Non-Orthogonal Multiple Access in Large-Scale 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*— In this paper, the potential benefits of applying non-orthogonal multiple access (NOMA) technique in K-tier hybrid heterogeneous networks (HetNets) is explored. A promising new transmission framework is proposed, in which NOMA is adopted in small cells and massive multiple-input multiple-output (MIMO) is employed in macro cells. For maximizing the biased average received power for mobile users, a NOMA and massive MIMO based user association scheme is developed. 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 network by deriving exact analytical expressions for NOMA enhanced small cells and a tractable lower bound for massive MIMO enabled macro cells. Finally, we investigate the energy efficiency of the hybrid HetNets. Our results demonstrate 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 network can be greatly improved by densely deploying NOMA enhanced small cell base stations; and 4) the proposed NOMA enhanced HetNets transmission scheme has superior performance compared with the orthogonal multiple access-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 Internet [2], which is brought by the emerging demanding applications such as high-definition videos, online games and virtual reality. Also, the rapid development of internet of things (IoT) requires for facilitating billions of devices to communicate with each other [3]. Such requirements pose new challenges for designing the fifth-generation (5G) networks. Driven by these challenges, non-orthogonal multiple access (NOMA), a promising technology for 5G networks, has attracted much attention for its potential ability to enhance spectrum efficiency [4] and improving user access [5], [6]. The key idea of NOMA[1] is to utilize a superposition coding (SC) technology at the transmitter and successive interference cancelation (SIC) technology at the receiver [7], and hence multiple access can be realized in power domain via different power levels for users in the same resource block. Some initial research investigations have been made in this field [8]–[11]. The system-level performance of the downlink NOMA with two users has been demonstrated in [8]. In [9], the performance of a general NOMA transmission has been evaluated in which one base station (BS) is able to communicate with several spatial randomly deployed users. As a further advance, the fairness issue of NOMA has been addressed in [10], by examining appropriate power allocation policies among the NOMA users. For multi-antenna NOMA systems, a two-stage multicast beamforming downlink transmission scheme has been proposed in [11], where the total transmitter power was optimized using closed-form expressions.

Heterogeneous networks (HetNets) and massive multiple-input multiple-output (MIMO), as two "big three" technologies [12], are seen as the fundamental structure for the 5G networks. The core idea of HetNets is to establish closer BS-user links by densely overlaying small cells. By doing so, promising benefits such as lower power consumption, higher throughput and enhanced spectrum spatial reuse can be experienced [13]. The massive MIMO regime enables tens of hundreds/thousands antennas at a BS, and hence it is capable of offering an unprecedented level of freedom to serve multiple mobile users [14]. Aiming to fully take advantage of both massive MIMO and HetNets, in [15], interference coordination issues found in massive MIMO enabled HetNets was addressed by utilizing the spatial blanking of macro cells. Ye *et al.* [16], 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, we therefore explore the potential performance enhancement brought

Manuscript received January 28, 2017; revised May 15, 2017; accepted May 22, 2017. Date of publication July 17, 2017; date of current version December 22, 2017. This work was supported by the U.K. Engineering and Physical Sciences Research Council under Grant EP/N029720/1. This paper was presented at the IEEE Global Communication Conference, Washington, DC, USA, Dec. 2016 [1]. *(Corresponding author: Zhijin Qin.)*

Y. Liu, M. Elkashlan, and A. Nallanathan are with the School of Electronic Engineering and Computer Science, Queen Mary University of London, London E1 4NS, U.K. (e-mail: yuanwei.liu@qmul.ac.uk; maged.elkashlan@qmul.ac.uk; arumugam.nallanathan@qmul.ac.uk).

Z. Qin is with the School of Computing and Communications, Lancaster University, LA1 4WA, U.K. (e-mail: zhijin.qin@lancaster.ac.uk).

J. A. McCann is with the Department of Computing, Imperial College London, London SW7 2AZ, U.K. (e-mail: j.mccann@imperial.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSAC.2017.2726718

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

by NOMA for the hybrid 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 [17]. Some research contributions with utilizing stochastic geometry approaches have been studied in the context of Hetnets and NOMA [18]–[24]. For HetNets scenarios, based on applying a flexible bias-allowed user association approach, the performance of multi-tier downlink HetNets has been examined in [18], where all BSs and users were assumed to be equipped with a single antenna. As a further advance, the coverage provability of the multi-antenna enabled HetNets has been investigated in [19], using a simple selection bias based cell selection policy. By utilizing massive MIMO enabled HetNets and a stochastic geometry model, the spectrum efficiency of uplinks and downlinks were evaluated in [20] and [21], respectively.

Regarding the literature of stochastic geometry based NOMA scenarios, an incentive user cooperation NOMA protocol was proposed in [22] to tackle spectrum and energy issues, by regarding near users as energy harvesting relays for improving the reliability of far users. By utilizing signal alignment technology, a new MIMO-NOMA design framework has been proposed in a stochastic geometry based model [23]. Driven by the security issues, two effective approaches—protection zone and artificial noise has been utilized to enhance the physical layer security for NOMA in large-scale networks in [24]. Very recently, the potential co-existence of two technologies, NOMA and millimeter wave (mmWave) has been examined in [25], in which the random beamforming technology is adopted.

Despite the ongoing research contributions having played a vital role for fostering HetNets and NOMA technologies, to the best of our knowledge, the impact of NOMA enhanced hybrid HetNets design has not been researched. Also, there is lack of complete systematic performance evaluation metrics, i.e., coverage probability and energy efficiency. Different from the conventional HetNets design [18], [20], NOMA enhanced HetNets design poses three additional challenges: i) NOMA technology brings additional co-channel interference from the superposed signal of the connected BS; ii) NOMA technology requires careful channel ordering design to carry out SIC operations at the receiver; and iii) the user association policy requires consideration of power sharing effects of NOMA. Aiming at tackling the aforementioned issues, developing a systematic mathematically tractable framework for intelligently investigating the effect of various types of interference on network performance is desired.

