Non-orthogonal Multiple Access in Large-Scale Hybrid Heterogeneous Networks

Yuanwei Liu, *Member, IEEE,* Zhijin Qin, *Member, IEEE,* Maged Elkashlan, *Member, IEEE,* Arumugam Nallanathan, *Fellow, IEEE,* and Julie A. McCann, *Member, IEEE*,

***Abstract*—The potential benefits of applying non-orthogonal multiple access (NOMA) technique into K-tier hybrid heterogeneous networks (HetNets) is exploited. A new promising transmission framework is proposed, in which NOMA technique is adopted in small cells and massive multiple-input multiple-output (MIMO) is employed in macro cells. For maximizing the biased average received power at mobile users, a NOMA and massive MIMO based user association scheme is developed. In an effort to evaluate the performance of the proposed framework, we first derive the analytical expressions for the coverage probability of NOMA enhanced small cells. We then examine the spectrum efficiency of the whole proposed networks, with deriving exact analytical expressions for NOMA enhanced small cells and a tractable lower bound for massive MIMO enabled macro cells. Lastly, we investigate the energy efficiency of the hybrid HetNets with the aid of a poplar energy consumption model. Our results confirm that: 1) The coverage probability of NOMA enhanced small cells is affected to a large extent by the targeted transmit rates and power sharing coefficients of two NOMA users; 2) Massive MIMO enabled macro cells are capable of significantly enhancing the spectrum efficiency by increasing the number of antennas; 3) The energy efficiency of the whole networks can be greatly improved by densely deploying NOMA enhanced small cell base stations (BSs); and 4) The proposed NOMA enhanced HetNets transmission scheme has superior performance compared to the orthogonal multiple access (OMA) based HetNets.**

***Index Terms*—HetNets, massive MIMO, NOMA, user association, stochastic geometry**

I. Introduction

The last decade has witnessed the escalating data explosion on the mobile Internets [1], which is brought by the emerging demanding applications such as high-definition videos, online games and virtue reality. Also, the rocket development of internet of things (IoT) requires to solve the user access issue for connecting hundreds/thousands devices simultaneously [2]. Such requirements pose new challenges for designing the fifth-generation (5G) networks. Driven by tackling these challenges, non-orthogonal multiple access (NOMA), as a promising technology in 5G networks, has attracted much attention for its potential ability for enhancing spectrum efficiency and improving the user access capability [3]. The key idea of NOMA[1] is to utilize superposition coding (SC) technique at the transmitter and successive interference cancelation (SIC) technique at the receiver [4], hence multiple access can be realized in power domain via different power levels. Some initial research investigations has been made in this field [5–8]. The system-level performance of a two user NOMA system in terms of downlink transmission was demonstrated in [5]. In [6], the performance of a general NOMA transmission was evaluated in which one BS is able to communicate with several spatial randomly deployed users. By examining appropriate power allocation policies among the NOMA users, the fairness issue of NOMA was addressed in [7]. For multi-antenna NOMA systems, a two-stage multicast beamforming downlink transmission scheme was proposed in [8], where the total transmitting power was optimized with providing closed-form expressions.

Heterogeneous networks (HetNets) and massive multiple-input multiple-output (MIMO), as two"big three" technologies [9], laid the fundamental structure for emerging 5G communication systems. The core idea of HetNets is to establish closer BS-user link by densely overlaying small cells. By doing so, promising benefits such as lower power consumption, higher throughput and enhanced spatial reuse of spectrum can be experienced [10]. The massive MIMO regime enables to equip tens of hundreds/thousands antennas at a BS, and hence is capable of offering an unprecedented level of freedom to serve multiple mobile users [11]. Aiming to fully take advantages of both massive MIMO and HetNets, in [12], the interference coordination issue of massive MIMO enabled HetNets was addressed by utilizing the spatial blanking of macro cells. In [13], the authors investigated a joint user association and interference management optimization problem in massive MIMO HetNets.

A. Motivation and Related Works

Sparked by the aforementioned potential benefits, it is promising to explore the potential performance enhancement brought by NOMA for HetNets. Stochastic geometry is an effective mathematical tool for capturing the topological randomness of networks. As such, it is capable of providing tractable analytical results in terms of average network behaviors [14]. Somewhat related performance evaluation research contributions with utilizing stochastic geometry approaches

Y. Liu and A. Nallanathan are with the Department of Informatics, King's College London, London WC2R 2LS, U.K. (email: {yuanwei.liu, arumugam.nallanathan}@kcl.ac.uk).

Z. Qin and J. McCann are with the Department of Computing, Imperial College London, London SW7 2AZ, U.K. (email: {z.qin, jamm}@imperial.ac.uk).

M. Elkashlan is with the School of Electronic Engineering and Computer Science, Queen Mary University of London, London E1 4NS, U.K. (email: maged.elkashlan@qmul.ac.uk).

[1]In this treatise, we use "NOMA" to refer to "power-domain NOMA" for simplicity.

have been studied in the context of Hetnets and NOMA [15–21]. For Hetnets scenarios, based on applying a flexible bias-allowed user association approach, the performance of multi-tier downlink HetNets was examined in [15], where all base stations (BSs) and users were assumed to equip with single antenna. As a further advance, the coverage provability of the multi-antenna enabled heterogeneous networks was investigated in [16], with using a simple selection bias based cell selection policy. By utilizing massive MIMO enabled HetNets stochastic geometry model, the spectrum efficiency of uplink and downlink were evaluated in [17] and [18], respectively.

Regarding the literature of stochastic geometry based NOMA scenarios, on the standpoint of tackling spectrum and energy issues, an incentive user cooperation NOMA protocol was proposed in [19], by regarding near users as energy harvesting relays for improving reliability of far users. With utilizing signal alignment technology, a new MIMO-NOMA design framework was proposed in a stochastic geometry based model [20]. Driven by dealing with the security issues, in [21] two effective approaches—protection zone and artificial noise were utilized for enhancing the physical layer security of NOMA in large-scale networks. Very recently, the potential co-existence of two hot technologies, NOMA and millimeter wave (mmWave) were examined in [22], with the aid of random beamforming design, where the locations of users are randomly deployed.

Despite the fact that there are ongoing research contributions having played a vital role for fostering HetNets and NOMA technologies with invoking stochastic geometry tools, to the best of our knowledge, the impact of NOMA enhanced hybrid HetNets design has not been researched yet and is still in its infancy. Also, there is lack of complete systematic performance evaluation metrics, i.e., coverage probability and energy efficiency. Different from the previous conventional HetNets design [15, 17, 18], note that considering NOMA enhanced HetNets design poses three mainly additional challenges: i) It brings additional co-channel interference from the superposed signal of the connected BS; ii) NOMA technology requires carefully channel ordering design for carrying out SIC operations at the receiver; and iii) The user association policy requires taking consideration of the effect of power sharing affected by NOMA. Aiming for tackling the aforementioned issues, developing a systematic mathematically tractable framework for intelligently investigating the effect of various types of interference on network performance is more than desired, which motivates us to contribute this treatise.

B. Contributions

We propose a new hybrid HetNets framework with NOMA enhanced small cells and massive MIMO aided macro cells. We believe that the novel structure design can be a new highly rewarding candidate, which will contribute to the design of a more promising 5G system due to the following key advantages:

- High spectrum efficiency: In NOMA enhanced HetNets, with employing higher BS densities, the NOMA enhanced BSs are capable of accessing the served users closer, which increase the transmit signal-to-interference-plus-noise ratio (SINR) by intelligently tracking the multi-category interference, such as inter/intra-tier interference and intra-BS interference.
- High compatibility and low complexity: NOMA is regarded as a promising "add-on" technology for the existing multiple access systems due to the gradually mature of SC and SIC technologies, and will not bring much implementation complexity. Additionally, with applying NOMA in the single-antenna based small cells, the complex cluster based precoding/detection design for MIMO-NOMA systems [23, 24] can be avoided.
- Fairness/throughput tradeoff: NOMA is capable of dealing with the fairness issue by allocating more power to weak users [4], which is of great significance for HetNets when investigating efficient resource allocation in the sophisticated large-scale multi-tier networks.

Different from the existing stochastic geometry based single cell research contributions in terms of NOMA [6, 19–22], we consider multi-cell multi-tier scenarios in this treatise, which is more challenging. In this framework, we consider a downlink K-tier HetNets, where macro BSs are equipped with large antenna arrays with linear zero-forcing beamforming (ZFBF) capability to serve multiple single-antenna users simultaneously, and small cells BSs are equipped with single antenna each to serve two single-antenna users simultaneously with NOMA transmission. Based on the proposed design, the primary theoretical contributions are summarized at least four folds as follows:

1) We develop a flexible biased association policy to address the impact of NOMA and massive MIMO on the maximum biased received power. With utilizing this policy, we first derive the exact analytical expressions for the coverage probability of a typical user associating to the NOMA enhanced small cells for the most general case. Additionally, we derive closed-form expressions in terms of coverage probability for the interference-limited case that each tier has the same path loss.
2) We derive the exact analytical expressions of the NOMA enhanced small cells in term of spectrum efficiency. Regarding the massive MIMO enabled macro cells, we provide a tractable analytical lower bound for the most general case and closed-form expressions for the case that each tier has the same path loss. Our analytical results illustrate that the spectrum efficiency can be greatly enhanced by increasing the scale of large antenna arrays.
3) We finally derive the energy efficiency of the whole networks with applying a popular power consumption model [25]. Our results reveal that NOMA enhanced small cells achieve high energy efficiency than macro cells. It is also shown that increasing antenna number at the macro cell BSs has the opposite effect on energy efficiency.
4) We show that NOMA enhanced small cells design has superior performance gains over conventional orthogonal multiple access (OMA) based small cells in

terms of coverage probability, spectrum efficiency and energy efficiency, which demonstrates the benefits of the proposed framework.

C. Paper Organization

The rest of the paper is organized as follows. In Section II, the network model for NOMA enhanced hybrid HetNets structure is introduced. In Section III, new analytical expressions for the coverage probability of the NOMA enhanced small cells are derived. Then spectral efficiency and energy efficiency are investigated in Section IV and Section V, respectively. Numerical results are presented in Section VI, which is followed by the conclusions in Section VI.

II. Network Model

A. Network Description

We focus on the downlink transmission scenarios. We consider a K-tier HetNets model, where the first tier represents the macro cells and the other tiers represent the small cells such as pico cells and femto cells. The positions of macro BSs and all the k-th tier ($k \in \{2, \cdots, K\}$) BSs are modeled as homogeneous poisson point processes (HPPPs) Φ_1 and Φ_k and with density λ_1 and λ_k, respectively. Motivated by the fact that it is common to overlay a high-power macro cell with successively denser and lower power small cell, we consider to apply massive MIMO technologies to macro cells and NOMA transmission to small cells in this work. As shown in Fig. 1, in macro cells, macro BSs are considered to be equipped with M antennas, each macro BS transmit signals to N users over the same resource block (e.g., time/frequency/code). We assume $M \gg N > 1$ and the linear ZFBF technology is applied at each macro BS with assigning equal power to N data streams [26]. In small cells, each small cell BS is considered to be equipped with single antenna. Such structure consideration is to avoid sophisticated MIMO-NOMA precoding/detection design in small cells. All users are considered to be equipped with single antenna each as well. We consider to adopt user pairing in each tier of small cells to implement NOMA for lowering the system complexity [19]. It is worth pointing out that in Long term evolution advanced (LTE-A), NOMA is also in a form of two-user case [27].

