

UDN: A Holistic Analysis of Multi-Piece Path Loss, Antenna Heights, Finite Users and BS Idle Modes

Ming Ding[‡], *Senior Member, IEEE*, David López Pérez[†], *Senior Member, IEEE*,

Youjia Chen^{*}, *Member, IEEE*, Guoqiang Mao^{‡†}, *Fellow, IEEE*,

Zihuai Lin[§], *Senior Member, IEEE*, Albert Y. Zomaya[§], *Fellow, IEEE*

[‡]Data61, CSIRO, Australia, [†]Nokia Bell Labs, Ireland

^{*}Fuzhou University, P.R.China, [‡]University of Technology Sydney, Australia

[§]The University of Sydney, Australia

Abstract—We discover a new capacity scaling law in ultra-dense networks under practical system assumptions, such as a general multi-piece path loss model, a non-zero base station to user equipment antenna height difference, and a finite user equipment density. The intuition and implication of this new capacity scaling law are completely different from that found in year 2011. That law indicated that the increase of the interference power caused by a denser network would be exactly compensated by the increase of the signal power due to the reduced distance between transmitters and receivers, and thus network capacity should grow *linearly* with network densification. However, we find that both the signal and interference powers become bounded in practical ultra-dense networks, which leads to a *constant* capacity scaling law. Moreover, our new discovery on the *constant* capacity scaling law indicates three network optimization problems respectively for base station deployment, user equipment scheduling and base station coordination. These three optimization problems are justified and solved in this paper, shedding new light on the deployment and optimization of ultra-dense networks.

Index Terms—capacity scaling law, coverage probability, area spectral efficiency, stochastic geometry, network densification

I. INTRODUCTION

Orthogonal deployments of dense small cell networks (SCNs), in which small cells and macrocells operate in different frequency bands (i.e., the 3rd Generation Partnership Project (3GPP) SCN Scenario #2a in [1]), have been identified as one of the most promising approaches to rapidly increase network capacity in the 4th-generation (4G) and the 5th-generation (5G) systems [2]. This deployment case is particularly useful because it provides a large spatial spectrum reuse and facilitates an easy network management due to the low interaction with the macrocell network, e.g., no inter-tier interference. In this paper, we focus on the analysis of these dense SCNs as they go ultra-dense (UD) in 5G systems, a.k.a. ultra-dense networks (UDNs), and shed new light on their capacity scaling law.

In Fig. 1, we plot the different capacity laws derived for an ultra-dense network (UDN) by the most advanced research works in the literature, from year 2011 to year 2017. In more detail,

- The green curve shows the signal to interference plus noise ratio (SINR) invariance law found in 2011 [3]. This law stated that the base station (BS) density would not

affect the per-BS coverage probability performance in an interference-limited¹ and fully-loaded² wireless network. As a result, the area spectral efficiency (ASE) would grow linearly with the BS density. The intuition behind this result was that the increase of the interference power caused by a denser network would be exactly compensated by the increase of the signal power originated by the reduced user to serving BS distance.

- The red curve shows the ASE crawl law found in 2016 [5]. This law stated, in contrast with the previous one, that the SINR — and thus the coverage probability, and subsequently the ASE — significantly degrades in the UDN regime. This is due to the increased interference caused by the transition of a large number of interference paths from non-line of sight (NLoS) to line of sight (LoS).
- The blue curve shows the ASE crash law found in 2017 [6]. This law stated that the SINR — and thus the coverage probability, and subsequently the ASE — suffers from an even severer degradation in the UDN regime than in the previous case due to the receive signal power cap imposed by the antenna height difference between the user and its serving BS antennas.
- The black curve shows the ASE take-off law found in 2017 [7]. This law, in contrast to the previous ones, indicates that the SINR — and thus the coverage probability, and subsequently the ASE — can be significantly enhanced when considering a finite active user density because of the interference mitigation provided by the idle mode capability deployed at the small cell BSs. It is important to note that this work did NOT consider the antenna height difference between the user and its serving BS antennas, and thus it did NOT touch the ASE crash effect.

Considering the above theoretical laws, a fundamental question arises: “Which capacity law best characterises an UDN?”. “Will the ASE crash have a larger weight in ASE performance than the ASE take-off, or the other way around?”. This paper

¹In an interference-limited network, the power of each BS is set to a value much larger than the noise power.

²In a fully-loaded network, all BSs are active. Such assumption implies that the user density is infinity or much larger than the BS density. According to the results in [4], the user density should be at least 10 times higher than the BS density to make sure that almost all BSs are active.

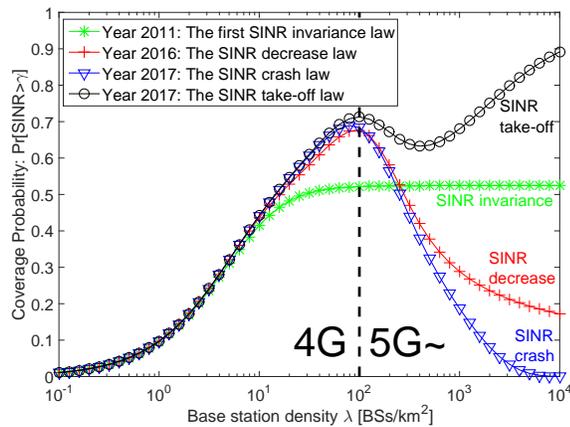


Fig. 1. Theoretical performance comparison of the coverage probability for an SINR threshold $\gamma = 0$ dB. Note that all the results are obtained using a practical 3GPP channel model [5], which will be introduced in detail later in this paper. Also note that the BS density regions for a typical 4G and 5G network have been illustrated in the figure, considering that the maximum BS density of a 4G SCN is in the order of 100 BSs/km^2 [1].

answers this question, and proposes a new capacity law — based on a new SINR invariance law in the UDN regime — different from those above. In more detail, this paper simultaneously considers for the first time the effects of *i)* the transition of a large number of interference paths from NLoS to LoS, *ii)* the antenna height difference between the user and its serving BS antennas, *iii)* a finite active user density, and *iv)* an idle mode capability deployed at the small cell BSs.

Moreover, insight for this new SINR capacity law, three basic network deployment problem are investigated. Firstly, should the network densification ever continue or stop in a certain point? Secondly, our new capacity scaling law points out that instead of letting the network capacity crash serving an infinite UE density, proactively choosing and serving a subset of UEs, leading to a finite UE density on each time/frequency resource block is an effective way to optimise network performance. And hence, the UE scheduling problem is investigated in this paper to find out the optimal user load in a network with density λ . Thirdly, there may be cases where the UE density may be tremendously large and it is needed to simultaneously serve a large number of UEs, e.g. Internet of Things (IoT) applications. In such cases we may not be able to simply decrease the active UE density to avoid the network capacity crash. How to optimize the network deployment with a density λ to serve more UEs at a time? In this paper, a BS coordination problem is investigated.

In particular, the main contributions of this paper are three-fold:

- We present and prove the existence of a new SINR invariance in UDNs under practical assumptions.
- Then, we present and prove a new capacity scaling law in UDNs, which is a constant capacity scaling law.
- Such capacity scaling law leads to three network optimization problems in UDNs, which are formally presented and solved in this paper.

Note that preliminary results of this work have been presented as a conference paper [8]. Compared with [8], the

additional contributions of this journal extension are: 1) The three network optimization problems respectively for BS deployment, UE scheduling and BS coordination are solved in this paper, which were not treated in [8]; 2) The mathematical proofs of the discovered capacity scaling law are provided in this paper, which were omitted in [8]; 3) The ASE map with respect to both the BS density and the UE density is presented, which shows a complete picture of the network behavior of dense SCNs; 4) Additional numerical results considering *Rician fading* are also provided to validate the discovered constant capacity scaling law.

The rest of this paper is structured as follows. Section II provides a brief review of the related work. Section III describes the network scenario and the system model considered in this paper. Section IV presents our theoretical results on the coverage probability and the ASE, followed by our discoveries of a new SINR invariance law and a new capacity scaling law in UDNs. The three network optimization problems are investigated in Section V. The numerical results are discussed in Section VI, with remarks shedding new light on the significance of our discoveries. The conclusions are drawn in Section VII.

