lambda_k P_k^2 m_{H^2}^{(k)} \int_0^\infty G_k^2(t)tdt \right)^\frac{3}{2}} \),

\[
c(x) = \min \left\{ 0.4785, \frac{31.935}{1 + |x|^3} \right\} \quad \text{and} \quad \Psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt,
\]

which is the standard normal CDF.
**Fig. 1.** Gaussian approximation bounds for the standardized AWI CDFs (upper figures). Comparison of the simulated standardized AWI CDFs with the standard normal CDF (lower figures). Rayleigh fading with unit mean power is assumed.

**Proof:** The proof follows from Theorem 1 by replacing $\mu_k(t)$ with $2\pi t 1_{\{t \geq 0\}}$.

When all network parameters are assumed to be the same, i.e., the same transmission power levels, fading distributions and BS distributions for all tiers, the HCN in question collapses to a single tier network. In this case, the Gaussian approximation result is given below.

**Corollary 1:** Assume $P_k = P$, $\mu_k(t) = 2\pi t 1_{\{t \geq 0\}}$, $G_k(t) = G(t)$, $\lambda_k = \lambda$, $m_{H2}^{(k)} = m_{H2}$ and $m_{H3}^{(k)} = m_{H3}$ for all $k = 1, \ldots, K$. Then, for all $x \in \mathbb{R}$, we have

$$\left| \Pr \left\{ \frac{I_{\lambda} - E[I_{\lambda}]}{\sqrt{\text{Var}[I_{\lambda}]} \leq x} \right\} - \Psi(x) \right| \leq \Xi \cdot c(x),$$

where $\Xi = \frac{1}{\sqrt{2\pi} \sqrt{K \lambda}} m_{H3}^{\frac{\lambda}{3}} \int_{0}^{\infty} G^3(t) dt$, and $c(x)$ and $\Psi(x)$ are as given in Theorem 1.

We note that this is the same result obtained in [28] as a special case of the network model studied in this paper.

**B. Numerical Verification for the Gaussian Approximation Results**

In this part, we will illustrate the analytical Gaussian approximation results derived for the standardized AWI distribution in Section IV-A for a specific three-tier HCN scenario. To this end, we will assume the same path-loss model $G_k(t) = \frac{1}{1+t^\alpha}$ for all tiers with various values of $\alpha > 2$. 

---
Similar conclusions continue to hold for other path-loss models. The BSs in each tier are distributed over the plane according to a homogeneous PPP, with BS intensity parameters given by $\lambda_1 = 0.1\kappa$, $\lambda_2 = \kappa$ and $\lambda_3 = 5\kappa$. Here, $\kappa$ is our unitless control parameter to control the average number of BSs interfering with the signal reception at the test user. The test user is assumed to be located at the origin without loss of any generality since we focus only on homogeneous PPPs in this numerical study. The random fading coefficients in all tiers are assumed to be independent and identically distributed random variables, drawn from a Rayleigh distribution with unit mean power gain. Our results are qualitatively the same for other fading distributions such as Nakagami and Rician fading distributions. The transmission power levels are set as $P_1 = 4P_2 = 16P_3$, where $P_2$ is assumed to be unity.

We only plot the CDF of AWI in Fig. 1 because Theorem 1 measures the deviations between the CDFs of standardized AWI and a standard normal random variable. In the upper figures in Fig. 1, we present the upper and lower bounds for the deviations between the standardized AWI and normal distributions, i.e., we plot the expressions $\Psi(x) + \Xi \cdot c(x)$ and $\Psi(x) - \Xi \cdot c(x)$ appearing in Theorem 1 with a variety of $\kappa$ values. Two different regimes are apparent in these figures. For the moderate values at which we want to estimate the CDF of standardized AWI, i.e., $\Pr\left\{ \frac{|X - E[I_x]|}{\sqrt{\text{Var}[I_x]}} \leq x \right\}$ with moderate $x$ values, our uniform Berry-Esseen bound, which is $\Xi \cdot 0.4785$, provides better estimations for the AWI distribution. On the other hand, for absolute values larger than $4.0359$ at which we want to estimate the CDF of standardized AWI, i.e., $\Pr\left\{ \frac{|X - E[I_x]|}{\sqrt{\text{Var}[I_x]}} \leq x \right\}$ with $|x|$ larger than $4.0359$, our non-uniform Berry-Esseen bound, which is $\Xi \cdot \frac{0.21.9075}{1 + |x|^3}$, is tighter. These figures also clearly demonstrate the effect of the BS intensity parameters $\lambda_k$ on our Gaussian approximation bounds. As suggested by Lemma 1, the KS distance between the standardized AWI and normal distributions approach zero at a rate at least as fast as $\frac{1}{\sqrt{\|\lambda\|_2^2}}$. Further, even if all BS intensity parameters are fixed, the distance between the upper and lower bounds in Theorem 1 disappears at a rate $O\left(|x|^{-3}\right)$ as $|x| \to \infty$ due to the non-uniform bound.

When we compare the upper lefthand side and righthand side figures in Fig. 1, we observe a better convergence behavior for smaller values of the path-loss exponent $\alpha$. This is due to the path-loss model dependent constants appearing in Theorem 1. For this particular choice of the path-loss model and BS distribution over the plane, our approximation results benefit from small values of path-loss exponent, although the difference between them becomes negligible for moderate to high values of $\kappa$.

We also performed Monte-Carlo simulations to compare simulated standardized AWI distributions with the normal distribution for $10^4$ random BS configurations. The lower figures in Fig. 1 provide further numerical evidence for the Gaussian approximation of AWI in HCNs. Surprisingly, there is a
good match between the simulated standardized AWI distribution and the standard normal CDF even for sparsely populated HCNs, i.e., $\kappa = 1$.

In the following section, we will make use of the analytical findings in this section to obtain various performance metrics in $K$-tier HCNs under general settings. In the light of these metrics, we will also gain insights into the impact of network densification on the outage and rate performance of HCNs under specific association policies including both biased and non-biased BS selection strategies.

V. OUTAGE AND RATE PERFORMANCE OF $K$-TIER HCNs

In this section, we will derive the performance bounds on the HCN capacity metrics, i.e., outage capacity, ergodic capacity and ASE, under two specific association policies: (i) a generic association policy and (ii) biased average received signal strength (BARSS) association policy. However, it should be noted that the analytical approach developed below is general enough for any association policy that preserves the Poisson distribution property for BS locations given the information of $X^\ast$. The validity of our analytical results will be numerically illustrated by simulations at the end of the current section.

A. Generic Association Policy

We start our discussion with the generic association policy. The generic association policy is the policy under which the test user is connected to a BS in tier-$k$ at a (deterministic) distance $r > 0$, and the locations of the rest of the (interfering) BSs in each tier form a non-homogeneous PPP over $\mathbb{R}^2 \setminus B(\mathbf{x}^{(o)}, d_i)$ with mean measure satisfying the above functional form $\Lambda_i \circ T^{-1} (S) = \lambda_i \int_S \mu_i(t) dt$ given this connection information for $i = 1, \ldots, K$, where $B(\mathbf{x}^{(o)}, d_i)$ is the planar ball centered at $\mathbf{x}^{(o)}$ with radius $d_i \geq 0$. $B(\mathbf{x}^{(o)}, d_i)$ can be thought to signify an exclusion region around the test user due to operation of the HCN network protocol stack.

The study of the generic association policy, which may seem a little artificial at first sight, will set the stage for us to analyze both outage and ergodic capacity performance of the BARSS association policy later in this section. The following lemma establishes the Gaussian approximation bounds for the distribution of (standardized) aggregate interference $I_\lambda$ at the test user under the generic association policy by specializing Theorem 1 to this case.

Lemma 3: Under the generic association policy described above,

$$\Pr \left\{ \frac{I_\lambda - \mathbb{E}[I_\lambda]}{\sqrt{\text{Var}[I_\lambda]}} \leq x \right\} - \Psi(x) \leq \Xi \cdot c(x)$$
for all \( x \in \mathbb{R} \), where \( \Psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt \) which is the standard normal CDF, \( c(x) = \min\left(0.4785, \frac{31.935}{1+|x|} \right) \), and 
\[
\Xi = \sum_{i=1}^{K} \frac{\lambda_i P(\mathcal{m}(i) \mu_i(t) dt)}{\sum_{i=1}^{K} \lambda_i P(\mathcal{m}(i) \mu_i(t) dt)}.
\]

Since the outage and ergodic capacity metrics given in Definitions 2 and 5 heavily depend on the level of AWI at the test user, the above Gaussian approximation bound plays a key role to obtain performance upper and lower bounds on the outage and rate performance of HCNs. Below, we start with the outage capacity metric.

**Theorem 3:** Let \( \zeta_k(h, \tau, r) = \frac{P_k(h G_k(c) - \text{SNR}_k^{-1}) \text{PG - E}[I_k]}{\sqrt{\text{Var}[I_k]}} \). Then, \( \Pr(\tau \text{-outage}) \) under the generic association policy is bounded above and below as

\[
1 - E[V^+_k(H_k, \tau, r)] \leq \Pr(\tau \text{-outage}) \leq 1 - E[V^-_k(H_k, \tau, r)],
\]

where \( H_k \) is a generic random variable with PDF \( q_k \), and the functions \( V^+_k \) and \( V^-_k \) are given as

\[
V^+_k(h, \tau, r) = \min \left\{ 1, \Psi(\zeta_k(h, \tau, r)) + \Xi \cdot c(\zeta_k(h, \tau, r)) \right\} 1_{\left\{ h \geq \frac{\text{SNR}_k^{-1}(e^{\tau}-1)}{\mu_k(r)} \right\}}
\]

and

\[
V^-_k(h, \tau, r) = \max \left\{ 0, \Psi(\zeta_k(h, \tau, r)) - \Xi \cdot c(\zeta_k(h, \tau, r)) \right\} 1_{\left\{ h \geq \frac{\text{SNR}_k^{-1}(e^{\tau}-1)}{\mu_k(r)} \right\}}.
\]

**Proof:** Please see Appendix C.

Using the bounds on \( \Pr(\tau \text{-outage}) \), we can bound \( C_o(\gamma) \) for the generic association policy as below.