B. Contributions and Organization

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

- High spectrum efficiency: With 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 multi-category interference, such as inter/intra-tier interference and intra-BS interference.
- Low complexity: By applying NOMA in single-antenna based small cells, the complex cluster based precoding/detection design for MIMO-NOMA systems [26], [27] can be avoided.
- Fairness/throughput tradeoff: NOMA is capable of addressing fairness issues by allocating more power to weak users [7], which is of great significance for HetNets when investigating efficient resource allocation in sophisticated large-scale multi-tier networks.

Different from most existing stochastic geometry based single cell research contributions in terms of NOMA [9], [22]–[25], 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 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 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. Utilizing this policy, we first derive the exact analytical expressions for the coverage probability of a typical user associating with 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 terms 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 network by applying a popular power consumption model [28]. Our results reveal that NOMA enhanced small cells achieve higher energy efficiency than macro cells. It is also shown that increasing antenna numbers at the macro cell BSs has the opposite effect on energy efficiency.
4) We show that the NOMA enhanced small cell design has superior performance 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.

The rest of the paper is organized as follows. In Section II, the network model for NOMA enhanced hybrid HetNets

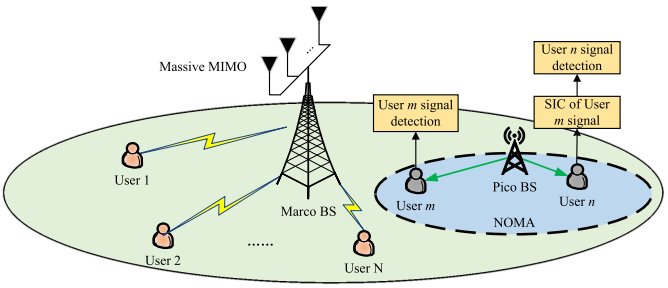


Fig. 1. Illustration of NOMA and massive MIMO based hybrid HetNets.

is introduced. In Section III, new analytical expressions for the coverage probability of the NOMA enhanced small cells are derived. Then spectrum 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 VII.

II. Network Model

A. Network Description

Focusing on 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. As it is common to overlay a high-power macro cell with successively denser and lower power small cells, we apply massive MIMO technologies to macro cells and NOMA to small cells in this work. As shown in Fig. 1, in macro cells, BSs are equipped with M antennas, each macro BS transmits signals to N users over the same resource block (e.g., time/frequency/code). We assume that $M \gg N > 1$ and linear ZFBF technique is applied at each macro BS assigning equal power to N data streams [29]. Perfect downlink channel state information (CSI) are assumed at the BSs. In small cells, each BS is equipped with single antenna. Such structure consideration is to avoid sophisticated MIMO-NOMA precoding/detection in small cells. All users are considered to be equipped with single antenna each. We adopt user pairing in each tier of small cells to implement NOMA to lower the system complexity [22]. It is worth pointing out that in long term evolution advanced (LTE-A), NOMA also implements a form of two-user case [30].

B. NOMA and Massive MIMO Based User Association

In this work, a user is allowed to access the BS of any tier, which provides the best coverage. We consider flexible user association 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 in OMA, NOMA exploits the power sparsity for multiple access by allocating different powers to different users. Due to the random spatial topology of the stochastic geometry model, the space information of users are not pre-determined. The user association policy for the NOMA enhanced small cells assumes that a near user is chosen as the typical one first. As such, at the i-th tier small cell, the averaged power received at users connecting to 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 a 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 a i-th tier BS, α_i is the path loss exponent of the i-th tier small cell, η is the frequency dependent factor, and B_i is the identical bias factor which are useful for offloading data traffic in HetNets.

2) Average Received Power in Massive MIMO Aided Macro Cells: In macro cells, as the macro BS is equipped with multiple antennas, macro cell users experience large array gains. By adopting the ZFBF transmission scheme, the array gain obtained at macro users is $G_M = M - N + 1$ [29], [31]. As a result, the average power received at users connecting to 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 a macro BS, $L\left(d_{\ell,1}\right) = \eta d_{\ell,1}^{-\alpha_1}$ is the large-scale path loss, $d_{\ell,1}$ is the distance between the user and a macro BS.

C. Channel Model

1) NOMA Enhanced Small Cell Transmission: In small cells, without loss of generality, we consider that each small cell BS is associated with one user in the previous round of user association process. Applying the NOMA protocol, we aim to squeeze a typical user into a same small cell to improve the spectral efficiency. For simplicity, we assume that the distances between the associated users and the connected small cell BSs are the same, which can be arbitrary values and are denoted as r_k, future work will relax this assumption. The distance between a typical user and the connected small cell BS is random. Due to the fact that the path loss is more stable and dominant compared to the instantaneous small-scale fading [32], we assume that the SIC operation always happens at the near user. We denote that d_{o,k_m} and d_{o,k_n} are the distances from the k-th tier small cell BS to user m and user n, respectively. Since it is not pre-determined that a typical user is a near user n or a far user m, we have the following near user case and far user case.

a) Near user case: When a typical user has smaller distance to the BS than the connected user ($x \leq r_k$, here x denotes the distance between the typical user and the BS), then we have $d_{o,k_m} = r_k$. Here we use m^* to represent the user which has been already connected to the BS in the last round of user association process, we use n to represent the typical user in near user case. User n will first decode the information of the connected user m^* to the same BS with the following SINR

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

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, $L\left(d_{o,k_n}\right) = \eta d_{o,k_n}^{-\alpha_i}$ is the large-scale path loss, $I_{M,k} = \sum_{\ell\in\Phi_1}\frac{P_1}{N}g_{\ell,1}L\left(d_{\ell,1}\right)$ 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} refer the small-scale fading coefficients and distance between a typical user and the BS in the k-th tier, $g_{\ell,1}$ and $d_{\ell,1}$ refer the small-scale fading coefficients and distance between a typical user and BS ℓ in the macro cell, respectively, $g_{j,i}$ and $d_{j,i}$ refers to the small-scale fading coefficients and distance between a typical user and its connected BS j except the serving BS $B_{o,k}$ in the i-th tier small cell, respectively. Here, $g_{o,k}$ and $g_{j,i}$ follow exponential distributions with unit mean. $g_{\ell,1}$ following Gamma distribution with parameters $(N, 1)$.

