Traffic Management for Heterogeneous Networks with Opportunistic Unlicensed Spectrum Sharing

Chun-Hung Liu and Hong-Cheng Tsai

Abstract

This paper studies how to maximize the per-user-based throughput in an M-tier heterogeneous wireless network (HetNet) by optimally managing traffic flows among the access points (APs) in the HetNet. The APs in the first $M-1$ tiers can use the licensed spectrum at the same time whereas they share the unlicensed spectrum with the APs in the Mth tier by the proposed opportunistic CSMA/CA protocol. We characterize the statistical property of the cell load and channel access probability of each AP using a general AP association scheme. For an AP in each tier, the tight bounds on its mean spectrum efficiencies in the licensed and unlicensed spectra are derived in a low-complexity form for general random channel gain and AP association weight models and they can give some insights on how channel gains, AP association weights and void AP probabilities affect the mean spectrum efficiencies. We define the per-user link throughput and per-user network throughput based on the derived the mean spectrum efficiencies and maximize them by proposing the decentralized and centralized traffic management schemes for the APs in the first $M-1$ tiers under the constraint that the per-user link throughput of the tier-M APs must be above some minimum required value. Finally, a numerical example of coexisting LTE and WiFi networks is provided to validate our derived results and findings.

I. Introduction

As more and more versatile services are offered over wireless networks and new generations of wireless smart handsets get wider and wider adoption, considerable data traffic flowing over spectrum-limited cellular networks is an inevitable phenomenon the network operators have to seriously face. To alleviate the spectrum crunch crisis of a cellular network, an effective means is to make a traditional

C.-H. Liu and H.-C. Tsai are with the Dept. of Electrical and Computer Engineering, National Chiao Tung University, Hsinchu, Taiwan. The contact author is Dr. Liu. (Email: chungliu@nctu.edu.tw) Manuscript date: September 20, 2018.

cellular network migrate to a heterogeneous cellular network in which many different kinds of base stations (BSs), such as macrocell, micro and small cell BSs, are densely deployed. Although heterogeneous cellular networks have a much higher network capacity compared to their traditional counterparts, their licensed spectrum is still very limited and their per-user link throughput may not be efficiently improved if the network has a huge user population. Accordingly, exploiting more available spectrum for heterogeneous cellular networks is the right track that should be followed, which fosters the idea of extending the service of the cellular BSs to the unlicensed spectrum.

If cellular BSs can access the licensed and unlicensed spectra at the same time, they can integrate all available spectrum resources by using the carrier aggregation technique [1]–[3]. However, extending cellular services to the unlicensed spectrum could severely impact the throughput performance of the existing access points (APs) using the unlicensed spectrum, such as WiFi APs. This coexisting problem in the unlicensed spectrum motivates us to investigate how to make different kinds of BSs and APs properly share the unlicensed spectrum and improve their total throughput is an important problem that needs to be investigated thoroughly. To generally and tractably analyze the throughput performance of a heterogeneous wireless network (HetNet) with unlicensed spectrum sharing, the HetNet considered in this paper has M tiers. All APs in the same tier of the HetNet are of the same type and performance and they follow an independent Poisson point process (PPP) with certain intensity. Specifically, the APs in the Mth tier only access the unlicensed spectrum whereas the APs in the first $M-1$ tiers can simultaneously access the licensed spectrum as well as the unlicensed spectrum if they have a chance. All APs use the proposed (slotted nonpersistant) opportunistic CSMA/CA with random backoff time protocol to contend the channel in the unlicensed spectrum. Such a HetNet model characterizes the coexisting impacts among different kinds of APs in the unlicensed spectrum while the CSMA/CA protocol is adopted.

A. Prior Works on Unlicensed Spectrum Sharing

Earlier studies on the coexisting interference problem in the unlicensed spectrum focused on how to make APs in different overlaid wireless networks share the unlicensed spectrum with certain fairness. For example, reference [4] proposed a game-theoretical approach to fairly sharing the unlicensed spectrum in multiple coexisting and interfering networks. The interference modeling and mitigation problems in the unlicensed spectrum are investigated in references [5]–[7]. These works use a small-scale and deterministic network model to formulate their problems and they do not study some coexisting performance metrics, such as coverage and network throughput. Although the authors in [8] developed a

more accurate interference analysis for large-scale networks based on the continuum field approximation; however, they did not investigate if the interference in the unlicensed spectrum can be effectively mitigated by using channel access protocols.

A more complicate coexistence problem in the unlicensed spectrum that recently attracts a lot of attentions is how to let the base stations (BSs) or APs originally use the licensed spectrum also be able to access the unlicensed spectrum and use the carrier aggregation technique to boost their overall throughput [2], [3]. A few recent works have already shown that LTE and WiFi networks coexisting in the unlicensed spectrum can significantly improve their entire network throughput [2], [3], [9]–[11]. However, how to manage the traffic flows between LTE BSs and WiFi APs to maximize the total or per-user link throughput in the licensed and unlicensed spectra is not addressed in these works. Although a recent work in [12] studied when to offload the traffic from the LTE network to the WiFi network and share the unlicensed spectrum in order to maximize the per-user link throughput, where their network, interference and spectrum sharing models are too much simple so that their observations may not be applicable to a large-scale stochastic network. Offloading traffic from an LTE network to another WiFi network may not increase the per-user link throughput since the offloaded users lose their licensed throughput. Hence, as long as the LTE and WiFi networks can coexist without causing severe interference, having them sharing the unlicensed spectrum is a good policy. In [13], [14], small cell BSs are shown to achieve a notable throughput gain if they can cleverly and properly access the unlicensed spectrum without causing much interference to the WiFi APs.

A stochastic-geometry-based framework in [10], [15]–[18] is applied to analyze the coexistence performance of large-scale LTE and WiFi networks, but the exact analyses of the coexisting throughputs of the APs in the licensed and unlicensed spectra were not studied in these works. Although few existing works indeed analyzed the throughput problem in coexisting LTE and WiFi networks, their throughput analysis is too simple to evaluate the link/network throughput very accurately. For example, reference [19] studied the network throughput by using the minimum required signal-to-interference power ratio (SIR) of each link and this network throughput may not be close to the real network throughput in that the link throughput of each AP could be significantly underestimated. In addition, these aforementioned prior works did not use a general network model to theoretically analyze the link throughput of each AP while the unlicensed spectrum is shared so that they are unable to offer some insights on how to maximize the per-user link throughput of each AP and overall network throughput by managing the traffic among all APs.

B. Contributions

The main contributions in this paper are summarized in the following:

- *A general M-tier HetNet architecture is proposed*: This M-tier network model is more general than the models in the aforementioned prior works so that it is able to generally characterize the licensed and unlicensed spectrum sharing problem among different kinds of APs using the opportunistic CSMA/CA protocol. Under this HetNet model, the coexisting interference models in the licensed and unlicensed spectra can be easily proposed.

- *A general AP association scheme is proposed*: It is able to cover several pathloss-based AP association schemes is adopted in the HetNet.

- *The void probability and association probability for an AP in each tier are accurately found based on the proposed AP association scheme*: We show that The void probability of densely-deployed APs is in general not small and thus cannot be ignored. However, many previous works overlooked this important fact and thus their analytical results could be very inaccurate.

- *The integral identity of the Shannon transformation is derived*: This identity can be applied to find the compact result of the mean spectrum efficiency of a wireless link in a Poisson wireless network.

- *The decentralized and centralized traffic management schemes are proposed*: The decentralized traffic management is able to maximize the per-user link throguhput of each AP with limited local information, whereas the the centralized traffic management scheme can maximize the per-user network throughput if the information of all APs can be processed jointly.

In this paper, we are only able to find the tight lower bounds on the mean spectrum efficiencies of the APs in each tier since the resulting transmitting APs in each tier are no longer a PPP due to performing AP association and CSMA to access the unlicensed channel. Nonetheless, with the aid of the integral identity of the Shannon transform, these tight lower bounds are derived without assuming any specific channel gain and AP association weight models. Most importantly, they are low-complexity so that they can give us some intuitions regarding how they are affected by channel gain impairments, AP association weights as well as void AP probabilities. This feature is absolutely important since we can easily judge whether the mean spectrum efficiency of the APs in a particular tier increases or decreases due to traffic loading or offloading of the APs in other tiers.

For the strategy of traffic management in this paper, our idea is to make the APs in the first $M - 1$ tiers that can access the licensed and unlicensed spectra achieve the per-user throughput as much as possible, whereas the APs in the Mth tier that only access the unlicensed spectrum just need to have their

per-user link throughput above some minimum required value. The *per-user link throughput* is defined based on the spectrum efficiencies of the APs in the licensed and unlicensed spectra. We propose a decentralized traffic manage scheme that maximizes the per-user link throughput of the APs in the first $M-1$ tiers and maintains the per-user link throughput requirement of the APs in the Mth tier. We also propose a centralized traffic management scheme that can be performed by the central processing unit of the core network of the HetNet to maximize the defined per-user network throughput by managing traffic offloading or loading of all APs. Moreover, the network model in this paper is built based on the assumption that all APs in each tier follows a homogeneous PPP so that the locations of all deployed APs are completely independent. Thus, the analytical results found in this paper can be used to do the worse-case performance evaluations for the LTE-U (LTE in Unlicensed spectrum) or LAA (Licensed-Assisted Access) system in 3GPP since the APs in a real LTE-U/LAA system may be deployed with location correlation [20]. Finally, a numerical simulation example of coexisting LTE BSs and WiFi APs is provided and the numerical results validate the accuracy and correctness of our analytical results and findings.

II. SYSTEM MODEL AND PRELIMINARIES

Consider a large-scale interference-limited HetNet consisting of M tiers of access points (APs). All the APs in the same tier are of the same type and performance. Specifically, the APs in the m-th tier, denoted by set $\mathcal{X}_m$, follow an independent marked Poisson point process (PPP) of intensity λ_m defined as follows

$$\mathcal{X}_m \triangleq \{(X_{m_i}, P_m, V_{m_i}) : X_{m_i} \in \mathbb{R}^2, P_m \in \mathbb{R}_+, V_{m_i} \in \{0,1\}, i \in \mathbb{N}_+\}, \, m \in \mathcal{M}, \tag{1}$$

where $\mathcal{M} \triangleq \{1, 2, \ldots, M\}$, X_{m_i} denotes AP i in the m-th tier and its location, P_m is the transmit power used by the APs in the m-th tier, and V_{m_i} is a Bernoulli random variable indicating whether AP X_{m_i} is void or not: if AP X_{m_i} is associated with at least one user (i.e., it is not void), then $V_{m_i} = 1$ and zero otherwise. Without loss of generality, we assume the APs in the Mth tier only use the unlicensed spectrum to deliver data, and all other APs in the first $M-1$ tiers primarily use the licensed spectrum and opportunistically use the unlicensed spectrum by carrier aggregation to transmit data if they have a chance to access the unlicensed spectrum. This network model with unlicensed spectrum sharing has a practical application context. For example, in a heterogeneous cellular network, LTE-U macrocell and small cell base stations (BSs) consisting of the APs in the first $M-1$ tiers can coexist and share the unlicensed spectrum with WiFi APs in the Mth tier if the LTE-U BSs can use the carrier aggregation technique to integrate the licensed and unlicensed spectrum resources [2], [3].

All users also follow an independent PPP $\mathcal{U}$ of intensity μ given by

$$\mathcal{U} \triangleq \{U_j : U_j \in \mathbb{R}^2, \forall j \in \mathbb{N}_+\} \tag{2}$$

and we assume there is typical user U_0 located at the origin without loss of generality. Our following location-dependent analyses will be based on typical user U_0 for simplicity since the analytical results do not depend where the typical user is located due to Slivnyak's theorem [21]. We consider a *downlink* transmission scenario in this paper and each user selects its serving AP X_o by adopting the following AP association scheme

$$X_o \triangleq \arg \sup_{X_{m_i} \in \bigcup_{m=1}^M \mathcal{X}_m} W_{m_i} \|X_{m_i}\|^{-\alpha} = \arg \inf_{X_{m_i} \in \bigcup_{m=1}^M \mathcal{X}_m} W_{m_i}^{\frac{1}{\alpha}} \|X_{m_i}\|, \tag{3}$$

where W_{m_i} is the random AP association weight with mean $\bar{w}_m$ for AP X_{m_i}, $\|X_i - X_j\|$ denotes the distance between nodes X_i and X_j for $i \neq j$, and $\|X_{m_i}\|^{-\alpha}$ is called the pathloss of AP X_{m_i} with pathloss exponent $\alpha > 2$. Furthermore, we assume that all W_{m_i}'s are independent, all $\frac{W_{m_i}}{\bar{w}_m}$'s are i.i.d. random variables with unit mean, and the a-fractional moment of W_{m_i} always exists for all $i \in \mathbb{N}_+$ and $m \in \mathcal{M}$, i.e., $\mathbb{E}[W_m^a] < \infty$ for all $a \in (0, 1)$. Note that the scheme in (3) makes users associate with an AP in any tier no matter which spectrum the AP primarily/only uses, and it can cover several different pathloss-based AP association schemes by changing the design of the AP association weights, such as the *biased nearest AP association* (BNA) scheme if W_m is a constant, the *biased mean strongest AP association* (BMSA) scheme if $W_{m_i} \equiv b_m P_m H_{m_i}^{(s)}$ for all $m \in \mathcal{M}$ where $b_m > 0$ is a constant bias and $H_{m_i}^{(s)}$ characterizes the large-scale channel gain of the tier-m APs such as shadowing, and other schemes etc. [22] [23].