**Theorem 4:** \( C_o(\gamma) \) under the generic association policy is bounded above and below as

\[
C_o(\gamma) \leq \sup \left\{ \tau \geq 0 : 1 - E[V^+_k(H_k, \tau, r)] \leq \gamma \right\}
\]

and

\[
C_o(\gamma) \geq \sup \left\{ \tau \geq 0 : 1 - E[V^-_k(H_k, \tau, r)] \leq \gamma \right\}.
\]

**Proof:** The proof follows from that the upper (lower) bound on the outage probability crosses the target outage probability \( \gamma \) earlier (later) than \( \Pr(\tau \text{-outage}) \) as \( \tau \) increases.

An important high level perspective on the detrimental effects of the network interference on the HCN outage performance can be obtained if we study the outage capacity bounds given in Theorem 4 as a function of \( \lambda \), i.e., for homogeneous PPPs. At each fading state \( H_k = h \), it can be shown that the outage capacity scales with the BS intensity parameters according to \( \Theta(\|\lambda\|_2^{-1}) \) as \( \|\lambda\|_2 \) grows to infinity. This observation is different from the scale-invariance property of SNR statistics with BS intensity observed in some previous work such as [10], [11], [25], [35]. The main reason is that the increase in \( I_A \) with denser HCN deployments cannot be counterbalanced by an increase in the received power levels for bounded path-loss models. From an HCN design perspective, this result implies that
it is imperative to set BS intensities at each tier appropriately for the proper delivery of data services with minimum required QoS to the end users.

Unlike the outage capacity metric, the ergodic capacity is more suitable for delay insensitive data traffic and obtained by averaging instantaneous data rates over long time intervals. In the following theorem, the bounds on the ergodic capacity are given, where no channel state information (CSI) is assumed at the transmitter side.

**Theorem 5:** \( C_{\text{erg}} \) under the generic association policy is bounded above and below as

\[
\int_0^\infty E \left[ V_k^- (H_k, \tau, r) \right] d\tau \leq C_{\text{erg}} \leq \int_0^\infty E \left[ V_k^+ (H_k, \tau, r) \right] d\tau.
\]

**Proof:** Please see Appendix D.

**B. BARSS Association Policy**

Now, we study the HCN outage and rate performance under the BARSS association policy, in which the test user associates itself to the BS \( X^* \) given by

\[
X^* = \arg \max_{X \in \Phi} \beta_X P_X G X (\|X\|_2),
\]

where it is understood that \( \beta_X = \beta_k \) if \( X \in \Phi_k \). Consider the event \( E_k(r) \) that \( A^* = k \) and \( R^* = r \), i.e., \( E_k(r) \) is the event that the test user is associated with a tier-\( k \) BS at a distance \( r \) under the BARSS association policy. Then, the locations of BSs in tier-\( i \) form a non-homogeneous PPP over \( \mathbb{R}^2 \setminus B \left( x^{(o)}, Q_i^k(r) \right) \) given the event \( E_k(r) \) for \( i = 1, \ldots, K \), where \( Q_i^k(r) = G_i^{-1} \left( \frac{\beta_i P_i}{\beta_k P_k} G_k(r) \right) \), and \( G_i^{-1}(y) = \inf \{ x \geq 0 : G_i(x) = y \} \) if \( y \in [0, G_i(0)] \) and zero otherwise. This observation puts us back into the generic association policy framework, and the derivation of the bounds for the conditional outage probability/capacity and ergodic capacity on the conditioned event \( E_k(r) \) proceeds as before. Averaging over the event \( E_k(r) \), we obtain bounds on the unconditional outage probability and capacity metrics. To this end, we need the following lemmas.

**Lemma 4:** Under the BARSS association policy described above, for all \( x \in \mathbb{R} \),

\[
\Pr \left\{ \frac{I_X - E[I_X]}{\sqrt{\text{Var}[I_X]}} \leq x \mid E_k(r) \right\} - \Psi(x) \leq \Xi_k(r) \cdot c(x),
\]

where \( \Xi_k(r) = \sum_{i=1}^K \frac{\lambda_i P_i^2 m_i(t) \int_0^\infty \frac{G_i^2(t) \mu_i(t)}{Q_i^k(r)} dt}{\sum_{i=1}^K \lambda_i P_i^2 m_i(t) \int_0^\infty \frac{G_i^2(t) \mu_i(t)}{Q_i^k(r)} dt} \), and \( c(x) \) and \( \Psi(x) \) are as defined in Lemma 3.

We note that this is almost the same result as in Lemma 3 except for a small change in the definition of the constant \( \Xi \) to show its dependence on the conditioned event \( E_k(r) \). In order to achieve averaging
over the event \( E_k(r) \), we need to know the connection probability to a tier-\( k \) BS and the conditional PDF of the connection distance given that the test user is associated with a tier-\( k \) BS. To this end, we first obtain the connection probability to a tier-\( k \) BS, which we denote by \( p^*_k \equiv \Pr \{ A^* = k \} \), in the next lemma for non-homogeneous PPPs.

**Lemma 5:** The probability that the test user is associated with tier-\( k \) is

\[
p^*_k = \int_0^\infty \prod_{i=1 \atop i \neq k}^K \exp \left( -\lambda_i \left( B \left( x^{(o)}, Q_i^{(k)}(u) \right) \right) \right) f_{R_k}(u) du \tag{5}
\]

where \( Q_i^{(k)}(u) = G_i^{-1} \left( \frac{\beta_k p_k}{\beta_i p_i} G_i(u) \right) \), \( \beta_k \) is the biasing factor for tier-\( k \) BSs and \( \Lambda_k(\cdot) \) is the mean measure of \( \Phi_k \), \( B \left( x^{(o)}, Q_i^{(k)}(u) \right) \) is the ball in \( \mathbb{R}^2 \) centered at \( x^{(o)} \) with radius \( Q_i^{(k)}(u) \), and \( f_{R_k}(u) = e^{\Lambda_k \left( B \left( x^{(o)}, u \right) \right)} \frac{d}{du} \Lambda_k \left( B \left( x^{(o)}, u \right) \right) \) is the nearest neighbour distance distribution for tier-\( k \).

**Proof:** Please see Appendix E. \( \blacksquare \)

In the following lemma, we focus on an important special case where BS locations follow a homogeneous PPP in each tier, which leads to Lemma 6 below.

**Lemma 6:** Let \( a_0 = 0, a_{K+1} = +\infty \) and \( a_i = \frac{\beta_i p_i}{\beta_k p_k} G_i(0) \) for \( i \in \{1, \ldots, K\} \setminus \{k\} \). Let \( \pi(i) \) be an enumeration of \( a_i \)’s in descending order, i.e., \( a_{\pi(i)} \geq a_{\pi(i+1)} \) for \( i = 0, \ldots, K-1 \). Let \( r_i = G_k^{-1} \left( a_{\pi(i)} \right) \) for \( i = 0, \ldots, K \). Then, \( p^*_k \) is given by

\[
p^*_k = 2\pi \lambda_k \sum_{j=1}^K \int_{r_{j-1}}^{r_j} u \exp \left( -\pi \left( \lambda_k u^2 + \sum_{i=1}^{j-1} \lambda_{\pi(i)} \left( Q_{\pi(i)}^{(k)}(u) \right)^2 \right) \right) du. \tag{6}
\]

**Proof:** Please see Appendix F. \( \blacksquare \)

Several important remarks are in order regarding Lemma 6. The integration in (6) is with respect to the nearest neighbor distance distribution for tier-\( k \) to which the test user is associated. Hence, the BSs in some tiers are inactive with regard to contributing to the association probability for different ranges of the nearest distance from \( \Phi_k \) to the origin, which is why we divide the integration limits into disjoint intervals from \( r_{j-1} \) to \( r_j \) for \( j = 1, \ldots, K \). This behavior is different than that observed in (11), which is again a manifestation of the bounded nature of the path-loss model. In the next lemma, we derive the PDF of \( R^* \) given \( A^* = k \) for a non-homogeneous PPP in each tier.

**Lemma 7:** The PDF of \( R^* \) given \( A^* = k \) is

\[
f_k(u) = \frac{1}{p^*_k} \prod_{i=1 \atop i \neq k}^K \exp \left( -\lambda_i \left( B \left( x^{(o)}, Q_i^{(k)}(u) \right) \right) \right) f_{R_k}(u) \tag{7}
\]

where \( f_{R_k}(u) \) is as given in Lemma 5.