II. RELATED WORK

In stochastic geometry, BS positions are typically modeled as a Homogeneous Poisson Point Process (HPPP) on the plane, and closed-form expressions of coverage probability can be found for some scenarios in single-tier cellular networks [3] and multi-tier cellular networks [9]. The major conclusion in [3, 9] is that neither the number of cells nor the number of cell tiers changes the coverage probability in interference-limited fully-loaded wireless networks. Recently, a few noteworthy studies have been carried out to further investigate the network performance analysis for dense and ultra-dense SCNs under more practical propagation models. The authors of [5, 10, 11] found that the coverage probability performance will start to decrease when the BS density is sufficiently large, but such decrease of coverage probability does *not* change the monotonic increase of the ASE as the BS density increases. The intuition behind this result is that as the BS density becomes larger than a threshold, the interference power increases faster than the signal power due to the transition of a large number of interference paths from NLoS to LoS [5]. Moreover, channel models with different LoS probability functions and different multi-path models have been analysed in [5] and [12], respectively, from which it can be concluded that the channel model changes quantitatively the results, but not qualitatively.

However, none of the above works [5, 10, 11] considered the antenna heights of BSs and UEs in the theoretical analysis. The first work setting the spotlight on the antenna height issue and showing the subsequent ASE crash can be found in [6, 13]. The results therein have been confirmed by another two very recent studies [14, 15], which investigated multiple non-zero BS-to-UE antenna height differences L [14], and $L = 4.5 \text{ m}$ [15], respectively.

Note that another type of ASE crash was recently reported in [16], which is caused by a bounded path loss in the near-field (NF) region. However, the non-zero BS-to-UE antenna

height difference renders the existence of the NF effect difficult, since in practice, the former one occurs when L is in the order of meters [6], while the latter one only emerges when the distance between a transmitter and a receiver is in the sub-meter region [16]. It is important to note that the authors of [17, 18] recently proposed a new approach of network performance analysis based on HPPP intensity matching. Such new approach may also be used to investigate the BS antenna height issue and the NF effect.

Furthermore, all of the above works ignored another important factor in practical networks. As the BS density increases, a large number of BSs can be put into idle mode without signal transmission, if there is no active UE within their coverage areas. This is a new network behavior arising from the surplus of BSs with respect to the finite number of UEs, i.e., it may occur that a significant number of BSs may not have any active UE in their coverage areas during certain time periods. Therefore, such BSs could mute their transmission to mitigate unnecessary inter-cell interference and reduce energy consumption [4, 18, 19].

Up to now, the existing work that did consider the IMC is limited. The only work that studied the IMC considering a general multi-piece path loss model with probabilistic LoS and NLoS transmissions can be found in [7]. Two important conclusions were drawn in it: (i) the active BS density with the mentioned probabilistic LoS/NLoS path loss model is lower-bounded by that with the simplistic single-slope path loss model derived in [4], and (ii) such lower bound, shown in [4], is tight, especially for UDNs. This shows a simple way of studying the IMC in UDNs.

In this paper, motivated by the above survey of the related work, we take a further step and investigate the combined performance impact of the following two phenomena on UDNs: (i) a non-zero BS-to-UE antenna height difference, L , and (ii) a non-fully-loaded network with a finite UE density, ρ , exploited by the IMC. Note that our study will consider the mentioned probabilistic LoS/NLoS path loss model recommended by the 3GPP, and reveal a new capacity scaling law for practical UDNs in 5G systems.

III. NETWORK SCENARIO AND SYSTEM MODEL

In this paper, we assume every user is connected to a small cell, for capacity boosting, and to a macrocell, for mobility purposes. Moreover, small cells and macrocells work in different frequency bands to avoid cross-tier interference. To study the impact of network densification, we focus on the small cell tier performance. That is, the BSs mentioned in this paper referred to the small cell BSs for brevity. In more detail, the urban scenario is considered according to [20]. We focus the impact of network densification using traditional sub-6GHz small cell technology, such as LTE small cell BS or 802.11 access points.

A. Network Scenario

We consider a downlink (DL) cellular network with BSs deployed on a plane according to a homogeneous Poisson point process (HPPP) Φ with a density of λ BSs/km². Active

DL UEs are also Poisson distributed in the considered network with a density of ρ UEs/km². Here, we only consider active UEs in the network because non-active UEs do not trigger any data transmission, and thus they can be safely ignored in the analysis. Note that the total UE number in cellular networks should be much higher than the number of the active UEs, but at a certain time slot and on a certain frequency band, the active UEs with data traffic demands are not too many. As indicated earlier, a typical UE density in populated scenarios is around $\rho = 300$ UEs/km² [2].

In practice, a BS will enter into idle mode, if there is no UE connected to it, which reduces the interference to UEs in neighboring BSs as well as the energy consumption of the network. Since UEs are randomly and uniformly distributed in the network, a widely accepted assumption is that the active BSs should follow another HPPP distribution $\tilde{\Phi}$ [4], the density of which is $\tilde{\lambda}$ BSs/km². Note that $\tilde{\lambda} \leq \lambda$ and $\tilde{\lambda} \leq \rho$, since one UE is served by at most one BS. Also note that a larger ρ requires more active BSs to serve the more active UEs, thus leading to a larger $\tilde{\lambda}$. From [4, 7], it was shown that the formula proposed in [4] to calculate $\tilde{\lambda}$ is accurate for UDNs, which is given by

$$\tilde{\lambda} = \lambda \left[1 - \frac{1}{\left(1 + \frac{\rho}{q\lambda}\right)^q} \right], \quad (1)$$

where an empirical value of 3.5 was suggested for q in [4].

B. Wireless System Model

The two-dimensional (2D) distance between a BS and a UE is denoted by r . Moreover, the absolute antenna height difference between a BS and a UE is denoted by L . Thus, the 3D distance between a BS and a UE can be expressed as

$$w = \sqrt{r^2 + L^2}. \quad (2)$$

Note that the value of L is in the order of several meters [21].

Following [5], we adopt a general path loss model, where the path loss $\zeta(w)$ is a multi-piece function of w written as

$$\zeta(w) = \begin{cases} \zeta_1(w), & \text{when } L \leq w \leq d_1 \\ \zeta_2(w), & \text{when } d_1 < w \leq d_2 \\ \vdots & \vdots \\ \zeta_N(w), & \text{when } w > d_{N-1} \end{cases}, \quad (3)$$

where each piece $\zeta_n(w)$, $n \in \{1, 2, \dots, N\}$ is modeled as

$$\zeta_n(w) = \begin{cases} \zeta_n^L(w) = A_n^L w^{-\alpha_n^L}, & \text{LoS: } \Pr_n^L(w) \\ \zeta_n^{NL}(w) = A_n^{NL} w^{-\alpha_n^{NL}}, & \text{NLoS: } 1 - \Pr_n^L(w) \end{cases}, \quad (4)$$

where

- $\zeta_n^L(w)$ and $\zeta_n^{NL}(w)$, $n \in \{1, 2, \dots, N\}$ are the n -th piece path loss functions for the LoS and the NLoS cases, respectively,
- A_n^L and A_n^{NL} are the path losses at a reference 3D distance $w = 1$ for the LoS and the NLoS cases, respectively,
- α_n^L and α_n^{NL} are the path loss exponents for the LoS and the NLoS cases, respectively.

Moreover, $\text{Pr}_n^L(w)$ is the n -th piece LoS probability function that a transmitter and a receiver separated by a 3D distance w has a LoS path, which is assumed to be a *monotonically decreasing function* with respect to w . Existing measurement studies have confirmed this assumption [21].