A. AP Association Probability and Cell Load Statistics

The AP association scheme in (3) can be reformulated to statistically represent the weighted pathloss of AP X_o given by

$$W_o \|X_o\|^{-\alpha} \stackrel{d}{=} \sup_{X_{m_i} \in \bigcup_{m=1}^M \mathcal{X}_m} W_{m_i} \|X_{m_i}\|^{-\alpha}, \tag{4}$$

where $W_o \in \{W_{m_i}, m \in \mathcal{M}, i \in \mathbb{N}_+\}$ is the AP association weight used by AP X_o and $\stackrel{d}{=}$ means the statistical equivalence in distribution. The statistical property of $W_o \|X_o\|^{-\alpha}$ is crucial for the following analysis and it is provided in the following lemma.

Lemma 1 (Revised from [24], [25]): Let $\|\widetilde{X}_o\|^{-\alpha}$ denote the weighted pathloss of the AP X_o given in (4), i.e., $\|\widetilde{X}_o\|^{-\alpha} = W_o \|X_o\|^{-\alpha}$. The probability density function (pdf) of $\|\widetilde{X}_o\|$ can be shown as

$$f_{\|\widetilde{X}_o\|}(x) = 2\pi x \widetilde{\lambda} \exp\left(-\pi \widetilde{\lambda} x^2\right), \tag{5}$$

where $\widetilde{\lambda} \triangleq \sum_{m=1}^{M} \widetilde{\lambda}_m$ and $\widetilde{\lambda}_m \triangleq \lambda_m \mathbb{E}\left[W_m^{\frac{2}{\alpha}}\right]$.

Proof: The proof of (5) can be referred to [24], [25]. ∎

According to the proof of Lemma 1, we essentially realize that $\|\widetilde{X}_o\| = W_o^{-1/\alpha}\|X_o\|$ can be statistically viewed as the distance from the nearest point in the PPP of $\bigcup_{m \in \mathcal{M}} \widetilde{X}_m$ to typical user U_0 [21], [23], [25]. This fact implies the probability that users associate with a tier-m AP is given by

$$\vartheta_m = \frac{\lambda_m \mathbb{E}\left[W_m^{\frac{2}{\alpha}}\right]}{\sum_{k=1}^{M} \lambda_k \mathbb{E}\left[W_k^{\frac{2}{\alpha}}\right]} = \frac{\widetilde{\lambda}_m}{\widetilde{\lambda}} = \frac{\lambda_m \bar{w}_m^{\frac{2}{\alpha}}}{\sum_{k=1}^{M} \lambda_k \bar{w}_k^{\frac{2}{\alpha}}}, \tag{6}$$

which can be used to characterize the distribution of the number of users associating with an AP in a particular tier as shown in the following lemma.

Lemma 2: Let $\mathcal{A}_{m_i}$ denote the cell area where all users associate with AP X_{m_i} by using the AP association scheme (3). If the cell load of a tier-m AP, denoted as $\mathcal{X}_m(\mathcal{A}_{m_i})$, is defined as the number of users associating a tier-m AP and $\mathbb{E}\left[W_m^{2/\alpha}\right] \mathbb{E}\left[W_m^{-2/\alpha}\right] < \infty$, then its probability mass function (pmf) is given by

$$\nu_{m,n} \triangleq \mathbb{P}\left[\mathcal{X}_m(\mathcal{A}_m) = n\right] = \frac{\Gamma(n + \zeta_m)}{n!\Gamma(\zeta_m)}\left(\frac{\zeta_m \lambda_m}{\zeta_m \lambda_m + \mu\vartheta_m}\right)^{n+\zeta_m}, \ n \in \mathbb{N}, \tag{7}$$

where $\Gamma(x) = \int_0^{\infty} t^{x-1}e^{-t}\mathrm{d}t$ is the Gamma function, $\zeta_m \triangleq \frac{7}{2}\mathbb{E}\left[W_m^{2/\alpha}\right]\mathbb{E}\left[W_m^{-2/\alpha}\right]$ and $\widetilde{\lambda}$ is already defined in Lemma 1.

Proof: Since users adopt the scheme in (3) to select their serving AP and $\widetilde{X}_o$ can be viewed as the nearest point in $\bigcup_{m=1}^{M} \widetilde{\mathcal{X}}_m$ to the typical user, cell load $\mathcal{X}_m(\mathcal{A}_{m_i})$ has the same distribution as $\widetilde{X}_m(\widetilde{\mathcal{A}}_{m_i})$ where $\widetilde{\mathcal{A}}_m$ is the cell area of AP $\widetilde{X}_{m_i} \in \widetilde{\mathcal{X}}_m$ and it is Voronoi-tessellated. Hence, it follows that

$$\nu_{m,n} = \mathbb{P}\left[\mathcal{X}_m(\mathcal{A}_m) = n\right] = \mathbb{P}\left[\widetilde{\mathcal{X}}_m(\widetilde{\mathcal{A}}_m) = n\right] = \mathbb{E}_{\widetilde{\mathcal{A}}_m}\left[\frac{(\mu\widetilde{\lambda}_m\widetilde{\mathcal{A}}_m/\widetilde{\lambda})^n}{n!}e^{-\mu\widetilde{\lambda}_m\widetilde{\mathcal{A}}_m/\widetilde{\lambda}}\right].$$

Since the exact pdf of a voronoi-tessellated area is still an open problem, its substitute expression accurately approximated by a Gamma random variable can be found as [22], [23]

$$f_{\widetilde{\mathcal{A}}_m}(x) = \frac{(\zeta_m\lambda_m x)^{\zeta_m}}{x\Gamma(\zeta_m)}e^{-\zeta_m\lambda_m x}.$$

Thus,

$$\nu_{m,n} = \frac{1}{n!}\left(\frac{\mu\widetilde{\lambda}_m}{\widetilde{\lambda}}\right)^n \frac{(\zeta_m\lambda_m)^{\zeta_m}}{\Gamma(\zeta_m)}\int_0^{\infty} x^{n+\zeta_m-1}e^{-(\mu\widetilde{\lambda}_m/\widetilde{\lambda}+\zeta_m\lambda_m)x}\,,$$

and substituting the result in (6) into $\nu_{m,n}$ yields the result in (7). ∎

The pmf of the cell load in Lemma 2 indicates that the probability of no users associating a tier-m AP is

$$\mathbb{P}[V_m = 0] = \nu_{m,0} = \left(1 + \frac{\mu \vartheta_m}{\zeta_m \lambda_m}\right)^{-\zeta_m}, \tag{8}$$

which is the void probability of a tier-m AP. As ζ_m (or $\mathbb{E}[W_m^{2/\alpha}]$) goes to infinity, it reduces to the following minimum:

$$\lim_{\zeta_m \to \infty} \nu_{m,0} = \exp\left(-\frac{\mu}{\lambda_m}\right). \tag{9}$$

Thus, the void probabilities of the APs cannot be ignored if the ratios of the user intensity to the AP intensities are not large, especially for the network with densely deployed APs. However, most prior works on the modeling of the multi-tier PPP-based HetNets overlook this important void AP issue. Also, the mean cell load of a tier-m AP is given by

$$\mathbb{E}\left[\mathcal{X}_m(\mathcal{A}_m)\right] = \frac{\widetilde{\lambda}_m}{\widetilde{\lambda}} \mu \times \frac{1}{\lambda_m} = \frac{\mu \mathbb{E}\left[W_m^{\frac{2}{\alpha}}\right]}{\sum_{k=1}^{M} \lambda_k \mathbb{E}\left[W_k^{\frac{2}{\alpha}}\right]} = \frac{\mu \bar{w}_m^{\frac{2}{\alpha}}}{\sum_{k=1}^{M} \lambda_k \bar{w}_k^{\frac{2}{\alpha}}}. \tag{10}$$

Hence, to completely balance the mean load between all tiers, all W_m's must be i.i.d. for all $m \in \mathcal{M}$. For example, the (unbiased) nearest AP association scheme that makes user associate their nearest AP can achieve a completely balanced and same mean cell load for all APs in different tiers in that all W_{m_i}'s in the scheme are the same constant.

B. Channel Access Protocols

In this paper, all APs are assumed to always have data to transmit to their tagged users. The channel access protocols for the licensed spectrum and unlicensed spectrum are quite different. All APs in the first $M - 1$ tiers share the entire licensed spectrum at the same time and they are synchronized when accessing the licensed channel[1]. Note that the APs in the Mth tier cannot access the licensed channel and they are only allowed to access the channel in the unlicensed spectrum. All APs have to use the (slotted non-persistent) *opportunistic* CSMA/CA (carrier sense multiple access with collision avoidance) protocol to access the unlicensed channel[2]. By adopting such an opportunistic CSMA/CA protocol, the APs whose channel gains are greater than some threshold are qualified and synchronized to contend

---

[1]Such a licensed channel access protocol is widely used in the cellular networks. In addition, we assume there is only one channel in the licensed spectrum for the ease of analysis.

[2]Like the case in the licensed spectrum, we assume there is only one available channel in the unlicensed spectrum in order to simplify our following analysis.

the unlicensed channel in the predesignated time slots. The feature of this opportunistic CSMA/CA protocol is able to make the unlicensed spectrum resource to be utilized effectively by the APs with good channel conditions so as to improve the spectrum sharing efficiency and throughput. Each AP in the m-th tier that performs the opportunistic CSMA/CA protocol has a sensing region $\mathcal{S}_m$ in which all unlicensed channel accessing activities can be detected by the AP. The channel access probability of the opportunistic CSMA/CA protocol is already derived in our previous works [9] [11], and it can be modified for the AP association scheme in (3) and rewritten as shown in the following lemma.

Lemma 3 (*Channel access probability in the unlicensed spectrum modified from [9] [11]*): If the random backoff time of a tier-m AP is uniformly distributed in $[0, \tau_m]$, then its channel access probability is

$$\rho_m = \frac{1 - e^{-\frac{\tau_M}{\mathbb{E}[W_m^{2/\alpha}]}\sum_{k=1}^{M} A_{m,k}\lambda_{k,M}^{\dagger}}}{\tau_m \sum_{k=1}^{M} A_{m,k}(\bar{w}_k/\bar{w}_m)^{\frac{2}{\alpha}}\lambda_{k,M}^{\dagger}} + \sum_{j=m}^{M} \frac{e^{-\tau_j \sum_{k=1}^{j} A_{m,k}(\frac{\bar{w}_k}{\bar{w}_m})^{\frac{2}{\alpha}}\lambda_{k,j}^{\dagger}} - e^{-\tau_{j+1}\sum_{k=1}^{j} A_{m,k}(\frac{\bar{w}_k}{\bar{w}_m})^{\frac{2}{\alpha}}\lambda_{k,j}^{\dagger}}}{\tau_m \sum_{k=1}^{j} A_{m,k}(\bar{w}_k/\bar{w}_m)^{\frac{2}{\alpha}}\lambda_{k,j}^{\dagger}}, \quad (11)$$

where $\tau_1 \geq \tau_2 \geq \cdots \geq \tau_M \geq 0$, $\tau_{M+1} \equiv 0$, $A_{m,k}$ is the mean area of region $\mathcal{S}_m$ where the tier-k APs are distributed (see [9] for the details of how to calculate $A_{m,k}$.), $\lambda_{k,j}^{\dagger} = \xi_k(1 - \nu_{k,0})\lambda_k(\frac{\tau_j - \tau_{j+1}}{\tau_k})$, $\xi_k \in [0,1]$ is the probability that the unlicensed channel (power) gain from a tier-k AP to its servicing user is greater than threshold $\delta > 0$. If all τ_m's are the same and equal to τ, (11) reduces to

$$\rho_m = \frac{1 - \exp\left(-\tau \sum_{k=1}^{M} A_{m,k}\xi_k q_{k,0}\lambda_k(\bar{w}_k/\bar{w}_m)^{\frac{2}{\alpha}}\right)}{\tau \sum_{k=1}^{M} A_{m,k}\xi_k q_{k,0}\lambda_k(\bar{w}_k/\bar{w}_m)^{\frac{2}{\alpha}}}, \quad (12)$$

where $q_{k,0} \triangleq 1 - \nu_{k,0}$ is the non-void probability of the tier-k APs.

The channel access probability in (11) indicates not only how much chance a tier-m AP can successfully access the unlicensed channel in a particular timeslot but also the fraction of time it can access the unlicensed channel in the long-term sense. Adjusting the random backoff time limit τ_m can make the tier-m APs have more/less priority or time fraction to access the unlicensed spectrum. For example, if the tier-M APs represent the WiFi APs, we can make its backoff time limit τ_M much shorter than those of the APs in the first $M-1$ tiers so that the throughput of the WiFi APs is guaranteed to remain at some level and not significantly reduced when the unlicensed spectrum is shared by many APs in other tiers at the same time. This is similarly implementing the ideas of the Listen-Before-Talk (LBT) with Carrier Sensing Adaptive Transmission (CSAT) and Licensed-Assisted Access (LAA) protocols proposed in the LTE-U [2], [3]. Another two characteristics of the result in (11) are including the probability of being void APs as well as the probability of opportunistically having a good channel state. They make the channel access probability more accurate and higher for the APs with good channels. In the following

analysis, we will see how the channel access probability plays a pivotal role in analyzing the throughput in the unlicensed spectrum.