B. NOMA and Massive MIMO Based User Association

In this work, a user is allowed to access any tier BS, which can provide the best coverage. We consider that the flexible user association is based on the maximum average received power of each tier.

1) Average received power in NOMA enhanced small cells: Different from the convectional user association of OMA based small cells, NOMA exploits the power sparsity for multiple access by allocating different power to different users. Due to the random spatial topology of our stochastic geometry model, the space information of users are not pre-determined. Our user association policy for the NOMA enhanced small cells is based on assuming the typical user as near user first. As such, in the i-th tier small cell, the averaged received power

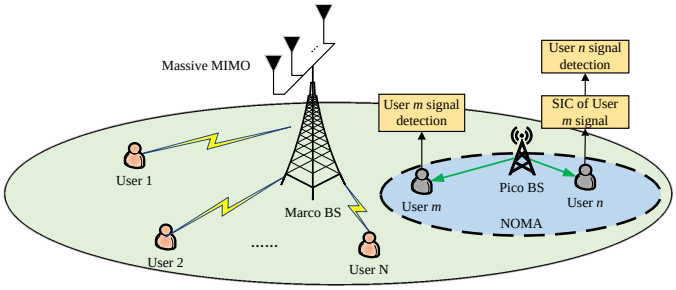

Fig. 1. Illustration of NOMA and massive MIMO based hybrid HetNets.

that users connect with the i-th tier BS j (where $j \in \Phi_i$) is given by

$$P_{r,i} = a_{n,i} P_i L\left(d_{j,i}\right) B_i, \tag{1}$$

where P_i is the transmit power of i-th tier BS, $a_{n,i}$ is the power sharing coefficient for the near user, $L\left(d_{j,i}\right) = \eta d_{j,i}^{-\alpha_i}$ is large-scale path loss, $d_{j,i}$ is the distance between the user and the i-th tier BS, α_i is the path loss exponent of the i-th tier small cells.

2) Average received power in massive MIMO aided macro cells: In macro cells, since the macro BS is equipped with multiple antennas, users in macro cells can experience large array gains. Adopting ZFBF transmission scheme, the array gain obtained at macro users is given by $G_M = M - N + 1$ [26, 28]. As a result, the average received power that users connect with macro BS ℓ (where $\ell \in \Phi_M$) is given by

$$P_{r,1} = G_M P_1 L\left(d_{\ell,1}\right) / N, \tag{2}$$

where P_1 is the transmit power of the massive MIMO aided macro BSs, $L\left(d_{\ell,1}\right) = \eta d_{\ell,1}^{-\alpha_1}$ is large-scale path loss, $d_{\ell,1}$ is the distance between users and macro BSs, η is the frequency dependent factor, B_i is the identical bias factor. It is noted that the biasing factor B_i is useful for offloading data traffic in HetNets [15].

C. Channel Model

1) NOMA enhanced small cell transmission: In small cells, without loss of generality, we consider that each small cell BS has already associated one user in the previous round of user association process. With applying NOMA protocol, we aim to squeeze a typical user into the same small cell to improve the spectral efficiency. For simplicity, we assume that the distances between the existing users and the connected small cell BSs are the same as r_k, future work will relax this assumption. The distance between the typical user and the connected small cell BS is a random value. We assume that the SIC operation always happened at the near user, due to the fact that the path loss is more stable and dominant compared to the instantaneous small-scale fading effects, which is quantitatively demonstrated in Chapter 2 of [29] by comparing small-scale fading and path loss. Since it is not pre-determined that the typical user is a near user n or a far user m. We denote d_{o,k_m} and d_{o,k_n} are the distance between the k-th tier small cell BS and user m and user n, respectively. As such, two possible cases can happen in the following.

Near user case: In the first case, we consider that the typical user is a near user n ($x \leq r_k$), then we have $d_{o,k_m} = r_k$. User n will first decode the information of user m^* with the following SINR

$$\gamma_{k_{n\to m*}} = \frac{a_{m,k} P_k g_{o,k} L\left(d_{o,k_n}\right)}{a_{n,k} P_k g_{o,k} L\left(d_{o,k}\right) + I_{M,k} + I_{S,k} + \sigma^2}, \quad (3)$$

where $a_{m,k}$ and $a_{n,k}$ are the power sharing coefficients for two users in the k-th layer, σ^2 is the additive white Gaussian noise (AWGN) power, $L\left(d_{o,k_n}\right) = \eta d_{o,k_n}^{-\alpha_i}$, $I_{M,k} = \sum_{\ell\in\Phi_1} \frac{P_1}{N} g_{\ell,1} L\left(d_{\ell,1}\right)$ is the interference from macro cells, $I_{S,k} = \sum_{i=2}^{K} \sum_{j\in\Phi_i\setminus B_{o,k}} P_i g_{j,i} L\left(d_{j,i}\right)$ is the interference from small cells, $g_{o,k}$ and d_{o,k_n} is the small-scale fading coefficients and distance between the typical user and the connected k-th tier BS, $g_{\ell,1}$ and $d_{\ell,1}$ are the small-scale fading coefficients and distance between a typical user and connected macro BS ℓ, respectively, $g_{j,i}$ and $d_{j,i}$ are the small-scale fading coefficients and distance between a typical user and connected i-th tier small cell BS j except the serving BS $B_{o,k}$, respectively. Here, $g_{o,k}$ and $g_{j,i}$ follow exponential distributions with unit mean. $g_{\ell,1}$ follows Gamma distribution with parameters $(N, 1)$.

If the above procedure is successful, user n then decode its own message. As such, the interference from the existing user can be canceled. As such, the received SINR that a typical user n connects with the k-th tier small cell can be expressed as

$$\gamma_{k_n} = \frac{a_{n,k} P_k g_{o,k} L\left(d_{o,k_n}\right)}{I_{M,k} + I_{S,k} + \sigma^2}. \quad (4)$$

For the connected far user m^*, it will directly decode its own message by treating the message of user n as interference. Therefore, the received SINR that for the existing user m^* in the k-th tier small cell can be expressed as

$$\gamma_{k_{m^*}} = \frac{a_{m,k} P_k g_{o,k} L\left(r_k\right)}{I_{k,n} + I_{M,k} + I_{S,k} + \sigma^2}, \quad (5)$$

where $I_{k,n} = a_{n,k} P_k g_{o,k} L\left(r_k\right)$ is the intra-BS interference from the connected k-th tier BS with superposition information of user n, and $L\left(r_k\right) = \eta {r_k}^{-\alpha_k}$.

Far user case: In the second case, we consider that the typical user is a far user m ($x > r_k$), then we have $d_{o,k_n} = r_k$. As such, for the connected near user n^*, it will first decode the information of user m with the following SINR

$$\gamma_{k_{n*\to m}} = \frac{a_{m,k} P_k g_{o,k} L\left(r_k\right)}{a_{n,k} P_k g_{o,k} L\left(r_k\right) + I_{M,k} + I_{S,k} + \sigma^2}, \quad (6)$$

It the above procedure is successful, it is capable of cancelling interference from the typical user m by applying SIC technology. Therefore, the received SINR of user n^* is given by

$$\gamma_{k_{n^*}} = \frac{a_{n,k} P_k g_{o,k} L\left(r_k\right)}{I_{M,k} + I_{S,k} + \sigma^2}. \quad (7)$$

Regarding the received SINR that user m connects with the k-th tier small cell, it can be expressed as

$$\gamma_{k_m} = \frac{a_{m,k} P_k g_{o,k} L\left(d_{o,k_m}\right)}{I_{k,n^*} + I_{M,k} + I_{S,k} + \sigma^2}, \quad (8)$$

where $I_{k,n^*} = a_{n,k} P_k g_{o,k} L\left(d_{o,k_m}\right)$ is the interference from the BS with superposition the information of existing user n^* of the k-th tier small cell with $L\left(d_{o,k_m}\right) = \eta d_{o,k_m}^{-\alpha_k}$, d_{o,k_n} is the distance between the typical user m and the connected k-th tier BS.

2) Massive MIMO aided macro cell transmission: Without loss of generality, we assume a typical user is located at the origin of an infinite two-dimension plane. Based on (1) and (2), the received SINR that a typical user connects with a macro BS at a random distance $d_{o,1}$ can be expressed as

$$\gamma_{r,1} = \frac{\frac{P_1}{N} h_{o,1} L\left(d_{o,1}\right)}{I_{M,1} + I_{S,1} + \sigma^2}, \quad (9)$$

where $I_{M,1} = \sum_{\ell\in\Phi_1\setminus B_{o,1}} \frac{P_1}{N} h_{\ell,1} L\left(d_{\ell,1}\right)$ is the interference from macro cells, $I_{S,1} = \sum_{i=2}^{K} \sum_{j\in\Phi_i} P_i h_{j,i} L\left(d_{j,i}\right)$ is the interference from small cells, $h_{o,1}$ is the small-scale fading coefficient between the typical user and the connected macro BS, $h_{\ell,1}$ and $d_{\ell,1}$ are the small-scale fading coefficients and distance between a typical user and connected macro BS ℓ except the serving macro BS $B_{o,1}$, respectively, $h_{j,i}$ and $d_{j,i}$ are the small-scale fading coefficients and distance between a typical user and connected i-th tier small cell BS j, respectively. Here, $h_{o,1}$ follows Gamma distribution with parameters $(M - N + 1, 1)$, $h_{\ell,1}$ follows Gamma distribution with parameters $(N, 1)$, and $h_{j,i}$ follows exponential distribution with unit mean.

III. Coverage Probability of Non-orthogonal Multiple Access Based Small Cells

In this section, we focus on analyzing the coverage probability of a typical user associated to the NOMA enhanced small cells, which is significantly different from the conventional OMA based small cells due to the channel ordering of two users. Regarding the analysis of coverage probability of a typical user associated to the massive MIMO aided macro cells is the same as the conventional massive MIMO aided OMA small cells, which has been investigated in [30]. As such, we skip this part in this treatise.

A. User Association Probability and Distance Distributions

As described in Section II-B, the user association of this proposed framework is based on maximizing the biased average received power at users. As such, based on (1) and (2), the user association of macro cells and small cells are given in the following. For simplicity, we denote $\tilde{B}_{ik} = \frac{B_i}{B_k}$, , $\tilde{\alpha}_{ik} = \frac{\alpha_i}{\alpha_k}$, $\tilde{\alpha}_{1k} = \frac{\alpha_1}{\alpha_k}$, $\tilde{\alpha}_{i1} = \frac{\alpha_i}{\alpha_1}$, $\tilde{P}_{1k} = \frac{P_1}{P_k}$, $\tilde{P}_{i1} = \frac{P_i}{P_1}$, and $\tilde{P}_{ik} = \frac{P_i}{P_k}$ in the following parts of this treatise.