For convenience, $\{\zeta_n^L(r)\}$ and $\{\zeta_n^{\text{NL}}(r)\}$ can be further stacked into piece-wise functions written as

$$\zeta^{\text{Path}}(r) = \begin{cases} \zeta_1^{\text{Path}}(r), & \text{when } L \leq r \leq d_1 \\ \zeta_2^{\text{Path}}(r), & \text{when } d_1 < r \leq d_2 \\ \vdots & \vdots \\ \zeta_N^{\text{Path}}(r), & \text{when } r > d_{N-1} \end{cases}, \quad (5)$$

where the string variable *Path* takes the value of “L” and “NL” for the LoS and the NLoS cases, respectively. Besides, $\{\text{Pr}_n^L(r)\}$ can also be stacked into a piece-wise function as

$$\text{Pr}^L(r) = \begin{cases} \text{Pr}_1^L(r), & \text{when } L \leq r \leq d_1 \\ \text{Pr}_2^L(r), & \text{when } d_1 < r \leq d_2 \\ \vdots & \vdots \\ \text{Pr}_N^L(r), & \text{when } r > d_{N-1} \end{cases}. \quad (6)$$

Note that the generality and the practicality of the adopted path loss model (3) have been well established in [5]. As a special case to show our analytical results in the following sections, we consider a practical two-piece path loss function and a two-piece exponential LoS probability function, defined by the 3GPP [21]. Specifically, we have $N = 2$, $\zeta_1^L(w) = \zeta_2^L(w) = A^L w^{-\alpha^L}$, $\zeta_1^{\text{NL}}(w) = \zeta_2^{\text{NL}}(w) = A^{\text{NL}} w^{-\alpha^{\text{NL}}}$, $\text{Pr}_1^L(w) = 1 - 5 \exp(-R_1/w)$, and $\text{Pr}_2^L(w) = 5 \exp(-w/R_2)$, where $R_1 = 156$ m, $R_2 = 30$ m, and $d_1 = \frac{R_1}{\ln 10} = 67.75$ m [21]. For clarity, this case is referred to as **the 3GPP Case** hereafter. Note that this 3GPP Case has been used to generate the results in Fig. 1 of Section I.

Moreover, we assume a practical user association strategy (UAS), in which each UE is connected to the BS giving the maximum average received signal strength (i.e., with the largest $\zeta(w)$) [5, 11]). It is very important to note that in our previous work [22] and some other existing work, e.g., [3, 10], it was assumed that each UE is associated with its closest BS. Such assumption is not appropriate for the considered path loss model in (3), because in practice a UE should connect to the BS offering the largest received signal strength. Such BS does not necessarily have to be the nearest one to the UE, and it could be a farther one with a strong LoS path.

As discussed in Section II, with the consideration of the IMC and the general path loss model in (3), it was concluded in [7] that it is accurate to characterize $\tilde{\lambda}$ using (1) with $q = 3.5$, especially for UDNs. For simplicity, in this paper, we use (1) to compute $\tilde{\lambda}$.

Finally, we assume that each BS’s transmission power is a constant value P , each BS/UE is equipped with an isotropic antenna, and the multi-path fading between a BS and a UE is modeled as independently identical distributed (i.i.d.) Rayleigh fading [5, 10, 11].

It is important to note that it has been shown in [23] that the analysis of a more accurate multi-path modeling with Rician fading is not urgent, as it does not change the qualitative conclusions of this type of performance analysis for single-antenna UDNs. Thus, we will first focus on presenting our most fundamental discoveries based on Rayleigh fading in the next section, and then study Rician fading later. In particular, a more practical Rician fading will be investigated in Section VI to show its minor impact on our conclusions. More specifically, for LoS transmissions, we adopt a practical Rician fading defined in the 3GPP [24], where the K factor in dB scale (the ratio between the power in the direct path and the power in the other scattered paths) is modeled as a function of distance, i.e., $K[\text{dB}] = 13 - 0.03w$, where w is defined in (2).

IV. MAIN RESULT

In this section, we study the coverage probability performance and the network capacity in terms of the ASE for a typical UE located at the origin o .

A. The Coverage Probability

First, we investigate the coverage probability that the SINR of a typical UE at the origin o is above a threshold γ :

$$p^{\text{cov}}(\lambda, \rho, \gamma) = \Pr[\text{SINR} > \gamma], \quad (7)$$

where the SINR is computed by

$$\text{SINR} = \frac{P\zeta(w)h}{I_{\text{agg}} + P_N}, \quad (8)$$

where h is the channel gain, which is modeled as an exponentially distributed random variable (RV) with a mean of one due to our consideration of Rayleigh fading (see Subsection III-B), P and P_N are the BS transmission power and the additive white Gaussian noise (AWGN) power at each UE, respectively, and I_{agg} is the cumulative interference given by

$$I_{\text{agg}} = \sum_{i: b_i \in \tilde{\Phi} \setminus b_o} P\beta_i g_i, \quad (9)$$

where b_o is the BS serving the typical UE, and b_i , β_i and g_i are the i -th interfering BS, the path loss from b_i to the typical UE and the multi-path fading channel gain associated with such link (also exponentially distributed RVs), respectively. Note that, when all BSs are assumed to be active, the set of all BSs Φ should be used in the expression of I_{agg} [5, 10, 11]. Here, in (9), only the BSs in $\tilde{\Phi} \setminus b_o$ inject effective interference into the network and only they are thus considered in the analysis, where $\tilde{\Phi}$ denotes the set of the active BSs. In other words, the BSs in idle mode are not taken into account in the computation of I_{agg} .

Note that the analytical results of $p^{\text{cov}}(\lambda, \rho, \gamma)$ can be directly obtained from [6], which is omitted here for brevity. In this paper, we focus on the performance scaling law, and we present our main result on the asymptotic performance of $p^{\text{cov}}(\lambda, \rho, \gamma)$ for UDNs, i.e., $\lim_{\lambda \rightarrow +\infty} p^{\text{cov}}(\lambda, \rho, \gamma)$, in Theorem 1.

Theorem 1. Considering the general path loss model in (3) and the adopted UAS, we can derive $\lim_{\lambda \rightarrow +\infty} p^{\text{cov}}(\lambda, \rho, \gamma)$ as

$$\begin{aligned} \lim_{\lambda \rightarrow +\infty} p^{\text{cov}}(\lambda, \rho, \gamma) &= \lim_{\lambda \rightarrow +\infty} \Pr \left[\frac{P\zeta_1^L(L)h}{I_{\text{agg}} + P_N} > \gamma \right] \\ &= \exp \left(-\frac{P_N\gamma}{P\zeta_1^L(L)} \right) \lim_{\lambda \rightarrow +\infty} \mathcal{L}_{I_{\text{agg}}}^L \left(\frac{\gamma}{P\zeta_1^L(L)} \right), \end{aligned} \quad (10)$$

where

$$\begin{aligned} \lim_{\lambda \rightarrow +\infty} \mathcal{L}_{I_{\text{agg}}}^L(s) &= \\ &\exp \left(-2\pi\rho \int_0^{+\infty} \frac{\Pr^L(\sqrt{u^2 + L^2})u}{1 + (sP\zeta^L(\sqrt{u^2 + L^2}))^{-1}} du \right) \times \\ &\exp \left(-2\pi\rho \int_0^{+\infty} \frac{[1 - \Pr^L(\sqrt{u^2 + L^2})]u}{1 + (sP\zeta^{\text{NL}}(\sqrt{u^2 + L^2}))^{-1}} du \right). \end{aligned} \quad (11)$$

Proof: Please refer to Appendix A. ■

From Theorem 1, we propose a new SINR invariance law in Theorem 2.

Theorem 2. A new SINR invariance law: If $L > 0$ and $\rho < +\infty$, then $\lim_{\lambda \rightarrow +\infty} p^{\text{cov}}(\lambda, \rho, \gamma)$ becomes a constant that is independent of λ in UDNs.

Proof: Please refer to Appendix B. ■

Theorem 2 dictates that (i) the SINR decrease effect due to the non-zero BS-to-UE antenna height difference L and (ii) the SINR increase due to the finite UE density ρ exploited by the IMC, cancel each other out in practical UDNs with $L > 0$ and $\rho < +\infty$. Note that the study on $\{L, \rho\}$ is finally complete with the availability of Theorem 2 because:

- The case of $L = 0$ and $\rho = +\infty$ has been studied in [5, 10, 11], showing that $\lim_{\lambda \rightarrow +\infty} p^{\text{cov}}(\lambda, \rho, \gamma)$ is a function of α_n^L .
- The case of $L > 0$ and $\rho = +\infty$ has been studied in [6], showing that $\lim_{\lambda \rightarrow +\infty} p^{\text{cov}}(\lambda, \rho, \gamma) = 0$, as illustrated in Fig. 1.
- The case of $L = 0$ and $\rho < +\infty$ has been studied in [7], showing that $\lim_{\lambda \rightarrow +\infty} p^{\text{cov}}(\lambda, \rho, \gamma) = 1$, as illustrated in Fig. 1.
- The case of $L > 0$ and $\rho < +\infty$ is characterized by Theorem 2, which reflects the most practical SCN deployment scenario among the above cases.