III. Shannon Transform, Mean Spectrum Efficiency and Per-User Throughput

In this section, we would like to study the mean spectrum efficiencies of a user in the unlicensed and licensed spectra. Finding the explicit expressions of these mean spectrum efficiencies is an important task since they provide the insights into how to manage traffic flow in order to achieve the throughput optimality. The prior approaches to deriving the mean spectrum efficiency in the literature are based on integrating the function of the coverage (success) probability under the Rayleigh fading channel model [24], [26], [27]. Hence, these approaches cannot characterize the mean spectrum efficiencies in a non-Rayleigh fading environment. In the following analysis, we will show how to derive the mean spectrum efficiency for any general channel gain and AP association weight models in a low-complexity expression. First, we need to introduce the Shannon transform of a nonnegative random variable and its integral identity since they are the key to obtaining the analytically tractable expression of the mean spectrum efficiency without making any specific modeling assumptions on the channel gains and AP association weights.

A. The Shannon Transform and Its Integral Identity

The Shannon transform of a nonnegative random variable Ψ for any nonnegative $\eta \in \mathbb{R}_{++}$ is defined as

$$\mathcal{S}_\Psi(\eta) = \mathbb{E}\left[\ln(1 + \eta\Psi)\right]. \tag{13}$$

The Shannon transform in (13) has an identity as shown in the following theorem.

Theorem 1 (The integral identity of the Shannon transform): Consider a nonnegative random variable Ψ and the Laplace transform of its reciprocal is always well-defined and exists, i.e., $\mathcal{L}_{\Psi^{-1}}(s) < \infty$ for $s \in \mathbb{R}_{++}$. Define $\mathcal{S}_\Psi(\eta) \triangleq \mathbb{E}[\ln(1 + \eta\Psi)]$ as the Shannon transform of random variable Ψ for a nonnegative constant $\eta \in \mathbb{R}_+$. If $\mathcal{S}_\Psi(\eta)$ exists for any $\eta \in \mathbb{R}_+$, then it has the following identity

$$\mathcal{S}_\Psi(\eta) = \int_{0+}^{\infty} \frac{(1 - e^{-\eta s})}{s} \mathcal{L}_{\Psi^{-1}}(s)\mathrm{d}s, \tag{14}$$

which always holds. Furthermore, we can have

$$\mathbb{E}\left[\mathcal{S}_\Psi(\eta)\right] = \int_{0+}^{\infty} \frac{[1 - \mathcal{L}_\eta(s)]}{s} \mathcal{L}_{\Psi^{-1}}(s)\mathrm{d}s \tag{15}$$

if η is a nonnegative random variable and its Laplace transform exists.

Proof: See Appendix A. ∎

The identity of the Shannon transformation in Theorem 1 provides a low-complexity means to find the mean spectrum efficiency of a link if $\eta\Psi$ is the SIR of the link and the Laplace transform of their reciprocal exists and can be explicitly found. As we will show in the following subsections, Theorem 1 facilitates the derivations of the mean spectrum efficiencies of an AP in the unlicensed and licensed spectra for any general random channel gain and AP association weight models.

B. Mean Spectrum Efficiency of the APs in the Unlicensed Spectrum

Consider the scenario that typical user U_0 associates with a tier-m AP by scheme (3) and is receiving data over the unlicensed channel. The mean spectrum efficiency (bps/Hz) the tier-m AP offers to the user in the unlicensed spectrum is defined as

$$R_{\mathsf{U}_m} \triangleq \mathbb{E}\left[\log_2\left(1 + \frac{P_m H_m}{I_{\mathsf{U}_m}\|X_o\|^\alpha}T_m\Xi_m\right)\right], \quad m \in \mathcal{M}, \tag{16}$$

where P_m is the transmit power used by AP $X_o \in \mathcal{X}_m$, H_m is the random channel (power) gain with mean $\bar{h}_m$ from AP X_o to typical user U_0, and I_{U_m} is the interference as shown in the following

$$I_{\mathsf{U}_m} \triangleq \sum_{X_{m_i} \in \bigcup_{m=1}^M \mathcal{X}_m \backslash X_o} P_m H_{m_i} V_{m_i} T_{m_i} \Xi_{m_i} \|X_{m_i}\|^{-\alpha}$$

in which H_{m_i} denotes the random channel gain with mean $\bar{h}_m$ from AP X_{m_i} to typical user U_0, $\Xi_{m_i} \in \{0,1\}$ is a Bernoulli random variable that equals to one if AP X_{m_i}'s channel gain is greater than threshold $\delta > 0$ and zero otherwise, $T_{m_i} \in \{0,1\}$ is also a Bernoulli random variable that equals to one if AP X_{k_i} can access the unlicensed channel and zero otherwise[3], $T_m \in \{T_{m_i}\}$ indicates whether AP X_o can access the unlicensed channel, and $\Xi_m \in \{\Xi_{m_i}\}$ indicates whether the channel gain of AP X_o in the unlicensed spectrum is greater than threshold δ. For the ease of analysis, we assume that all $\frac{H_{m_i}}{\bar{h}_m}$'s are i.i.d. for all $i \in \mathbb{N}_+$ and $m \in \mathcal{M}$ throughout this paper.

Letting $\Psi_{\mathsf{U}_m} \triangleq P_m H_m T_m \Xi_m / I_{\mathsf{U}_m}\|X_o\|^\alpha$ and using the definition of the Shannon transform in the previous subsection, R_{U_m} can be further expressed as

$$R_{\mathsf{U}_m} = \frac{\rho_m \xi_m}{\ln(2)}\int_{0+}^\infty \frac{1 - \mathcal{L}_{H_m}(s)}{s}\mathcal{L}_{P_m^{-1}\|X_o\|^\alpha I_{\mathsf{U}_m}}(s)\mathrm{d}s \tag{17}$$

and the explicit expression of R_{U_m} derived by using Theorem 1 is shown in the following theorem.

---

[3]Namely, the probability that Ξ_{m_i} is equal to one is $\mathbb{P}[\Xi_m = 1] = \xi_m$, and the probability that T_{m_i} is equal to one is $\mathbb{P}[T_m = 1] = \rho_m$ for all $m \in \mathcal{M}$.

Theorem 2: Suppose users adopt the AP association scheme in (3). If all non-void APs use the opportunistic CSMA/CA protocol to access the unlicensed channel, the mean spectrum efficiency of the user in the unlicensed spectrum in (17) can be explicitly lower bounded by

$$
R_{\mathsf{U}_m} \geq \frac{\rho_m \xi_m}{\ln(2)} \int_{0^+}^{\infty} \frac{\left[1 - \mathcal{L}_{\widehat{H}}(u)\right] \mathrm{d}u}{u \left(\sum_{k=1}^{M} q_{k,0} \xi_k \rho_k \vartheta_k \ell_{\widehat{H}}\left(\frac{\bar{w}_m \bar{h}_k P_k}{\bar{w}_k \bar{h}_m P_m} u, \frac{2}{\alpha}\right) + 1\right)}, \tag{18}
$$

where $\widehat{H} \triangleq \frac{H_m \bar{w}_m}{W_m \bar{h}_m}$ is a random variable with unit mean and $\ell_Z(x, y)$ for $y \in (0, 1)$ is defined as

$$
\ell_Z(x, y) \triangleq x^y \Gamma(1 - y) \mathbb{E}[Z^y] + \int_0^1 \mathcal{L}_Z\left(xt^{-\frac{1}{y}}\right) \mathrm{d}t - 1. \tag{19}
$$

If all W_{m_i}'s are deterministic and equal to constant $\bar{w}_m$ and all H_{m_i}'s are i.i.d. random variables with unit mean for all $m \in \mathcal{M}$, then (18) reduces to

$$
R_{\mathsf{U}_m} \geq \frac{\rho_m \xi_m}{\ln(2)} \int_{0^+}^{\infty} \frac{\left[1 - \mathcal{L}_H(u)\right] \mathrm{d}u}{u \left(\sum_{k=1}^{M} q_{k,0} \xi_k \rho_k \vartheta_k \ell_H\left(\frac{\bar{w}_m \bar{h}_k P_k}{\bar{w}_k \bar{h}_m P_m} u, \frac{2}{\alpha}\right) + 1\right)}, \tag{20}
$$

where H has the same distribution as all H_{m_i}'s.

Proof: See Appendix B. ∎

It is worth mentioning a few features of the lower bounds in (18) and (20). First of all, the lower bound is in general *fairly tight* since the location correlations between the non-void APs due to AP association and opportunistic CSMA/CA are usually very weak. These location correlations will be weakened and thus R_{U_m} will be very close to its lower bound when either the user intensity or the channel gain threshold for opportunistic CSMA/CA increases. The tightness of the lower bound in (18) will be verified by the numerical results presented in Section V-B. Second, the lower bound in (18) is valid for all random models of channel gains and AP association weights as long as the Laplace transforms of the channel gains and the AP association weights exist, which is never derived in the literatures. This is a very important feature since we are able to realize how different channel and AP association models affect R_{U_m} and gain some insights about how to improve R_{U_m} by appropriately designing the AP association weights in order to manage the traffic flows among different tiers of the APs. Third, according to Jensen's inequality, we know $\mathcal{L}_Z(u) \geq \exp(-u\mathbb{E}[Z])$, which means if all H_m's and/or W_m's are not random the lower bound in (18) will increase (i.e., R_{U_m} will increase). For example, in the special case that there are no random channel gain impairments and users adopt the BNA scheme (i.e., $H_m \equiv 1$ and all W_{m_i}'s are equal to constant w_m for all $i \in \mathbb{N}_+$) or in the special

case of the AP association scheme with $W_{m_i} = b_m P_m H_{m_i}$ for all $m \in \mathcal{M}$, the lower bound in (20) becomes

$$R_{\mathsf{U}_m} \geq \frac{\rho_m \xi_m}{\ln(2)} \int_{0^+}^{\infty} \frac{(1 - e^{-u}) \, \mathrm{d}u}{u \left(\ell(u, \frac{2}{\alpha}) \sum_{k=1}^{M} q_{k,0} \xi_k \rho_k \vartheta_k + 1 \right)}, \tag{21}$$

where $\ell(x, y) \triangleq x^y \Gamma(1 - y) + \int_0^1 e^{-xt^{-\frac{1}{y}}} \mathrm{d}t - 1$. And this is the maximum lower bound achieved by the biased AP association schemes in that the randomness of the channel gains does not exist in the R_{U_m}.

C. Mean Spectrum Efficiency of the APs in the Licensed Spectrum

For the mean spectrum efficiency of a user associating with a tier-m AP in the licensed spectrum, its formal definition can be written as

$$R_{\mathsf{L}_m} \triangleq \mathbb{E}\left[\log_2 \left(1 + \frac{P_m H_m}{I_{\mathsf{L}_m} \|X_o\|^\alpha} \right) \right], \tag{22}$$

where $X_o \in \mathcal{X}_m$ and I_{L_m} is given by

$$I_{\mathsf{L}_m} \triangleq \sum_{X_{m_i} \in \bigcup_{m=1}^{M-1} \mathcal{X}_m \backslash X_o} P_m H_{m_i} V_{m_i} \|X_{m_i}\|^{-\alpha}. \tag{23}$$

Note that all channel gains in (22) are evaluated in the licensed spectrum. The explicit result of R_{L} is shown in the following theorem.