**Proof:** Please see Appendix G. \( \blacksquare \)

In the next lemma, we again consider the important special case of homogeneous PPPs.
Lemma 8: Let \( a_0 = 0, a_{K+1} = +\infty \) and \( a_i = \frac{\beta_i P_i}{\beta_i P_k} G_1(0) \) for \( i \in \{1, \ldots, K\} \setminus \{k\} \). Let \( \pi(i) \) be an enumeration of \( a_i \)'s in descending order, i.e., \( a_{\pi(i)} \geq a_{\pi(i+1)} \) for \( i = 0, \ldots, K-1 \). Let \( r_i = G_k^{-1}(a_{\pi(i)}) \) for \( i = 0, \ldots, K \). Then, the conditional PDF \( f_k(u) \) of \( R^k \) given \( A^* = k \) is given as

\[
f_k(u) = \frac{2\pi \lambda_k}{p_k} \sum_{j=1}^{K} u \exp \left( -\pi \left( \lambda_k u^2 + \frac{1}{2} \sum_{i=1}^{j-1} \lambda_{\pi(i)} \left( \frac{Q_{\pi(i)}^k(u)}{Q_{\pi(i)}^k(0)} \right)^2 \right) \right) 1_{\{u \in [r_{j-1}, r_j]\}}.
\]

Proof: Please see Appendix [H]. \(\blacksquare\)

The conditional connection PDF \( f_k(u) \) given in (8) can be significantly simplified for small numbers of tiers. A reduced expression for one particular but important case of a two-tier HCN is given by the following corollary.

Corollary 2: Assume \( K = 2, \beta_1 P_1 G_1(0) \leq \beta_2 P_2 G_2(0) \) and \( u^* = G_2^{-1} \left( \frac{\beta_1 P_1}{\beta_2 P_2} G_1(0) \right) \). Then,

\[
f_1(u) = \frac{2\pi \lambda_1}{p_1} u \exp \left( -\pi \left( \lambda_1 u^2 + \lambda_2 \left( \frac{Q_2^1(u)}{Q_2^1(0)} \right)^2 \right) \right) 1_{\{u \geq 0\}}
\]

and

\[
f_2(u) = \frac{2\pi \lambda_2}{p_2} u \exp \left( -\pi \lambda_2 u^2 \right) 1_{\{u < u^*\}} + \frac{2\pi \lambda_2}{p_2} u \exp \left( -\pi \left( \lambda_2 u^2 + \lambda_1 \left( \frac{Q_1^2(u)}{Q_1^2(0)} \right)^2 \right) \right) 1_{\{u \geq u^*\}}.
\]

Using these preliminary results, the performance bounds on the outage probability, outage capacity, ergodic capacity and ASE under the BARSS association policy are given in theorems below.

Theorem 6: Let \( V_k^\pm(h, \tau, r) \) be defined as in (3) and (4), respectively, by replacing \( \Xi \) by \( \Xi_k(r) \) given in Lemma 4. Then, \( \Pr \{ \tau\text{-outage} \} \) under the BARSS association policy is bounded below and above as

\[
\Pr \{ \tau\text{-outage} \} \geq 1 - \sum_{k=1}^{K} p_k^* \int_0^\infty f_k(r) \mathbb{E} \left[ \hat{V}_k^+(H_k, \tau, r) \right] dr
\]

and

\[
\Pr \{ \tau\text{-outage} \} \leq 1 - \sum_{k=1}^{K} p_k^* \int_0^\infty f_k(r) \mathbb{E} \left[ \hat{V}_k^-(H_k, \tau, r) \right] dr.
\]

Proof: The proof follows from calculating these bounds for \( \Pr \{ \tau\text{-outage} \mid E_k(r) \} \) using Theorem 3 and then averaging them by using (5) and (7). \(\blacksquare\)

Theorem 7: \( C_0(\gamma) \) under the BARSS association policy is bounded above and below as

\[
C_0(\gamma) \leq \sup \left\{ \tau \geq 0 : 1 - \sum_{k=1}^{K} p_k^* \int_0^\infty f_k(r) \mathbb{E} \left[ \hat{V}_k^+(H_k, \tau, r) \right] dr \leq \gamma \right\}
\]

and

\[
C_0(\gamma) \geq \sup \left\{ \tau \geq 0 : 1 - \sum_{k=1}^{K} p_k^* \int_0^\infty f_k(r) \mathbb{E} \left[ \hat{V}_k^-(H_k, \tau, r) \right] dr \leq \gamma \right\}.
\]
**Proof:** The proof follows from that the upper (lower) bound on the outage probability crosses the target outage probability $\gamma$ earlier (later) than $\Pr(\tau\text{-outage})$ as $\tau$ increases.

Similar to Theorem 5, the performance bounds for achievable ergodic capacity are given in the next theorem.

*Theorem 8:* The ergodic capacity achievable under the BARSS association policy is bounded above and below as

$$C_{\text{erg}} \leq \sum_{k=1}^{K} p_k^* \int_0^\infty \int_0^\infty E[\hat{V}_k^+(H, \tau, r)] d\tau f_k(r) dr$$

and

$$C_{\text{erg}} \geq \sum_{k=1}^{K} p_k^* \int_0^\infty \int_0^\infty E[\hat{V}_k^-(H, \tau, r)] d\tau f_k(r) dr$$

where $\hat{V}_k^+(H_k, \tau, r)$ and $\hat{V}_k^-(H_k, \tau, r)$ are defined as in Theorem 6.

*Proof:* The ergodic capacity can be written as $C_{\text{erg}} = \sum_{k=1}^{K} p_k^* \int_0^\infty C_{\text{erg}}(k, r) f_k(r) dr$, where $C_{\text{erg}}(k, r)$ is the conditional ergodic capacity given that the user is connected to the BS located in tier-$k$ at a distance $r$ from the test user under the BARSS association policy. Then, similar to Appendix D, the proof follows from calculating these bounds for $1 - \Pr\{\tau\text{-outage} \mid E_k(r)\}$ using Theorem 8 and then averaging them by using (5) and (7) for non-homogeneous PPPs, or by using (6) and (8) for homogeneous PPPs, respectively.

Available spectrum for wireless networks is often very limited. Hence, it is of prime importance to investigate the ASE of a multi-tier HCN defined as the sum of the maximum bit rates per second per hertz per unit area. In the next theorem, we provide upper and lower bounds for ASE in a multi-tier HCN, which requires the calculation of conditional outage capacity in each tier with individual target outage probability $\gamma_k$ employed to provide flexibility of setting a balance between link reliability and rate. We only consider homogeneous PPPs in Theorem 9 to avoid technical complexities arising in the case of non-homogeneous PPPs. Further, since ASE characterizes the collective network performance, rather than the one observed at a specific point as above, it is assumed that all the BSs are serving a user.

*Theorem 9:* ASE under the BARSS association policy is bounded above and below as

$$\text{ASE}(\lambda, \gamma) \leq \sum_{k=1}^{K} \lambda_k (1 - \gamma_k) C_{o}^+(k, \gamma_k)$$

and

$$\text{ASE}(\lambda, \gamma) \geq \sum_{k=1}^{K} \lambda_k (1 - \gamma_k) C_{o}^-(k, \gamma_k)$$

where $C_{o}^+(k, \gamma_k) \overset{\Delta}{=} \sup \{\tau \geq 0 : \rho_k^+ \leq \gamma_k\}$ and $C_{o}^-(k, \gamma_k) \overset{\Delta}{=} \sup \{\tau \geq 0 : \rho_k^- \leq \gamma_k\}$ are conditional outage capacity upper and lower bounds in tier-$k$, $\lambda = [\lambda_1, \ldots, \lambda_K]^T$, $\gamma = [\gamma_1, \ldots, \gamma_K]^T$, and $\rho_k^+ \overset{\Delta}{=} \rho_k^-$.
\[1 - \int_0^\infty f_k(r) \mathbb{E} \left[ \tilde{V}_k^+ \left( H_k, \tau, r \right) \right] dr\] and \(\rho_k^- = 1 - \int_0^\infty f_k(r) \mathbb{E} \left[ \tilde{V}_k^- \left( H_k, \tau, r \right) \right] dr\) are conditional outage probabilities given that the test user is connected to a tier-\(k\) BS, \(\tilde{V}_k^+ \left( H_k, \tau, r \right)\) and \(\tilde{V}_k^- \left( H_k, \tau, r \right)\) are given as in Theorem 6 specialized to the homogeneous PPPs, and \(f_k(r)\) is given in (8).

\[\text{Proof:}\] The proof easily follows from Theorems 6 and 7.

C. Numerical Results

In this part, we present our simulation results illustrating the upper and lower bounds on the HCN capacity metrics derived in Section V. In particular, we will investigate \(C_o(\gamma), C_{\text{erg}}\) and ASE \((\lambda, \gamma)\) under the BARSS association policy. \(N_0\) is set to zero and all fading coefficients are independently drawn from Nakagami-\(m\) distribution with unit mean power gain and \(m = 5\). The path-loss function is taken to be \(G(x) = \frac{1}{1+x^\alpha}\) for all tiers. The transmission powers are set as \(P_1 = 10P_2 = 50P_3\), while we set BS locations according to homogeneous PPPs with intensities given as \(\lambda_1 = 0.1\kappa, \lambda_2 = \kappa\) and \(\lambda_3 = 5\kappa\). Here, \(\kappa\) is our (unitless) control parameter to control the average number of BSs per unit area. For the 2-tier scenario, only \(\{P_k, \lambda_k\}_{k=1}^2\) are considered. The target outage probability is 0.15 for Fig. 2 and PG is set to 25 for both figures. For the sake of simplicity, we assume that network layer queues at BSs are fully-loaded, an extension of which to the lightly loaded case will be considered as a future work [40].