From Theorem 2, it is trivial to show that for a given $\{L, \rho\}$, $\lim_{\lambda \rightarrow +\infty} p^{\text{cov}}(\lambda, \rho, \gamma)$ decreases as γ increases since a higher SINR requirement implies a lower coverage probability. Thus, in Lemmas 1 and 2, we only address how $\lim_{\lambda \rightarrow +\infty} p^{\text{cov}}(\lambda, \rho, \gamma)$ varies with L and ρ , respectively.

Lemma 1. For a given $\{\rho, \gamma\}$, $\lim_{\lambda \rightarrow +\infty} p^{\text{cov}}(\lambda, \rho, \gamma)$ decreases as L increases.

Proof: Please refer to Appendix C. ■

Lemma 2. For a given $\{L, \gamma\}$, $\lim_{\lambda \rightarrow +\infty} p^{\text{cov}}(\lambda, \rho, \gamma)$ decreases as ρ increases, according to a power law with respect to ρ .

More specifically, we have

$$\lim_{\lambda \rightarrow +\infty} p^{\text{cov}}(\lambda, \rho, \gamma) = c(\gamma)g^{\rho}(\gamma), \quad (12)$$

where $c(\gamma)$ and $g(\gamma)$ are expressed as

$$c(\gamma) = \exp \left(-\frac{P_N\gamma}{P\zeta_1^L(L)} \right), \quad (13)$$

and

$$\begin{aligned} g(\gamma) &= \exp \left(-2\pi \int_0^{+\infty} \frac{\Pr^L(\sqrt{u^2 + L^2})u}{1 + (sP\zeta^L(\sqrt{u^2 + L^2}))^{-1}} du \right) \\ &\times \exp \left(-2\pi \int_0^{+\infty} \frac{[1 - \Pr^L(\sqrt{u^2 + L^2})]u}{1 + (sP\zeta^{\text{NL}}(\sqrt{u^2 + L^2}))^{-1}} du \right), \end{aligned} \quad (14)$$

where $s = \frac{\gamma}{P\zeta_1^L(L)}$.

Proof: Please refer to Appendix D. ■

The intuitions of Lemmas 1 and 2 are explained as follows:

- The signal power becomes bounded in UDNs due to the lower-bound on the BS-to-UE distance, as a UE cannot be closer than L to a BS. Moreover, a larger L implies a tighter bound on the signal power, leading to the decrease of $\lim_{\lambda \rightarrow +\infty} p^{\text{cov}}(\lambda, \rho, \gamma)$ in Lemma 1.
- The interference power becomes bounded in UDNs due to the activation of a finite density of BSs (i.e., $\tilde{\lambda}$ BSs/km²) to serve a finite density of UEs (i.e., ρ UEs/km²). Moreover, a larger ρ results in a larger $\tilde{\lambda}$, relaxing the bound on the interference power, which leads to the decrease of $\lim_{\lambda \rightarrow +\infty} p^{\text{cov}}(\lambda, \rho, \gamma)$ in Lemma 2. Such decrease follows a power law with respect to ρ , because an HPPP distribution of UEs with ρ UEs/km² can be decomposed into ρ independent HPPP ones with 1 UEs/km², and the coverage criterion (7) should be satisfied for every one of these HPPP distributions, which yields a power law with respect to ρ .

B. The Area Spectral Efficiency

Next, we investigate the network capacity performance in terms of the ASE in bps/Hz/km², which is defined as [5]

$$A^{\text{ASE}}(\lambda, \rho, \gamma_0) = \tilde{\lambda} \int_{\gamma_0}^{+\infty} \log_2(1 + \gamma) f_{\Gamma}(\lambda, \rho, \gamma) d\gamma, \quad (15)$$

where $\tilde{\lambda}$ is calculated from (1), γ_0 is the minimum working SINR of a practical SCN, and $f_{\Gamma}(\lambda, \rho, \gamma)$ is the probability density function (PDF) of the SINR γ observed at the typical UE for a particular pair of values $\{\lambda, \rho\}$. Based on the definition of $p^{\text{cov}}(\lambda, \rho, \gamma)$ in (7) and the partial integration theorem shown in [18], (15) can be reformulated as

$$\begin{aligned} A^{\text{ASE}}(\lambda, \rho, \gamma_0) &= \frac{\tilde{\lambda}}{\ln 2} \int_{\gamma_0}^{+\infty} \frac{p^{\text{cov}}(\lambda, \rho, \gamma)}{1 + \gamma} d\gamma \\ &+ \tilde{\lambda} \log_2(1 + \gamma_0) p^{\text{cov}}(\lambda, \rho, \gamma_0). \end{aligned} \quad (16)$$

From (1), we have that $\tilde{\lambda}$ is a finite value since $\rho < +\infty$ and $\tilde{\lambda}$ represents the spatial spectrum reuse in UDNs. Thus,

$\tilde{\lambda}$ is used in the expression of $A^{\text{ASE}}(\lambda, \rho, \gamma_0)$ because only the active BSs make an effective contribution to the ASE. The ASE definition in (15) captures the dependence of the transmission rate on SINR, but it is less tractable to analyze, as it requires one more fold of numerical integral compared with the ASE definition in [10].

C. A New Capacity Scaling Law

From Theorem 3 and the expression of the ASE in (16), we propose a new capacity scaling law in Theorem 3.

Theorem 3. *A constant capacity scaling law: If $L > 0$ and $\rho < +\infty$, then $\lim_{\lambda \rightarrow +\infty} A^{\text{ASE}}(\lambda, \rho, \gamma_0)$ becomes a constant that is independent of λ in UDNs. In more detail, $\lim_{\lambda \rightarrow +\infty} A^{\text{ASE}}(\lambda, \rho, \gamma_0)$ is given by*

$$\lim_{\lambda \rightarrow +\infty} A^{\text{ASE}}(\lambda, \rho, \gamma_0) = \frac{\rho}{\ln 2} \int_{\gamma_0}^{+\infty} \frac{\lim_{\lambda \rightarrow +\infty} p^{\text{cov}}(\lambda, \rho, \gamma)}{1 + \gamma} d\gamma + \rho \log_2(1 + \gamma_0) \lim_{\lambda \rightarrow +\infty} p^{\text{cov}}(\lambda, \rho, \gamma_0), \quad (17)$$

where $\lim_{\lambda \rightarrow +\infty} p^{\text{cov}}(\lambda, \rho, \gamma)$ is obtained from Theorem 1, and it is independent of λ .

Proof: Please refer to Appendix E. ■

V. INSIGHTS OF OUR CONSTANT CAPACITY LAW

The implications of our *new capacity scaling law* — a constant scaling law — derived in Theorem 3 are profound, and they are discussed in more detail in the following.

A. The Limits of Network Densification: The BS Deployment/Activation Problem

From this new capacity scaling law, it can be concluded that, for a given active UE density ρ , network densification should not be abused indefinitely, and instead should be *stopped* at a certain BS density. This is because both the coverage probability and the ASE will asymptotically reach a maximum constant value, and any network densification beyond such level is a waste of both money and energy.

As a result, and to find the optimum active BS density λ^* , we propose the following BS deployment/activation problem, which is formulated as follows. For a given UE density ρ , there exists an optimal active BS density λ^* that can achieve an ASE $A^{\text{ASE}}(\lambda, \rho, \gamma_0)$ that is within a gap, ϵ -percent, from the asymptotic maximum ASE value $\lim_{\lambda \rightarrow +\infty} A^{\text{ASE}}(\lambda, \rho, \gamma_0)$, i.e.,

$$\begin{aligned} & \underset{\lambda}{\text{maximize}} && 1 \\ & \text{s.t.} && \frac{\left| \lim_{\lambda \rightarrow +\infty} A^{\text{ASE}}(\lambda, \rho, \gamma_0) - A^{\text{ASE}}(\lambda, \rho, \gamma_0) \right|}{\lim_{\lambda \rightarrow +\infty} A^{\text{ASE}}(\lambda, \rho, \gamma_0)} = \epsilon. \end{aligned} \quad (18)$$

Note that the solution λ^* to this BS deployment/activation problem (18) answers the fundamental question of “for a given UE density ρ , how dense an UDN should be?”. As one would

expect, the answer to such question depends on the active UE density ρ .

Overall, and as illustrated in Fig. 1, an operator should *stop* its network densification at BS density λ^* , since network capacity saturates at such BS density λ^* , and thus any further densification will not provide more than an ϵ -percent gain with respect to $\lim_{\lambda \rightarrow +\infty} A^{\text{ASE}}(\lambda, \rho, \gamma_0)$.