Theorem 3: If users adopt the AP association scheme in (3), then the mean spectrum efficiency of a user in the licensed spectrum defined in (22) can be shown as

$$R_{\mathsf{L}_m} \geq \frac{1}{\ln(2)} \int_{0^+}^{\infty} \frac{\left[1 - \mathcal{L}_{\widehat{H}}(u) \right] \mathrm{d}u}{u \left(\sum_{k=1}^{M-1} q_{k,0} \vartheta_k \ell_{\widehat{H}}\big(\frac{\bar{w}_m \bar{h}_k P_k}{\bar{w}_k \bar{h}_m P_m} u, \frac{2}{\alpha} \big) + 1 \right)}, \quad m \in \{1, 2, \dots, M-1\}. \tag{24}$$

If all W_{m_i}'s are deterministic and equal to $\bar{w}_m$ and all H_{m_i}'s are i.i.d. random variables with unit mean for all $m \in \mathcal{M}$, (24) reduces to

$$R_{\mathsf{L}_m} \geq \frac{1}{\ln(2)} \int_{0^+}^{\infty} \frac{\left[1 - \mathcal{L}_{H}(u) \right] \mathrm{d}u}{u \left(\sum_{k=1}^{M-1} q_{k,0} \vartheta_k \ell_{H}\big(\frac{\bar{w}_m \bar{h}_k P_k}{\bar{w}_k \bar{h}_m P_m} u, \frac{2}{\alpha} \big) + 1 \right)}. \tag{25}$$

Proof: First, consider the case that all W_{m_i}'s are random variables. Since all APs in the first $M-1$ tiers access the licensed spectrum without using the opportunistic CSMA/CA protocol, the mean spectrum efficiency for a user associating with a tier-m AP can be explicitly obtained as

$$\mathbb{E}\left[\log_2 \left(1 + \frac{P_m H_m}{I_{\mathsf{L}_m} \|X_o\|^\alpha} \right) \right] = \frac{1}{\ln(2)} \int_{0^+}^{\infty} \frac{\left[1 - \mathcal{L}_{\widehat{H}_m}(u) \right] \mathrm{d}u}{u \left(\sum_{k=1}^{M-1} q_{k,0} \vartheta_k \ell_{\widehat{H}_k}\big(\frac{P_k \bar{w}_m \bar{h}_k}{P_m \bar{w}_k \bar{h}_m} u, \frac{2}{\alpha} \big) + 1 \right)} \tag{26}$$

by applying the result of R_{U_m} for the mean spectrum efficiency in the m-th tier with $\rho_k = \xi_k = 1$ for all $k \in \mathcal{M}$. Substituting (26) into (22) results in (24). ∎

The results in Theorem 3 are valid for any channel gain and AP association weight models and they are never derived in the literature as well. Prior results on the spectrum efficiency in the licensed spectrum are only derived for Rayleigh fading channels and do not reflect the impact of the void APs. The lower bound on R_{L_m} in (24) is obtained by assuming the non-void correlated APs form M independent thinning homogeneous PPPs and in general it is also very tight, like the lower bounds on R_{U_m} in Theorem 2, since the location correlations of the non-void APs are fairly weak. Thus, as the user intensity goes to infinity, $q_{k,0}$ converges to one so that R_{L_m} exactly reduces to the lower bound given by

$$\lim_{\mu \to \infty} R_{\mathsf{L}_m} = \frac{1}{\ln(2)} \int_{0^+}^{\infty} \frac{\left[1 - \mathcal{L}_{\widehat{H}_m}(u)\right] \mathrm{d}u}{u \left(\sum_{k=1}^{M-1} \vartheta_k \ell_{\widehat{H}_k} \left(\frac{P_k \bar{w}_m \bar{h}_k}{P_m \bar{w}_k h_m} u, \frac{2}{\alpha} \right) + 1 \right)}, \tag{27}$$

which is the lowest mean spectrum efficiency of the tier-m APs in the licensed spectrum. This indicates that *the mean spectrum efficiency is significantly underestimated in a dense network if the void AP impact is not considered in the interference model*. Also, if the AP association scheme in (3) with $W_{m_i} = b_m P_m H_{m_i}$ can characterize the channel gains, the lower bound in (24) becomes

$$R_{\mathsf{L}_m} \geq \frac{1}{\ln(2)} \int_{0^+}^{\infty} \frac{(1 - e^{-u}) \, \mathrm{d}u}{u \left(\ell(u, \frac{2}{\alpha}) \sum_{k=1}^{M-1} q_{k,0} \vartheta_k + 1 \right)}, \tag{28}$$

which is the maximum lower bound achieved by AP association and also the lower bound for the scenario that channels do not suffer random impairments and users select their APs by using the BNA scheme.

D. Per-User Throughput Characterization

We have characterized the mean spectrum efficiencies of the licensed and unlicensed spectra in the previous subsections, which can be used to characterize the total (licensed and unlicensed) link throughput of a tier-m AP. Assume the bandwidths of the licensed spectrum and unlicensed spectrum are denoted by B_{L} and B_{U}, respectively. Accordingly, the total link throughput of a tier-m AP can be expressed as

$$C_m = B_{\mathsf{L}} R_{\mathsf{L}_m} \mathbb{1}(m \neq M) + B_{\mathsf{U}} R_{\mathsf{U}_m}, \ m \in \mathcal{M}, \tag{29}$$

where $\mathbb{1}(\mathcal{E})$ is an indicator function which is one if event $\mathcal{E}$ is true and zero otherwise. Note that the link throughput of a tier-M AP, only in the unlicensed spectrum, is $C_M = B_{\mathsf{U}} R_{\mathsf{U}_M}$. The total link

throughput of each AP highly depends on how the AP association weights in (3) are designated. For instance, when $\mathbb{E}[W_m^{2/\alpha}]$ becomes larger, more users associate the tier-m APs (i.e., more traffic is offloaded to the tier-m APs) so that the mean spectrum efficiencies would change very likely due to interference variations since the void probability of the tier-m APs reduces and correspondingly the void probabilities of the APs in other tiers increase. In this case, whether the total link throughput of the tier-m APs increases (or decreases) is dependent upon whether the interferences in the licensed and unlicensed spectra decrease (or increase). The channel access probability of the APs in the unlicensed spectrum decreases (or increases) due to offloading traffic to the tier-m APs.

Recall that the mean cell load of a tier-m AP given in (10), denoted by $\mathbb{E}[\mathcal{X}_m(\mathcal{A}_m)]$, is the mean number of users associating a tier-m AP. By assuming all users equally share the spectrum resources, the per-user link throughput of the tier-m APs is given by

$$c_m \triangleq \frac{C_m}{\mathbb{E}[\mathcal{X}_m(\mathcal{A}_m)|V_m = 1]} = \frac{q_{m,0}\lambda_m}{\mu\vartheta_m}[B_{\mathsf{L}}R_{\mathsf{L}_m}\mathbb{1}(m \neq M) + B_{\mathsf{U}}R_{\mathsf{U}_m}]. \tag{30}$$

In general, this per-user link throughput decreases as the traffic offloaded to the tier-m APs increases in that $\mathbb{E}[\mathcal{X}_m(\mathcal{A}_m)]$ increases but C_m may not increase. Furthermore, we can define the per-user network throughput as $\sum_{m=1}^{M} c_m\vartheta_m$ since it can be explicitly written as

$$\sum_{m=1}^{M} c_m\vartheta_m = \frac{B_{\mathsf{L}}}{\mu} \sum_{m=1}^{M} q_{m,0}\lambda_m \left[R_{\mathsf{L}_m}\mathbb{1}(m \neq M) + \frac{B_{\mathsf{U}}}{B_{\mathsf{L}}}R_{\mathsf{U}_m} \right] = \frac{1}{\mu} \sum_{m=1}^{M} q_{m,0}\lambda_m C_m, \tag{31}$$

and the term $\frac{1}{\mu} \sum_{m=1}^{M} q_{m,0}\lambda_m C_m$ characterizes how much network throughput a user can obtain when the void cell issue is considered and the unlicensed spectrum is opportunistically shared by all APs. In the following section, we will study how to maximize this per-user network throughput by optimizing the AP association weights in (3).

IV. TRAFFIC MANAGEMENT FOR COEXISTING LICENSED AND UNLICENSED APs

The underlying HetNet considered in this paper models the scenario that licensed and unlicensed APs coexist and opportunistically share the unlicensed spectrum, which corresponds a realistic situation in which WiFi APs in the Mth tier share their unlicensed spectrum with licensed LTE APs in the first $M-1$ tiers. The traffic management problem that interests us here is how to offload or load traffic between two orthogonal spectrum domains such that the per-user-based throughput increases. Since the APs in the first $M-1$ tiers could simultaneously assess the unlicensed channel, their transmitting behaviors in the unlicensed spectrum definitely affect the throughput of the APs in the Mth tier. Accordingly, the primary premise of managing the traffic in different tiers is to make the per-user link throughput of the

APs in the Mth tier higher than some minimum required value. Namely, assuming the per-user link throughput of the tier-M APs must be at least greater than some minimum value $c_{\min}$, i.e., $c_M \geq c_{\min}$ and it is

$$\frac{q_{M,0}B_\mathsf{U}}{c_{\min}\mu}\left(\sum_{m=1}^{M}\lambda_m\bar{w}_m^{\frac{2}{\alpha}}\right) \geq \frac{\bar{w}_M^{\frac{2}{\alpha}}}{R_{\mathsf{U}_M}}. \tag{32}$$

With this constraint on $\bar{w}_M^{\frac{2}{\alpha}}$, we are able to study how to maximize the per-user link throughput of the APs in any particular tier and the per-user network throughput by optimizing the designs of the AP association weights.

A. Decentralized Traffic Management

According to Theorems 2, 3 and the per-user link throughput of the tier-m APs given in (30), c_m is significantly affected by all $\{\bar{w}_m^{2/\alpha}\}$. Now our interest here is to gain some insights on when an AP in a particular tier should independently determine to offload its traffic to APs in other tiers (i.e., reduce its AP association weight) or load traffic from APs in other tiers (i.e., increase its AP association weight) so that its per-user link throughput increases. This is essentially a "decentralized" traffic management problem since the traffic loading or offloading decision is independently made by each AP from the perspective of the per-user link throughput. In other words, this decentralized traffic management problem is to study how to increase or even maximize c_m by unilaterally changing or even optimizing the value $\bar{w}_m^{2/\alpha}$ of a tier-m AP under the constraint (32). That is, if possible, we would like to solve the following optimization problem of $\bar{w}_m^{2/\alpha}$:

$$\begin{cases} \max\limits_{\omega_m>0} & c_m(\omega_m) \\ \text{s.t.} & \left(\sum_{k=1}^{M}\lambda_k\omega_k\right)\frac{R_{\mathsf{U}_M}}{\omega_M} \geq \frac{c_{\min}\mu}{q_{M,0}B_\mathsf{U}} \end{cases} \tag{33}$$

in which we define $\omega_k \triangleq \bar{w}_k^{2/\alpha}$ to simplify the notation in the problem. The solution of this optimization problem exists as shown in the following lemma.

Lemma 4: Let Ω_m be the feasible set of ω_m with the constraint (32), i.e., it is

$$\Omega_m \triangleq \left\{\omega_m \in \mathbb{R}_{++} : \left(\sum_{k=1}^{M}\lambda_k\omega_k\right)\frac{R_{\mathsf{U}_M}}{\omega_M} \geq \frac{c_{\min}\mu}{q_{M,0}B_\mathsf{U}}, \omega_k > 0, k \in \mathcal{M} \setminus m\right\}. \tag{34}$$

If Ω_m is nonempty, there exists a maximizer ω_m^* of c_m over Ω_m for all $m \in \{1, 2, \ldots, M-1\}$.

Proof: First, we would like to show that Ω_m is a compact set if it is nonempty. According to (11) and (18), we can infer the following facts

$$\lim_{\omega_m\to 0}\left(\sum_{k=1}^{M}\lambda_k\omega_k\right)\frac{R_{\mathsf{U}_M}}{\omega_M} < \infty \text{ and } \lim_{\omega_m\to\infty}\left(\sum_{k=1}^{M}\lambda_k\omega_k\right)\frac{R_{\mathsf{U}_M}}{\omega_M} = 0.$$

Since $\left(\sum_{k=1}^M \lambda_k \omega_k\right) \frac{R_{\mathsf{U}_M}}{\omega_M}$ is bounded, there must exist a positive number $R_{\mathsf{U}_M}^\dagger$ at $\omega_m^\dagger$ such that

$$R_{\mathsf{U}_M}^\dagger(\omega_m^\dagger) = \sup_{\omega_m > 0} \left\{\left(\sum_{k \in \mathcal{M}} \lambda_k \omega_k\right) \frac{R_{\mathsf{U}_M}}{\omega_M}\right\} \geq \frac{c_{\min}\mu}{q_{M,0}B_{\mathsf{U}}}$$

if Ω_m is nonempty. For $\omega_m \in (\omega_m^\dagger, \infty)$, $\left(\sum_{k=1}^M \lambda_k \omega_k\right) \frac{R_{\mathsf{U}_M}}{\omega_M}$ is a monotonically decreasing function of ω_m. Accordingly, there must exist a $\omega_m^\ddagger$ such that $\left(\sum_{k=1}^M \lambda_k \omega_k\right) \frac{R_{\mathsf{U}_M}}{\omega_M} \geq \frac{c_{\min}\mu}{q_{M,0}B_{\mathsf{U}}}$ for $\omega_m \in (0, \omega_m^\ddagger]$. Hence $\Omega_m = (0, \omega_m^\ddagger]$ is closed and bounded. Furthermore, c_m is a continuous function of ω_m over Ω_m. According to the Weierstrass theorem [28], the maximizer ω_m^* must exist. ∎

Lemma 4 reveals that $\omega_m^* \in \{\arg\sup_{\omega_m \in \Omega_m} c_m(\omega_m)\}$ and $\frac{\partial c_m}{\partial \omega_m}\big|_{\omega_m=\omega_m^*} = 0$. However, finding ω_m^* needs the information of other ω_k's which is in general unknown for the APs in the m-th tier in the decentralized context.

The two fundamental traffic management rules for a tier-m AP can be easily realized as

$$\frac{\partial c_m}{\partial \omega_m} < 0 \Leftrightarrow \begin{cases} \text{loading traffic reduces } c_m \\ \text{offloading traffic increases } c_m \end{cases} \tag{35}$$

and

$$\frac{\partial c_m}{\partial \omega_m} > 0 \Leftrightarrow \begin{cases} \text{loading traffic increases } c_m \\ \text{offloading traffic reduces } c_m \end{cases}. \tag{36}$$

These two rules indicate that *the APs in the first $M-1$ tiers need to offload traffic if $\frac{\partial c_m}{\partial \omega_m} < 0$ and load traffic if $\frac{\partial c_m}{\partial \omega_m} > 0$ under the constraint that the APs in the Mth tier need to maintain their per-user link throughput above the threshold value* $c_{\min}$. According to the facts in (35) and (36), we develop a decentralized traffic management scheme for the APs in each tier as shown in the following theorem.