We plot the bounds in Theorem 7 on \(C_o(\gamma)\) for 2-tier and 3-tier HCNs as a function of \(\kappa\) in Fig. 2. Two different values of \(\alpha\) are used. As this figure shows, both upper and lower bounds approximate \(C_o(\gamma)\) within 0.06 Nats/Sec/Hz for \(\alpha = 2.7\) and within 0.15 Nats/Sec/Hz for \(\alpha = 3.3\) in the 2-tier scenario. They are tighter for the 3-tier scenario due to denser HCN deployment. The heuristic rate curve, which is the arithmetic average of the upper and lower bounds, almost perfectly track \(C_o(\gamma)\) for all cases considered in Fig. 2.

An interesting observation is the monotonically decreasing nature of \(C_o(\gamma)\) with \(\kappa\). This is in accordance with the discussion on the \(\Theta(\|\lambda\|_2^{-1})\)-type scaling behavior of outage capacity in Section V. Hence, we cannot improve the downlink data rates indefinitely in an HCN by adding more BS infrastructure. We must either mitigate interference more efficiently or find the optimum BS intensity per tier maximizing delivered data rates per unit area.

We plot \(C_{\text{erg}}\) and the corresponding bounds given in Theorem 8 for the 2-tier HCN scenario as a function of \(\kappa\) in Fig. 3. We set path-loss exponent \(\alpha\) to 3. For moderate values of \(\kappa\), we observe that while our upper bound approximates \(C_{\text{erg}}\) within one Nats/Sec/Hz, the lower bound is closer than 0.3 Nats/Sec/Hz. Our bounds become very close to the simulated rate for moderate to high values of \(\kappa\), and especially the gaps among the bounds and the \(C_{\text{erg}}\) become negligibly small for high values of \(\kappa\).
Finally, Fig. 4 depicts the changes of ASE for a 2-tier HCN scenario as a function of $\kappa$, where BS intensities are set as $\lambda_1 = 0.1$ and $\lambda_2 = \kappa$. Note that we fix $\lambda_1$ at 0.1 to see how ASE changes by adding more unplanned infrastructure to the network as $\lambda_2$ grows. Biasing coefficient for tier-1, i.e., $\beta_1$, is assumed to be unity while the biasing value $\beta_2$ for tier-2 is set to 3. Target outage probabilities in both tiers are assumed to be identical, i.e., $\gamma_k = \gamma$, and taken as 0.15 for maintaining the same link reliability in each tier.

We firstly observe that our upper and lower bounds again accurately characterize the simulated ASE curve. Further, the heuristic ASE curve, which is taken to be the arithmetic average of the upper and lower bounds, almost perfectly tracks the simulated ASE. We note three different forms of behavior
of ASE as a function of $\kappa$. For sparse to moderate values of $\kappa$, i.e., between 1 and 6, ASE increases rapidly due to efficient utilization of the spectrum which is made available by tier-2 BSs. For $\kappa$ from moderate to high values, i.e., between 6 and 12, improvement in ASE starts to slow down due to the pressure of growing AWI. For $\kappa$ values higher than 12, ASE almost stops improving due to the overwhelming growth in AWI. Especially, the contribution of macro BSs to the overall ASE is much more negatively affected than that of micro BSs by the elevated AWI levels in the dense HCN regime, as illustrated by the macro and micro contribution curves. These observations point out that significant ASE gains can be achievable through biasing along with applying an efficient interference management and suppression in hyper-dense multi-tier HCNs. Otherwise, the network density should be carefully adjusted to maintain target ASE.

VI. CONCLUSIONS AND FUTURE WORK

This paper has examined various performance metrics of interest in downlink $K$-tier HCNs under general settings including general PPPs by introducing a principled methodology. To this end, we have first investigated the Gaussian approximation for the AWI distribution in the downlink of $K$-tier HCNs under a general set-up. Analytical bounds measuring the Kolmogorov-Smirnov distance between these two distributions have been obtained. The derived bounds have also been illustrated numerically through simulations of a particular three-tier HCN scenario. A good statistical fit between the simulated (centralized and normalized) AWI distribution and the standard normal distribution has been observed even for moderate values of BS intensities.

Secondly, we have examined the downlink capacity metrics of $K$-tier HCNs under general signal propagation models, allowing for the use of general bounded path-loss functions, arbitrary fading distributions and general PPPs. Tight upper and lower bounds on the outage capacity, ergodic capacity and area spectral efficiency have been obtained for two specific association policies - the generic association policy and the BARSS association policy. The validity of our analytical results has also been confirmed by simulations. The proposed approach can be extended to other association policies, and has the potential of understanding the HCN performance and design beyond specific selections of the path-loss model and the fading distribution. Utilizing these results, the future plans of the authors include the development of novel spectrum management techniques for HCNs, an investigation of multi-slope path-loss models for emerging millimeter wave communications under general settings, and an investigation of provably near-optimum network control mechanisms such as optimum hybrid-access and cell-range expansion.
APPENDIX A

AUXILIARY LEMMAS FOR THE PROOF OF THEOREM 1

In this appendix, we will provide five lemmas to construct the proof of Theorem 1 in the next appendix. We start our analysis by showing that AWI has a non-degenerate probability distribution function. By using Laplace functionals of Poisson processes (i.e., please refer to [37] for details), we can find the Laplace transform for $I_\Lambda$ as follows:

$$L_{I_\Lambda}(s) = \mathbb{E} [e^{-sI_\Lambda}] = \prod_{k=1}^{K} \exp \left( -\lambda_k \int_0^{\infty} \int_0^{\infty} \left( 1 - e^{-sP_k h G_k(t)} \right) \mu_k(t) q_k(h) \, dt \, dh \right),$$

where $s \geq 0$. The following lemma establishes that $I_\Lambda$ is of a non-degenerate distribution.

**Lemma 9:** For all $s \geq 0$, $\int_0^{\infty} \int_0^{\infty} \left( 1 - e^{-sP_k h G_k(t)} \right) \mu_k(t) q_k(h) \, dt \, dh < \infty$.

**Proof:** Recall that $G_k(t) = O(t^{-\alpha_k})$ as $t \to \infty$. Hence, we can find constants $B_1 > 0$ and $\vartheta > 0$ such that $G_k(t) \leq \vartheta t^{-\alpha_k}$ for all $t \geq B_1$. This implies that

$$\int_0^{\infty} \int_0^{\infty} \left( 1 - e^{-sP_k h G_k(t)} \right) \mu_k(t) q_k(h) \, dt \, dh \leq \int_0^{\infty} \int_0^{B_1} \left( 1 - e^{-sP_k h G_k(t)} \right) \mu_k(t) q_k(h) \, dt \, dh + \int_0^{\infty} \int_{B_1}^{\infty} \left( 1 - e^{-sP_k h \vartheta^{-\alpha_k}} \right) \mu_k(t) q_k(h) \, dt \, dh.$$

The first integral in the last line in (9) is finite since $\Lambda_k$ is locally finite. To show the finiteness of the second integral, we divide it into two parts as follows:

$$\int_0^{\infty} \int_0^{\infty} \left( 1 - e^{-sP_k h \vartheta^{-\alpha_k}} \right) \mu_k(t) q_k(h) \, dt \, dh =$$

$$\int_0^{\frac{1}{sP_k \vartheta^{-\alpha_k}}} \int_0^{\infty} \left( 1 - e^{-sP_k h \vartheta^{-\alpha_k}} \right) \mu_k(t) q_k(h) \, dt \, dh + \int_{\frac{1}{sP_k \vartheta^{-\alpha_k}}}^{\infty} \int_{B_1}^{\infty} \left( 1 - e^{-sP_k h \vartheta^{-\alpha_k}} \right) \mu_k(t) q_k(h) \, dt \, dh.$$

(10)

The first integral above can be bounded as

$$\int_0^{\frac{1}{sP_k \vartheta^{-\alpha_k}}} \int_0^{\infty} \left( 1 - e^{-sP_k h \vartheta^{-\alpha_k}} \right) \mu_k(t) q_k(h) \, dt \, dh \leq \int_{B_1}^{\infty} \left( 1 - e^{-t^{-\alpha_k}} \right) \mu_k(t) \, dt,$$

which is finite since $1 - e^{-t^{-\alpha_k}} = O(t^{-\alpha_k})$ and $\mu_k(t) = O(t^{\alpha_k-1})$ as $t \to \infty$. Hence, proving the finiteness of $\int_{\frac{1}{sP_k \vartheta^{-\alpha_k}}}^{\infty} \int_{B_1}^{\infty} \left( 1 - e^{-sP_k h \vartheta^{-\alpha_k}} \right) \mu_k(t) q_k(h) \, dt \, dh$ will complete the proof. To this end, we need the following lemma.

**Lemma 10:** $1 - e^{-at^{-\alpha_k}} \leq 2a \left( 1 - e^{-a} \right) t^{-\alpha_k}$ for all $a \geq 1$ and $t$ large enough.
Proof: We let \( f_t(a) = 1 - e^{-at} \) and \( g_t(a) = 2a (1 - e^{-a}) t^{-\alpha_k} \). For \( a = 1 \), we have \( \lim_{t \to \infty} \frac{f_t(1)}{t^{-\alpha_k}} = 1 \) and \( \lim_{t \to \infty} \frac{g_t(1)}{t^{-\alpha_k}} = 2 (1 - e^{-1}) > 1 \). Hence, there exists a constant \( B_2 > 0 \) such that \( g_t(1) > f_t(1) \) for all \( t \geq B_2 \). We now fix an arbitrary \( t \) greater than \( B_2 \). Then,

\[
\frac{df_t(a)}{da} = t^{-\alpha_k} e^{-at} \quad \text{and} \quad \frac{dg_t(a)}{da} = 2t^{-\alpha_k} (1 + ae^{-a} - e^{-a}).
\]

Thus, \( g_t(a) \) grows faster than \( f_t(a) \), implying that \( g_t(a) \geq f_t(a) \) for all \( a \geq 1, t \geq B_2 \).