As shown by (18), the BS deployment/activation problem is not complex, and its solution can be easily found by numerical search over the ASE, $A^{\text{ASE}}(\lambda, \rho, \gamma_0)$, the details of which are omitted for brevity.

B. The Network-wide UE Scheduling Problem

As mentioned in Section I and II, concerns about a network capacity crash in UDNs emerged in the literature, but were unfounded, as all those studies assumed an unrealistic infinite UE density in UDNs. As shown by Theorem 3, the ASE crash should not occur in practical networks with *i)* a reasonable finite UE density and *ii)* small cell BSs equipped with an idle mode capability.

Having said that, it is important to notice that even in the presence of a very large or infinite active UE density, the ASE crash can still be avoided by using intelligent scheduling decisions. Time division multiple access (TDMA) or frequency division multiple access (FDMA) can be used to divide the very large set of active UEs into smaller *groups of UEs*, each with a finite and moderate UE density. Through properly dimensioning such scheduling groups, an optimum network capacity can be achieved, as indicated by Theorem 3. Such strategy involves a network-wide-aware scheduling, and its implementation is feasible in practice. In other words, instead of letting the network capacity crash serving an infinite active UE density, our new capacity scaling law points out another way of proactively choosing and serving a subset of UEs, leading to a finite active UE density on each time/frequency resource block. This avoids the ASE crash and leads to an optimised network performance.

To find such solution, we further investigated (17), and observe that the asymptotic maximum ASE value $\lim_{\lambda \rightarrow +\infty} A^{\text{ASE}}(\lambda, \rho, \gamma_0)$ is a concave function with respect to the active UE density ρ . This implies, in line with our previous discussion, an optimal active UE density ρ^* that can maximise $\lim_{\lambda \rightarrow +\infty} A^{\text{ASE}}(\lambda, \rho, \gamma_0)$. This is because

- Lemma 2 states that $\lim_{\lambda \rightarrow +\infty} p^{\text{cov}}(\lambda, \rho, \gamma)$ decreases as ρ increases,
- while such increase of ρ linearly scales the terms in (17) too.

To find the optimum active UE density λ^* , we propose the following UE scheduling problem, which is formulated as follows. For a given BS density λ , there exists an optimal active UE density ρ^* that can maximise the ASE $A^{\text{ASE}}(\lambda, \rho, \gamma_0)$, i.e.,

$$\begin{aligned} & \underset{\rho}{\text{maximize}} && A^{\text{ASE}}(\lambda, \rho, \gamma_0) \\ & \text{s.t.} && 0 < \rho \leq \lambda. \end{aligned} \quad (19)$$

Note that the solution ρ^* to this UE scheduling problem (19) answers the fundamental question of “for a given BS density λ , what is the optimal UE load that can maximise the ASE?”.

1) *The Optimal Solution:* First, we consider the case where the BS density tends to infinity, i.e. $\lambda \rightarrow +\infty$, and recall Lemma 2 to rewrite (17) as

$$\lim_{\lambda \rightarrow +\infty} A^{\text{ASE}}(\lambda, \rho, \gamma_0) = \frac{\rho}{\ln 2} \int_{\gamma_0}^{+\infty} \frac{c(\gamma) g^\rho(\gamma)}{1 + \gamma} d\gamma + \log_2(1 + \gamma_0) \rho c(\gamma_0) g^\rho(\gamma_0). \quad (20)$$

Taking the derivative of $\lim_{\lambda \rightarrow +\infty} A^{\text{ASE}}(\lambda, \rho, \gamma_0)$ with respect to the active UE density ρ , and denoting such derivative function by $D^{\text{ASE}}(\rho, \gamma_0)$, we have that

$$\begin{aligned} D^{\text{ASE}}(\rho, \gamma_0) &\triangleq \frac{\partial \left[\lim_{\lambda \rightarrow +\infty} A^{\text{ASE}}(\lambda, \rho, \gamma_0) \right]}{\partial \rho} \\ &= \frac{\partial \left[\frac{\rho}{\ln 2} \int_{\gamma_0}^{+\infty} \frac{c(\gamma) g^\rho(\gamma)}{1 + \gamma} d\gamma + \log_2(1 + \gamma_0) \rho c(\gamma_0) g^\rho(\gamma_0) \right]}{\partial \rho} \\ &= \frac{1}{\ln 2} \int_{\gamma_0}^{+\infty} \frac{c(\gamma) g^\rho(\gamma) (1 + \rho \ln g(\gamma))}{1 + \gamma} d\gamma \\ &\quad + \log_2(1 + \gamma_0) c(\gamma_0) g^\rho(\gamma_0) (1 + \rho \ln g(\gamma_0)). \end{aligned} \quad (21)$$

According to the convex optimisation theory [25], the maximum of $\lim_{\lambda \rightarrow +\infty} A^{\text{ASE}}(\lambda, \rho, \gamma_0)$ is obtained when $D^{\text{ASE}}(\rho, \gamma_0) = 0$. Since $D^{\text{ASE}}(\rho, \gamma_0)$ has a closed-form expression, shown in (21), we can obtain the optimal ρ^* using a bisection search [26]. In paper, we propose Algorithm 1 to calculate ρ^* , which is illustrated in the following.

Algorithm 1 The proposed algorithm to find ρ^*

Step 1: Initialization

- Set $\rho^{\text{left}} = 0$, $\rho^{\text{right}} = \lambda$, $\rho^{\text{mid}} = \frac{\rho^{\text{left}} + \rho^{\text{right}}}{2}$.

Step 2: Iteration

- Compute $D^{\text{ASE}}(\rho^{\text{left}}, \gamma_0)$, $D^{\text{ASE}}(\rho^{\text{right}}, \gamma_0)$, and $D^{\text{ASE}}(\rho^{\text{mid}}, \gamma_0)$ using (21).
- If $D^{\text{ASE}}(\rho^{\text{mid}}, \gamma_0) > 0$, update $\rho^{\text{left}} = \rho^{\text{mid}}$; Else, update $\rho^{\text{right}} = \rho^{\text{mid}}$.

Step 3: Termination

- If $|D^{\text{ASE}}(\rho^{\text{mid}}, \gamma_0)| < \delta_0$, where δ_0 is a small value, which sets a precision condition to terminate the numerical search, go to Step 4; Else, return to Step 2.

Step 4: Output

- $\rho^* = \rho^{\text{mid}}$.
-

Unfortunately, the above calculation only holds when $\lambda \rightarrow +\infty$. For a general case, in which λ adopts a finite value, $D^{\text{ASE}}(\rho, \gamma_0)$ cannot be expressed by a simple closed-form expression, as in (21). This is because $p^{\text{cov}}(\lambda, \rho, \gamma)$ has a much more complicated expression, requiring two folds of integrations. The ASE definition in (16) makes things even more complex, adding another fold of integration. However, it

is important to note that, for a given λ , we can still numerically evaluate $\frac{\partial [A^{\text{ASE}}(\lambda, \rho, \gamma_0)]}{\partial \rho}$, as $A^{\text{ASE}}(\lambda, \rho, \gamma_0)$ has been obtained in semi-closed-form expressions in (16). Therefore, for a general case, in which λ adopts a finite value, we can reuse Algorithm 1 to find ρ^* , with $D^{\text{ASE}}(\rho, \gamma_0)$ replaced with $\frac{\partial [A^{\text{ASE}}(\lambda, \rho, \gamma_0)]}{\partial \rho}$.

C. The BS coordination problem

There may be cases/applications where there is a tremendously large UE density and the need to simultaneously schedule many of them, e.g. Internet of Things (IoT). In such scenarios, it may not be enough to simply decrease the active UE density to avoid the ASE crash, as previously described. Instead, in this section, we propose a BS coordination method able to serve more UEs at a time at expense of a reduce ASE. The method builds in the following fundamental observation. This method builds on time-domain (TD) or frequency domain (FD) inter-cell interference coordination (ICIC). In more detail, this method partitions the set of BSs into N^{BSG} BS groups (BSGs), and allows each BSG to use a disjoint subset of time/frequency resources. This reduces the spatial reuse, but mitigates inter-cell interference, as the number of interferers seen by any given BS decreases. Note that number N^{BSG} is also widely known as the resource reuse factor, which usually takes an integer value.