Lemma 1. *The user association probability that a typical user connects with NOMA enhanced small cell BSs in the k-th tier and with macro BSs can be calculated as*

$$A_k = 2\pi\lambda_k \int_0^\infty r \exp\left[-\pi \sum_{i=2}^{K} \lambda_i \left(\tilde{P}_{ik}\tilde{B}_{ik}\right)^{\delta_i} r^{\frac{2}{\tilde{\alpha}_{ik}}} - \pi\lambda_1 \left(\frac{\tilde{P}_{1k} G_M}{N a_{n,k} B_k}\right)^{\delta_1} r^{\frac{2}{\tilde{\alpha}_{1k}}}\right] dr., \quad (10)$$

and

$$A_1 = 2\pi\lambda_1 \int_0^\infty r \exp\left[-\pi\sum_{i=2}^{K}\lambda_i\left(\frac{a_{n,i}\tilde{P}_{i1}B_iN}{G_M}\right)^{\delta_i} r^{\frac{2}{\tilde{\alpha}_{i1}}} - \pi\lambda_1 r^2\right] dr., \tag{11}$$

respectively, where $\delta_1 = \frac{2}{\alpha_1}$ *and* $\delta_i = \frac{2}{\alpha_i}$.

Proof: Using the similar method as Lemma 1 of [15], (10) *and* (10) *can be easily obtained* ■

Corollary 1. *For the special case that each tier has the same path loss exponent, i.e.,* $\alpha_1 = \alpha_k = \alpha$*, the user association probability of the NOMA enhanced small cells in the* k*-th tier and macro cells can be expressed in closed form as*

$$\tilde{A}_k = \frac{\lambda_k}{\sum_{i=2}^{K}\lambda_i\left(\tilde{P}_{ik}\tilde{B}_{ik}\right)^{\delta} + \lambda_1\left(\frac{\tilde{P}_{1k}G_M}{Na_{n,k}B_k}\right)^{\delta}}, \tag{12}$$

and

$$\tilde{A}_1 = \frac{\lambda_1}{\sum_{i=2}^{K}\lambda_i\left(\frac{a_{n,i}\tilde{P}_{i1}B_iN}{G_M}\right)^{\delta} + \lambda_1}, \tag{13}$$

respectively, where $\delta = \frac{2}{\alpha}$.

Remark 1. *The derived results in* (12) *and* (13) *demonstrate that that by increasing the number of antennas at the macro cell BSs, the user association probability to the macro cells is increases and the user association probability to the small cells decreases. This is due to the large array gains bringing by the macro cells to the served users. It is also worth noting that increasing the power sharing coefficient of* a_n *results in higher association probability to the NOMA enhanced small cells. As* $a_n \to 1$*, the user association becomes conventional OMA based approach.*

Then we consider the probability density function (PDF) of the distance between a typical user and the connected k-th tier small cell BS. Based on (10), we obtain

$$f_{d_{o,k}}(x) = \frac{2\pi\lambda_k x}{A_k}\exp\left[-\pi\sum_{i=2}^{K}\lambda_i\left(\tilde{P}_{ik}\tilde{B}_{ik}\right)^{\delta_i} x^{\frac{2}{\tilde{\alpha}_{ik}}} - \pi\lambda_1\left(\frac{\tilde{P}_{1k}G_M}{Na_{n,k}B_k}\right)^{\delta_1} x^{\frac{2}{\tilde{\alpha}_{1k}}}\right]. \tag{14}$$

We then calculate the PDF of the distance between a typical user and the connected macro BS. Based on (11), we obtain

$$f_{d_{o,1}}(x) = \frac{2\pi\lambda_1 x}{A_1}\exp\left[-\pi\sum_{i=2}^{K}\lambda_i\left(\frac{a_{n,i}\tilde{P}_{i1}B_iN}{G_M}\right)^{\delta_i} x^{\frac{2}{\tilde{\alpha}_{i1}}} - \pi\lambda_1 x^2\right]. \tag{15}$$

B. Laplace Transform of Interferences

Our next step is to derive the Laplace transform of a typical user. We denote $I_k = I_{S,k} + I_{M,k}$ as the total interference to the typical user in the k-th layer. The laplace transform of I_k with utilizing it as $\mathcal{L}_{I_k}(s) = \mathcal{L}_{I_{S,k}}(s)\,\mathcal{L}_{I_{M,k}}(s)$. We first calculate the first part in the following Lemma.

Lemma 2. *The Laplace transform of interferences from the small cell BSs to a typical user can be expressed as*

$$\mathcal{L}_{I_{S,k}}(s) = \exp\left\{-s\sum_{i=2}^{K}\frac{\lambda_i 2\pi P_i\eta(\omega_{i,k}(x_0))^{2-\alpha_i}}{\alpha_i(1-\delta_i)}\times {}_2F_1\left(1,1-\delta_i;2-\delta_i;-sP_i\eta(\omega_{i,k}(x_0))^{-\alpha_i}\right)\right\}, \tag{16}$$

where ${}_2F_1(\cdot,\cdot;\cdot;\cdot)$ *is the is the Gauss hypergeometric function [31, Eq. (9.142)], and* $\omega_{i,k}(x_0) = \left(\tilde{B}_{ik}\tilde{P}_{ik}\right)^{\frac{\delta_i}{2}} x_0^{\frac{1}{\tilde{\alpha}_{ik}}}$ *is the nearest distance allowed between the typical user associated to the* k*-th tier small cell.*

Proof: See Appendix A. ■

Then we turn our attention to calculating the second part of laplace transform of interference from the macro cell for the typical user, which is given in the following Lemma.

Lemma 3. *The Laplace transform of interference from the macro cell BSs to a typical user can be expressed as*

$$\mathcal{L}_{I_{M,k}}(s) = \exp\left[-\lambda_1\pi\delta_1\sum_{p=1}^{N}\binom{N}{p}\left(s\frac{P_1}{N}\eta\right)^{p}\left(-s\frac{P_1}{N}\eta\right)^{\delta_1-p} \times B\left(-s\frac{P_1}{N}\eta[\omega_{1,k}(x_0)]^{-\alpha_1}; p-\delta_1, 1-N\right)\right], \tag{17}$$

where $B(\cdot;\cdot,\cdot)$ *is the is the incomplete Beta function [31, Eq. (8.319)], and* $\omega_{1,k}(x_0) = \left(\frac{\tilde{P}_{1k}G_M}{a_{n,k}B_kN}\right)^{\frac{\delta_1}{2}} x^{\frac{1}{\tilde{\alpha}_{1k}}}$ *is the nearest distance allowed between the typical user associated to the the macro cell BS.*

Proof: See Appendix B. ■

C. Coverage Probability

The coverage probability is defined as the selected typical pair of users can successfully transmit with targeted data rate R_t and R_c, for the typical user and existing connected user, respectively. According to the distances, two cases are considered in the following.

Near user case: For the near user case $x_0 < r_k$, the success decoding will happen on the condition that the following two events are both satisfied.

1) The first one is that the typical user can decode the message of the connected user.
2) The second one is that after the SIC process, the typical user can decode the message of its own message.

As such, the coverage probability of the typical user on the condition of the distance x_0 in the k-th tier is:

$$P_{cov,k}(\tau_c,\tau_t,x_0)|_{x_0\le r_k} = \Pr\left\{\gamma_{k_{n\to m^*}} > \tau_c, \gamma_{k_n} > \tau_t\right\}, \tag{18}$$

where $\tau_t = 2^{R_t} - 1$ and $\tau_c = 2^{R_c} - 1$.

We first calculate the conditional coverage probability of a typical user for near user case in the following Lemma.

Lemma 4. *Conditioned on $x_0 \leq r_k$, assuming the condition $a_{m,k} - \tau_c a_{n,k} \geq 0$ holds, the coverage probability of a typical user for the near user case is expressed as*

$$P_{cov,k}\left(\tau_c,\tau_t,x_0\right)\big|_{x_0\leq r_k} = \exp\left\{-\frac{\varepsilon^*\left(\tau_c,\tau_t\right)x_0^{\alpha_k}\sigma^2}{P_k\eta}\right.$$
$$-\lambda_1\delta_1\pi\Big(\tilde{P}_{1k}\varepsilon^*\left(\tau_c,\tau_t\right)/N\Big)^{\delta_1}x_0^{\frac{2}{\tilde{\alpha}_{1k}}}Q_{1,t}^n\left(\tau_c,\tau_t\right)$$
$$\left.-\sum_{i=2}^{K}\frac{\lambda_i\delta_i\pi\Big(\tilde{B}_{ik}\Big)^{\frac{2}{\alpha_i}-1}\Big(\tilde{P}_{ik}\Big)^{\frac{2}{\alpha_i}}x_0^{\frac{2}{\tilde{\alpha}_{ik}}}}{1-\delta_i}Q_{i,t}^n\left(\tau_c,\tau_t\right)\right\}, \quad (19)$$

otherwise, $P_{cov,k}\left(\tau_c,\tau_t,x_0\right)\big|_{x_0\leq r_k} = 0$. *Here* $\varepsilon_t^n = \frac{\tau_t}{a_{n,k}}$, $\varepsilon_c^f = \frac{\tau_c}{a_{m,k}-\tau_c a_{n,k}}$, $\varepsilon^*\left(\tau_c,\tau_t\right) = \max\left\{\varepsilon_c^f,\varepsilon_t^n\right\}$, $Q_{i,t}^n\left(\tau_c,\tau_t\right) = \varepsilon^*\left(\tau_c,\tau_t\right){}_2F_1\left(1,1-\delta_i;2-\delta_i;-\frac{\varepsilon^*(\tau_c,\tau_t)}{\tilde{B}_{ik}}\right)$, *and* $Q_{1,t}^n\left(\tau_c,\tau_t\right) = \sum_{p=1}^{N}\binom{N}{p}(-1)^{\delta_1-p}\times B\left(-\frac{\varepsilon^*(\tau_c,\tau_t)a_{n,k}B_k}{G_M};p-\delta_1,1-N\right)$.