By using Lemma 10, we can upper bound the second integral in (10) as

\[
\int_{\frac{1}{sP_k^\sigma}}^\infty \int_{B_1}^\infty \left( 1 - e^{-sP_k h t^{-\alpha_k}} \right) \mu_k(t) q_k(h) dt dh 
\leq \int_{B_1}^B \mu_k(t) dt + \int_{B_3}^\infty \int_{\frac{1}{sP_k^\sigma}}^\infty 2sP_k h \left( 1 - e^{-sP_k h \alpha_k} \right) q_k(h) t^{-\alpha_k} \mu_k(t) dh dt
\]

(11)

for some positive constant \( B_3 \) large enough. The first integral in (11) is finite due to local finiteness of \( \Lambda_k \). The second integral in (11) can be upper bounded by \( 2sP_k h m_H^{(k)} \int_{B_3}^\infty t^{-\alpha_k} \mu_k(t) dt \), which is finite since \( m_H^{(k)} < \infty \) and \( \mu_k(t) = O(t^{\alpha_k - 1} - \varepsilon) \) as \( t \to \infty \). This completes the proof of Lemma 9.

The following lemma shows that the probability distribution of \( I_n \) can be approximated by the limit distribution of a sequence of random variables \( I_n \), i.e., \( I_n \xrightarrow{d} I_\lambda \) as \( n \to \infty \).

**Lemma 11:** For each \( n \), let \( U_{1,n}^{(k)}, \ldots, U_{n,n}^{(k)} \) be a sequence of i.i.d. random variables with a common probability density function \( f_k(t) = \frac{\lambda_k}{\Lambda_{n,k}} 1_{0 \leq t \leq a} \) for tier-\( k \), where \( \Lambda_{n,k} = \lambda_k \int_0^a \mu_k(t) dt \) and \( \lfloor \cdot \rfloor \) is the smallest integer greater than or equal to its argument. Let

\[
I_n = \sum_{k=1}^K I_n^{(k)},
\]

(12)

where \( I_n^{(k)} = P_k \sum_{i=1}^{\lfloor \Lambda_{n,k} \rfloor} H_i^{(k)} G_k \left( U_{i,n}^{(k)} \right) \) and \( \left\{ H_i^{(k)} \right\}_{i=1}^{\infty} \) is an i.i.d. collection of random variables with the common probability density function \( q_k(h) \) for \( k = 1, \ldots, K \). Then \( I_n \) converges in distribution to \( I_\lambda \), which is shown as \( I_n \xrightarrow{d} I_\lambda \), as \( n \to \infty \).

**Proof:** It is enough to show that \( \mathcal{L}_{I_n(s)} \) converges to \( \mathcal{L}_{I_\lambda(s)} \) pointwise as \( n \) tends to infinity. Observing that the random variables \( I_n^{(k)} \) for \( k = 1, \ldots, K \) are independent, we can write the Laplace transform of \( I_n \) as

\[
\mathcal{L}_{I_n(s)} = \prod_{k=1}^K E \left[ e^{-st_k^{(k)}} \right] = \prod_{k=1}^K \mathcal{L}_{I_n^{(k)}(s)},
\]

where \( \mathcal{L}_{I_n^{(k)}(s)} \) is the Laplace transform of \( I_n^{(k)} \), which is given by

\[
\mathcal{L}_{I_n^{(k)}(s)} = \left( 1 - \frac{\lambda_k}{\Lambda_{n,k}} \int_0^\infty \int_0^a \left( 1 - e^{-sP_k h G_k(t)} \right) \mu_k(t) q_k(h) dh dt \right)^{\lfloor \Lambda_{n,k} \rfloor}.
\]
As \( n \) grows to infinity, \( \int_0^\infty \int_0^n \left( 1 - e^{-\lambda_k \mu_k(t) h} \right) \mu_k(t) q_k(h) \, dt \, dh \) converges to

\[
\int_0^\infty \int_0^\infty \left( 1 - e^{-\lambda_k \mu_k(t) h} \right) \mu_k(t) q_k(h) \, dt \, dh
\]

and \( \int_0^\infty \int_0^\infty \left( 1 - e^{-\lambda_k \mu_k(t) h} \right) \mu_k(t) q_k(h) \, dt \, dh < \infty \) by Lemma 9. This observation leads to the identity \( \lim_{n \to \infty} \mathcal{L}_{I_n}(s) = \exp \left( -\lambda_k \int_0^\infty \int_0^\infty \left( 1 - e^{-\lambda_k \mu_k(t) h} \right) \mu_k(t) q_k(h) \, dt \, dh \right) \), which is exactly the Laplace transform of the AWI at the test user coming from tier-\( k \) BSs alone. Utilizing this result, we have

\[
\lim_{n \to \infty} \mathcal{L}_{I_n}(s) = \lim_{n \to \infty} \prod_{k=1}^K \mathcal{L}_{I_n(k)}(s)
\]

\[
= \prod_{k=1}^K \lim_{n \to \infty} \mathcal{L}_{I_n(k)}(s)
\]

\[
= \prod_{k=1}^K \exp \left( -\lambda_k \int_0^\infty \int_0^\infty \left( 1 - e^{-\lambda_k \mu_k(t) h} \right) \mu_k(t) q_k(h) \, dt \, dh \right)
\]

\[
= \mathcal{L}_{I_\lambda}(s),
\]

which completes the proof.

The next lemma shows that the mean value and variance of \( I_\lambda \) can also be approximated by the mean value and variance of \( I_n \).

**Lemma 12:** Let \( I_n \) be defined as in (12). Then,

\[
\lim_{n \to \infty} E[I_n] = E[I_\lambda]
\]

and

\[
\lim_{n \to \infty} \text{Var}[I_n] = \text{Var}[I_\lambda]
\]

**Proof:** Using Campbell’s Theorem [37], we can express \( E[I_\lambda] \) and \( \text{Var}[I_\lambda] \) as

\[
E[I_\lambda] = \sum_{k=1}^K \lambda_k P_k m_{\lambda k} \int_0^\infty G_k(t) \mu_k(t) dt
\]

and

\[
\text{Var}[I_\lambda] = \sum_{k=1}^K \lambda_k P_k^2 m_{\lambda k}^2 \int_0^\infty G_k^2(t) \mu_k(t) dt.
\]

We note that our modeling assumptions ensure that \( E[I_{\lambda k}] \) and \( \text{Var}[I_{\lambda k}] \) are both finite numbers. Let the random variables \( U_{i,n}^{(k)} \), \( H_i^{(k)} \) and \( I_n^{(k)} \) be as defined in Lemma 11. Further, let \( m_{i,n}^{(k)} = E[P_k H_i^{(k)} G_k(U_{i,n}^{(k)})] \) and \( \sigma_{i,n}^{(k)} = \sqrt{\text{Var}[P_k H_i^{(k)} G_k(U_{i,n}^{(k)})]} \). We first observe that

\[
E[I_{n}^{(k)}] = \left[ \Lambda_{n,k} \right] m_{i,n}^{(k)} \quad \text{and} \quad \text{Var}[I_{n}^{(k)}] = \left[ \Lambda_{n,k} \right] \left( \sigma_{i,n}^{(k)} \right)^2.
\]
Furthermore, we can express \( m_{i,n}^{(k)} \) as \( m_{i,n}^{(k)} = \lambda_k P_k m_{H}^{(k)} \int_0^\infty G_k(t) \mu_k(t) dt \), which implies that \( \lim_{n \to \infty} \mathbb{E} \left[ I_n^{(k)} \right] = \lambda_k P_k m_{H}^{(k)} \int_0^\infty G_k(t) \mu_k(t) dt \). Using this result, we have

\[
\lim_{n \to \infty} \mathbb{E} \left[ I_n \right] = \sum_{k=1}^K \lim_{n \to \infty} \mathbb{E} \left[ I_n^{(k)} \right]
\]

\[
= \sum_{k=1}^K \lambda_k P_k m_{H}^{(k)} \int_0^\infty G_k(t) \mu_k(t) dt
\]

\[
= \mathbb{E} \left[ I_\lambda \right].
\]

Repeating the similar steps and using the identity

\[
\left( \sigma_i^{(k)} \right)^2 = \frac{\lambda_k P_k^2 m_{H}^{(k)} \int_0^n G_k^2(t) \mu_k(t) dt}{\Lambda_{n,k}} - \frac{\lambda_k^2 P_k^2 \left( m_{H}^{(k)} \right)^2}{\Lambda_{n,k}^2} \left( \int_0^n G_k(t) \mu_k(t) dt \right)^2,
\]

we also obtain \( \lim_{n \to \infty} \text{Var} \left[ I_n \right] = \text{Var} \left[ I_\lambda \right] \).

**Lemma 13:** Let \( \xi_1, \ldots, \xi_m \) be a sequence of independent and real-valued random variables such that \( \mathbb{E} \left[ \xi_i \right] = 0 \) and \( \sum_{i=1}^m \mathbb{E} \left[ \xi_i^2 \right] = 1 \). Let \( \chi = \sum_{i=1}^m \mathbb{E} \left[ |\xi_i^3| \right] \). Then,

\[
\left| \Pr \left\{ \sum_{i=1}^m \xi_i \leq x \right\} - \Psi(x) \right| \leq \chi \min \left( 0.4785, \frac{31.935}{1 + |x|^3} \right)
\]

for all \( x \in \mathbb{R} \).