To formally introduce this approach, suppose that the required UE density to be served is $\rho_0 \leq \lambda$, then we propose the following BS coordination problem: For a given BS density λ , serving a UE density of at least ρ_0 , there exists an optimal N^{BSG^*} that can maximise the ASE. The larger ρ_0 , the larger the spatial reuse, but the smaller the maximum ASE, due to the excessive inter-cell interference. The solution N^{BSG^*} to this problem should strike the right balance between spatial reuse and inter-cell interference mitigation for such given scenario.

This BS coordination problem can be formulated as

$$\begin{aligned} \underset{N^{\text{BSG}} \in \mathbb{N}}{\text{maximize}} \quad & N^{\text{BSG}} \times \frac{1}{N^{\text{BSG}}} \times A^{\text{ASE}} \left(\frac{\lambda}{N^{\text{BSG}}}, \rho, \gamma_0 \right) \\ & = A^{\text{ASE}} \left(\frac{\lambda}{N^{\text{BSG}}}, \rho, \gamma_0 \right) \\ \text{s.t.} \quad & \rho \geq \frac{\rho_0}{N^{\text{BSG}}}, \\ & 1 \leq N^{\text{BSG}} \leq \lambda, \end{aligned} \quad (22)$$

where

- the first term in the objective function, i.e. N^{BSG} , indicates that the ASE is contributed by N^{BSG} BSGs,
- the second term in the objective function, i.e. $\frac{1}{N^{\text{BSG}}}$, indicates that each BSG only uses $\frac{1}{N^{\text{BSG}}}$ of the time/frequency resource,
- the third term in the objective function, i.e. $A^{\text{ASE}} \left(\frac{\lambda}{N^{\text{BSG}}}, \rho, \gamma_0 \right)$, calculates the achievable ASE when the BS density is $\frac{\lambda}{N^{\text{BSG}}}$ and the UE density is ρ ,
- the first constraint fulfils the requirement of that the total UE density to be served (i.e., $N^{\text{BSG}} \rho$) should be no less than ρ_0 , and
- the second constraint is to make sure that at least one BS is present in each BSG.

Note that for a given N^{BSG} , the optimal UE density ρ^* can be found by solving the UE scheduling problem (19). If such ρ^* does not satisfy the first constraint in (22), we should then increase the UE density to $\frac{\rho_0}{N^{\text{BSG}}}$ and recalculate the achievable ASE in (22). Using this logic, an exhaustive search over all possible values of N^{BSG} , while solving (19), on condition of the BS density being $\frac{\lambda}{N^{\text{BSG}}}$, can be used to find the optimal solution $N^{\text{BSG}*}$. Such optimal solution $N^{\text{BSG}*}$ of the BS coordination problem (22) answers the fundamental question of “for a given BS density λ , serving a UE density of at least ρ_0 , what is the optimal resource reuse factor that can maximise the ASE?”.

To solve (22), we propose Algorithm 2 to calculate $N^{\text{BSG}*}$. This algorithm performs an exhaustive search over all possible values of N^{BSG} . In essence, each iteration executes Algorithm 1, checks the feasibility of the solution $\frac{\rho_0}{N^{\text{BSG}}}$ and performs the comparison to find the optimal solution.

Algorithm 2 The proposed algorithm to find $N^{\text{BSG}*}$

Step 1: Initialization

- Set $A_{\text{max}}^{\text{ASE}} = 0$, $N^{\text{BSG}*} = 0$.

Step 2: Iteration

- For $N^{\text{BSG}} = \{1, 2, \dots, \lambda\}$, do the following
 - Use algorithm 1 to find ρ^* on condition of the BS density being $\frac{\lambda}{N^{\text{BSG}}}$.
 - If $\rho^* \geq \frac{\rho_0}{N^{\text{BSG}}}$, continue;
 - Else, set $\rho^* = \frac{\rho_0}{N^{\text{BSG}}}$.
 - If $A^{\text{ASE}}(\frac{\lambda}{N^{\text{BSG}}}, \rho^*, \gamma_0) > A_{\text{max}}^{\text{ASE}}$, update $A_{\text{max}}^{\text{ASE}} = A^{\text{ASE}}(\frac{\lambda}{N^{\text{BSG}}}, \rho^*, \gamma_0)$ and $N^{\text{BSG}*} = N^{\text{BSG}}$;
 - Else, continue.

Step 3: Output

- The optimal solution is $N^{\text{BSG}*}$ and the maximum ASE is $A_{\text{max}}^{\text{ASE}}$.
-

VI. SIMULATION AND DISCUSSION

In this section, we present numerical results to validate the accuracy of our analysis. According to Tables A.1-3~A.1-7 of [21] and [24], we adopt the following parameters for the 3GPP Case: $\alpha^{\text{L}} = 2.09$, $\alpha^{\text{NL}} = 3.75$, $A^{\text{L}} = 10^{-10.38}$, $A^{\text{NL}} = 10^{-14.54}$, $P = 24$ dBm, $P_{\text{N}} = -95$ dBm (with a noise figure of 9 dB). Finally, a very wide BS density ranging from 10^{-1} BSs/km² all the way up to 10^6 BSs/km² is studied.

Aiming at analysing the potential capacity gains that can be obtained by using different levels of network densification, we investigate network performance with respect to various network densities, from quite sparse networks to very dense ones. Inspired by the network density of 10^3 BSs/km² implemented in the field test in 2014 [27], and the even higher one suggested by the 3GPP for potential urban scenario use cases in [20], we set the upper bound network density to 10^6 BSs/km² in our simulations. Although such a high density of BSs is currently not realistic due to techno-economic reasons, our analysis is targeted at providing a fundamental understanding of the different tradeoffs that may arise when considering all network densification levels, from a theoretical perspective.

A. Validation of the Coverage Probability Performance

In Fig. 2, we display the coverage probability for the 3GPP Case with $\gamma = 0$ dB and various values of ρ and L . Here, solid lines, markers, and dash lines represent analytical results, simulation results, and $\lim_{\lambda \rightarrow +\infty} p^{\text{cov}}(\lambda, \rho, \gamma)$ derived in Theorem 1, respectively. Note that the analytical results of

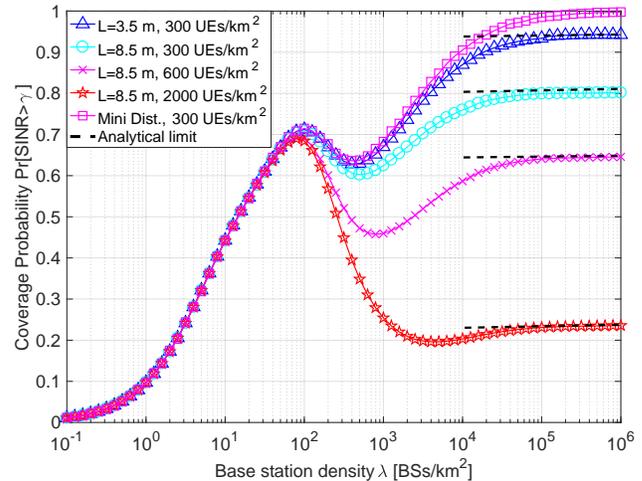


Fig. 2. The coverage probability $p^{\text{cov}}(\lambda, \rho, \gamma)$ vs. λ for the 3GPP Case with $\gamma = 0$ dB and various values of ρ and L .