If the information of user m^* can be decoded successfully, user n then decodes its own message. As such, the SINR at a typical user n, which connects with the k-th tier small cell, can be expressed as

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

For the connected far user m^* served by the same BS, the signal can be decoded by treating the message of user n as interference. Therefore, the SINR that for the connected user m^* to the same BS in the k-th tier small cell can be expressed as

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

where $I_{k,n} = a_{n,k}P_k g_{o,k}L\left(r_k\right)$, and $L\left(r_k\right) = \eta {r_k}^{-\alpha_k}$.

b) Far user case: When a typical user has a larger distance to the BS than the connected user($x > r_k$), we have $d_{o,k_n} = r_k$. Here we use n^* to represent the user which has been already connected to the BS in the last round of user association process, we use m to represent the typical user in far user case. As such, for the connected near user n^*, it will first decode the information of user m with the following SINR

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

Once user m is decoded successfully, the interference from a typical user m can be canceled, by applying SIC technique. Therefore, the SINR at the connected user n^* to the same BS in the k-th tier small cell is given by

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

For user m that connects to the k-th tier small cell, the SINR can be expressed as

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

where $I_{k,n^*} = a_{n,k}P_k g_{o,k}L\left(d_{o,k_m}\right)$, $L\left(d_{o,k_m}\right) = \eta d_{o,k_m}^{-\alpha_k}$, d_{o,k_n} is the distance between a typical user m and the connected BS in the k-th tier.

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

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

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 the 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 the small cells; $h_{o,1}$ is the small-scale fading coefficient between a typical user and the connected macro BS, $h_{\ell,1}$ and $d_{\ell,1}$ refer to the small-scale fading coefficients and distance between a typical user and the connected macro BS ℓ except for the serving BS $B_{o,1}$ in the macro cell, respectively, $h_{j,i}$ and $d_{j,i}$ refer to the small-scale fading coefficients and distance between a typical user and BS j in the i-th tier small cell, 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 our attention on analyzing the coverage probability of a typical user associated to the NOMA enhanced small cells, which is different from the conventional OMA based small cells due to the channel ordering of two users.

A. User Association Probability and Distance Distributions

As described in Section II-B, the user association of the proposed framework is based on maximizing the biased average received power at the users. As such, based on (1) and (2), the user association of macro cells and small cells are given by 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 work.

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

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

and

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

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

Proof: Using a similar method to Lemma 1 of [18], (10) and (11) 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 the macro cells can be expressed in closed form as

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

and

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

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

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

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

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

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

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

B. The Laplace Transform of Interference

The next step is to derive the Laplace transform of a typical user. We denote that $I_k = I_{S,k} + I_{M,k}$ is the total interference to the typical user in the k-th tier. The laplace transform of I_k is $\mathcal{L}_{I_k}(s) = \mathcal{L}_{I_{S,k}}(s)\,\mathcal{L}_{I_{M,k}}(s)$. We first calculate the Laplace transform of interference from the small cell BS to a typical user $\mathcal{L}_{I_{S,k}}(s)$ in the following Lemma.

Lemma 2: The Laplace transform of interference from the small cell BSs to a typical user can be expressed as

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

where ${}_2F_1(\cdot,\cdot;\cdot;\cdot)$ is the is the Gauss hypergeometric function [33, 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 and its connected small cell BS in the k-th tier.

Proof: See Appendix A. ∎

Then we calculate the laplace transform of interference from the macro cell to a typical user $\mathcal{L}_{I_{M,k}}(s)$ in the following Lemma.

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

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

where $B(\cdot;\cdot,\cdot)$ is the is the incomplete Beta function [33, eq. (8.319)], and $\omega_{1,k}(x_0) = \left(\frac{\tilde{P}_{1k}G_M}{a_{n,k}B_k N}\right)^{\frac{\delta_1}{2}} x^{\frac{1}{\tilde{\alpha}_{1k}}}$ is the nearest distance allowed between a typical user and its connected BS in the macro cell.

Proof: See Appendix B. ∎

C. Coverage Probability

The coverage probability is defined as that a typical user can successfully transmit signals with a targeted data rate R_t. According to the distances, two cases are considered in the following.

1) Near User Case: For the near user case, $x_0 < r_k$, successful decoding will happen when the following two conditions hold:

1) The typical user can decode the message of the connected user served by the same BS.
2) After the SIC process, the typical user can decode 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:

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

where $\tau_t = 2^{R_t} - 1$ and $\tau_c = 2^{R_c} - 1$. Here R_c is the targeted data rate of the connected user served by the same BS.

Based on (18), for the near user case, we can obtain the expressions for the conditional coverage probability of a typical user in the following Lemma.

Lemma 4: If $a_{m,k} - \tau_c a_{n,k} \ge 0$ holds, the conditional coverage probability of a typical user for the near user case is expressed in closed-form as

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

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

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

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

Then by plugging (16) *and* (17) *into* (20)*, we obtain the conditional coverage probability for the near user case in* (19)*. The proof is complete.* ∎

2) Far User Case: For the far user case, $x_0 > r_k$, successful decoding will happen if the typical user can decode its own message by treating the connected user served by the same BS as noise. The conditional coverage probability of a typical user for the far user case is calculated in the following Lemma.

Lemma 5: If $a_{m,k} - \tau_t a_{n,k} \ge 0$ *holds, the coverage probability of a typical user for the far user case is expressed in closed-form as*

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

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

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

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

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

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

Theorem 1: The coverage probability of a typical user associated to the k-th tier small cells is expressed as

$$\begin{aligned} P_{cov,k}(\tau_c,\tau_t) &= \int_0^{r_k} P_{cov,k}(\tau_c,\tau_t,x_0)\big|_{x_0\le r_k} f_{d_{o,k}}(x_0)\,dx_0 \\ &\quad + \int_{r_k}^{\infty} P_{cov,k}(\tau_t,x_0)\big|_{x_0>r_k} f_{d_{o,k}}(x_0)\,dx_0, \end{aligned} \tag{23}$$

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

Proof: Based on (19) *and* (21)*, considering the 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) has provided the exact analytical expression for the coverage probability of a typical user, it is difficult to directly obtain insights from this expression. Driven by this, we provide one special case that considers each tier with the same path loss exponents. 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.[2] Then based on (23), we can obtain the closed-form coverage probability of a typical user in the following Corollary.