Proof: Substituting (3) *and* (4) *into* (18)*, we obtain*

$$P_{cov,k}\left(\tau_c,\tau_t,x_0\right)\big|_{x_0\leq r_k} = \Pr\left\{\frac{g_{o,k_n}P_k\eta}{x_0^{\alpha_i}\left(I_k+\sigma^2\right)} > \varepsilon^*\left(\tau_c,\tau_t\right)\right\}$$
$$= e^{-\frac{\varepsilon^*(\tau_c,\tau_t)x_0^{\alpha_k}\sigma^2}{P_k\eta}}\mathbb{E}_{I_k}\left\{e^{-\frac{\varepsilon^*(\tau_c,\tau_t)x_0^{\alpha_k}y}{P_k\eta}}\right\}$$
$$= e^{-\frac{\varepsilon^*(\tau_c,\tau_t)x_0^{\alpha_k}\sigma^2}{P_k\eta}}\mathcal{L}_{I_k}\left(\frac{\varepsilon^*\left(\tau_c,\tau_t\right)}{P_k\eta}x_0^{\alpha_k}\right), \quad (20)$$

Substituting (16) *and* (17) *into* (20)*, we obtain the successful probability for the near user case on the condition of the distance x_0 in the k-th tier. The proof is completed.* ■

Far user case: For the far user case $x_0 > r_k$, the success decoding will happen on the condition that the typical user can decode the message of itself by treating the connected user as noise. We then calculate the conditional coverage probability of a typical user for far user case in the following Lemma.

Lemma 5. *Conditioned on $x_0 > r_k$, assuming the condition $a_{m,k} - \tau_t a_{n,k} \geq 0$ holds, the coverage probability of a typical user for the far user case is expressed as*

$$P_{cov,k}\left(\tau_t,x_0\right)\big|_{x_0>r_k} = \exp\left\{-\frac{\varepsilon_t^f x_0^{\alpha_k}\sigma^2}{P_k\eta}\right.$$
$$-\lambda_1\delta_1\pi\Big(\tilde{P}_{1k}\varepsilon_t^f/N\Big)^{\delta_1}x_0^{\frac{2}{\tilde{\alpha}_{1k}}}Q_{1,t}^f\left(\tau_t\right)$$
$$\left.-\sum_{i=2}^{K}\frac{\lambda_i\delta_i\pi\Big(\tilde{B}_{ik}\Big)^{\frac{2}{\alpha_i}-1}\Big(\tilde{P}_{ik}\Big)^{\frac{2}{\alpha_i}}x_0^{\frac{2}{\tilde{\alpha}_{ik}}}}{1-\delta_i}Q_{i,t}^f\left(\tau_t\right)\right\}, \quad (21)$$

otherwise, $P_{cov,k}\left(\tau_t,x_0\right)\big|_{x_0>r_k} = 0$. *Here* $\varepsilon_t^f = \frac{\tau_t}{a_{m,k}-\tau_t a_{n,k}}$, *and* $Q_{1,t}^f\left(\tau_t\right) = \sum_{p=1}^{N}\binom{N}{p}(-1)^{\delta_1-p}B\left(-\frac{\varepsilon_t^f a_{n,k}B_k}{G_M};p-\delta_1,1-N\right)$ $Q_{i,t}^f\left(\tau_t\right) = \varepsilon_t^f{}_2F_1\left(1,1-\delta_i;2-\delta_i;-\frac{\varepsilon_t^f}{\tilde{B}_{ik}}\right)$.

Proof: Based on (8)*, we have*

$$P_{cov,k}\left(\tau_t,x_0\right)\big|_{x_0>r_k} = \Pr\left\{g_{o,k_m} > \frac{\varepsilon_t^f x_0^{\alpha_i}\left(I_k+\sigma^2\right)}{P_k\eta}\right\}. \quad (22)$$

Following the similar procedure to obtain (19)*, with interchanging $\varepsilon^*\left(\tau_c,\tau_t\right)$ with ε_t^f, we obtain the desired results in* (21)*. The proof is completed.* ■

Based on **Lemma 4** and **Lemma 5**, we can calculate the coverage probability of the typical user in the following Theorem.

Theorem 1. *Conditioned on the HPPPs, the coverage probability of a typical user associated to the k-th small cells can be expressed as*

$$P_{cov,k}\left(\tau_c,\tau_t\right) = \int_0^{r_k}P_{cov,k}\left(\tau_c,\tau_t,x_0\right)\big|_{x_0\leq r_k}f_{d_{o,k}}\left(x_0\right)dx_0$$
$$+\int_{r_k}^{\infty}P_{cov,k}\left(\tau_t,x_0\right)\big|_{x_0>r_k}f_{d_{o,k}}\left(x_0\right)dx_0, \quad (23)$$

where $P_{cov,k}\left(\tau_c,\tau_t,x_0\right)\big|_{x_0\leq r_k}$ *is given in* (19) *and* $P_{cov,k}\left(\tau_t,x_0\right)\big|_{x_0>r_k}$ *is given in* (21)*. Here,* $f_{d_{o,k}}\left(x_0\right)$ *is given in* (14)*.*

Proof: Based on the derived results of (19) *and* (21)*, taking considerations of distant distributions of a typical user associated to the k-th user small cells, we can easily obtain the desired results in* (23)*. The proof is complete.* ■

Although (23) have provided the exact analytical expression for the coverage probability of typical user, it is hard to directly obtain insights from this expression. Driven by this, we provide one special case with considering that each tier is with the same path loss exponents on each tier. As such, we have $\tilde{\alpha}_{1k} = \tilde{\alpha}_{ik} = 1$. In addition, we consider the interference limited case, where the thermal noise can be neglected. Actually, this is a common assumption in stochastic geometry based large-scale networks [15, 32]. Then based on (23), we can obtain the coverage probability of a typical user in closed-form in the following Corollary.

Corollary 2. *Conditioned on HPPP and $\alpha_1 = \alpha_k = \alpha$, the coverage probability of a typical user can be expressed in closed-form as follows:*

$$\tilde{P}_{cov,k}\left(\tau_c,\tau_t\right) = \frac{b_k\left(1-e^{-\pi\left(b_k+c_1^n(\tau_c,\tau_t)+c_2^n(\tau_c,\tau_t)\right)r_k^2}\right)}{b_k+c_1^n\left(\tau_c,\tau_t\right)+c_2^n\left(\tau_c,\tau_t\right)}$$
$$+\frac{b_k e^{-\pi\left(b_k+c_1^f(\tau_t)+c_2^f(\tau_t)\right)r_k^2}}{b_k+c_1^f\left(\tau_t\right)+c_2^f\left(\tau_t\right)}, \quad (24)$$

where $b_k = \sum_{i=2}^{K}\lambda_i\Big(\tilde{P}_{ik}\tilde{B}_{ik}\Big)^{\delta} + \lambda_1\Big(\frac{\tilde{P}_{1k}G_M}{Na_{n,k}B_k}\Big)^{\delta}$, $c_1^n\left(\tau_c,\tau_t\right) = \lambda_1\delta_1\Big(\frac{\tilde{P}_{1k}\varepsilon^*(\tau_c,\tau_t)}{N}\Big)^{\delta}\tilde{Q}_{1,t}^n\left(\tau_c,\tau_t\right)$, $c_2^n\left(\tau_c,\tau_t\right) = \sum_{i=2}^{K}\frac{\lambda_i\delta_i\left(\tilde{B}_{ik}\right)^{\frac{2}{\alpha}-1}\left(\tilde{P}_{ik}\right)^{\frac{2}{\alpha}}}{1-\delta_i}\tilde{Q}_{i,t}^n\left(\tau_c,\tau_t\right)$, $c_1^f\left(\tau_t\right) = \lambda_1\delta_1\Big(\frac{\tilde{P}_{1k}\varepsilon_t^f}{N}\Big)^{\delta_1}\tilde{Q}_{1,t}^f\left(\tau_t\right)$, *and* $c_2^f\left(\tau_t\right) = \sum_{i=2}^{K}\frac{\lambda_i\delta_i\left(\tilde{B}_{ik}\right)^{\frac{2}{\alpha}-1}\left(\tilde{P}_{ik}\right)^{\frac{2}{\alpha}}}{1-\delta}\tilde{Q}_{i,t}^f\left(\tau_t\right)$. *Here,*

$\tilde{Q}^n_{1,t}(\tau_c,\tau_t)$, $\tilde{Q}^n_{i,t}(\tau_c,\tau_t)$, $\tilde{Q}^f_{1,t}(\tau_t)$, and $\tilde{Q}^f_{i,t}(\tau_t)$ are based on interchanging the same path loss exponents, i.e. $\alpha_1=\alpha_k=\alpha$, for each tier from $Q^n_{1,t}(\tau_c,\tau_t)$, $Q^n_{i,t}(\tau_c,\tau_t)$, $Q^f_{1,t}(\tau_t)$, and $Q^f_{i,t}(\tau_t)$.

Proof: When $\alpha_1=\alpha_k=\alpha$, (10) *can be rewritten as*

$$\tilde{A}_k=\frac{\lambda_1}{b_k}, \tag{25}$$

Then we have

$$\tilde{f}_{d_{o,k}}(x)=2\pi b_k x\exp\left(-\pi b_k x^2\right). \tag{26}$$

Then substituting (26) *into* (23) *and after some mathematical manipulations, we obtain the desired results in* (24). ■

Remark 2. *The derived results in* (24) *demonstrate that the coverage probability of a typical user is determined by the target rate of itself as well as the target rate of the connected user. Additionally, inappropriate power allocation such as $a_{m,k}-\tau_t a_{n,k}<0$ will lead to the coverage probability always be zero.*

IV. Spectrum Efficiency

In an effort to evaluate the spectrum efficiency of the proposed NOMA enhanced hybrid HetNets framework, we calculate the spectrum efficiency of each tier in the following subsections.

A. Spectrum Efficiency of NOMA enhanced Small Cells

The aim of this work is to apply NOMA transmission in small cells to further improve the spectrum efficiency. Note that different from calculating the coverage probability in where the targeted rate are fixed, the achievable ergodic rate is opportunistically determined by the channel conditions of users. Also it is easily to verify that if the far user can decode the information of itself, the near user can definitely decode the information of far user since it has a better channel condition [6]. Recall that the distance order between the connected BS and the two users are not predetermined (as aforementioned in Section II), as such, in this subsection, we calculate the achievable ergodic rate of small cells both for the near user case and far user case in **Lemma 6** and **Lemma 7** in the following respectively.