**Proof:** Please refer to [28].

**APPENDIX B**

**PROOF OF THEOREM**

In this appendix, we provide the proof for our main Gaussian approximation result given in Theorem 1. To this end, we let \( \xi_{i,n}^{(k)} = \frac{P_k H_i^{(k)} G_k \left( U_{i,n}^{(k)} \right) - m_{i,n}^{(k)}}{\sigma_n} \) for \( k = 1, \ldots, K \), \( n \geq 1 \) and \( 1 \leq i \leq \left\lceil \Lambda_{n,k} \right\rceil \), where \( \sigma_n = \sqrt{\text{Var} \left[ I_n \right]} \), and \( I_n, U_{i,n}^{(k)}, \Lambda_{n,k} \) and \( m_{i,n}^{(k)} \) are as defined in Appendix A. We note that \( \mathbb{E} \left[ \xi_{i,n}^{(k)} \right] = 0 \) and \( \sum_{k=1}^K \sum_{i=1}^{\left\lceil \Lambda_{n,k} \right\rceil} \mathbb{E} \left[ \xi_{i,n}^{(k)} \right]^2 = 1 \). Hence, the collection of random variables
\[ \bigcup_{k=1}^{K} \left\{ \xi^{(k)}_{i,n} : i = 1, \ldots, \left[ \Lambda_{n,k} \right] \right\} \] is in the correct form to apply Lemma 13. We need to calculate \( \chi_n = \sum_{k=1}^{K} \sum_{i=1}^{\left[ \Lambda_{n,k} \right]} \left| \xi^{(k)}_{i,n} \right|^3 \) to complete the proof. We can upper bound \( \chi_n \) as

\[
\chi_n \leq \frac{1}{\sigma_n^3} \sum_{k=1}^{K} \left[ \Lambda_{n,k} \right] E \left[ \left| P_k H_1^{(k)} G_k \left( U_{1,n}^{(k)} \right) + m_{1,n}^{(k)} \right|^3 \right]
\]

\[
= \frac{1}{\sigma_n^3} \sum_{k=1}^{K} \left[ \Lambda_{n,k} \right] \left[ P_k^{3} \left( H_1^{(k)} \right)^{3} \left( G_k \left( U_{1,n}^{(k)} \right) \right)^{3} + 3 P_k^{2} \left( H_1^{(k)} \right)^{2} \left( G_k \left( U_{1,n}^{(k)} \right) \right)^{2} m_{1,n}^{(k)} \\
+ 3 P_k H_1^{(k)} G_k \left( U_{1,n}^{(k)} \right) \left( m_{1,n}^{(k)} \right)^2 + \left( m_{1,n}^{(k)} \right)^3 \right] \]

\[
= \frac{1}{\sigma_n^3} \sum_{k=1}^{K} \left[ \Lambda_{n,k} \right] \left( \frac{P_k^{3} m_{H_3}^{(k)} \lambda_k}{\Lambda_{n,k}} \int_{0}^{n} \left( G_k(t) \right)^{3} \mu_k(t)dt + 3 \frac{P_k^{2} m_{H_2}^{(k)} \lambda_k}{\Lambda_{n,k}} \int_{0}^{n} \left( G_k(t) \right)^{2} \mu_k(t)dt \cdot m_{1,n}^{(k)} \\
+ 3 \frac{P_k m_{H_1}^{(k)} \lambda_k}{\Lambda_{n,k}} \int_{0}^{n} G_k(t) \mu_k(t)dt \cdot \left( m_{1,n}^{(k)} \right)^2 + \left( m_{1,n}^{(k)} \right)^3 \right). \]

We note that \( m_{1,n}^{(k)} = o(1) \) and \( \left[ \Lambda_{n,k} \right] \left( m_{1,n}^{(k)} \right)^3 = o(1) \) as \( n \to \infty \), i.e., see the proof of Lemma 12. Furthermore, we know that \( \sigma_n^2 \) converges to \( \text{Var} \left[ I_\lambda \right] \) as \( n \to \infty \) by Lemma 12. Using these results, we have

\[
\limsup_{n \to \infty} \chi_n \leq \frac{1}{(\text{Var} \left[ I_\lambda \right])^{\frac{3}{2}}} \sum_{k=1}^{K} P_k^{3} m_{H_3}^{(k)} \lambda_k \int_{0}^{\infty} \left( G_k(t) \right)^{3} \mu_k(t)dt. \tag{14}
\]

After substituting the expression for \( \text{Var} \left[ I_\lambda \right] \) (i.e., see the proof of Lemma 12) in (14), we obtain

\[
\limsup_{n \to \infty} \chi_n \leq \sum_{k=1}^{K} \frac{\lambda_k P_k^{3} m_{H_3}^{(k)} \int_{0}^{\infty} G_k(t) \mu_k(t)dt}{\left( \sum_{k=1}^{K} \lambda_k P_k^{2} m_{H_2}^{(k)} \int_{0}^{\infty} G_k(t) \mu_k(t)dt \right)^{\frac{3}{2}}}. \tag{15}
\]

By using Lemma 13, we have

\[
\left| \text{Pr} \left\{ \sum_{k=1}^{K} \sum_{i=1}^{\left[ \Lambda_{n,k} \right]} \xi^{(k)}_{i,n} \leq x \right\} - \Psi(x) \right| \leq \chi_n \min \left( 0.4785, \frac{31.935}{1 + |x|^3} \right) \tag{16}
\]

for all \( n \geq 1 \) and \( x \in \mathbb{R} \). Further, Lemmas 11 and 12 imply that

\[
\sum_{k=1}^{K} \sum_{i=1}^{\left[ \Lambda_{n,k} \right]} \xi^{(k)}_{i,n} \xrightarrow{d} \frac{I_\lambda - \mathbb{E}[I_\lambda]}{\sqrt{\text{Var}[I_\lambda]}} \quad \text{as} \quad n \to \infty. \tag{17}
\]
Hence, using \(15\) and taking the \(\lim\sup\) of both sides in \(16\) we have
\[
\begin{align*}
\limsup_{n \to \infty} \left| \frac{1}{n} \sum_{k=1}^{K} \sum_{i=1}^{n} \xi_{i,k} \leq x \right| - \Psi(x) \\
= \left| \Pr \left\{ \frac{I_{\lambda} - E[I_{\lambda}]}{\sqrt{\text{Var}[I_{\lambda}]} \leq x} \right\} - \Psi(x) \right| \\
\leq \sum_{k=1}^{K} \frac{\lambda_k P_k^3 m_h^3 \int_0^\infty G_k^3(t) \mu_k(t) dt}{(\sum_{k=1}^{K} \lambda_k P_k^2 m_h^2 \int_0^\infty G_k^2(t) \mu_k(t) dt)} \min \left\{ \frac{0.4785}{1 + |x|^3}, \frac{31.935}{1 + |x|^3} \right\},
\end{align*}
\]
which completes the proof.

APPENDIX C

THE PROOF OF THEOREM 3

In this appendix, we will provide the proof for Theorem 3 establishing the outage capacity bounds for the generic association policy. Given that the test user is associated to a BS at a distance \(r\) in tier-\(k\), we can express the \(\tau\)-outage probability as
\[
\Pr(\tau\text{-outage}) = \Pr \{ \log (1 + \text{SNR}_k) < \tau \} \\
= 1 - \int_{\text{SNR}_k^{-1}(\tau-\epsilon)}^{\infty} \Pr \left\{ I_{\lambda} \leq P_k \left( \frac{h G_k(r)}{e^\tau - 1} - \text{SNR}_k^{-1} \right) P_G \right\} q_k(h) dh,
\]
where the last equality follows from the fact that \(I_{\lambda}\) is a positive random variable, and we have
\[
P_k \left( \frac{h G_k(r)}{e^\tau - 1} - \text{SNR}_k^{-1} \right) P_G < 0 \text{ if and only if } h < \frac{(\epsilon^\tau - 1)\text{SNR}_k^{-1}}{G_k(r)}.
\]
By using Lemma 3 and the natural bounds 0 and 1 on the probability, we can upper and lower bound \(\Pr(\tau\text{-outage})\) as
\[
\begin{align*}
\Pr(\tau\text{-outage}) &\leq 1 - \int_{\text{SNR}_k^{-1}(\tau-\epsilon)}^{\infty} \max \left\{ 0, \Psi(\zeta_k(h, \tau, r)) - \Xi \cdot c(\zeta_k(h, \tau, r)) \right\} q_k(h) dh \\
&= 1 - \mathbb{E} \left[ \max \left\{ 0, \Psi(\zeta_k(H_k, \tau, r)) - \Xi \cdot c(\zeta_k(H_k, \tau, r)) \right\} 1\{H_k \geq \frac{\text{SNR}_k^{-1}(\tau-\epsilon)}{G_k(r)}\} \right] \\
&= 1 - \mathbb{E} \left[ V_k^-(H_k, \tau, r) \right]
\end{align*}
\]
and
\[
\begin{align*}
\Pr(\tau\text{-outage}) &\geq 1 - \int_{\text{SNR}_k^{-1}(\tau-\epsilon)}^{\infty} \min \left\{ 1, \Psi(\zeta_k(h, \tau, r)) + \Xi \cdot c(\zeta_k(h, \tau, r)) \right\} q_k(h) dh \\
&= 1 - \mathbb{E} \left[ \min \left\{ 1, \Psi(\zeta_k(H_k, \tau, r)) + \Xi \cdot c(\zeta_k(H_k, \tau, r)) \right\} 1\{H_k \geq \frac{\text{SNR}_k^{-1}(\tau-\epsilon)}{G_k(r)}\} \right] \\
&= 1 - \mathbb{E} \left[ V_k^+(H_k, \tau, r) \right],
\end{align*}
\]
where \(\Xi\) and \(c(x)\) are as given in Lemma 3, \(\Psi(x)\) is the standard normal CDF and \(1\{\cdot\}\) is the indicator function.
APPENDIX D