$p^{\text{cov}}(\lambda, \rho, \gamma)$ in Fig. 2 are obtained from [6] with λ replaced with λ . From this figure, we can observe that:

- As already shown in Fig. 1, when the BS density is around $\lambda \in [10^{-1}, 10^2]$ BSs/km², the network is noise-limited, and thus $p^{\text{cov}}(\lambda, \rho, \gamma)$ increases with λ as the network is lightened up with more BSs and the signal power benefits from LoS transmissions.
- As already shown in Fig. 1, when the BS density is around $\lambda \in [10^2, 10^3]$ BSs/km², $p^{\text{cov}}(\lambda, \rho, \gamma)$ decreases as λ increases. This is due to the transition of a large number of interfere paths from NLoS to LoS, which accelerates the growth of the aggregate inter-cell interference [5, 10].
- When $\lambda \in [10^3, 10^5]$ BSs/km², $p^{\text{cov}}(\lambda, \rho, \gamma)$ continuously increases thanks to the IMC [7], i.e., the signal power continues increasing with the network densification, while the interference power becomes bounded, as only BSs with active UEs are turned on, and thus the number of interfering BSs is limited by the number of active UEs.
- When $\lambda > 10^5$ BSs/km², $p^{\text{cov}}(\lambda, \rho, \gamma)$ gradually reaches its limit characterized by Theorem 1, which verifies the SINR invariance law in Theorem 2. Numerically speaking, the gap between the analytical results of $p^{\text{cov}}(\lambda, \rho, \gamma)$ and those of $\lim_{\lambda \rightarrow +\infty} p^{\text{cov}}(\lambda, \rho, \gamma)$ are less than 0.5 % for all of the investigated cases when $\lambda = 10^6$ BSs/km², which validates the accuracy of Theorem 1.
- As predicted by Lemma 1, when $\rho = 300$ UEs/km², the limit of $p^{\text{cov}}(\lambda, \rho, \gamma)$ with $L = 3.5$ m is larger than that with $L = 8.5$ m.
- As predicted by Lemma 2, $\lim_{\lambda \rightarrow +\infty} p^{\text{cov}}(\lambda, \rho, \gamma)$ decreases as ρ grows due to more interference generated by more

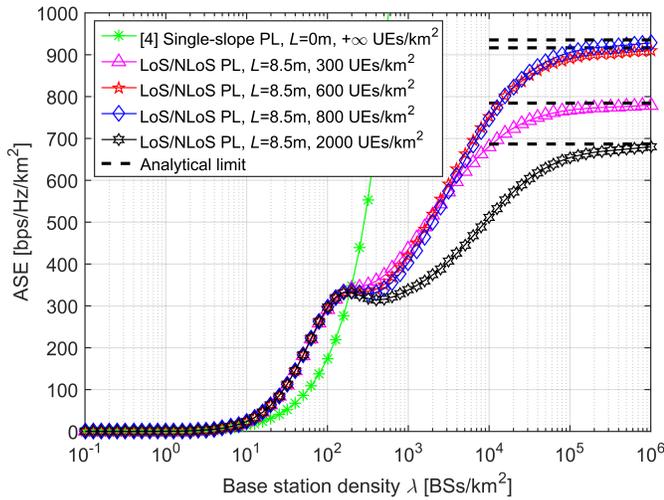


Fig. 3. The ASE $A^{\text{ASE}}(\lambda, \rho, \gamma_0)$ vs. λ for the 3GPP Case with $\gamma_0 = 0$ dB, $L = 8.5$ m and various values of ρ .

active BSs. When $L = 8.5$ m, $c(\gamma) = 1$ in Lemma 1 and $\lim_{\lambda \rightarrow +\infty} p^{\text{cov}}(\lambda, \rho, \gamma)$ with $\rho = 300$ UEs/km² is 0.806, while that with $\rho = 600$ UEs/km² is 0.65, which equals to the square of 0.806, thus verifying the power law of $\lim_{\lambda \rightarrow +\infty} p^{\text{cov}}(\lambda, \rho, \gamma)$ with respect to ρ .

Since a constraint on the minimum horizontal UE-to-BS distance also imposes a cap on the signal power, to show the performance impact of such minimum UE-to-BS distance on UDNs, we plot the coverage probability with $R_0 = \min(5\text{m}, 0.1 \times \text{ISD})$ (ISD: inter-site distance) in Fig. 2, where R_0 regulates that a UE is not allowed to be within a minimum distance to the BS. Note that, we define R_0 as a decreasing function of the BS density to reflect the fact that users are allowed to approach small BSs more closely in a denser network. If R_0 is set to a constant that does not change with the BS density, then in a UDN each BS will create a black hole around it, forbidding UE deployment inside such black hole. At a certain level of high BS densities, the aforementioned black holes will eat up the entire network scenario due to the high density of them. From the presented results, the coverage probability increases towards 1 with the BS density, which is basically the same as the conclusion in [7]. That is, when the BS tends to infinity, $\lambda \rightarrow +\infty$, the user to serving BS distance tends to zero, $r \rightarrow 0$, and the signal power continuously increases.

B. Validation of the New Capacity Scaling Law

In Fig. 3, we plot the ASE results for the 3GPP Case with $\gamma_0 = 0$ dB, $L = 8.5$ m and various values of ρ . Since $A^{\text{ASE}}(\lambda, \rho, \gamma_0)$ is calculated from the results of $p^{\text{cov}}(\lambda, \rho, \gamma)$ using (16), and because the analysis on $p^{\text{cov}}(\lambda, \rho, \gamma)$ has been validated in Subsection VI-A, we only show the analytical results of $A^{\text{ASE}}(\lambda, \rho, \gamma_0)$ in Fig. 3. From this figure, we can observe that:

- Due to its simplistic assumptions on channel modeling and UE density, the linear capacity scaling law in [3] shows an optimistic but unrealistic future for 5G UDNs.

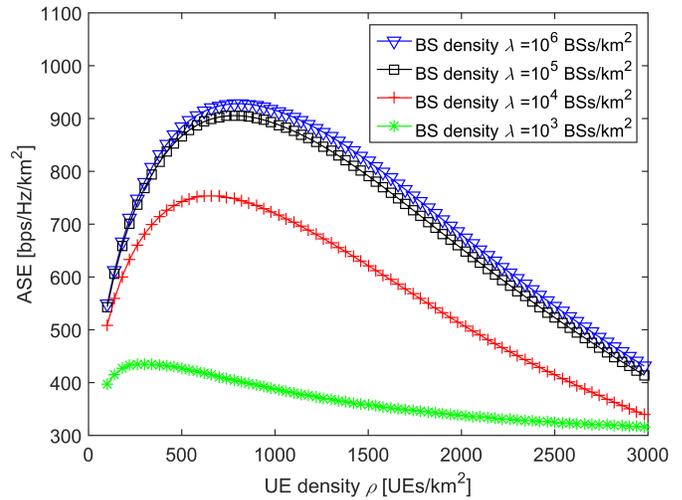


Fig. 4. The ASE $A^{\text{ASE}}(\lambda, \rho, \gamma_0)$ vs. ρ for the 3GPP Case with $\gamma_0 = 0$ dB, $L = 8.5$ m and various values of λ .

- The constant capacity scaling law in Theorem 3 is validated for UDNs with a non-zero L and a finite ρ , showing a practical future for 5G UDNs,
- The ASE crawls (not increasing quickly) when $\lambda \in [10^2, 10^3]$ BSs/km², which is due to the degradation of the coverage probability (see Fig. 2) caused by the transition of a large number of interfere paths from NLoS to LoS [5].
- For a given ρ , e.g., $\rho = 300$ UEs/km², the value of $A^{\text{ASE}}(\lambda, \rho, \gamma_0)$ saturates as $\lambda \rightarrow +\infty$, which justifies the BS deployment problem (18) addressed in Subsection V-A.
- For a given λ , e.g., $\lambda = 10^5$ BSs/km², it is interesting to see that $A^{\text{ASE}}(\lambda, \rho, \gamma_0)$ is indeed a concave function of ρ , and it achieves its maximum value when $\rho \in [600, 2000]$ UEs/km², which justify the UE scheduling problem (19) addressed in Subsection V-B. To illustrate such concave function more clearly, in Fig. 4 we plot the ASE results versus ρ for the 3GPP Case with $\gamma_0 = 0$ dB, $L = 8.5$ m and various values of λ .
- Based on Fig. 3 and Fig. 4, it is of great interest to visualize the ASE performance with respect to both the BS density and the UE density. To this end, in Fig. 5 we display the heat map of $A^{\text{ASE}}(\lambda, \rho, \gamma_0)$ versus ρ and λ for the 3GPP Case with $\gamma_0 = 0$ dB and $L = 8.5$ m. In this figure, the ASE performance is represented by colors, the redder, the higher ASE. As can be seen from Fig. 5, for a given UE density, the ASE performance saturates in UDNs (Subsection V-A); while for a given BS density, the UE density should be carefully chosen to maximize the ASE (Subsection V-B).
- For a given λ , e.g., $\lambda = 10^5$ BSs/km², it is interesting to see that those BSs can be partitioned into, say $N^{\text{BSG}} = 10$ groups, each group with 10^4 BSs/km² having access to only one tenth of the resource blocks. For instance, when $\rho = 800$ UEs/km², the total ASE of the case with BS partition (i.e., 10 groups of 10^4 BSs/km²) is around $10 \times \frac{1}{10} \times 746.6 = 746.6$ bps/Hz/km², which is less than