Lemma 6. *Conditioned on the HPPPs, the achievable ergodic rate of the k-th tier small cell for the near user case can be expressed as follows:*

$$\tau^n_k=\frac{2\pi\lambda_k}{A_k\ln 2}\left[\int_0^{\frac{a_{m,k}}{a_{n,k}}}\frac{\bar{F}_{k_{m^*}}(z)}{1+z}dz+\int_0^{\infty}\frac{\bar{F}_{k_n}(z)}{1+z}dz\right], \tag{27}$$

where $\bar{F}_{k_{m^*}}(z)$ and $\bar{F}_{k_n}(z)$ are given in the following equations:

$$\bar{F}_{k_{m^*}}(z)=\int_0^{r_k}x\exp\left[-\frac{\sigma^2 z{r_k}^{\alpha_k}}{(a_{m,k}-a_{n,k}z)P_k\eta}-\Theta\left(\frac{z{r_k}^{\alpha_k}}{(a_{m,k}-a_{n,k}z)P_k\eta}\right)+\Lambda(x)\right]dx, \tag{28}$$

and

$$\bar{F}_{k_n}(z)=\int_0^{r_k}x\exp\left[\Lambda(x)-\frac{\sigma^2 z x^{\alpha_k}}{a_{n,k}P_k\eta}-\Theta\left(\frac{z x^{\alpha_k}}{a_{n,k}P_k\eta}\right)\right]dx. \tag{29}$$

Here $\Lambda(x)=-\pi\sum_{i=2}^{K}\lambda_i\left(\tilde{P}_{ik}\tilde{B}_{ik}\right)^{\delta_i}x^{\frac{2}{\tilde{\alpha}_{ik}}}-\pi\lambda_1\left(\frac{\tilde{P}_{1k}G_M}{Na_{n,k}B_k}\right)^{\delta_1}x^{\frac{2}{\tilde{\alpha}_{1k}}}$ *and* $\Theta(s)$ *in* (29) *and* (28) *is given by*

$$\begin{aligned}\Theta(s)=&\lambda_1\pi\delta_1\sum_{p=1}^{N}\binom{N}{p}\left(s\frac{P_1}{N}\eta\right)^p\left(-s\frac{P_1}{N}\eta\right)^{\delta_1-p}\\&\times B\left(-s\frac{P_1}{N}\eta[\omega_{1,k}(x)]^{-\alpha_1};p-\delta_1,1-N\right)\\&+s\sum_{i=2}^{K}\frac{\lambda_i 2\pi P_i\eta(\omega_{i,k}(x))^{2-\alpha_i}}{\alpha_i(1-\delta_i)}\\&\times{}_2F_1\left(1,1-\delta_i;2-\delta_i;-sP_i\eta(\omega_{i,k}(x))^{-\alpha_i}\right)\Big].\end{aligned} \tag{30}$$

Proof: See Appendix C. ■

Lemma 7. *Conditioned on the HPPPs, the achievable ergodic rate of the k-th tier small cell for the far user case can be expressed as follows:*

$$\tau^f_k=\frac{2\pi\lambda_k}{A_k\ln 2}\left[\int_0^{\infty}\frac{\bar{F}_{k_{n^*}}(z)}{1+z}dz+\int_0^{\frac{a_{m,k}}{a_{n,k}}}\frac{\bar{F}_{k_m}(z)}{1+z}dz\right], \tag{31}$$

where $\bar{F}_{k_{n^*}}(z)$ and $\bar{F}_{k_m}(z)$ are given in the following equations:

$$\begin{aligned}F_{k_m}(z)=\int_{r_k}^{\infty}x\exp\Bigg[&-\frac{\sigma^2 z x^{\alpha_k}}{P_k\eta(a_{m,k}-a_{n,k}z)}\\&-\Theta\left(\frac{z x^{\alpha_k}}{P_k\eta(a_{m,k}-a_{n,k}z)}\right)+\Lambda(x)\Bigg]dx,\end{aligned} \tag{32}$$

and

$$\bar{F}_{k_{n^*}}(z)=\int_{r_k}^{\infty}x\exp\left[\Lambda(x)-\frac{\sigma^2 z{r_k}^{\alpha_k}}{P_k\eta a_{n,k}}-\Theta\left(\frac{z{r_k}^{\alpha_k}}{P_k\eta a_{n,k}}\right)\right]dx. \tag{33}$$

Proof: The proof procedure is similar to the approach of obtaining (27)*, which is detailed introduced in Appendix C.* ■

Theorem 2. *Conditioned on the HPPPs, the achievable ergodic rate of the small cells can be expressed as follows:*

$$\tau_k=\tau^n_k+\tau^f_k, \tag{34}$$

where τ^n_k and τ^f_k are obtained from (27) *and* (31).

Proof: Combining the derived results in terms of achievable ergodic rate for both the near user case and the far user case, we obtain the desired results. ■

Note that the derived results in (34) is a double integral form, since even for some special cases, it is challenging to obtain closed form solutions. However, the derived expression is still much more efficient and also more accurate compared to using the approach of Monte Carlo simulations, which highly depends on the repeated iterations of random sampling.

B. Spectrum Efficiency of Macro cells

In massive MIMO aided macro cells, the achievable ergodic rate can be significantly improved due to multiple-antenna array gains, but with undertaking more power consumption and high complexity. However, note that the exact analytical results require high order derivatives of laplace transform with the aid of Faa Di Bruno's formula [33]. When the number of antennas goes large, it becomes mathematical intractable to calculate the derivatives due to the unacceptable complexity. In order to evaluate the spectrum efficiency of the whole system, we provide a tractable lower bound of throughput of macro cells derived in the following theorem.

Theorem 3. *Conditioned on the HPPPs, the lower bound of achievable ergodic rate of the macro cells can be expressed as follows:*

$$\tau_{1,L} = \log_2\left(1 + \frac{P_1 G_M \eta}{N\int_0^\infty \left(Q_1(x)+\sigma^2\right)x^{\alpha_1} f_{d_{o,1}}(x)\,dx}\right), \tag{35}$$

where $f_{d_{o,1}}(x)$ *is given in* (15), $Q_1(x) = \frac{2P_1\eta\pi\lambda_1}{\alpha_1-2}x^{2-\alpha_1} + \sum_{i=2}^{K} 2\pi\lambda_i\left(\frac{P_i\eta}{\alpha_i-2}\right)\left[\omega_{i,1}(x)\right]^{2-\alpha_i}$, *and* $\omega_{i,1}(x) = \left(\frac{a_{n,i}\tilde{P}_{i1}B_iN}{G_M}\right)^{\frac{\delta_i}{2}} x^{\frac{1}{\alpha_{i1}}}$ *is denoted as the nearest distance allowed between* i*-th tier small cell BS and the typical user associated to the macro cell.*

Proof: See Appendix C. ∎

Corollary 3. *Conditioned on HPPP and* $\alpha_1 = \alpha_k = \alpha$*, the lower bound of achievable ergordic rate of the macro cell is given by in closed-form as*

$$\tilde{\tau}_{1,L} = \log_2\left(1 + \frac{P_1 G_M \eta/N}{\psi(\pi b_1)^{-1} + \sigma^2\Gamma\left(\frac{\alpha}{2}+1\right)(\pi b_1)^{-\frac{\alpha}{2}}}\right), \tag{36}$$

where $\psi = \frac{2P_1\eta\pi\lambda_1}{\alpha-2} + \sum_{i=2}^{K}\left(\frac{2\pi\lambda_i P_i\eta}{\alpha-2}\right)\left(\frac{a_{n,i}\tilde{P}_{i1}B_iN}{G_M}\right)^{\delta-1}$ *and* $b_1 = \sum_{i=2}^{K}\lambda_i\left(\frac{a_{n,i}\tilde{P}_{i1}B_iN}{G_M}\right)^{\delta} + \lambda_1$.

Proof: When $\alpha_1 = \alpha_k = \alpha$, (11) *can be rewritten as*

$$\tilde{A}_1 = \frac{\lambda_1}{b_1}, \tag{37}$$

Then we have

$$\tilde{f}_{d_{o,1}}(x) = 2\pi b_1 x \exp\left(-\pi b_1 x^2\right). \tag{38}$$

Substituting the (39) *into* (35)*, we can obtain*

$$\tilde{\tau}_{1,L} = \log_2\left(1 + \frac{P_1 G_M \eta/N}{\int_0^\infty \left(\tilde{Q}_1(x)+\sigma^2\right)x^{\alpha}\tilde{f}_{d_{o,1}}(x)\,dx}\right), \tag{39}$$

where $\tilde{Q}_1(x) = \frac{2P_1\eta\pi\lambda_1}{\alpha-2}x^{2-\alpha} + \sum_{i=2}^{K}\left(\frac{2\pi\lambda_i P_i\eta}{\alpha-2}\right)\left(\frac{a_{n,i}\tilde{P}_{i1}B_iN}{G_M}\right)^{\delta-1}x^{2-\alpha} + \sigma^2$. *Then with the aid of [31, Eq. (3.326.2)], we obtain the desired closed-form expression as* (3)*. The proof is completed.* ∎

Remark 3. *The derived results in* (36) *demonstrated that that achievable ergordic rate of the macro cell can be enhanced by increasing the number of antennas at the macro cell BSs. This is because by doing so, large array gains are obtained by the served users in the macro cells.*

C. Spectrum Efficiency of the Proposed Hybrid Hetnets

Based on the analysis of last two subsections, a tractable lower bound of spectrum efficiency can be given by in the following Proposition.

Proposition 1. *The spectrum efficiency of the proposed hybrid Hetnets is as follows:*

$$\tau_{\mathrm{SE,L}} = A_1 N\tau_{1,L} + \sum\nolimits_{k=2}^{K} A_k\tau_k, \tag{40}$$

where $N\tau_1$ *and* τ_k *are the low bound spectrum efficiency of macro cells and exact spectrum efficiency of the* k*-th tier small cells, respectively. Here,* A_1 *and* A_k *are obtained from* (11) *and* (10)*,* $\tau_{1,L}$ *and* τ_k *are obtained from* (35) *and* (34)*.*

V. Energy Efficiency

In this section, we proceed to investigated the performance of the proposed frameworks from the perspective of energy efficiency, due to the fact that energy efficiency is an important performance metric in 5G systems.

A. Power Consumption Model

In order to calculate the energy efficiency of the considered networks, we first require to model the power consumption parameter of both small cell BSs and macro cell BSs. The power consumption of small cell BSs is given by

$$P_{i,total} = P_{i,static} + \frac{P_i}{\varepsilon_i}, \tag{41}$$

where $P_{i,static}$ is the static hardware power consumption of small cell BSs in the i-th tier, and ε_i is the efficiency factor for the power amplifier of small cell BSs in the i-th tier.

The power consumption of macro cell BSs is given by

$$P_{1,total} = P_{1,static} + \sum_{a=1}^{3}\left(N^a\Delta_{a,0} + N^{a-1}M\Delta_{a,1}\right) + \frac{P_1}{\varepsilon_1}, \tag{42}$$

where $P_{1,static}$ is the static hardware power consumption of macro cell BSs, ε_1 is the efficiency factor for the power amplifier of macro cell BSs, and $\Delta_{a,0}$ and $\Delta_{a,1}$ are the practical parameters which are depended on the chains of transceivers, precoding, coding/decoding, etc[2].

[2]The power consumption parameters applied in this treatise are based on a general massive MIMO model which are proposed in [25, 34]. The parameter settings are detailed in Table I.

B. Energy Efficiency of NOMA enhanced Small Cells and Macro Cells

The definition of energy efficiency is given by

$$\Theta_{\mathrm{EE}} = \frac{\text{Total data rate}}{\text{Total energy consumption}}. \tag{43}$$

Therefore, based on (43) and the power consumption model for small cell which we have provided in (41), the energy efficiency of the k-th tier of NOMA enhanced small cells is expressed as

$$\Theta_{\mathrm{EE}}^{k} = \frac{\tau_k}{P_{k,total}}, \tag{44}$$

where τ_k is obtained from (34).