Theorem 4: For the APs in the m-th tier and $m \in \{1, 2, \ldots, M-1\}$, the following decentralized traffic management scheme can maximize their per-user link throughput under the constraint in (32)

$$\omega_m(n+1) = \frac{c^*(n)N_m(n)\omega_m(n)}{C_m(n)}, \ \omega_m(0) > 0, \ n \in \mathbb{N}, \tag{37}$$

where $N_m(n) = \frac{1}{n}\sum_{i=0}^{n-1} N_m(i)$ denotes the average number of the users associating with a tier-m AP at time n and $c^*(n) \triangleq \max\{c_{\min}, c^*(n-1), C_m(n)/N_m(n)\}$. In addition, as n goes to infinity this scheme makes ω_m converge to ω_m^* that is the fixed point of the function $\Upsilon_m(x)$ given by

$$\Upsilon_m(x) = \frac{q_{m,0}(x)C_m(x)}{\mu c_{\min}}\left(\sum_{k \in \mathcal{M}\backslash m} \lambda_k \omega_k + \lambda_m x\right). \tag{38}$$

Proof: See Appendix C. ∎

Since a tier-m AP can estimate $R_{\mathsf{L}_m}(n)$ and $R_{\mathsf{U}_m}(n)$ and other parameters in (37) are locally available to the tier-m AP, the scheme in (37) can be easily implemented by the AP. Function $\Upsilon_m(x)$ in (38) can help us roughly determine the initial value $\omega_m(0)$ of $\omega_m(n)$ provided that each tier-m AP initially knows all other ω_k's and this would shorten the process of $\omega_m(n)$ converging to ω_m^*. Note that in general the per-user link throughput achieved by the scheme in (37) in the steady state is just a suboptimal result because other $M - 1$ parameters $\{\omega_k, k \in \mathcal{M} \setminus m\}$ are not optimized. In the following subsection, a centralized traffic management approach to maximizing the per-user network throughput is proposed and studied.

B. Centralized Traffic Management

If all the APs in the HetNet can send their information to their backhaul processing unit in their core network, the centralized traffic management is implementable[4]. Under this circumstance, we can maximize the per-user network throughput defined in (31) under the constraint in (32) by optimizing all $\{\omega_m\}$. That is, we can solve the following optimization problem of the centralized traffic management

$$
\begin{cases}
\max\limits_{\{\omega_m\} > 0} & \sum_{m=1}^{M} q_{m,0} \lambda_m \left[R_{\mathsf{L}_m} \mathbb{1}(m \neq M) + \frac{B_{\mathsf{U}}}{B_{\mathsf{L}}} R_{\mathsf{U}_m} \right] \\
\text{s.t.} & \left(\sum_{m=1}^{M} \lambda_m \omega_m \right) \frac{R_{\mathsf{U}_M}}{\omega_M} \geq \frac{c_{\min} \mu}{q_{M,0} B_{\mathsf{U}}} \\
& \omega_m > 0
\end{cases}
\tag{39}
$$

to find the optimal M-tuple vector of $(\omega_1, \omega_2, \ldots, \omega_M)$. Note that the objective function of the optimization problem in (39) is the per-user network throughput normalized by constant $\frac{\mu}{B_{\mathsf{L}}}$ which does not affect the optimization solutions. The solution of this optimization problem exists as shown in the following theorem.

Lemma 5: Let Ω be the feasible set of the M-tuple vector $(\omega_1, \omega_2, \ldots, \omega_M)$ with the constraint (32) and it is expressed as

$$
\Omega \triangleq \left\{ (\omega_1, \omega_2, \ldots, \omega_M) \in \mathbb{R}_{++}^M : \left(\sum_{m=1}^{M} \lambda_m \omega_m \right) \frac{R_{\mathsf{U}_M}}{\omega_M} \geq \frac{c_{\min} \mu}{q_{M,0} B_{\mathsf{U}}} \right\}.
\tag{40}
$$

If Ω is nonempty, there exists an optimal M-tuple vector $(\omega_1^*, \omega_2^*, \ldots, \omega_M^*)$ that maximizes the per-user network throughput $\sum_{m=1}^{M} c_m \vartheta_m$.

Proof: According to Lemma 4, we know there exists an M-tuple vector $(\omega_1, \ldots, \omega_m^*, \ldots, \omega_M)$ in set Ω_m if Ω_m is nonempty. In other words, there must exist a vector $(\omega_1^*, \omega_2^*, \ldots, \omega_M^*) \in \Omega_1 \times \Omega_2 \times \cdots \times \Omega_M \triangleq$

---

[4]For example, in a cloud-RAN architecture the core network is able to know the information of all BSs so that it can perform the following proposed centralized traffic management scheme to maximize the per-user network throughput.

TABLE I

NETWORK PARAMETERS FOR SIMULATION

Parameter \ AP Type (Tier #)	Macrocell (1)	Picocell (2)	Femtocell (3)	WiFi (4)
Power P_m (W)	20	1	0.2	0.1
Intensity λ_m (APs/m^2)	5×10^{-6}	5×10^{-5}	2.5×10^{-4}	5×10^{-4}
Maximum Backoff Time τ_m	∞	2		1
Sensing Area $\mathcal{S}_m$ (m^2)	N/A	900π		
CSMA Threshold δ	N/A	4.481		
$H_{m_i} = H_m^{(f)} \times H_m^{(s)}$	$\sim \exp(1,1) \times \ln\mathcal{N}(0,3(\text{dB}))$			
Unlicensed Bandwidth B_U	160 MHz			
Licensed Bandwidth B_L	100 MHz			
Pathloss Exponent α	4			
User Intensity μ (users/Km^2)	500			

$\prod_{m=1}^{M} \Omega_m$ if $\prod_{m=1}^{M} \Omega_m$ is nonempty since all Ω_m's are compact. Let Ω be an M-dimensional closed ball that encloses $\prod_{m=1}^{M} \Omega_m$, i.e., $\prod_{m=1}^{M} \Omega_m \subseteq \Omega$. Hence, Ω is compact as well as nonempty if $\prod_{m=1}^{M} \Omega_m$ is nonempty. Also, $\sum_{m=1}^{M} c_m \vartheta_m$ is continuous over $\prod_{m=1}^{M} \Omega_m$ in that c_m is continuous over Ω_m for all $m \in \mathcal{M}$ and $\sum_{m=1}^{M} c_m \vartheta_m$ is a linear combination of all c_m's, which follows that $\sum_{m=1}^{M} c_m \vartheta_m$ is also continuous over Ω. Since $\sum_{m=1}^{M} c_m \vartheta_m$ is continuous over Ω and Ω is compact, there must exist an optimal M-tuple vector $(\omega_1^*, \omega_2^*, \ldots, \omega_M^*)$ that maximizes the per-user network throughput $\sum_{m=1}^{M} c_m \vartheta_m$ according to the Weierstrass theorem. ∎

Although the optimal solution vector for the optimization problem in (39) cannot be found in closed-form, we can resort to a numerical technique to acquire it. In addition, since the per-user network throughput can also be interpreted as the "mean" per-user link throughput of the HetNet as mentioned before, the solution of the optimization in (39) essentially also maximizes the mean per-user link throughput of the HetNet.

V. NUMERICAL EXAMPLE FOR COEXISTING LTE AND WIFI NETWORKS

In this section, we provide some numerical results by simulating a scenario that there are four tiers in the HetNet consisting of LTE BSs and WiFi APs. The first tier of the HetNet consists of the macrocellcell BSs that do not access the unlicensed spectrum, the second and third tiers consist of picocell and femtocell BSs, and the fourth tier consists of the WiFi APs. The BSs and APs in the last three tiers use the opportunistic CSMA/CA protocol to access the unlicensed spectrum. All channels suffer Rayleigh fading and log-normal shadowing and all users adopt the BMSA scheme defined in
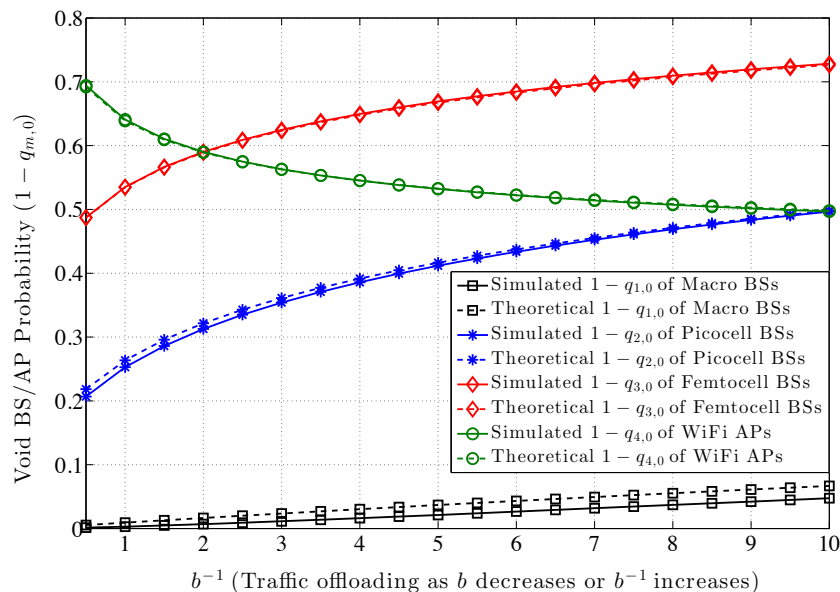
Fig. 1. Simulation results of the void probabilities of the BSs and APs in the four different tiers.

Section II to associate their serving BSs or APs. Specifically, the AP association weight in the scheme (3) for the BSs in the first three tiers is designated as $W_{m_i} = b P_m H_{m_i}^{(s)}$ for $m \in \{1, 2, 3\}$ in which $b > 0$ is a constant bias and $H_{m_i}^{(s)}$ characterizes the channel gain due to log-normal shadowing[5], whereas the AP association weight for the WiFi APs in the fourth tier is $W_{4_i} = H_{4_i}^{(s)}$, i.e., adopting the unbiased MSA scheme. Note that we have $\omega_m = (b P_m \mathbb{E}[H_m^{(s)}])^{\frac{2}{\alpha}}$ for $m \in \{1, 2, 3\}$ and $\omega_4 = (P_4 \mathbb{E}[H_4^{(s)}])^{\frac{2}{\alpha}}$. All the network parameters for simulation are listed in Table I. In the following, we first provide the simulation results of the void BS/AP probability and the channel access probability for the proposed opportunistic CSMA/CA protocol. Next, the simulation results of the mean spectrum efficiencies in the licensed and unlicensed spectra are given to demonstrate the tightness of the derived lower bounds on the mean spectrum efficiencies in Section III. Finally, we show the simulation results of traffic offloading from the LTE network to the WiFi network in order to validate the previous discussions on the per-user link throughput of an AP and per-user network throughput with the decentralized and centralized traffic management schemes.
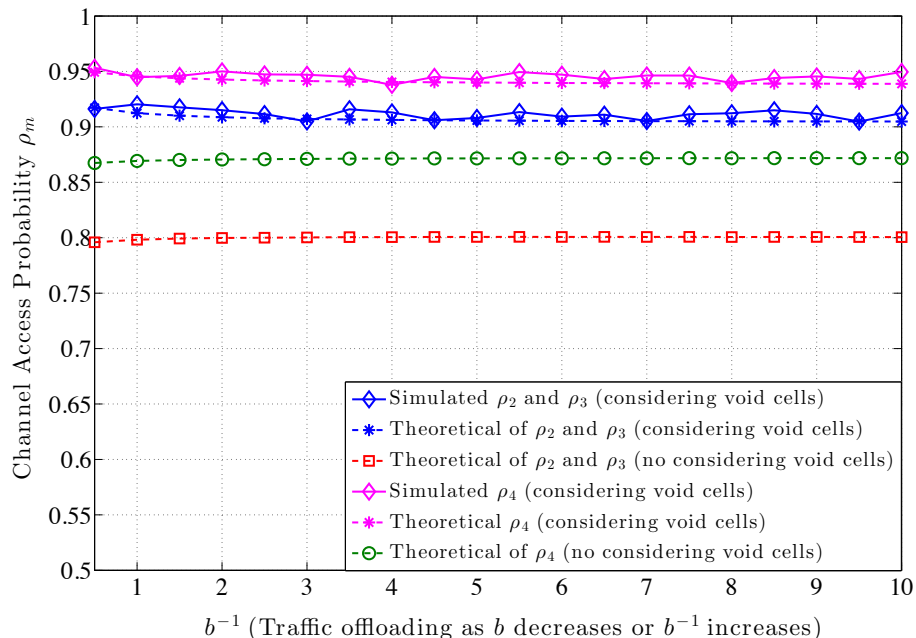
Fig. 2. Simulation results of the channel access probabilities of the BSs and APs in the unlicensed spectrum when the opportunistic CSMA/CA is adopted.