Corollary 2: With $\alpha_1 = \alpha_k = \alpha$ *and* $\sigma^2 = 0$*, the coverage probability of a typical user can be expressed in closed-form as follows:*

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

where $b_k = \sum_{i=2}^{K}\lambda_i\left(\tilde{P}_{ik}\tilde{B}_{ik}\right)^{\delta} + \lambda_1\left(\frac{\tilde{P}_{1k}G_M}{Na_{n,k}B_k}\right)^{\delta}$, $c_1^n(\tau_c,\tau_t) = \lambda_1\delta_1\left(\frac{\tilde{P}_{1k}\varepsilon^*(\tau_c,\tau_t)}{N}\right)^{\delta}\tilde{Q}_{1,t}^n(\tau_c,\tau_t)$, $c_2^n(\tau_c,\tau_t) = \sum_{i=2}^{K}\frac{\lambda_i\delta_i\left(\tilde{B}_{ik}\right)^{\frac{2}{\alpha}-1}\left(\tilde{P}_{ik}\right)^{\frac{2}{\alpha}}}{1-\delta_i}\tilde{Q}_{i,t}^n(\tau_c,\tau_t)$, $c_1^f(\tau_t) = \lambda_1\delta_1\left(\frac{\tilde{P}_{1k}\varepsilon_t^f}{N}\right)^{\delta_1}\tilde{Q}_{1,t}^f(\tau_t)$, *and* $c_2^f(\tau_t) = \sum_{i=2}^{K}\frac{\lambda_i\delta_i\left(\tilde{B}_{ik}\right)^{\frac{2}{\alpha}-1}\left(\tilde{P}_{ik}\right)^{\frac{2}{\alpha}}}{1-\delta}\tilde{Q}_{i,t}^f(\tau_t)$. *Here,* $\tilde{Q}_{1,t}^n(\tau_c,\tau_t)$, $\tilde{Q}_{i,t}^n(\tau_c,\tau_t)$, $\tilde{Q}_{1,t}^f(\tau_t)$, *and* $\tilde{Q}_{i,t}^f(\tau_t)$ *are based on interchanging the same path loss exponents, i.e.* $\alpha_1 = \alpha_k = \alpha$*, for each tier from* $Q_{1,t}^n(\tau_c,\tau_t)$, $Q_{i,t}^n(\tau_c,\tau_t)$, $Q_{1,t}^f(\tau_t)$, *and* $Q_{i,t}^f(\tau_t)$.

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

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

Then we have

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

[2]This is a common assumption in stochastic geometry based large-scale networks [18], [34].

Then by plugging (26) *into* (23) *and after some mathematical manipulations, we can 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 both the target rate of itself and the target rate of the connected user served by the same BS. Additionally, inappropriate power allocation such as,* $a_{m,k} - \tau_t a_{n,k} < 0$*, will lead to the coverage probability always being zero.*

IV. Spectrum Efficiency

To evaluate the spectrum efficiency of the proposed NOMA enhanced hybrid HetNets framework, we calculate the spectrum efficiency of each tier in this section.

A. Ergodic Rate of NOMA Enhanced Small Cells

Rather than calculating the coverage probability of the case with fixed targeted rate, the achievable ergodic rate for NOMA enhanced small cells is opportunistically determined by the channel conditions of users. It is easy to verify that if the far user can decode the message of itself, the near user can definitely decode the message of far user since it has better channel conditions [9], [35]. Recall that the distance order between the connected BS and the two users are not predetermined, as such, we calculate the achievable ergodic rate of small cells both for the near user case and far user case in the following Lemmas.

Lemma 6: The achievable ergodic rate of the k-th tier small cell for the near user case can be expressed as follows:

$$\tau_k^n = \frac{2\pi\lambda_k}{A_k \ln 2}\left[\int_0^{\frac{a_{m,k}}{a_{n,k}}} \frac{\bar{F}_{\gamma_{km*}}(z)}{1+z}dz + \int_0^{\infty} \frac{\bar{F}_{\gamma_{kn}}(z)}{1+z}dz\right], \quad (27)$$

where $\bar{F}_{\gamma_{km*}}(z)$ *and* $\bar{F}_{\gamma_{kn}}(z)$ *are given by*

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

and

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

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

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

Proof: See Appendix C. ■

Lemma 7: The achievable ergodic rate of the k-th tier small cell for the far user case can be expressed as follows:

$$\tau_k^f = \frac{2\pi\lambda_k}{A_k \ln 2}\left[\int_0^{\infty} \frac{\bar{F}_{\gamma_{kn*}}(z)}{1+z}dz + \int_0^{\frac{a_{m,k}}{a_{n,k}}} \frac{\bar{F}_{\gamma_{km}}(z)}{1+z}dz\right], \quad (31)$$

where $\bar{F}_{\gamma_{km}}(z)$ *and* $\bar{F}_{\gamma_{kn*}}(z)$ *are given by*

$$\bar{F}_{\gamma_{km}}(z) = \int_{r_k}^{\infty} x \exp\left[-\frac{\sigma^2 z x^{\alpha_k}}{P_k \eta\left(a_{m,k} - a_{n,k}z\right)} - \Theta\left(\frac{z x^{\alpha_k}}{P_k \eta\left(a_{m,k} - a_{n,k}z\right)}\right) + \Lambda(x)\right] dx, \quad (32)$$

and

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

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:

$$\tau_k = \tau_k^n + \tau_k^f, \quad (34)$$

where τ_k^n *and* τ_k^f *are obtained from* (27) *and* (31).

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 Monte Carlo simulations, which highly depends on the repeated iterations of random sampling.

B. Ergodic Rate 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 more power consumption and high complexity. However, the exact analytical results require high order derivatives of the Laplace transform with the aid of Faa Di Bruno's formula [36]. When the number of antennas goes large, it becomes mathematically intractable to calculate the derivatives due to the unacceptable complexity. In order to evaluate spectrum efficiency for the whole system, we provide a tractable lower bound of throughput for macro cells in the following theorem.

Theorem 3: The lower bound of the achievable ergodic rate of the macro cells can be expressed as follows:

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

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* $\omega_{i,1}(x) = \left(\frac{a_{n,i}\tilde{P}_{i1}B_i N}{G_M}\right)^{\frac{\delta_i}{2}} x^{\frac{1}{\tilde{\alpha}_{i1}}}$ *is denoted as the nearest*

distance allowed between the i-th tier small cell BS and the typical user that is associated with the macro cell.