Then we turn to our attention to calculating the energy efficiency of macro cells. Based on (43) and (42), the energy efficiency of macro cell is expressed as

$$\Theta_{\mathrm{EE}}^{1} = \frac{\tau_{1,L}}{P_{1,total}}, \tag{45}$$

where $\tau_{1,L}$ is obtained from (35).

C. Energy Efficiency of the Proposed Hybrid Hetnets

According to the derived results of energy efficiency of NOMA enhanced small cells and macro cells, we can express the energy efficiency in the following Proposition.

Proposition 2. *The energy efficiency of the proposed hybrid Hetnets is as follows:*

$$\Theta_{\mathrm{EE}}^{\mathrm{Hetnets}} = A_1 N \Theta_{\mathrm{EE}}^{1} + \sum\nolimits_{k=2}^{K} A_k \Theta_{\mathrm{EE}}^{k}, \tag{46}$$

where A_1 and A_k are obtained from (11) *and* (10)*, Θ_{EE}^{1} and Θ_{EE}^{k} are obtained from* (45) *and* (44)*.*

VI. Numerical Results

In this section, numerical results are presented to facilitate the performance evaluations of NOMA enhanced hybrid K-tier HetNets. The noise power is $\sigma^2 = -170 + 10 \times \log_{10}(BW) + N_f$. The power allocation coefficients of NOMA for each tier are assumed to be the same as $a_{m,k} = a_m$ and $a_{n,k} = a_n$ for simplicity. Monte Carlo simulations marked as '$\circ$' are provided to verify the accuracy of our analysis. Table I summarizes the the Monte Carlo simulation parameters used in this section.

TABLE I
Table of Parameters

Monte Carlo simulations repeated	10^5 times
The radius of the plane	10^4 m
Carrier frequency	1 GHz
The BS density of macro cells	$\lambda_1 = \left(500^2 \times \pi\right)^{-1}$
Pass loss exponent	$\alpha_1 = 3.5$, $\alpha_k = 4$
The noise figure	$N_f = 10$ dB
The noise power	$\sigma^2 = -90$ dBm
Static hardware power consumption	$P_{1,total} = 4$ W, $P_{i,total} = 2$ W
Power amplifier efficiency factor	$\varepsilon_1 = \varepsilon_i = 0.4$
Precoding power consumption	$\Delta_{1,0} = 4.8, \Delta_{2,0} = 0$
——	$\Delta_{3,0} = 2.08 \times 10^{-8}$
——	$\Delta_{1,1} = 1, \Delta_{2,1} = 9.5 \times 10^{-8}$
——	$\Delta_{3,1} = 6.25 \times 10^{-8}$

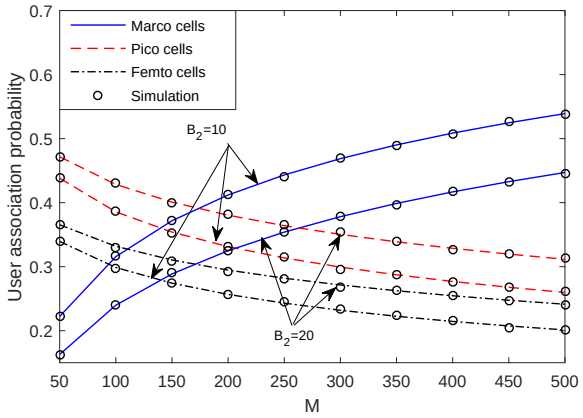

Fig. 2. User association probability of the considered network, with $K = 3$, $N = 15$, $P_1 = 40$ dBm, $P_2 = 30$ dBm and $P_3 = 20$ dBm, $r_k = 50$ m, $a_m = 0.6$, $a_n = 0.4$, $\lambda_2 = \lambda_3 = 20 \times \lambda_1$, and $B_3 = 20B_2$.

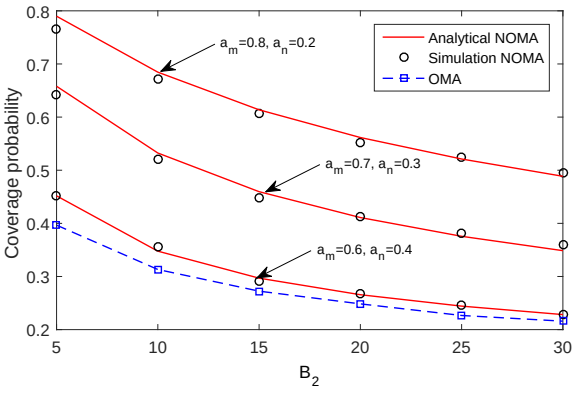

Fig. 3. Comparison of NOMA enhanced and OMA based small cells in terms of coverage probability of typical user, with $K = 2$, $M = 200$, $N = 15$, $\lambda_2 = 20 \times \lambda_1$, $R_t = R_c = 1$ BPCU, $r_k = 10$ m $P_1 = 40$ dBm, and $P_2 = 20$ dBm.

A. User Association Probability and Coverage Probability

Fig. 2 shows the effect of M and bias factor on user association probability, where the tiers of HetNets are set to be $K = 3$, including macro cells and two tiers of small cells. The analytical curves representing macro cells and small cells are from (11) and (10), respectively. One can observe that as the number of antennas at macro BSs increases, more users are likely to associate to macro cells. This is because that the massive MIMO aided macro cells are capable of providing larger array gain, which in turns enhance the average received power for the connected users. This observations are also consistent with **Remark 1** we provided in Section III. Another observation is that increasing the bias factor can encourage more users to connect to the small cells, which is an efficient method to extend the coverage of small cells or control loading balance among each tier of HetNets.

Fig. 3 plots the coverage probability of a typical user associated to the k-tier NOMA enhanced small cells versus

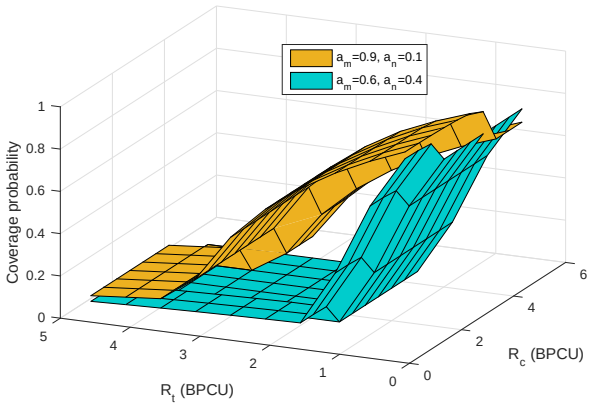

Fig. 4. Successful probability of typical user versus R_t and R_c, with $K = 2$, $M = 200$, $N = 15$, $\lambda_2 = 20 \times \lambda_1$, $r_k = 15$ m, $B_2 = 5$, $P_1 = 40$ dBm, and $P_2 = 20$ dBm.

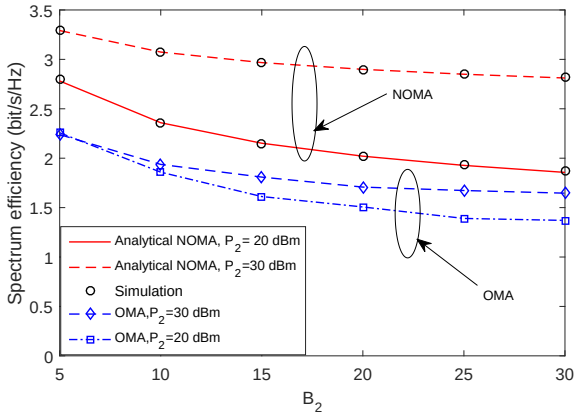

Fig. 5. Comparison of NOMA enhanced and OMA based small cells in terms of spectrum efficiency, with $K = 2$, $M = 200$, $N = 15$, $r_k = 50$ m, $a_m = 0.6$, $a_n = 0.4$, $\lambda_2 = 20 \times \lambda_1$, and $P_1 = 40$ dBm.

bias factor. The solid curves representing the analytical results of NOMA are from (23). One can observe is that the coverage probability decreased as bias factor increases, which means that the unbiased user association outperforms the biased one, i.e., when $B_2 = 1$, the scenario becomes unbiased user association. This is because by invoking biased user association, users cannot be always associated to the BS which provides the highest received power. But the biased user association is capable of offering more flexibility for users as well as the whole networks, especially for the case that cells are fully over load. We also demonstrate that NOMA has superior behavior over OMA scheme[3]. What is worth pointing out is that power sharing between two NOMA users has significant effect on coverage probability, and optimizing the power sharing coefficient can further enlarge the performance gap over the OMA based scheme [24], which is out of the scope of this paper.

Fig. 4 plots the coverage probability of a typical user associated to the k-tier NOMA enhanced small cells versus both the targeted rates of itself and the connected user. We observe that there is a cross between these two plotted surface, which means that there exists optimal power sharing allocation for the given targeted rate. In contrast, for fixed power sharing coefficients, e.g., $a_m = 0.9, a_n = 0.1$, there also exists an optimal point for the targeted rates of two users in terms of coverage probability. This figure also illustrates that for inappropriate power and targeted rate selection, the coverage probability is always be zero, which also verified our obtained insights in **Remark 2**.

B. Spectrum Efficiency

Fig. 5 plots the spectrum efficiency of NOMA enhanced and OMA based small cells versus bias factor with different transmit power of small cell BSs. The curves representing the performance of NOMA enhanced small cells are from (34). The performance of conventional OMA based small cells is illustrated as a benchmark to demonstrate the effectiveness of our proposed framework. We observe that the spectrum efficiency of small cells decreases as the bias factor increases. This behavior can be explained as follows: larger bias factor makes more macro users with low SINR are associated to small cells, which in turn degrades the spectrum efficiency of small cells. It is also worth noting that the performance of NOMA enhanced small cells outperforms the conventional OMA based small cells, which in turn can enhance the spectrum efficiency of the whole HetNets.

[3]The OMA benchmark adopted in this treatise is that by dividing the two users in equal time/frequency slots.

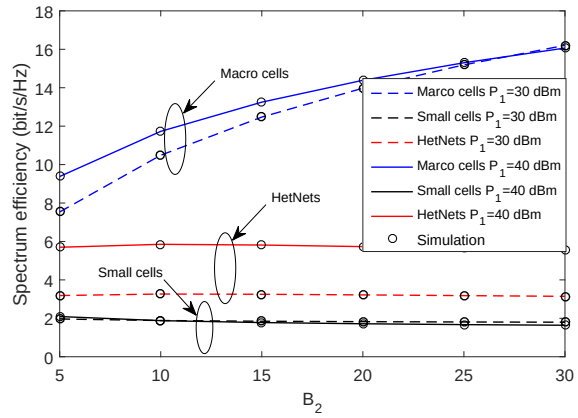

Fig. 6. Spectrum efficiency of the proposed framework, $r_k = 50$ m, $a_m = 0.6$, $a_n = 0.4$,with $K = 2$, $M = 50$, $N = 5$, $P_2 = 20$ dBm, and $\lambda_2 = 100 \times \lambda_1$.