A. Simulation Results of the Void BS/AP Probability and the Channel Access Probability

The void probability of an AP in a particular tier is already derived in (8). To verify the accuracy of (8) and illustrate this void AP issue that cannot be overlooked in our HetNet setting here, the simulation results of the void probabilities of the BSs and APs in all tiers are presented in Fig. 1. As shown in Fig. 1, the theoretical and simulated void probabilities are almost the same for the BSs or APs in any particular tiers, which verifies the correctness and accuracy of the derived pmf in (7) as well as the void AP probability in (8). In addition to the void probability of macrocell BSs, all the void probabilities of the BSs and APs in the last three tiers are actually not small at all since their intensities are not very small compared with the user intensity so that the voidness issue of the dense-deployed BSs or APs should not be carefully considered in modeling and analysis. Also, we can see that all the void probabilities of the BSs in the first three tiers increase whereas the void probability of the WiFi APs decreases while offloading traffic from the LTE network to the WiFi network. Thus, offloading or loading traffic also significantly affects the void probabilities especially for those of BSs and APs with high intensity,

---

[5]In this section, the channel gain H_{m_i} is equal to $H_{m_i}^{(f)} \times H_{m_i}^{(s)}$ where $H_{m_i}^{(f)}$ characterizes the channel gain due to Rayleigh fading and $H_{m_i}^{(s)}$, as already specified, characterizes the channel gain due to log-normal shadowing.

and this gives rise to strong impacts on the mean spectrum efficiency and per-user link throughput as shown in the following subsections. The simulation results of the channel access probabilities of the BSs and APs in all tiers are shown in Fig. 2. They not only verify the correctness of the channel access probability given in (11) but also indicate that overlooking the void cells while modeling the contending channel behaviors among the BSs and APs can also make the channel access probabilities be seriously underestimated. Furthermore, Fig. 2 also shows that all channel access probabilities almost remain unchanged as b^{-1} increases since all small BSs originally can access the unlicensed channel and the intensity of the macrocell BSs is small so that offloading traffic does not make the total intensity of BSs and APs contending the unlicensed channel alter much.

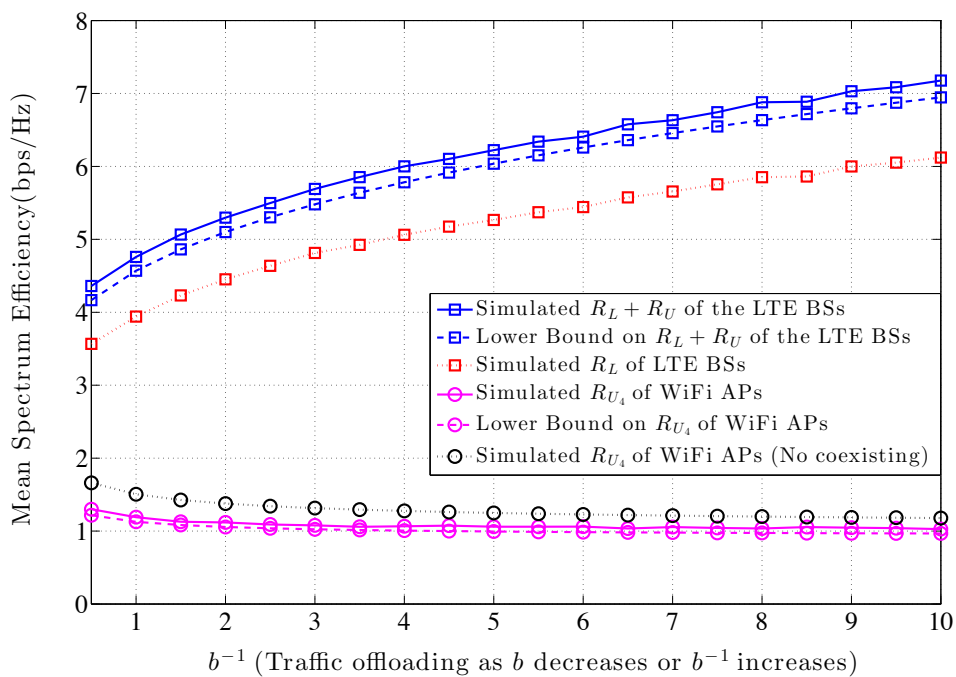*B. Simulation Results of Mean Spectrum Efficiency*

Fig. 3. Simulation results of the mean spectrum efficiencies in licensed and unlicensed spectra.

The simulation results of the mean spectrum efficiencies in the licensed and unlicensed spectra are shown in Fig. 3. Since all BSs in the first three tiers use the BMSA scheme with the same bias, their mean spectrum efficiencies in the licensed spectrum are the same, i.e., $R_{L_1} = R_{L_2} = R_{L_3} = R_L$. Also, note that $R_{U_1} = 0$ since macrocell BSs do not access the unlicensed spectrum, and the BSs in the second and third tiers have the same spectrum efficiency in the unlicensed spectrum, i.e., $R_{U_2} = R_{U_3} = R_U$,
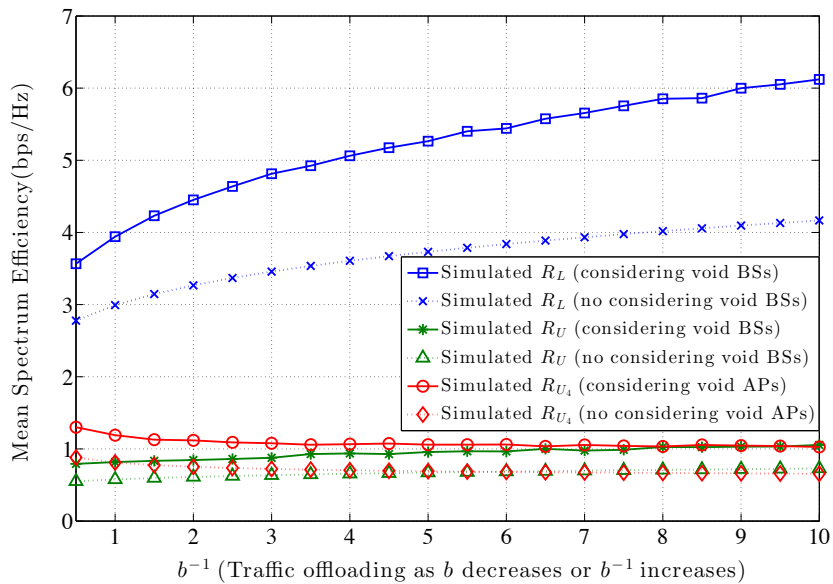
Fig. 4. Simulation results of the mean spectrum efficiencies with and without considering the void cell phenomenon in the interference model of the licensed and unlicensed spectra.

because they have the same channel access probability in the unlicensed spectrum. As a result, the BSs in the second and third tiers have the same sum of the mean spectrum efficiencies in the licensed and unlicensed spectra, i.e., $R_L + R_U$. From Fig. 3, we can gain a few important observations. First, the theoretical lower bound on $R_L + R_U$ is very tight to the simulated result of $R_L + R_U$, and the lower bound on R_{U_4} is also very close to the simulated result of R_{U_4}. Thus, the derived lower bounds in (18) and (24) are fairly tight, as we already emphasized this point in the previous section. Second, when LTE BSs offload their traffic, R_L significantly increases and R_U slightly increases so that $R_L + R_U$ significantly increases, as expected, whereas the mean spectrum efficiency R_{U_4} of the WiFi APs just slightly reduces. Hence, letting LTE small cell BSs and WiFi APs coexist and them share the unlicensed spectrum indeed improves their total mean spectrum efficiency in the unlicensed spectrum. Third, since the mean spectrum efficiency of the WiFi APs just slightly reduces as more traffic is offloaded from the LTE network to the WiFi network, offloading the traffic from the LTE network to the WiFi network is the best traffic management strategy for the BSs in this network setting. In Fig. 4, we show the simulation results of the mean spectrum efficiencies with and without considering the void AP/BS phenomenon in the interference model of the licensed and unlicensed spectra. As can be observed, R_L is significantly underestimated when the void BSs are not modeled in the interference. This validates our previous

claim that the void BSs and APs should be considered in a densely deployed HetNet model. Similarly, in the unlicensed spectrum the mean spectrum efficiencies of the LTE BSs and the WiFi APs are also underestimated when void BSs and APs are not considered. As the traffic is offloaded from the LTE network to the WiFi network, the inaccuracy of the mean spectrum efficiencies of the LTE BSs without considering void BSs is exacerbated in that the void cell probabilities of the LTE BSs increase. On the contrary, the voidness impact on the mean spectrum efficiency of the WiFi network is alleviated since the offloaded traffic helps reduce the void probability of the WiFi APs.

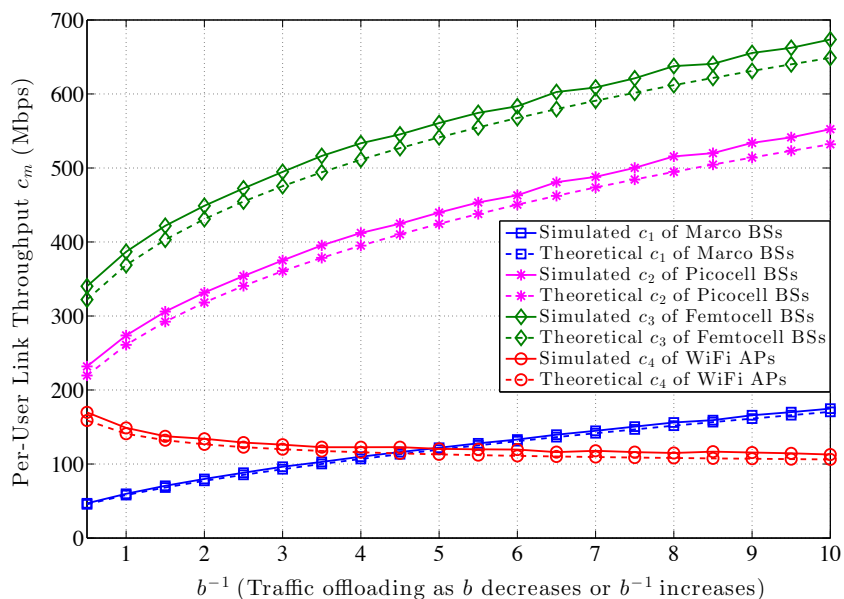*C. Simulation Results of the Per-User Throughput and Per-User Network Throughput*

Fig. 5. Simulation results of the per-user link throughputs of the BSs and APs in the four different tiers. Note that the simulated result is obtained when the decentralized traffic management scheme is performed to update/decrease $(\omega_1, \omega_2, \omega_3)$ (so as to increase $b^{-1} = P_m \mathbb{E}[H_m^{(s)}]/\omega_m^{\frac{\alpha}{2}}$.). The decentralized traffic management scheme actually updates b^{-1} not much as b^{-1} approaches 5 since the per-user link throughput of the WiFi APs slowly reduces to $c_{\min} = 100$ Mbps after $b^{-1} \approx 5$.

The simulation results of the per-user link throughputs of the APs in the four different tiers are shown in Fig. 5 when the decentralized traffic management scheme in (37) is performed. The minimum required per-user link throughput of an WiFi AP is $c_{\min} = 100$ Mbps. Initially, the unlicensed per-user link throughput of the WiFi APs is much higher than $c_{\min}$ so that all BSs start to offload their traffic as shown in Fig. 5. As can be seen, offloading traffic from the LTE network to the WiFi network largely improves the per-user link throughput of the LTE BSs since $\frac{\partial c_m}{\partial \omega_m} < 0$ holds in this context. Although the
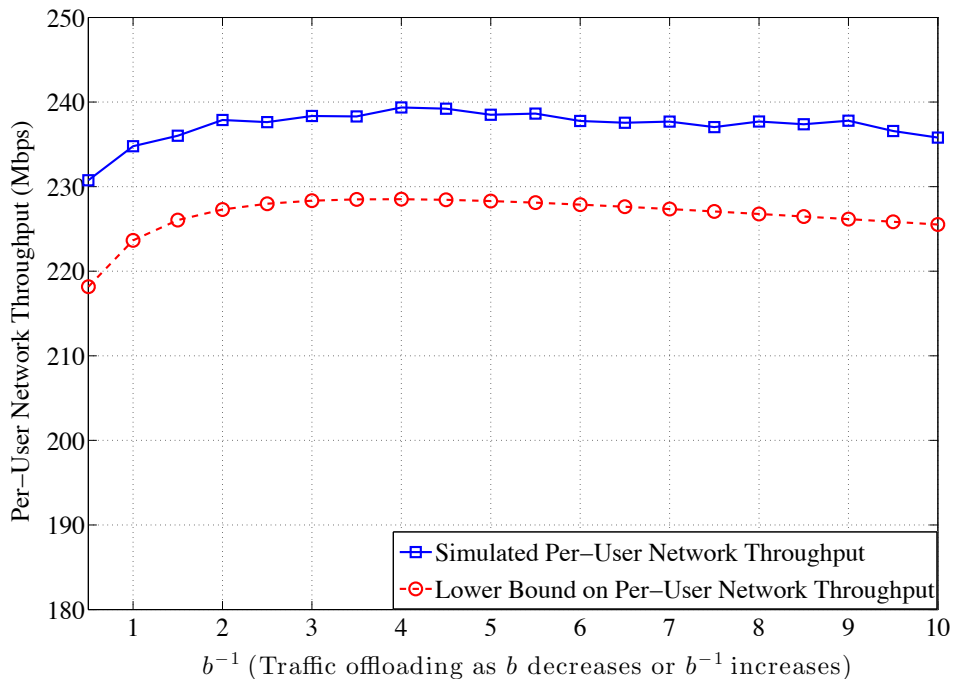
Fig. 6. Simulation results of the per-user network throughputs of the BSs and APs in the four different tiers. Note that the simulated result is obtained when the centralized traffic management scheme is performed to update/decrease $(\omega_1, \omega_2, \omega_3)$ (so as to increase $b^{-1} = P_m \mathbb{E}[H_m^{(s)}]/\omega_m^{\frac{\alpha}{2}}$.). The centralized traffic management scheme actually stops to update b^{-1} as b^{-1} is close to 4 since the per-user network throughput does not increase (starts to reduce) after $b^{-1} \approx 4$.