THE PROOF OF THEOREM 5

The achievable ergodic capacity can be simply written as
\[ C_{\text{erg}} = \mathbb{E}_{\text{SINR}_A} [\log (1 + \text{SINR}_A)] = \int_0^\infty 1 - \Pr (\tau\text{-outage}) d\tau, \]
where (a) follows from the fact that \( \log (1 + \text{SINR}_A) \) is a positive random variable and hence its expectation is equal to \( \int_0^\infty \Pr \{ \log (1 + \text{SINR}_A) > \tau \} d\tau \). Using this observation, we can easily obtain the upper and lower bounds on the ergodic capacity as
\[ C_{\text{erg}} \leq \int_0^\infty \mathbb{E} [V_k^+ (H_k, \tau, r)] d\tau \text{ and } C_{\text{erg}} \geq \int_0^\infty \mathbb{E} [V_k^- (H_k, \tau, r)] d\tau \]
by using the steps similar to those in Appendix C.

APPENDIX E

THE PROOF OF LEMMA 5

In this appendix, we will derive the connection probability of the test user to a serving BS in tier-\( k \) under the BARSS association policy, which is denoted by \( p_k^* \overset{\Delta}{=} \Pr \{ A^* = k \} \). Let \( R_i \) be the nearest distance from \( \Phi_i \) to the test user for \( i = 1, \ldots, K \). Then, utilizing the structure of the BARSS association policy, this probability can be written as
\[ p_k^* = \int_0^\infty \prod_{i=1}^K \Pr \{ \beta_i P_i G_i (R_i) \leq \beta_k P_k G_k (u) \mid R_k = u \} f_{R_k}(u) du \]
\[ \overset{(b)}{=} \int_0^\infty \prod_{i=1}^K \Pr \{ \beta_i P_i G_i (R_i) \leq \beta_k P_k G_k (u) \} f_{R_k}(u) du, \quad (18) \]
where the identity (a) follows from the conditional independence of the events
\[ \{ \beta_i P_i G_i (R_i) \leq \beta_k P_k G_k (R_k) \} \text{ for } i \in \{1, \ldots, K\} \setminus \{k\} \]
given any particular realization of \( R_k \), and the identity (b) follows from the independence of the nearest neighbour distances from different tiers. Each probability term in (18) can be written as
\[ \Pr \{ \beta_i P_i G_i (R_i) \leq \beta_k P_k G_k (u) \} = \Pr \left\{ R_i \geq Q_i^{(k)} (u) \right\} = \exp \left( -\Lambda_i \left( \mathbb{B} \left( X^{(0)}, Q_i^{(k)} (u) \right) \right) \right), \quad (19) \]
where the last equality follows from the nearest neighbour distance distribution for \( R_i \). Using (19), we can write \( p_k^* \) as
\[ p_k^* = \int_0^\infty \prod_{i=1}^K \exp \left( -\Lambda_i \left( \mathbb{B} \left( X^{(0)}, Q_i^{(k)} (u) \right) \right) \right) f_{R_k}(u) du. \]
APPENDIX F
THE PROOF OF LEMMA 6

For the case of spatial homogeneous PPPs, we just need to replace \( \Lambda_i \left( \Phi \left( X^{(0)}, Q^{(k)}(u) \right) \right) \) with \( \pi \lambda_{i} \left( Q^{(k)}(u) \right)^{2} \). Without loss of generality, we assume that the location of the test user is at the origin, i.e., \( X^{(0)} = (0,0) \). Thus, using Lemma 5 we can write \( p^*_k \) as

\[
p^*_k = \int_{0}^{\infty} \prod_{i=1}^{K} \exp \left( -\pi \lambda_{i} \left( Q^{(k)}(u) \right)^{2} \right) f_{R_k}(u) du. \tag{20}
\]

We note that exponent in (20) can be written as a summation \( \sum_{i=1, i \neq k}^{K} \lambda_{i} \left( Q^{(k)}(u) \right)^{2} \), and some terms inside the summation may not be active for some particular values of \( u \) if \( G_k(u) \geq \frac{\beta_k P_i}{\beta_k P_k} G_i(0) \). Recalling the definition of \( a_i \leq \frac{\beta_i P_i}{\beta_k P_k} G_i(0) \), we observe that the condition \( G_k(u) \geq \frac{\beta_i P_i}{\beta_k P_k} G_i(0) \) holds if and only if \( u \leq G_k^{-1}(a_j) \). Introducing \( a_0 = 0 \) and \( a_{K+1} = +\infty \) to have the integration limits from 0 to \( \infty \), and enumerating \( a_i \)'s in descending order for \( i \neq k \), we finally arrive the desired result \( p^*_k = \sum_{j=1}^{K} \int_{r_{j-1}}^{r_j} \exp \left( -\pi \sum_{i=1}^{j-1} \lambda_{\pi(i)} \left( Q^{(k)}(u) \right)^{2} \right) f_{R_k}(u) du \), where \( \pi(i) \) is an enumeration of \( a_i \)'s in descending order, i.e., \( a_{\pi(i)} \geq a_{\pi(i+1)} \) for \( i = 0, \ldots, K-1 \) and \( r_i = G_k^{-1}(a_{\pi(i)}) \) for \( i = 0, \ldots, K \).

APPENDIX G
THE PROOF OF LEMMA 7

In this appendix, we will derive the conditional PDF of the connection distance \( R^* \) given the event \( \{ A^* = k \} \). To this end, we will first calculate the conditional CDF of \( R^* \) given \( \{ A^* = k \} \), which will be denoted by \( F_{R^*|\{A^* = k\}}(r) \). Let \( R_i \) be the nearest distance from \( \Phi_i \) to the test user for \( i = 1, \ldots, K \). Then,

\[
F_{R^*|\{A^* = k\}}(r) = \frac{1}{p_k} \text{Pr} \{ R^* \leq r \text{ and } A^* = k \} = \frac{1}{p_k} \int_{0}^{r} \text{Pr} \left\{ \bigcap_{i=1, i \neq k}^{K} \{ \beta_i P_i G_i(R_i) \leq \beta_k P_k G_k \left( R_k \right) \} \bigg| R_k = u \right\} f_{R_k}(u) du. \tag{21}
\]

Using the conditional independence of the events \( \{ \beta_i P_i G_i \left( R_i \right) \leq \beta_k P_k G_k \left( R_k \right) \} \) for \( i \in \{1, \ldots, K\} \setminus \{k\} \) for any given particular realization of \( R_k \) and the independence of the nearest neighbour distances
from different tiers, we can further simplify (21) as

$$F_{R^*|\{A^* = k\}}(r) = \frac{1}{p_k^*} \int_0^r \Pr \left\{ \bigcap_{i=1}^K \{ \beta_i P_i G_i(r_i) \leq \beta_k P_k G_k(u) \} \mid R_k = u \right\} f_{R_k}(u) du$$

$$= \frac{1}{p_k^*} \int_0^r \prod_{i=1}^K \exp \left( -\pi \lambda_i \left( Q_i^{(k)}(r) \right)^2 \right) f_{R_k}(u) du. \quad (22)$$

We obtain the conditional PDF of $R^*$ given $A^* = k$ by differentiating (22) with respect to $r$. This leads to

$$f_k(r) = \frac{1}{p_k^*} \prod_{i \neq k}^K \exp \left( -\Lambda_i \left( B(\mathbf{x}^{(i)}, Q_i^{(k)}(r)) \right) \right) f_{R_k}(r),$$

where $f_{R_k}(r) = e^{-\Lambda_k B(\mathbf{x}^{(o)},r)} \frac{d}{dr} \Lambda_k \left( B(\mathbf{x}^{(o)}, r) \right)$ is the nearest neighbour distance distribution for tier-$k$.

**APPENDIX H**

**THE PROOF OF LEMMA 8**

Similar to Appendix F, the mean measure $\Lambda_i \left( B(\mathbf{x}^{(i)}, Q_i^{(k)}(r)) \right)$ reduces to $\pi \lambda_i \left( Q_i^{(k)}(r) \right)^2$ when only homogeneous PPPs are considered. Without loss of generality, the location of the test user is assumed to be at the origin, i.e., $\mathbf{x}^{(o)} = (0, 0)$. Thus, we can rewrite $f_k(r)$ as

$$f_k(r) = \frac{1}{p_k^*} \exp \left( -\pi \sum_{i=1}^K \lambda_i \left( Q_i^{(k)}(r) \right)^2 \right) f_{R_k}(r) \quad \text{for } r \geq 0. \quad (23)$$

We observe that the summation term appearing in (23) is exactly the same one appeared in (20). Hence, the same enumeration step can be carried out to arrive at the final result

$$f_k(r) = \frac{1}{p_k^*} \sum_{j=1}^K \exp \left( -\pi \sum_{i=1}^{j-1} \lambda_i \left( Q_i^{(k)}(r) \right)^2 \right) f_{R_k}(r) 1_{\{r \in [r_{j-1}, r_j)\}},$$

where $a_0 = 0$, $a_{K+1} = +\infty$, $a_i = \frac{\beta_i P_i}{\beta_k P_k} G_i(0)$ for $i \in \{1, \ldots, K\} \setminus \{k\}$, $\pi(i)$ is an enumeration of $a_i$’s in descending order, i.e., $a_{\pi(i)} \geq a_{\pi(i+1)}$ for $i = 0, \ldots, K-1$ and $r_i = G_k^{-1} \left( a_{\pi(i)} \right)$ for $i = 0, \ldots, K$.