Fig. 6 plots the spectrum efficiency of the proposed whole heterogeneous networks versus bias factor with different transmit power . The curves representing the spectrum efficiency of small cells, macro cells and HetNets are from (34), (35) and (40), respectively. We can observe that macro cells can achieve higher spectrum efficiency compared to small cells. This is attributed to the fact that macro BSs are able to serve multiple users simultaneously with offering promising array gains to each user, which has also been analytically demonstrated in **Remark 3**. It is also noted that the spectrum efficiency of

macro cells improves as bias factor increases. The reason is again that more low SINR macro cell users are associated to small cells, which in turn makes the spectrum efficiency of macro cells enhance.

C. Energy Efficiency

Fig. 7 plots the energy efficiency of the proposed whole heterogeneous networks versus bias factor with different transmit antenna of macro cell BSs. Several observations are as follows: 1) One can observe is that the energy efficiency of the macro cells decreases as the number of antenna increases. This behavior can be explained in the following. Although enlarging the number of antenna at the macro BSs is capable of offering a larger array gain, which in turn enhances the spectrum efficiency. Such operations also bring significant power consumption from the baseband signal processing of massive MIMO, which results in decreased energy efficiency. 2) Another observation is that NOMA enhanced small cells can achieve higher energy efficiency than the massive MIMO aided macro cells. It means that on the stand point of the energy consumption perspective, densely deploying NOMA enhanced small cell BSs is a more effective approach. 3) It is also worth points that the number of antennas at the macro cell BSs almost has no effect on the energy efficiency of the NOMA enhanced small cells. 4) It also demonstrates that NOMA enhanced small cells has superior behavior than conventional OMA based small cells in terms of energy efficiency. Such observations above demonstrate the benefits of proposed NOMA enhanced hybrid heterogeneous networks and provide insightful guidelines for designing the practical large scale networks.

VII. Conclusions

In this paper, a novel hybrid NOMA enhanced massive MIMO enabled HetNets framework has been designed. A flexible NONA and massive MIMO based user association policy was considered. Stochastic geometry was employed to model the networks and evaluate the corresponding performance. Analytical expressions for coverage probability of NOMA enhanced small cells were derived. It was analytically demonstrated that the inappropriate power allocation among two users will result the 'always ZERO' coverage probability. Moreover, analytical results for the spectrum efficiency and energy efficiency of the whole networks were obtained. It was interesting to observe that the number of antenna at the macro BSs has weakly effects on the energy efficiency of NOMA enhanced small cells. It has been demonstrated that NOMA enhanced small cells were able to well-coexist with the current HetNets structure and were capable of achieving superior behavior in terms of coverage probability, spectrum efficiency and energy efficiency. A promising future direction is to optimize the power sharing coefficients among NOMA users to further enhance the performance of the proposed framework.

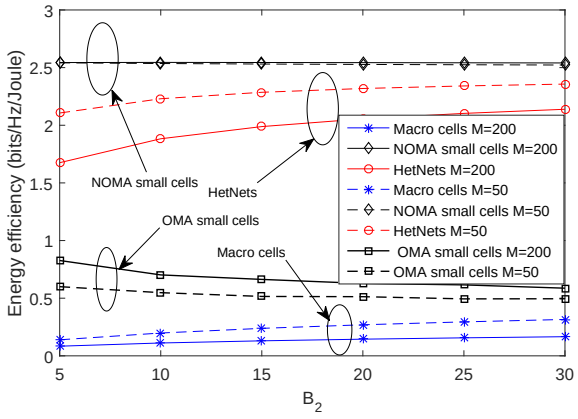

Fig. 7. Energy efficiency of the proposed framework, with $K = 2$, $r_k = 10$ m, $a_m = 0.6$, $a_n = 0.4$, $N = 15$, $P_1 = 30$ dBm, $P_2 = 20$ dBm, and $\lambda_2 = 20 \times \lambda_1$.

Appendix A: Proof of Lemma 2

Based on (3), the laplace transform of the interference from small cell BSs can be expressed as follows:

$$\begin{aligned}
\mathcal{L}_{I_{S,k}}(s) &= \mathbb{E}_{I_{S,k}}\left[e^{-sI_{S,k}}\right] \\
&\overset{(a)}{=} \mathbb{E}_{\Phi_i}\left[\sum\nolimits_{i=2}^{K}\prod_{j\in\Phi_i\setminus B_{o,k}}\mathbb{E}_{g_{j,i}}\left[e^{-sP_i g_{j,i}\eta d_{j,i}^{-\alpha_i}}\right]\right] \\
&\overset{(b)}{=} \exp\left(-\sum_{i=2}^{K}\lambda_i 2\pi\int_{\omega_{i,k}(x_0)}^{\infty}\left(1-\mathbb{E}_{g_{j,i}}\left[e^{-\frac{g_{j,i}sP_i\eta}{r^{\alpha_i}}}\right]\right)rdr\right) \\
&= \exp\left(-\sum_{i=2}^{K}\lambda_i 2\pi\int_{\omega_{i,k}(x_0)}^{\infty}\left(1-\mathcal{L}_{g_{j,i}}\left(sP_i\eta r^{-\alpha_i}\right)\right)rdr\right) \\
&\overset{(c)}{=} \exp\left(-\sum_{i=2}^{K}\lambda_i 2\pi\int_{\omega_{i,k}(x_0)}^{\infty}\left(1-\left(1+sP_i\eta r^{-\alpha_i}\right)^{-1}\right)rdr\right),
\end{aligned} \tag{A.1}$$

where (a) is resulted from applying Campbell's theorem, (b) is obtained by using the generating functional of PPP, (c) is obtained by $g_{j,i}$ follows exponential distribution with unit mean. Then applying [31, Eq. (3.194.2)], we can obtain the laplace transform in an more elegant form in (16). The proof is completed.

Appendix B: Proof of Lemma 3

Based on (3), the laplace transform of the interference from macro cell BSs can be expressed as follows:

$$\begin{aligned}
\mathcal{L}_{I_{M,k}}(s) &= \mathbb{E}_{I_{M,k}}\left[\exp\left(-s\sum_{\ell\in\Phi_1}\frac{P_1}{N}g_{\ell,1}L\left(d_{\ell,1}\right)\right)\right] \\
&= \mathbb{E}_{\Phi_1}\left[\prod_{\ell\in\Phi_1}\mathbb{E}_{g_{\ell,1}}\left[\exp\left(-s\frac{P_1}{N}g_{\ell,1}\eta d_{\ell,1}^{-\alpha_1}\right)\right]\right] \\
&\overset{(a)}{=} \exp\left(-\lambda_1 2\pi\int_{\omega_{i,1}(x)}^{\infty}\left(1-\mathbb{E}_{g_{\ell,1}}\left[e^{-\frac{sP_1 g_{\ell,1}\eta}{Nr^{\alpha_1}}}\right]\right)rdr\right),
\end{aligned} \tag{B.1}$$

where (a) is obtained with the aid of invoking generating functional of poisson point process (PPP). Recall that the $g_{\ell,1}$ follows Gamma distribution with parameter $(N,1)$. With the aid of Laplace transform for the Gamma distribution, we obtain $E_{g_{\ell,1}}\left[\exp\left(-s\frac{P_1}{N}g_{\ell,1}\eta r^{-\alpha_1}\right)\right]=\mathcal{L}_{g_{\ell,1}}\left(s\frac{P_1}{N}\eta r^{-\alpha_1}\right)=\left(1+s\frac{P_1}{N}\eta r^{-\alpha_1}\right)^{-N}$. As such, we can rewrite (B.1) as

$$\begin{aligned}
&\mathcal{L}_{I_{M,k}}(s)=\\
&\exp\left(-\lambda_1 2\pi\int_{\omega_{1,k}(x)}^{\infty}\left(1-\left(1+\frac{sP_1\eta}{Nr^{\alpha_1}}\right)^{-N}\right)rdr\right)\\
&\overset{(a)}{=}\exp\left(-2\pi\lambda_1\sum_{p=1}^{n}\binom{n}{p}\left(\frac{s\eta P_1}{N}\right)^{p}\int_{\omega_{1,k}(x_0)}^{\infty}\frac{r^{-\alpha_1 p+1}}{\left(1+\frac{s\eta P_1}{r^{\alpha}N}\right)^{N}}dr\right)\\
&\overset{(b)}{=}\exp\left[-\pi\lambda_1\delta_1\left(\frac{s\eta P_1}{N}\right)^{\delta_1}\sum_{p=1}^{N}\binom{N}{p}(-1)^{\delta_1-p}\right.\\
&\quad\left.\times\int_{0}^{-\omega_{1,k}(x)^{-\alpha}s\eta P_1/N}\frac{t^{p-\delta_1-1}}{(1-t)^{N}}dt\right],
\end{aligned}\tag{B.2}$$

where (a) is obtained by applying binomial expression and after some mathematical manipulations, (b) is obtained by using $t=-s\eta r^{-\alpha_1}P_1/N$. Then with the aid of [31, Eq. (8.391)], we obtain the Laplace transform of $I_{M,k}$ as (17). The proof is complete.

Appendix C: Proof of Lemma 6

For small cells, the achievable ergodic rate of near user case for the k-th tier can be expressed as

$$\begin{aligned}
\tau_k^n&=E\left\{\log_2(1+\gamma_{k_{m^*}})+\log_2(1+\gamma_{k_n})\right\}\\
&=\frac{1}{\ln 2}\int_0^{\infty}\frac{\bar{F}_{k_{m^*}}(z)}{1+z}dz+\frac{1}{\ln 2}\int_0^{\infty}\frac{\bar{F}_{k_n}(z)}{1+z}dz.
\end{aligned}\tag{C.1}$$

We first need to obtain the expressions for $F_{k_n}(z)$. Based on (4), we can obtain

$$\begin{aligned}
F_{k_n}(z)&=\int_0^{r_k}\Pr\left[\frac{a_{n,k}P_k g_{o,k}\eta x^{-\alpha_k}}{I_{M,k}+I_{S,k}+\sigma^2}>z\right]f_{d_{o,k}}(x)\,dx\\
&=\int_0^{r_k}\exp\left(-\frac{\sigma^2 z x^{\alpha_k}}{a_{n,k}P_k\eta}\right)\mathcal{L}_{I_k}\left(\frac{zx^{\alpha_k}}{a_{n,k}P_k\eta}\right)f_{d_{o,k}}(x)\,dx,
\end{aligned}\tag{C.2}$$

Then combining (17) and (16), we can obtain the Laplace transform of I_{k*} as $\mathcal{L}_{I_{k*}}(s)=\exp(-\Theta(s))$, where $\Theta(s)$ is given in (30). Then plugging (14) and $\mathcal{L}_{I_{k*}}(s)$ into (C.2), we obtain the complete cumulative distribution function (CCDF) of $\bar{F}_{k_n}(z)$ in (29). Then we turn to our attention to derive the CCDF of $\bar{F}_{k_{m^*}}(z)$. Based on (5), we can obtain $\bar{F}_{k_{m^*}}(z)$ as

$$\begin{aligned}
&\bar{F}_{k_{m^*}}(z)=\int_0^{r_k}f_{d_{o,k}}(x)\times\\
&\Pr\left[(a_{m,k}-a_{n,k}z)\,g_{o,k}>\frac{\left(I_{M,k}+I_{S,k}+\sigma^2\right)z}{P_k\eta r_k{}^{-\alpha_k}}\right]dx.
\end{aligned}\tag{C.3}$$

Note that for the case $z\geq\frac{a_{m,k}}{a_{n,k}}$, it is easy to observe that $\bar{F}_{k_{m^*}}(z)=0$. For the case $z\leq\frac{a_{m,k}}{a_{n,k}}$, following the similar procedure of deriving (29), we can obtain the ergodic rate of the existing user for the near user case as (28). The proof is complete.