Proof: See Appendix C. ∎

Corollary 3: If $\alpha_1 = \alpha_k = \alpha$ holds, the lower bound of the achievable ergordic rate of the macro cell is given by in closed-form as

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

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

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

$$\tilde{A}_1 = \frac{\lambda_1}{b_1}, \quad (37)$$

Then we have

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

By substituting the (38) *into* (35)*, we can obtain*

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

where $\tilde{Q}_1(x) = \frac{2P_1\eta\pi\lambda_1}{\alpha-2}x^{2-\alpha} + \sum_{i=2}^{K}\left(\frac{2\pi\lambda_i P_i\eta}{\alpha-2}\right)\left(\frac{a_{n,i}\tilde{P}_{i1}B_i N}{G_M}\right)^{\delta-1} x^{2-\alpha} + \sigma^2$*. Then with the aid of [33, eq. (3.326.2)], we obtain the desired closed-form expression as* (36)*. The proof is complete.* ∎

Remark 3: The derived results in (36) *demonstrate that the achievable ergordic rate of the macro cell can be enhanced by increasing the number of antennas at the macro cell BSs. This is because the users in the macro cells can experience larger array gains.*

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 in the following Proposition.

Proposition 1: The spectrum efficiency of the proposed hybrid Hetnets is

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

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

V. Energy Efficiency

In this section, we proceed to investigate the performance of the proposed hybrid HetNets framework from the perspective of energy efficiency, due to the fact that energy efficiency is an important performance metric for 5G systems.

A. Power Consumption Model

To calculate the energy efficiency, we first need 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

$$P_{i,total} = P_{i,static} + \frac{P_i}{\varepsilon_i}, \quad (41)$$

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

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

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.[3]

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

The energy efficiency is defined as

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

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

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

where τ_k is obtained from (34).

Based on (42) and (43), the energy efficiency of macro cell is expressed as

$$\Theta_{\mathrm{EE}}^{1} = \frac{N\tau_{1,L}}{P_{1,total}}, \quad (45)$$

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:

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

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

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	$\Delta_{1,0} = 4.8, \Delta_{2,0} = 0$
——	$\Delta_{3,0} = 2.08 \times 10^{-8}$
——	$\Delta_{1,1} = 1, \Delta_{2,1} = 9.5 \times 10^{-8}$
——	$\Delta_{3,1} = 6.25 \times 10^{-8}$

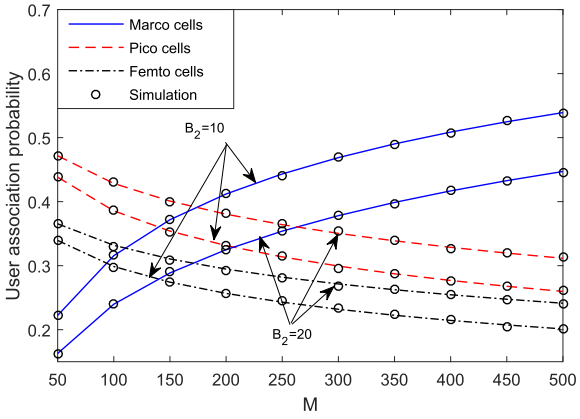


Fig. 2. User association probability versus antenna number with different bias factor, with $K = 3$, $N = 15$, $P_1 = 40$ dBm, $P_2 = 30$ dBm and $P_3 = 20$ dBm, $r_k = 50$ m, $a_m = 0.6$, $a_n = 0.4$, $\lambda_2 = \lambda_3 = 20 \times \lambda_1$, and $B_3 = 20 \times B_2$.

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 sharing coefficients of NOMA for each tier are same as $a_{m,k} = a_m$ and $a_{n,k} = a_n$ for simplicity. BPCU is short for bit per channel use. Monte Carlo simulations marked as '$\circ$' are provided to verify the accuracy of our analysis. Table I summarizes the the simulation parameters used in this section.

A. User Association Probability and Coverage Probability

Fig. 2 shows the effect of the number of antennas equipped at each macro BS, M, and the bias factor on the 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 small cells and macro cells are from (10) and (11), respectively. One can observe that as the number of antennas at each macro BS increases, more users are likely to associate to macro cells. This is because the massive MIMO aided macro cells are capable of providing larger array gain, which in turn enhances the average received power for the connected users. This observation is consistent with **Remark 1** 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 way to extend the coverage of small cells or control the load 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 the bias factor. The solid curves representing the analytical results of NOMA are from (23). One can observe that the coverage probability decreases as the 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 network, especially for the case that cells are fully over load. We also demonstrate that NOMA has superior behavior over OMA scheme.[4] Actually, the OMA based HetNets scheme has been analytically investigated in the previous research contributions such as [18], the OMA benchmark adopted in this treatise is generated by numerical approach. It is worth pointing out that power sharing between two NOMA users has a significant effect on coverage probability, and optimizing the power sharing coefficients can further enlarge the performance gap over OMA based schemes [27], which is out of the scope of this paper.

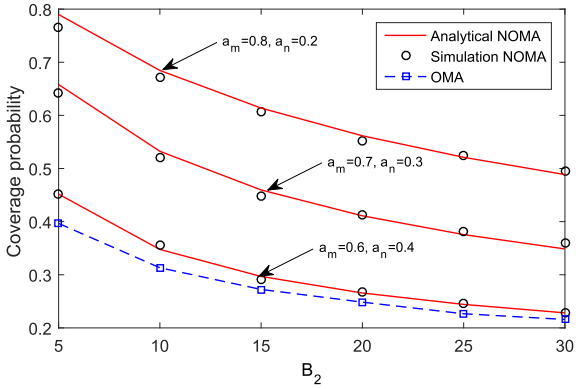


Fig. 3. Coverage probability comparison of NOMA and OMA based small cells. $K = 2$, $M = 200$, $N = 15$, $\lambda_2 = 20 \times \lambda_1$, $R_t = R_c = 1$ BPCU, $r_k = 10$ m $P_1 = 40$ dBm, and $P_2 = 20$ dBm.