per-user link throughput of the WiFi APs also reduces, the throughput loss of the WiFi APs is actually not much. Accordingly, *offloading the traffic from LTE to WiFi as much as possible can significantly benefit the per-user link throughput of the LTE BSs as long as $\frac{\partial c_m}{\partial \omega_m} < 0$ holds and the required per-user link throughput of the WiFi APs is maintained*, as expected. However, we should notice that offloading too much traffic from the LTE network to the WiFi network could eventually give rise to the reduction in c_m because R_{L_m} and R_{U_m} could both reduce in this case. Furthermore, Fig. 5 also illustrates that the derived lower bound on c_m is a very tight to the simulated c_m and its accuracy improves as the intensity of the APs is low (see the c_4 curve for the macrocell BSs in the figure). Finally, we provide the simulation results of the per-user network throughput of the HetNet in Fig. 6 assuming the centralized traffic management scheme can be performed. Like the case of the per-user link throughput, offloading traffic from the LTE network to the WiFi network initially improves the per-user network throughput, but it eventually leads to the reduction in the per-user network throughput due to too much offloading. We can see the maximum of the per-user network throughput achieves at $b^{-1} \approx 4$. Hence, there indeed

exists an optimal 4-tuple vector $(\omega_1^*, \omega_2^*, \omega_3^*, \omega_4^*)$ the maximizes the per-user network throughput as shown in Lemma 4.

VI. CONCLUSION

In this paper, we consider an M-tier HetHet in which all APs in any particular tier form an independent PPP and the APs in the first $M - 1$ tiers can simultaneously access the licensed spectrum and use the opportunistic CSMA/CA protocol to share the unlicensed spectrum with the APs in the Mth tier. The distribution of the cell load is studied first since it can be used to find the mean cell load, the association probability of the APs in each tier. Most importantly, it characterizes the void probability of an AP in each tier that impacts the channel access probability of an AP in the unlicensed spectrum as well as the interference model especially in a dense-deployed HetNet. A novel approach is devised to find the tight lower bounds on the mean spectrum efficiencies of an AP in the licensed and unlicensed spectra for any general channel gain and AP association weight models. The per-user link throughput of an AP and the per-user network throughput are proposed and they are used to develop the decentralized and centralized traffic management schemes, respectively. These two traffic management schemes are shown to have the capability of maximizing the per-user link throughput of an AP and per-user network throughput under the constraint posed on the per-user link throughput of the APs in the Mth tier.

APPENDIX

A. Proof of Theorem 1

The Shannon transformation of random variable Z can be rewritten as

$$\mathcal{S}_Z(\eta) = \int_0^1 \mathbb{E}\left[\frac{\eta Z}{1 + y\eta Z}\right] \mathrm{d}y = \int_0^1 \mathbb{E}\left[\frac{1}{1/\eta Z + y}\right] \mathrm{d}y.$$

Since $\mathcal{L}_{Z^{-1}}(s)$ always exists, for any $y \in [0, 1]$ we have

$$\mathbb{E}\left[\frac{1}{1/\eta Z + y}\right] = \int_0^\infty e^{-uy}\mathbb{E}\left[e^{-u/\eta Z}\right] \mathrm{d}u = \int_0^\infty e^{-uy}\mathcal{L}_{Z^{-1}}(u/\eta)\mathrm{d}u = \int_0^\infty e^{-\eta sy}\mathcal{L}_{Z^{-1}}(s)\eta\mathrm{d}s$$

and then substituting this result into $\mathcal{S}_Z(\eta)$ yields

$$\mathcal{S}_Z(\eta) = \eta \int_0^\infty \int_0^1 e^{-\eta sy}\mathcal{L}_{Z^{-1}}(s)\mathrm{d}y\mathrm{d}s = \int_{0+}^\infty \int_0^\infty \frac{(1 - e^{-\eta s})}{se^{\frac{s}{z}}} f_Z(z)\mathrm{d}z\mathrm{d}s \text{ (Letting } s = \eta z).$$

$$= \int_{0+}^\infty \frac{(1 - e^{-\eta s})}{s}\mathcal{L}_{Z^{-1}}(s)\mathrm{d}s,$$

which is exactly the result in (14). The result in (15) readily follows from (14) and the definition of the Laplace transform of a nonnegative random variable.

B. Proof of Theorem 2

We first show the case that all W_{m_i}'s are random. Let $\Delta_{\mathsf{U}_m} = P_m H_m / I_{\mathsf{U}_m} \|X_o\|^\alpha$ be the signal-to-interference ratio (SIR) of the typical user when AP X_o is from the m-th tier. Thus, $R_{\mathsf{U}_m} = \mathbb{E}[\log_2(1 + T_m \Xi_m \Delta_{\mathsf{U}_m})] = \xi_m \rho_m \mathbb{E}[\log_2(1 + \Delta_{\mathsf{U}_m})]$ and by using the identity of the Shannon transform in Theorem 14 it can be expressed as

$$R_{\mathsf{U}_m} = \frac{\xi_m \rho_m}{\ln(2)} \mathbb{E}[\ln(1 + \Delta_{\mathsf{U}_m})] = \frac{\xi_m \rho_m}{\ln(2)} \mathcal{S}_{\Delta_{\mathsf{U}_m}}(1) = \frac{\xi_m \rho_m}{\ln(2)} \int_0^\infty \int_0^1 e^{-sy} \mathcal{L}_{\Delta_{\mathsf{U}_m}^{-1}}(s) \mathrm{d}y \mathrm{d}s \qquad (41)$$

The Laplace transform of $\Delta_{\mathsf{U}_m}^{-1}$ can be further explicitly expressed as

$$\mathcal{L}_{\Delta_{\mathsf{U}_m}^{-1}}(s) = \mathbb{E}\left[\exp\left(-\sum_{X_{k_i} \in \bigcup_{k=1}^M \mathcal{X}_k \backslash X_o} s \frac{P_k H_{k_i} V_{k_i} T_{k_i} \Xi_{k_i} \|X_o\|^\alpha}{P_m H_m \|X_{k_i}\|^\alpha}\right)\right]$$

$$\overset{(a)}{=} \mathbb{E}\left[\exp\left(-\sum_{\widetilde{X}_{k_i} \in \bigcup_{k=1}^M \widetilde{\mathcal{X}}_k \backslash \widetilde{X}_o} s \frac{P_k H_{k_i} V_{k_i} T_{k_i} \Xi_{k_i} W_m \|\widetilde{X}_o\|^\alpha}{P_m H_m W_{k_i} \|\widetilde{X}_{k_i}\|^\alpha}\right)\right]$$

$$\overset{(b)}{=} \mathbb{E}\left[\exp\left(-\sum_{\widetilde{X}_{k_i} \in \bigcup_{k=1}^M \widetilde{\mathcal{X}}_k \backslash \widetilde{X}_o} S_{m,k_i} \frac{\|\widetilde{X}_o\|^\alpha}{\|\widetilde{X}_{k_i}\|^\alpha}\right)\right]$$

where $\widetilde{X}_o$, $\widetilde{X}_{m_i}$ and $\widetilde{\mathcal{X}}_m$ are all already defined in the Proof of Lemma 1, (a) follows from the result in Lemma 1 and (b) is obtained by letting $S_{m,k_i} \triangleq \frac{s P_k H_{k_i} W_m}{P_m H_m W_{k_i}} V_{k_i} T_{k_i} \Xi_{k_i}$. Now be aware that all T_{k_i}'s are independent since we assume all channel gains are independent, all V_{k_i}'s are not completely and mutually independent since cell association could induce the location correlations between the non-void APs [22] [23], all Ξ_{k_i}'s are not completely and mutually independent as well since all non-void APs use CSMA/CA to access the unlicensed channel and the resulting APs accessing the channel is not a PPP any more and instead they become a Martén hard-core point process (MHPP) [29] [9].

A tractable lower bound on $\mathcal{L}_{\Delta_{\mathsf{U}_m}^{-1}}$ can be obtained by treating all V_{k_i}'s (Ξ_{k_i}'s) are independent and i.i.d. for same k so that all non-void APs that have good channels and successfully access to unlicensed channel become a thinning homogeneous PPP which induces larger interference. Accordingly, we have

$$\mathcal{L}_{\Delta_{\mathsf{U}_m}^{-1}}(s) \geq \mathbb{E}\left[\exp\left\{-\pi \sum_{k=1}^M \widetilde{\lambda}_k \mathbb{E}_{S_{m,k}}\left[\int_0^\infty \left(1 - e^{-S_{m,k}\left(1 + \frac{x}{\|\widetilde{X}_o\|^2}\right)^{-\frac{\alpha}{2}}}\right) \mathrm{d}x\right]\right\}\right]$$

by following the proof of Proposition 2 in [23] since $\widetilde{X}_o$ is the nearest point in the point process of $\bigcup_{m=1}^M \widetilde{X}_m$ to the typical user. Also, letting Y be an exponential random variable with unit mean, i.e.,

$Y \sim \exp(1)$, leads to the following results:

$$\int_0^\infty \left(1 - e^{-S_{m,k}\left(1 + \frac{x}{\|\widetilde{X}_o\|^2}\right)^{-\frac{\alpha}{2}}}\right) \mathrm{d}x = \int_0^\infty \mathbb{P}\left[Y \leq S_{m,k}\left(1 + \frac{x}{\|\widetilde{X}_o\|^2}\right)^{-\frac{\alpha}{2}}\right] \mathrm{d}x$$

$$= \int_0^\infty \mathbb{P}\left[1 + \frac{x}{\|\widetilde{X}_o\|^2} \leq \left(\frac{S_{m,k}}{Y}\right)^{\frac{2}{\alpha}}\right] \mathrm{d}x$$

$$= \|\widetilde{X}_o\|^2 \left(\int_0^\infty \mathbb{P}\left[u \leq \left(\frac{S_{m,k}}{Y}\right)^{\frac{2}{\alpha}}\right] \mathrm{d}u - \int_0^1 \mathbb{P}\left[u \leq \left(\frac{S_{m,k}}{Y}\right)^{\frac{2}{\alpha}}\right] \mathrm{d}u\right)$$

$$= \|\widetilde{X}_o\|^2 \left[S_{m,k}^{\frac{2}{\alpha}}\Gamma\left(1 - \frac{2}{\alpha}\right) + \int_0^1 e^{-S_{m,k}u^{-\frac{\alpha}{2}}} \mathrm{d}u - 1\right].$$

Hence, we have

$$\mathbb{E}_{S_{m,k}}\left[\int_0^\infty \left(1 - e^{-S_{m,k}\left(1 + \frac{x}{\|\widetilde{X}_o\|^2}\right)^{-\frac{\alpha}{2}}}\right) \mathrm{d}x\right] = \|\widetilde{X}_o\|^2(1 - p_{k,0})\xi_k\rho_k \times$$

$$\left\{\left(\frac{sP_k\bar{w}_m\bar{h}_k}{P_m\widehat{H}\bar{w}_k\bar{h}_m}\right)^{\frac{2}{\alpha}} \mathbb{E}\left[\left(\frac{H_k/\bar{h}_k}{W_k/\bar{w}_k}\right)^{\frac{2}{\alpha}}\right]\Gamma\left(1 - \frac{2}{\alpha}\right) + \int_0^1 \mathcal{L}_{\widehat{H}}\left(\frac{sP_k\bar{w}_m\bar{h}_k u^{-\frac{\alpha}{2}}}{P_m\bar{w}_k\bar{h}_m\widehat{H}}\right) \mathrm{d}u - 1\right\}$$

$$= \|\widetilde{X}_o\|^2(1 - p_{k,0})\xi_k\rho_k \ell_{\widehat{H}}\left(\frac{P_k\bar{w}_m\bar{h}_k s}{P_m\bar{w}_k\bar{h}_m\widehat{H}}, \frac{2}{\alpha}\right)$$

and this follows that

$$\mathcal{L}_{\Delta_{\mathsf{U}_m}^{-1}}(s) \geq \mathbb{E}_{\widehat{H}}\left[\exp\left\{-\pi\|\widetilde{X}_o\|^2\sum_{k=1}^M q_{k,0}\xi_k\rho_k\widetilde{\lambda}_k\ell_{\widehat{H}}\left(\frac{sP_k\bar{w}_m\bar{h}_k}{P_m\bar{w}_k\bar{h}_m\widehat{H}}, \frac{2}{\alpha}\right)\right\}\right]$$

$$= \mathbb{E}_{\widehat{H}}\left[\int_0^\infty \pi\widetilde{\lambda}\exp\left\{-\pi x\widetilde{\lambda}\left(\sum_{k=1}^M q_{k,0}\xi_k\rho_k\vartheta_k\ell_{\widehat{H}}\left(\frac{sP_k\bar{w}_m\bar{h}_k}{P_m\bar{w}_k\bar{h}_m\widehat{H}}, \frac{2}{\alpha}\right) + 1\right)\right\} \mathrm{d}x\right]$$

$$= \mathbb{E}_{\widehat{H}}\left[\left(\sum_{k=1}^M q_{k,0}\xi_k\rho_k\vartheta_k\ell_{\widehat{H}}\left(\frac{sP_k\bar{w}_m\bar{h}_k}{P_m\bar{w}_k\bar{h}_m\widehat{H}}, \frac{2}{\alpha}\right) + 1\right)^{-1}\right].$$