Appendix D: Proof of Theorem 3

With the aid of Jensen's inequality, we can obtain the lower bound of the achievable ergodic rate of the macro cell as

$$\mathbb{E}\left\{\log_2(1+\gamma_{r,1})\right\}\geq\tau_{1,L}=\log_2\left(1+\left(\mathbb{E}\left\{(\gamma_{r,1})^{-1}\right\}\right)^{-1}\right)\tag{D.1}$$

By invoking the law of large numbers, we have $h_{o,1}\approx G_M$. Then based on (9), $\tau_{1,L}$ can be approximated as follows:

$$\begin{aligned}
&\mathbb{E}\left\{(\gamma_{r,1})^{-1}\right\}\approx\frac{N}{P_1G_M\eta}\mathbb{E}\left\{\left(I_{M,1}+I_{S,1}+\sigma^2\right)x^{\alpha_1}\right\}\\
&=\frac{N}{P_1G_M\eta}\int_0^{\infty}\left(\mathbb{E}\left\{I_{M,1}+I_{S,1}|\,d_{o,1}=x\right\}+\sigma^2\right)\\
&\times x^{\alpha_1}f_{d_{o,1}}(x)\,dx.
\end{aligned}\tag{D.2}$$

Then we turn to our attention to the expectation, denote $Q_1(x)=\mathbb{E}\left\{I_{M,1}+I_{S,1}|\,d_{o,1}=x\right\}$, with the aid of Campbell's Theorem, we obtain

$$\begin{aligned}
Q_1(x)&=\mathbb{E}\left\{\sum_{\ell\in\Phi_1\setminus B_{o,1}}\frac{P_1}{N}h_{\ell,1}L(d_{\ell,1})\middle|\,d_{o,1}=x\right\}\\
&+\mathbb{E}\left\{\sum\nolimits_{i=2}^{K}\sum\nolimits_{j\in\Phi_i}P_i h_{j,i}L(d_{j,i})\middle|\,d_{o,1}=x\right\}\\
&=\frac{2P_1\eta\pi\lambda_1}{\alpha_1-2}x^{2-\alpha_1}+\sum_{i=2}^{K}2\pi\lambda_i\left(\frac{P_i\eta}{\alpha_i-2}\right)[\omega_{i,1}(x)]^{2-\alpha_i},
\end{aligned}\tag{D.3}$$

We first calculate the first part of (D.3) as

$$\begin{aligned}
&\mathbb{E}\left\{\sum_{\ell\in\Phi_1\setminus B_{o,1}}\frac{P_1}{N}h_{\ell,1}L(d_{\ell,1})\middle|\,d_{o,1}=x\right\}\\
&\overset{(a)}{=}\frac{P_1}{N}\eta\mathbb{E}\left\{h_{\ell,1}\right\}\lambda_1\int_R r^{-\alpha_1}dr\\
&\overset{(b)}{=}2\pi P_1\eta\lambda_1\int_x^{\infty}r^{1-\alpha_1}dr\\
&=\frac{2P_1\eta\pi\lambda_1}{\alpha_1-2}x^{2-\alpha_1},
\end{aligned}\tag{D.4}$$

where (a) is obtained by applying Campbell's theorem, (b) is obtained since the expectation of $h_{\ell,1}$ is N. Then we turn to our attention to the second part of (D.3), with using the similar approach, we obtain

$$\begin{aligned}
&\mathbb{E}\left\{\sum\nolimits_{i=2}^{K}\sum\nolimits_{j\in\Phi_i}P_i h_{j,i}L(d_{j,i})\middle|\,d_{o,1}=x\right\}\\
&=\sum_{i=2}^{K}\left(\frac{2\pi\lambda_i P_i\eta}{\alpha_i-2}\right)[\omega_{i,1}(x)]^{2-\alpha_i}.
\end{aligned}\tag{D.5}$$

Then substituting (D.4) and (D.5) into (D.2), we obtain the desired results in (35). The proof is complete.

REFERENCES

[1] Cisco, "Cisco visual networking index: Global mobile data traffic forecast update 2014-2019 white paper," Dec. 2015.

[2] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of things: A survey on enabling technologies, protocols, and applications," *IEEE Commun. Surveys Tutorials*, vol. 17, no. 4, pp. 2347–2376, Fourth quarter 2015.

[3] L. Dai, B. Wang, Y. Yuan, S. Han, C. l. I, and Z. Wang, "Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, Sep. 2015.

[4] Z. Ding, Y. Liu, J. Choi, Q. Sun, M. Elkashlan, C.-L. I, and H. V. Poor, "Application of non-orthogonal multiple access in LTE and 5G networks," *IEEE Commun. Mag.*, Feb. 2017.

[5] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in *Proc. Vehicular Technology Conference (VTC Spring)*, June Dresden, Germany, Jun. 2013, pp. 1–5.

[6] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users," *IEEE Signal Process. Lett.*, vol. 21, no. 12, pp. 1501–1505, Dec. 2014.

[7] S. Timotheou and I. Krikidis, "Fairness for non-orthogonal multiple access in 5G systems," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1647–1651, Oct. 2015.

[8] J. Choi, "Minimum power multicast beamforming with superposition coding for multiresolution broadcast and application to NOMA systems," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 791–800, Mar. 2015.

[9] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, 2014.

[10] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, "Femtocell networks: a survey," *IEEE Commun. Mag.*, vol. 46, no. 9, pp. 59–67, Sep. 2008.

[11] H. Xie, F. Gao, and S. Jin, "An overview of low-rank channel estimation for massive MIMO systems," *IEEE Access*, vol. 4, pp. 7313–7321, 2016.

[12] A. Adhikary, H. S. Dhillon, and G. Caire, "Massive-MIMO meets hetnet: Interference coordination through spatial blanking," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 6, pp. 1171–1186, Jun. 2015.

[13] Q. Ye, O. Y. Bursalioglu, H. C. Papadopoulos, C. Caramanis, and J. G. Andrews, "User association and interference management in massive MIMO hetnets," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 2049–2065, May 2016.

[14] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic geometry and its applications*. John Wiley & Sons, 2013.

[15] 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, 2012.

[16] A. K. Gupta, H. S. Dhillon, S. Vishwanath, and J. G. Andrews, "Downlink multi-antenna heterogeneous cellular network with load balancing," *IEEE Trans. Commun.*, vol. 62, no. 11, pp. 4052–4067, Nov. 2014.

[17] W. Liu, S. Jin, C. K. Wen, M. Matthaiou, and X. You, "A tractable approach to uplink spectral efficiency of two-tier massive MIMO cellular hetnets," *IEEE Commun. Lett.*, vol. 20, no. 2, pp. 348–351, Feb. 2016.

[18] A. He, L. Wang, M. Elkashlan, Y. Chen, and K. K. Wong, "Spectrum and energy efficiency in massive MIMO enabled hetnets: A stochastic geometry approach," *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2294–2297, Dec. 2015.

[19] Y. Liu, Z. Ding, M. Elkashlan, and H. V. Poor, "Cooperative non-orthogonal multiple access with simultaneous wireless information and power transfer," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, April 2016.

[20] Z. Ding, R. Schober, and H. V. Poor, "A general MIMO framework for NOMA downlink and uplink transmission based on signal alignment," *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 4438–4454, June 2016.

[21] Y. Liu, Z. Qin, M. Elkashlan, Y. Gao, and L. Hanzo, "Enhancing the physical layer security of non-orthogonal multiple access in large-scale networks," *IEEE Trans. Wireless Commun.*, vol. PP, no. 99, pp. 1–1, 2017.

[22] Z. Ding, P. Fan, and H. V. Poor, "Random beamforming in millimeter-wave NOMA networks," 2016. [Online]. Available: https://arxiv.org/abs/1607.06302

[23] Z. Ding, F. Adachi, and H. V. Poor, "The application of MIMO to non-orthogonal multiple access," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 537–552, Jan. 2015.

[24] Y. Liu, M. Elkashlan, Z. Ding, and G. K. Karagiannidis, "Fairness of user clustering in MIMO non-orthogonal multiple access systems," *IEEE Commun. Lett.*, vol. 20, no. 7, pp. 1465–1468, July 2016.

[25] E. Bjornson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Designing multi-user MIMO for energy efficiency: When is massive MIMO the answer?" in *Proc. Wireless Commun. and Networking Conf. (WCNC)*, Apr. 2014, pp. 242–247.

[26] H. Huh, A. M. Tulino, and G. Caire, "Network MIMO with linear zero-forcing beamforming: Large system analysis, impact of channel estimation, and reduced-complexity scheduling," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 2911–2934, 2012.

[27] 3rd Generation Partnership Project (3GPP), "Study on downlink multiuser superposition transmission for LTE," Mar. 2015.

[28] K. Hosseini, W. Yu, and R. S. Adve, "Large-scale MIMO versus network MIMO for multicell interference mitigation," vol. 8, no. 5, pp. 930–941, 2014.

[29] R. Steele and L. Hanzo, *Mobile Radio Communications: Second and Third Generation Cellular and WATM Systems: 2nd*, 1999.

[30] A. He, L. Wang, Y. Chen, M. Elkashlan, and K. K. Wong, "Massive MIMO in k-tier heterogeneous cellular networks: Coverage and rate," in *IEEE Proc. of Global Commun. Conf. (GLOBECOM)*, Dec. 2015, pp. 1–6.

[31] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, 6th ed. New York, NY, USA: Academic Press, 2000.

[32] Y. Liu, L. Wang, S. Zaidi, M. Elkashlan, and T. Duong, "Secure D2D communication in large-scale cognitive cellular networks: A wireless power transfer model," *IEEE Trans. Commun.*, vol. 64, no. 1, pp. 329–342, Jan. 2016.

[33] H. S. Dhillon, M. Kountouris, and J. G. Andrews, "Downlink MIMO hetnets: Modeling, ordering results and performance analysis," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 5208–5222, Oct. 2013.

[34] E. Bjornson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?" *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3059–3075, Jun. 2015.