Fig. 4 plots the coverage probability of a typical user associated to the k-tier NOMA enhanced small cells versus both R_t and R_c. We observe that there is a cross between these two plotted surfaces, which means that there exists an optimal power sharing allocation scheme 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 optimal targeted rates of two users for coverage probability. This figure also illustrates that for inappropriate power and targeted rate selection, the coverage probability is always zero, which also verifies our obtained insights in **Remark 2**.

[3]The power consumption parameters applied in this treatise are based on an established massive MIMO model proposed in [28] and [37].

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

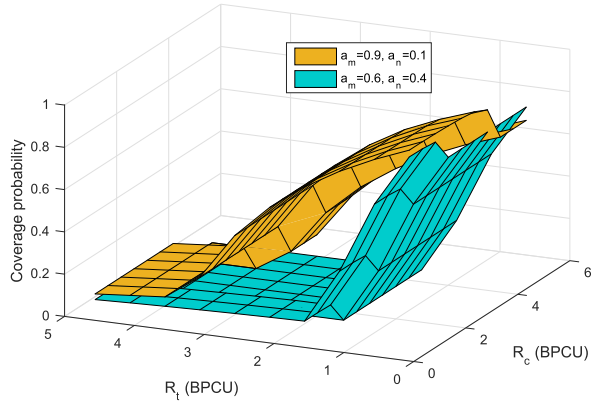


Fig. 4. Successful probability of typical user versus targeted rates of R_t and R_c, with $K = 2$, $M = 200$, $N = 15$, $\lambda_2 = 20 \times \lambda_1$, $r_k = 15$ m, $B_2 = 5$, $P_1 = 40$ dBm, and $P_2 = 20$ dBm.

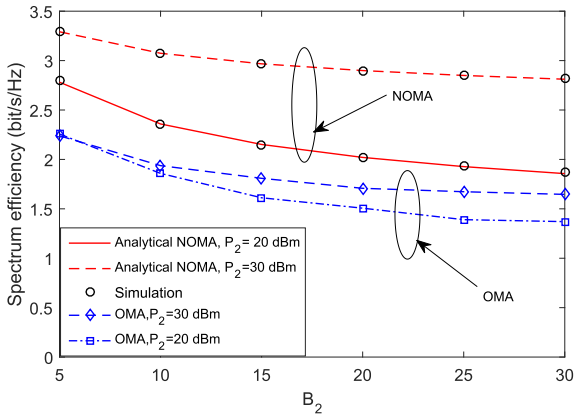


Fig. 5. Spectrum efficiency comparison of NOMA and OMA based small cells. $K = 2$, $M = 200$, $N = 15$, $r_k = 50$ m, $a_m = 0.6$, $a_n = 0.4$, $\lambda_2 = 20 \times \lambda_1$, and $P_1 = 40$ dBm.

B. Spectrum Efficiency

Fig. 5 plots the spectrum efficiency of small cells with NOMA and OMA versus bias factor, B_2, with different transmit powers of small cell BSs, P_2. 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 associates more macro users with low SINR 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 enhances the spectrum efficiency of the whole HetNets.

Fig. 6 plots the spectrum efficiency of the proposed hybrid HetNets versus bias factor, B_2, with different transmit powers, P_1. The curves representing the spectrum efficiency of small cells, macro cells and HetNets are from (40). 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 offering promising array gains to each user, which has been

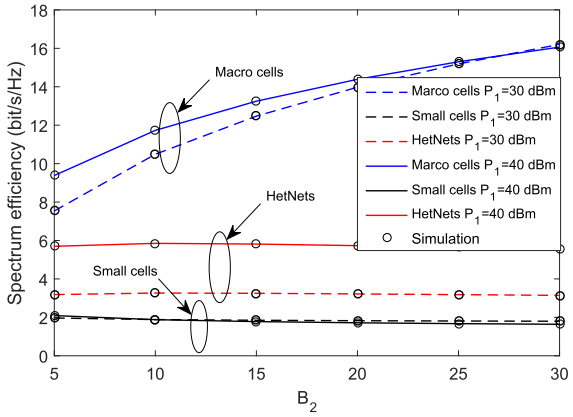


Fig. 6. Spectrum efficiency of the proposed framework. $r_k = 50$ m, $a_m = 0.6$, $a_n = 0.4$, $K = 2$, $M = 50$, $N = 5$, $P_2 = 20$ dBm, and $\lambda_2 = 100 \times \lambda_1$.

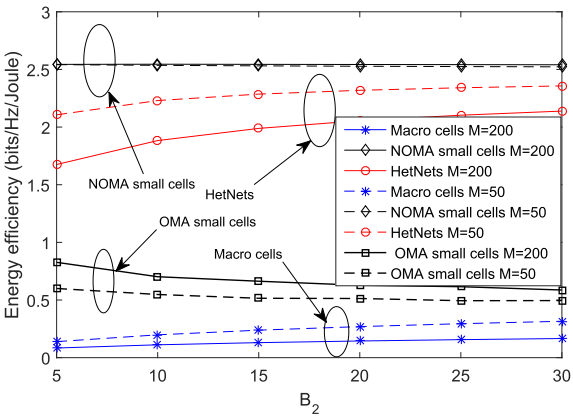


Fig. 7. Energy efficiency of the proposed framework. $K = 2$, $r_k = 10$ m, $a_m = 0.6$, $a_n = 0.4$, $N = 15$, $P_1 = 30$ dBm, $P_2 = 20$ dBm, and $\lambda_2 = 20 \times \lambda_1$.

analytically demonstrated in **Remark 3**. It is also shown that the spectrum efficiency of macro cells improves as the bias factor increases. The reason is again that when more low SINR macro cell users are associated to small cells, the spectrum efficiency of macro cells can be enhanced.

C. Energy Efficiency

Fig. 7 plots the energy efficiency of the proposed hybrid HetNets versus the bias factor, B_2, with different numbers of transmit antenna of macro cell BSs, M. Several observations are as follows: 1) The energy efficiency of the macro cells decrease as the number of antenna increases. Although enlarging the number of antenna at the macro BSs offers a larger array gains, 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 from the perspective of energy consumption, densely deploying BSs in NOMA enhanced small cell is a more effective approach. 3) It is also worth noting 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 performance than conventional OMA based small cells in terms of energy efficiency. Such observations above demonstrate the benefits of the proposed NOMA enhanced hybrid HetNets and provide insightful guidelines for designing the practical large scale networks.