**REFERENCES**

[1] S. Ak, H. Inaltekin and H. V. Poor, “Gaussian approximation for the downlink interference in heterogeneous cellular networks,” in Proc. IEEE International Symposium on Information Theory (ISIT), Barcelona, Spain, Jul. 2016.

[2] S. Ak, H. Inaltekin and H. V. Poor, “Downlink outage performance of heterogeneous cellular networks,” in Proc. IEEE International Symposium on Information Theory (ISIT), Barcelona, Spain, Jul. 2016.
[3] J. Andrews, H. Claussen, M. Dohler, S. Rangan and M. Reed, “Femtocells: Past, present, and future,” IEEE J. Sel. Areas Commun., vol. 30, no. 3, pp. 497-508, Apr. 2012.

[4] 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 Trans. Commun. Mag., vol. 50, no. 6, pp. 54–64, June 2012.

[5] I. Hwang, B. Song and S. S. Soliman, “A holistic view on hyper-dense heterogeneous and small cell networks,” IEEE Commun. Mag., vol. 51, no. 6, pp. 20-27, June 2013.

[6] D. Lopez-Perez, M. Ding, H. Claussen and A. H. Jafari, “Towards 1 Gbps/UE in cellular systems: Understanding ultra-dense small cell deployments,” IEEE Commun. Surveys Tuts., vol. 17, no. 4, pp. 2078–2101, Nov. 2015.

[7] O. Galinina, A. Pyattaev, S. Andreevy, M. Dohler and Y. Koucheryavy, “5G multi-RAT LTE-WiFi ultra-dense small cells: Performance dynamics, architecture, and trends,” IEEE J. Sel. Areas Commun., vol. 33, no. 6, pp. 1224–1240, May 2015.

[8] D. Liu, L. Wang, Y. Chen, M. Elkashlan, K.-K. Wong, R. Schober and L. Hanzo, “User association in 5G networks: A survey and an outlook,” IEEE Commun. Surveys Tuts., to appear, 2016.

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

[10] 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 Commun., vol. 30, no. 3, pp. 550–560, Apr. 2012.

[11] 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.

[12] T. Samarasinghe, H. Inaltekin and J. S. Evans, “On the outage capacity of opportunistic beamforming with random user locations,” IEEE Trans. Commun., vol. 62, no. 8, pp. 3015-3026, Aug. 2014.

[13] 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 Tuts., vol. 15, no. 3, pp. 996–1019, June 2013.

[14] V. H. Macdonald, “Advanced mobile phone service: The cellular concept,” Bell Syst. Tech. J., vol. 58, no. 1, pp.15–41, Jan. 1979.

[15] S. Dep, P. Monogioudis, J. Miernik and J. Seymour, “Algorithms for enhanced inter-cell interference coordination (eICIC) in LTE HetNets,” IEEE/ACM Trans. Networking, vol. 22, no. 1, pp. 137–150, Feb. 2014.

[16] E. Hossain, M. Rasti, H. Tabassum and A. Abdelnasser, “Evolution toward 5G multi-tier cellular wireless networks: An interference management perspective,” IEEE Wireless Commun., vol. 21, no. 3, pp.118–127, Jun. 2014.

[17] M. Haenggi and R. K. Ganti, “Interference in large wireless networks,” Foundations and Trends in Networking, vol. 3, no. 2, pp. 127-248, 2009.

[18] S. Deng, M. K. Samimi and T. S. Rappaport, “28GHz and 73GHz millimeter-wave indoor propagation measurements and path loss models,” in Proc. 2015 ICC Workshop. IEEE, pp. 1244-1250, Jun. 2015.

[19] X. Zhang and J. G. Andrews, “Downlink cellular network analysis with multi-slope path loss models,” IEEE Trans. Commun., vol. 63, no. 5, pp. 1881-1894, May 2015.

[20] E. S. Sousa and J. A. Silvester, “Optimum transmission ranges in a direct-sequence spread-spectrum multihop packet radio network,” IEEE Journal on Selected Areas in Communications, vol. 8, no. 5, pp. 762-771, Jun 1990.

[21] J. Ilow and D. Hatzinakos, “Analytic alpha-stable noise modeling in a Poisson field of interferers or scatterers,” IEEE Trans. Signal Process., vol. 46, no. 6, pp. 1601-1611, June 1998.

[22] C. C. Chan and S. V. Hanly, “Calculating the outage probability in a CDMA network with spatial Poisson traffic,” IEEE Trans. Veh. Technol., vol. 50, no. 1, pp. 183-204, Jan 2001.

[23] S. P. Weber and J. G. Andrews, “Bounds on the SIR distribution for a class of channel models in ad hoc networks,” in Proc. 49th IEEE Global Telecommunications Conference, San Francisco, CA, Nov. 2006.
[24] M. Z. Win, P. C. Pinto and L. A. Shepp, “A mathematical theory of network interference and its applications,” Proc. IEEE, vol. 97, no. 2, pp. 205-230, Feb. 2009.

[25] P. Madhusudhanan, J. G. Restrepo, Y. Liu, T. X. Brown and K. R. Baker, “Downlink performance analysis for a generalized shotgun cellular system,” IEEE Trans. Wireless Commun., vol. 13, no. 12, pp. 6684-6696, Dec. 2014.

[26] T. Samarasinghe, H. Inaltekin and J. S. Evans, “On optimal downlink coverage in Poisson cellular networks with power density constraints,” IEEE Trans. Commun., vol. 62, no. 4, pp. 1382-1392, Apr. 2014.

[27] M. Aljuaid and H. Yanikomeroglu, “Investigating the Gaussian convergence of the distribution of the aggregate interference power in large wireless networks,” IEEE Trans. Veh. Technol., vol. 59, no. 9, pp. 4418-4424, Nov. 2010.

[28] H. Inaltekin, “Gaussian approximation for the wireless multi-access interference distribution,” IEEE Trans. Signal Process., vol. 60, no. 11, pp. 6114-6120, Nov. 2012.

[29] H. Inaltekin, “Gaussian approximation for the wireless multi-access interference distribution and its applications,” Technical Report, Antalya Int. Univ., 2012. [Online]. Available: http://arxiv.org/abs/1201.3812.

[30] M. D. Renzo, A. Guidotti and G. E. Corazza, “Average rate of downlink heterogeneous cellular networks over generalized fading channels: A stochastic geometry approach,” IEEE Trans. Commun., vol. 61, no. 7, pp. 3050-3071, Jul. 2013.

[31] A. Shojaeifard, K. A. Hamdi, E. Alsusa, D. K. C. So and J. Tang, “Exact SINR statistics in the presence of heterogeneous interferers,” IEEE Trans. Inf. Theory, vol. 61, no. 12, pp. 6759-6773, Dec. 2015.

[32] R. W. Heath, M. Kountouris and T. Bai, “Modeling heterogeneous network interference with using Poisson point processes,” IEEE Trans. Signal Process., vol. 61, no. 16, Aug. 2013.

[33] S. Mukherjee, “Distribution of downlink SINR in heterogeneous cellular networks,” IEEE J. Sel. Areas Commun., vol. 30, no. 3, pp. 575-585, Apr. 2012.

[34] T. Samarasinghe, H. Inaltekin and J. S. Evans, “Modeling and analysis of opportunistic beamforming for Poisson wireless networks,” IEEE Trans. Wireless Commun., vol. 15, no. 5, pp. 3732-3745, May 2016.

[35] P. Madhusudhanan, J. G. Restrepo, Y. Liu and T. X. Brown, “Analysis of downlink connectivity models in a heterogeneous cellular network via stochastic geometry,” IEEE Trans. on Wireless Commun., vol. 15, no. 6, pp. 3895-3907, Jun. 2016.

[36] B. Blaszczyszyn and H. P. Keeler, “Studying the SINR process of the typical user in Poisson networks by using its factorial moment measures,” IEEE Trans. Inf. Theory, vol. 61, pp. 6774-6794, Dec. 2015.

[37] J. F. C. Kingman, Poisson Processes, Clarendon Press, Oxford, 1993.

[38] L. H. Y. Chen and Q. M. Shao, “Steins method for normal approximation,” IMS Lecture Notes Series, National University of Singapore, vol. 4, pp. 1-59, Feb. 2005.

[39] H. Inaltekin, M. Chiang, H. V. Poor, and S. B. Wicker, “On unbounded path-loss models: effect of singularity on wireless network performance,” IEEE J. Sel. Areas Commun., vol. 27, pp. 1078-1092, Sep. 2009.

[40] H. Dhillon, R. Ganti and J. G. Andrews, “Load-aware modeling and analysis of heterogeneous cellular networks,” IEEE Trans. Wireless Commun., vol. 12, no. 4, pp. 1666–1677, Apr. 2013.