Then we know

$$\int_0^\infty \mathcal{L}_{\Delta_{\mathsf{U}_m}^{-1}}(s)\mathrm{d}s \geq \int_0^\infty \int_0^\infty \frac{e^{-sy}f_{\widehat{H}}(x)}{\sum_{k=1}^M q_{k,0}\xi_k\rho_k\vartheta_k\ell_{\widehat{H}}\left(\frac{sP_k\bar{w}_m\bar{h}_k}{xP_m\bar{w}_k\bar{h}_m}, \frac{2}{\alpha}\right) + 1}\mathrm{d}s\mathrm{d}x$$

$$= \int_0^\infty \frac{\int_0^\infty e^{-uxy}f_{\widehat{H}}(x)x\mathrm{d}x}{\sum_{k=1}^M q_{k,0}\xi_k\rho_k\vartheta_k\ell_{\widehat{H}}\left(\frac{P_k\bar{w}_m\bar{h}_k}{P_m\bar{w}_k\bar{h}_m}u, \frac{2}{\alpha}\right) + 1}\mathrm{d}u.$$

According to (41), R_{U_m} can be lower bounded as

$$R_{\mathsf{U}_m} = \frac{\xi_m\rho_m}{\ln(2)}\int_0^\infty \int_0^1 e^{-sy}\mathcal{L}_{\Delta_{\mathsf{U}_m}^{-1}}(s)\mathrm{d}y\mathrm{d}s \geq \frac{\xi_m\rho_m}{\ln(2)}\int_0^\infty \frac{\int_0^\infty \left(\int_0^1 e^{-uxy}\mathrm{d}y\right)f_{\widehat{H}}(x)x\mathrm{d}x}{\sum_{k=1}^M q_{k,0}\xi_k\rho_k\vartheta_k\ell_{\widehat{H}}\left(\frac{P_k\bar{w}_m\bar{h}_k}{P_m\bar{w}_k\bar{h}_m}u, \frac{2}{\alpha}\right) + 1}\mathrm{d}u,$$

which is exactly the lower bound in (18) by carrying out those inner double integrals for variables x and y. For the case that all W_{m_i}'s are constant $\bar{w}_m$ for all $m \in \mathcal{M}$, letting all W_m leave in function $\ell(\cdot, \cdot)$ yields the result in (20) with $\widehat{H} = H_m/\bar{h}_m$ for all $m \in \mathcal{M}$.

C. Proof of Theorem 4

First notice that the coefficient of $\omega_m(n)$ in (37) has the physical meaning that $c^*(n)$ is normalized by the per-user link throughput of a tier-m AP in the unlicensed spectrum. The tier-m APs will offload (load) its traffic in the next time if this coefficient at time n is smaller (greater) than one. In other words, the per-user link throughput of the tier-m APs can continuously increase (decrease) under the traffic management scheme in (37) as long as it is smaller (larger) than $c^*(n)$ for all $n \in \mathbb{N}$. Now we want to show that $\omega_m(n)$ converges to a steady-state value ω_m^* as n goes to infinity. Consider a Lyapunov function $V(n) \triangleq (\omega_m(n) - \omega_m^*/2)^2$ and we can have the following result

$$
\begin{aligned}
V(n+1) - V(n) &= (\omega_m(n+1) - \omega_m^*/2)^2 - (\omega_m(n) - \omega_m^*/2)^2 \\
&= [\omega_m(n+1) - \omega_m(n)] [\omega_m(n+1) + \omega_m(n) - \omega_m^*] \\
&= \left[\frac{c^*(n)N_m(n)}{C_m(n)} - 1 \right] \left[\frac{c^*(n)N_m(n)}{C_m(n)} + 1 - \frac{\omega_m^*}{\omega_m(n)} \right] \omega_m^2(n) < 0,
\end{aligned}
$$

which yields the following two constraints

$$
1 < \frac{c^*(n)N_m(n)}{C_m(n)} < \frac{\omega_m^*}{\omega_m(n)} - 1 \quad \text{and} \quad \frac{\omega_m^*}{\omega_m(n)} - 1 < \frac{c^*(n)N_m(n)}{C_m(n)} < 1.
$$

As long as $w_m(n)$ and $\frac{c^*(n)N_m(n)}{C_m(n)}$ satisfy these two constraints, $\omega_m(n)$ converges to ω_m^* as n goes to infinity based on the Foster-Lyapunov criterion [30]. In other words, as n goes to infinity, $c^*(n)$ will converge to $\sup_n\{C_m(n)/N_m(n)\}$. Also, these two constraints always hold since $\frac{c^*(n)N_m(n)}{C_m(n)} < 1$ (i.e. $C_m/N_m(n) > c^*(n)$) makes $\omega_m(n)$ reduce and approach to ω_m^* if $\omega_m(n) > \omega_m^*/2$ and $\frac{c^*(n)N_m(n)}{C_m(n)} > 1$ (i.e. $C_m/N_m(n) > c^*(n)$) makes $\omega_m(n)$ reduce and approach to ω_m^* if $\omega_m(n) < \omega_m^*/2$. Therefore, the decentralized traffic management scheme in (37) will make $\omega_m(n)$ converge to ω_m^* as time goes to infinity and thus we must have $0 < \frac{c^*(n)N_m(n)}{C_m(n)} < 1$, i.e., $C_m(\omega_m^*)/N_m(\infty) > c^*(\infty)$ for all $m \in \mathcal{M}$, and $c_M \geq c_{\min}$ is surely satisfied. In addition, since $N_m(\infty) = \mu\vartheta_m(\omega_m^*)/\lambda_m$, in the steady state ω_m^* must satisfy the following constraint

$$
\frac{\mu c^* \vartheta_m(\omega_m^*)}{\lambda_m C_m(\omega_m^*)} = 1 \Rightarrow \omega_m^* = \frac{C_m(\omega_m^*)}{\mu c^*} \left(\sum_{k \in \mathcal{M}\backslash m} \lambda_k \omega_k + \lambda_m \omega_m^* \right),
$$

which indicates ω_m^* is the fixed point of $\Upsilon_m(x)$ in (38).

REFERENCES

[1] M. Bennis, M. Simsek, A. Czylwik *et al.*, "When cellular meets WiFi in wireless small cell networks," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 44–50, Jun. 2013.

[2] H. Zhang, X. Chu, W. Guo, and S. Wang, "Coexistence of Wi-Fi and heterogeneous small cell networks sharing unlicensed spectrum," *IEEE Commun. Mag.*, vol. 53, no. 3, pp. 158–164, Mar. 2015.

[3] R. Zhang, M. Wang, L. X. Cai, Z. Zheng, X. Shen, and L.-L. Xie, "LTE-unlicensed: the future of spectrum aggregation for cellular networks," *IEEE Wireless Commun. Mag.*, vol. 22, no. 3, pp. 150–159, Jun. 2015.

[4] R. Etkin, A. Parekh, and D. Tse, "Spectrum sharing for unlicensed bands," *IEEE J. Select. Areas Commun.*, vol. 25, no. 3, pp. 517 – 528, Apr. 2007.

[5] H. Yomo, P. Popovski, H. C. Nguyen, and R. Prasad, "Adaptive frequency rolling for coexistence in the unlicensed band," *IEEE Trans. Wireless Commun.*, vol. 6, no. 2, pp. 598–608, Oct. 2007.

[6] Y. M. Shobowale and K. A. Hamdi, "A unified model for interference analysis in unlicensed frequency bands," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4004–4013, Aug. 2009.

[7] J. B. Ernst, N. Nasser, and J. Rodrigues, "Co-channel interference modelling between RATs in heterogeneous wireless networks," in *IEEE Int. Conf. on Commun.*, Jun. 2012, pp. 5321–5325.

[8] J. Jeon, Q. Li, H. Niu, A. Papathanassiou, and G. Wu, "LTE in the unlicensed spectrum: A novel coexistence analysis with WLAN systems," in *IEEE Global Commun. Conf.*, Dec. 2014, pp. 3459–3464.

[9] C.-H. Liu and H.-C. Tsai, "On the limits of coexisting coverage and capacity in multi-RAT heterogeneous networks," *submitted to IEEE Trans. Wireless Commun.*, 2016. [Online]. Available: http://arxiv.org/abs/1602.02250

[10] X. Ding, C.-H. Liu, L.-C. Wang, and X. Zhao, "Coexisting success and throughput of multi-rat wireless networks with unlicensed band access," *IEEE Wireless Comm. Letters*, vol. 5, no. 1, pp. 4–7, Feb. 2016.

[11] H.-C. Tsai, C.-H. Liu, and L.-C. Wang, "An analytical approach to coexisting evaluation in Multi-RAT heterogeneous networks with opportunistic CSMA/CA," in *IEEE Int. Conf. on Commun.*, May 2016, pp. 1–6.

[12] Q. Chen, G. Yu, H. Shan, A. Maaref, G. Y. Li, and A. Huang, "Cellular meets WiFi: Traffic offloading or resource sharing?" *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3354–3367, May 2016.

[13] A. R. Elsherif, W.-P. Chen, A. Ito, and Z. Ding, "Resource allocation and inter-cell interference management for dual-access small cells," *IEEE J. Select. Areas Commun.*, vol. 33, no. 6, pp. 1082–1096, Jun. 2015.

[14] F. Liu, E. Bala, E. Erkip, M. C. Beluri, and R. Yang, "Small-cell traffic balancing over licensed and unlicensed bands," *IEEE Trans. Veh. Technol.*, vol. 64, no. 12, pp. 5850–5865, Dec. 2015.

[15] A. Bhorkar, C. Ibars, and P. Zong, "Performance analysis of LTE and WiFi in unlicensed band using stochastic geometry," in *IEEE Globecom Workshop on Heterogeneous and Small Cell Networks*, Sep. 2014, pp. 1310–1314.

[16] S. Sagari, I. Seskar, and D. Raychaudhuri, "Modeling the coexistence of LTE and WiFi heterogeneous networks in dense deployment scenarios," in *IEEE ICC workshop on LTE in unlicensed bands: potentials and challenges*, Jun. 2015, pp. 2301–2306.

[17] C.-H. Liu and L.-C. Wang, "Modeling and analysis of coexisting multiple radio access technologies in heterogeneous wireless networks," in *IEEE Int. Conf. on Computing, Networking and Commun.*, Feb. 2016, pp. 1–5.

[18] Y. Li, F. Baccelli, J. G. Andrews, T. D. Novlan, and J. C. Zhang, "Modeling and analyzing the coexistence of Wi-Fi and LTE in unlicensed spectrum," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6310–6326, Dec. 2016.

[19] X. Wang, T. Q. S. Quek, M. Sheng, and J. Li, "Throughput and fairness analysis of Wi-Fi and LTE-U in unlicensed band," *IEEE J. Select. Areas Commun.*, vol. 35, no. PP, accepted and to appear in 2017.

[20] 3GPP TR 36.889 v13, "Study on licensed-assisted access to unlicensed spectrum," May 2015.

[21] D. Stoyan, W. Kendall, and J. Mecke, *Stochastic Geometry and Its Applications*, 2nd ed. New York: John Wiley and Sons, Inc., 1996.

[22] C.-H. Liu and L.-C. Wang, "Random cell association and void probability in poisson-distributed cellular networks," in *IEEE Int. Conf. on Commun.*, Jun. 2015, pp. 2816–2821.

[23] ——, "Optimal cell load and throughput in green small cell networks with generalized cell association," *IEEE J. Select. Areas Commun.*, vol. 34, no. 5, pp. 1058–1072, May 2016.

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

[25] C.-H. Liu and K. L. Fong, "Fundamentals of the downlink green coverage and energy efciency in heterogeneous networks," *IEEE J. Select. Areas Commun.*, vol. 34, no. 12, pp. 3271–3287, Dec. 2016.

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

[27] P. Xia, C.-H. Liu, and J. G. Andrews, "Downlink coordinated multi-point with overhead modeling in heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 4025–4037, Jun. 2013.

[28] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA: Athena Scientific, 1999.

[29] F. Baccelli and B. Błaszczyszyn, "Stochastic geometry and wireless networks: Volume II Applications," *Foundations and Trends in Networking*, vol. 3, no. 3-4, pp. 249–449, 2010.

[30] S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability*. New York: Springer-Verlag, 1993.