VII. Conclusions

In this paper, a novel hybrid 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 its performance. Analytical expressions for the coverage probability of NOMA enhanced small cells were derived. It was analytically demonstrated that the inappropriate power allocation among two users will result 'always ZERO' coverage probability. Moreover, analytical results for the spectrum efficiency and energy efficiency of the whole network was obtained. It was interesting to observe that the number of antenna at the macro BSs has weak effects on the energy efficiency of NOMA enhanced small cells. It has been demonstrated that NOMA enhanced small cells were able to coexist well with the current HetNets structure and were capable of achieving superior performance compared to OMA based small cells. Note that applying NOMA scheme also brings hardware complexity and processing delay to the existing HetNets structure, which should be taken into considerations. A promising future direction is to optimize power sharing coefficients among NOMA users to further enhance the performance of the proposed framework.

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-function of PPP, and (c) is obtained by $g_{j,i}$ follows exponential distribution with unit mean. By applying [33, eq. (3.194.2)], we can obtain the Laplace transform in an more elegant form in (16). The proof is complete.

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-function of 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, and (b) is obtained by using $t=-s\eta r^{-\alpha_1}P_1/N$. Based on [33, eq. (8.391)], we can obtain the Laplace transform of $I_{M,k}$ as given in (17). The proof is complete.

Appendix C: Proof of Lemma 6

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

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

We need to obtain the expressions for $\bar{F}_{\gamma_{k_n}}(z)$ first. Based on (4), we can obtain

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

By 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). By plugging (14) and $\mathcal{L}_{I_{k*}}(s)$ into (C.2), we obtain the complete cumulative distribution function (CCDF) of γ_{k_n} in (29). In the following, we turn to our attention to derive the CCDF of $\gamma_{k_{m*}}$. Based on (5), we can obtain $\bar{F}_{\gamma_{k_{m*}}}(z)$ as

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

Note that for the case $z \geq \frac{a_{m,k}}{a_{n,k}}$, it is easy to observe that $\bar{F}_{\gamma_{k_{m*}}}(z) = 0$. 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 cells as

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

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:

$$\begin{aligned}\mathbb{E}\left\{\left(\gamma_{r,1}\right)^{-1}\right\} &\approx \frac{N}{P_1 G_M \eta}\mathbb{E}\left\{\left(I_{M,1}+I_{S,1}+\sigma^2\right)x^{\alpha_1}\right\} \\ &= \frac{N}{P_1 G_M \eta}\int_0^{\infty}\left(\mathbb{E}\left\{I_{M,1}+I_{S,1}\middle|\, d_{o,1}=x\right\}+\sigma^2\right) \\ &\quad \times x^{\alpha_1} f_{d_{o,1}}(x)\,dx. \quad \text{(D.2)}\end{aligned}$$

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

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

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

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

where (a) is obtained by applying Campbell's theorem, and (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

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

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

References

[1] Y. Liu, Z. Qin, M. Elkashlan, Y. Gao, and N. Arumugam, "Non-orthogonal multiple access in massive MIMO aided heterogeneous networks," in *Proc. Global Commun. Conf. (GLOBECOM)*, Washington, DC, USA, Dec. 2016, pp. 1–6.

[2] "Cisco visual networking index: Global mobile data traffic forecast update 2014–2019," Cisco, San Jose, CA, USA, White Paper, Dec. 2015.

[3] 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 Tuts.*, vol. 17, no. 4, pp. 2347–2376, 4th Quart., 2015.

[4] 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.

[5] M. Shirvanimoghaddam, M. Dohler, and S. Johnson. (Dec. 2016). "Massive non-orthogonal multiple access for cellular IoT: Potentials and limitations." [Online]. Available: https://arxiv.org/abs/1612.00552

[6] Y. Cai, Z. Qin, F. Cui, G. Y. Li, and J. A. McCann. (Feb. 2017). "Modulation and multiple access for 5G networks." [Online]. Available: https://arxiv.org/abs/1702.07673

[7] Z. Ding *et al.*, "Application of non-orthogonal multiple access in LTE and 5G networks," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 185–191, Feb. 2017.

[8] 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. 77th IEEE VTC-Spring*, Dresden, Germany, Jun. 2013, pp. 1–5.

[9] 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.

[10] 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.

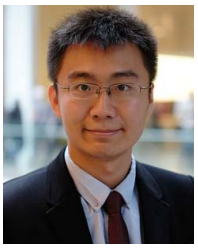
[11] J. Choi, "Minimum power multicast beamforming with superposition coding for multiresolution broadcast and application to NOMA systems," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 791–800, Mar. 2015.

[12] J. G. Andrews *et al.*, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

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

[14] 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.

[15] 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.
[16] 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.
[17] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic Geometry and its Applications*. Hoboken, NJ, USA: Wiley, 2013.
[18] H.-S. Jo, Y. J. Sang, P. Xia, and J. G. Andrews, "Heterogeneous cellular networks with flexible cell association: A comprehensive downlink SINR analysis," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3484–3495, Oct. 2012.
[19] 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.
[20] 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.
[21] 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.
[22] 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, pp. 938–953, Apr. 2016.
[23] 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, Jun. 2016.
[24] 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. 16, no. 3, pp. 1656–1672, Mar. 2017.
[25] Z. Ding, P. Fan, and H. V. Poor, "Random beamforming in millimeter-wave NOMA networks," *IEEE Access*, vol. 5, pp. 7667–7681, Sep. 2017.
[26] 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.
[27] 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, Jul. 2016.
[28] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Designing multi-user MIMO for energy efficiency: When is massive MIMO the answer?" in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2014, pp. 242–247.
[29] 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, May 2012.
[30] *3rd Generation Partnership Project Study on Downlink Multiuser Superposition Transmission for LTE*, 3GPP, Mar. 2015.
[31] K. Hosseini, W. Yu, and R. S. Adve, "Large-scale MIMO versus network MIMO for multicell interference mitigation," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 930–941, Oct. 2014.
[32] R. Steele and L. Hanzo, *Mobile Radio Communications: Second and Third Generation Cellular and WATM Systems*, 2nd ed. New York, NY, USA: Wiley, 1999.
[33] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, 6th ed. New York, NY, USA: Academic, 2000.
[34] 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.
[35] Y. Liu, Z. Ding, M. Elkashlan, and J. Yuan, "Non-orthogonal multiple access in large-scale underlay cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 10152–10157, Dec. 2016.
[36] 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.
[37] E. Björnson, 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.



Yuanwei Liu (S'13–M'16) received the B.S. and M.S. degrees from the Beijing University of Posts and Telecommunications in 2011 and 2014, respectively, and the Ph.D. degree in electrical engineering from the Queen Mary University of London, U.K., in 2016. He has been a Lecturer (Assistant Professor) with the School of Electronic Engineering and Computer Science, Queen Mary University of London, since 2017. He was with the Department of Informatics, King's College London, from 2016 to 2017, where he was a Post-Doctoral Research Fellow.

His research interests include 5G wireless networks, Internet of Things, stochastic geometry, and matching theory. He received the Exemplary Reviewer Certificate of the IEEE WIRELESS COMMUNICATION LETTERS in 2015 and the IEEE TRANSACTIONS ON COMMUNICATIONS in 2017. He has served as a TPC Member for many IEEE conferences, such as GLOBECOM and ICC. He currently serves as an Editor of the IEEE COMMUNICATIONS LETTERS and the IEEE ACCESS.

Zhijin Qin (S'13–M'16) received the bachelor's degree from the Beijing University of Posts and Telecommunications, China, in 2012, and the Ph.D. degree in electronic engineering from the Queen Mary University of London, U.K., in 2016. She was with the Department of Computing, Imperial College London, as a Research Associate, from 2016 to 2017. She has been a Lecturer (Assistant Professor) with the School of Computing and Communications, Lancaster University, since 2017. Her research interests include low power wide area network in Internet of Things, compressive sensing, and non-orthogonal multiple access in 5G networks.

She served as a TPC member for various IEEE conferences, such as GLOBECOM'17, ICC'16, and VTC'14-15. She was a recipient of the Best Paper Award at the Wireless Technology Symposium, London, U.K., in 2012. She currently serves as an Editor of the IEEE ACCESS.

Maged Elkashlan (M'06) received the Ph.D. degree in electrical engineering from The University of British Columbia, Canada, in 2006. From 2006 to 2007, he was with the Laboratory for Advanced Networking, The University of British Columbia. From 2007 to 2011, he was with the Wireless and Networking Technologies Laboratory, Commonwealth Scientific and Industrial Research Organization, Australia. During this time, he held an adjunct appointment at the University of Technology Sydney, Australia. In 2011, he joined the School of Electronic Engineering and Computer Science, Queen Mary University of London, U.K., as an Assistant Professor. He also holds visiting faculty appointments at the University of New South Wales, Australia, and the Beijing University of Posts and Telecommunications, China. His research interests fall into the broad areas of communication theory, wireless communications, and statistical signal processing for distributed data processing and heterogeneous networks.

Dr. Elkashlan received the Best Paper Award at the IEEE International Conference on Communications in 2014, the International Conference on Communications and Networking in China in 2014, and the IEEE Vehicular Technology Conference in 2013. He received the Exemplary Reviewer Certificate of the IEEE COMMUNICATIONS LETTERS in 2012. He currently serves as an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, and the IEEE COMMUNICATIONS LETTERS. He also serves as the Lead Guest Editor for the Special Issue on Green Media: The Future of Wireless Multimedia Networks of the *IEEE Wireless Communications Magazine*, the Lead Guest Editor for the Special Issue on Millimeter Wave Communications for 5G of the *IEEE Communications Magazine*, a Guest Editor for the Special Issue on Energy Harvesting Communications of the *IEEE Communications Magazine*, and a Guest Editor for the Special Issue on Location Awareness for Radios and Networks of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS.

Arumugam Nallanathan (S'97–M'00–SM'05–F'17) was an Assistant Professor with the Department of Electrical and Computer Engineering, National University of Singapore, from 2000 to 2007. He was with the Department of Informatics, King's College London, from 2007 to 2017, where he was a Professor of wireless communications from 2013 to 2017. He has been a Professor of wireless communications with the School of Electronic Engineering and Computer Science, Queen Mary University of London, since 2017. He has authored or co-authored over 350 technical papers in scientific journals and international conferences. His research interests include 5G wireless networks, Internet of Things, and molecular communications.

He was a co-recipient of the Best Paper Award at the IEEE International Conference on Communications 2016 (ICC2016) and the IEEE International Conference on Ultra-Wideband 2007 (ICUWB 2007). He is an IEEE Distinguished Lecturer. He received the IEEE Communications Society SPCE Outstanding Service Award 2012 and the IEEE Communications Society RCC Outstanding Service Award 2014. He was selected as a Web of Science (ISI) Highly Cited Researcher in 2016. He is an Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS and the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. He was an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS (2006–2011), the IEEE WIRELESS COMMUNICATIONS LETTERS, and the IEEE SIGNAL PROCESSING LETTERS. He served as the Chair of the Signal Processing and Communication Electronics Technical Committee of the IEEE Communications Society and the Technical Program Chair and member of Technical Program Committees in numerous IEEE conferences.

Julie A. McCann currently directs the Adaptive Emergent/Embedded/Ephemeral Systems Engineering Research Group, which explores how wireless sensing-based actuator systems interact with physical world and how this knowledge can inform protocols, edge-analytics, and delay tolerant control systems. She also co-directs the Intel Collaborative Research Institute in Urban IoT systems and the cross Imperial Smart Connected Futures Centre who explore IoT, cyber-physical systems, crowdsourcing, and related topics. She is an active program committee member for many of the adaptive computing, sensor network, and communications journals and conferences. She has also co-chaired conferences, such as the IEEE SECON and the ACM Self-Adaptive and Self-Organising Systems. She has invited to give many keynotes, distinguished seminars, and has lead and participated in academic and industrial panel discussions. She has been called to advise government on aspects concerning IoT, procurement, and security.