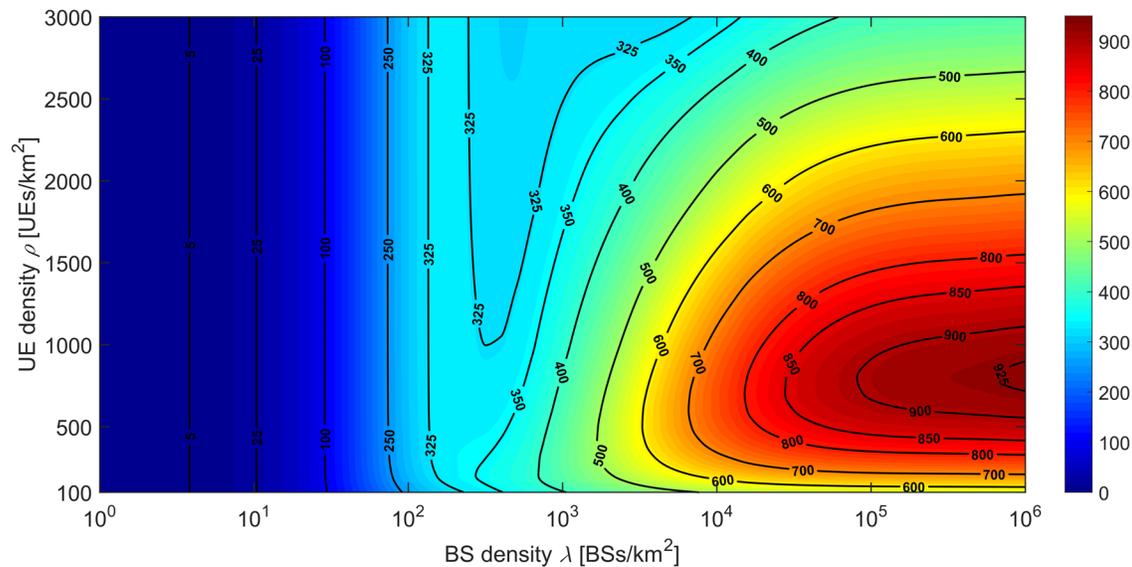


Fig. 5. The ASE $A^{\text{ASE}}(\lambda, \rho, \gamma_0)$ vs. ρ and λ for the 3GPP Case with $\gamma_0 = 0$ dB and $L = 8.5$ m.

that of the case without BS partition (i.e., 1 group of 10^5 BSs/km²), i.e., 905.5 bps/Hz/km². However, the total density of served UEs can be increased by 10 times from 800 UEs/km² to 8000 UEs/km², which justify the BS coordination problem (22).

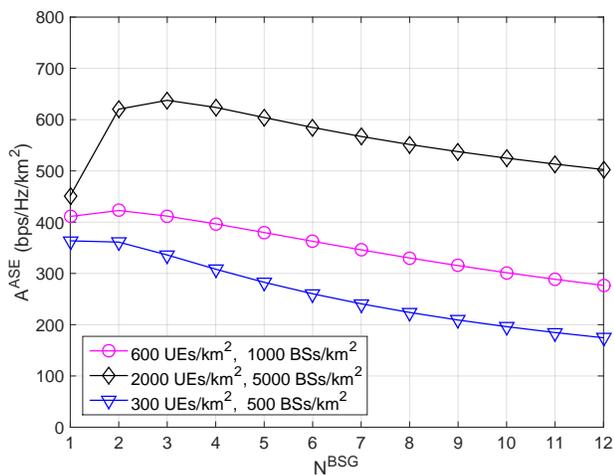


Fig. 6. The performance impact of N^{BSG} in dense networks and UDNs.

C. Numerical Results for the Proposed Optimization Problems

In this section, we present numerical results for the proposed optimization problems:

- As discussed in Subsection V-A, the BS deployment problem is trivial and its solution can be easily obtained by numerical search over $A^{\text{ASE}}(\lambda, \rho, \gamma_0)$ in Fig. 3. For example, for the following set of parameter values: $\rho = 300$ UEs/km², $L = 8.5$ m and $\gamma_0 = 0$ dB, we can calculate $\lim_{\lambda \rightarrow +\infty} A^{\text{ASE}}(\lambda, \rho, \gamma_0)$ using Theorem 3 and obtain its value as 784.4 bps/Hz/km². Considering a performance gap of $\epsilon = 5$ percent (i.e., a target ASE

of 745.2 bps/Hz/km²), it is easy to find the solution for problem (18) as $\lambda^* = 33420$ BSs/km². Such BS density means that any network densification beyond this level will generate no more than 5% of the maximum ASE.

- As discussed in Subsection V-B, the UE scheduling problem is non-trivial, but its solution can be numerically found by means of the proposed Algorithm 1 on Fig. 4. For example, for the following set of parameter values: $\lambda = 10^6$ BSs/km², $L = 8.5$ m and $\gamma_0 = 0$ dB, we can find the solution for problem (19) as $\rho^* = 804$ UEs/km² with a maximum ASE of 928.2 bps/Hz/km².
- From Fig. 6, we can see that the optimal number of BS groups $N^{\text{BSG}*}$ is i) $N^{\text{BSG}} = 2$ for the case with 600 UEs/km² and 1000 BSs/km², ii) $N^{\text{BSG}} = 3$ for the case with 2000 UEs/km² and 5000 BSs/km² and iii) $N^{\text{BSG}} = 1$ for the case with 300 UEs/km² and 500 BSs/km². These results clearly show that the tradeoff between the increased SINR and the decreased available bandwidth in each BS group determines the ASE performance. When the network deployment results in low SINRs, the multi-group coordination strategy can greatly boost the ASE due to the enhanced SINR. This is in line with the Shannon theorem, where capacity gains due to SINR are logarithmic in nature, and thus are more noticeable in the low SINR regime.

D. Performance Impact of Rician Fading on the Constant Capacity Scaling Law

In Fig. 7, we investigate the performance of $A^{\text{ASE}}(\lambda, \rho, \gamma_0)$ under the assumptions of Rician fading [24] for LoS transmissions and Rayleigh fading for NLoS ones. As discussed at the end of Section III, here we adopt a practical model of Rician fading in [24], where the K factor increases as the transmitter to receiver distance decreases. As discussed in [23], the analysis for Rician fading is challenging and not urgent. Therefore, we only display simulation results for Rician fading in Fig. 7. As can be seen from this figure, the ASE results

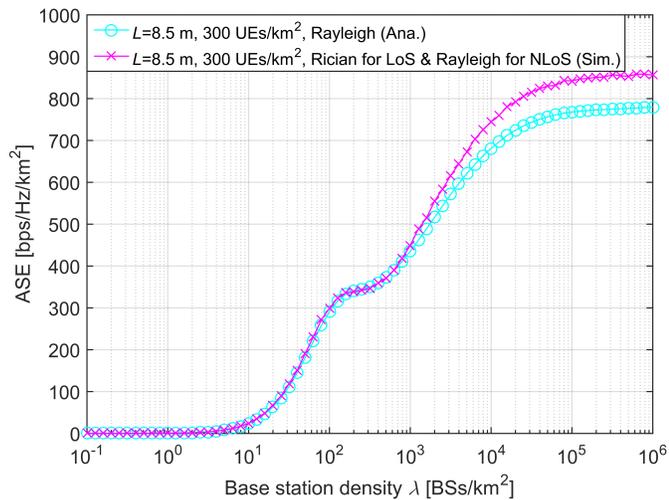


Fig. 7. The ASE $A^{\text{ASE}}(\lambda, \rho, \gamma_0)$ vs. λ for the 3GPP Case with $\gamma_0 = 0$ dB, $\rho = 300$ UEs/km², $L = 8.5$ m, Rician fading for LoS transmissions and Rayleigh fading for NLoS transmissions.

with the assumption of Rician fading verify the correctness of our fundamental discovery on the constant capacity scaling law. The fading model only has some quantitative, but not qualitative impact on network performance. More specifically, due to its less channel fading, Rician fading leads to a better performance compared with Rayleigh fading in UDNs.

VII. CONCLUSION

A new constant capacity scaling law has been presented for UDNs considering a general multi-piece path loss model, a non-zero BS to UE antenna height difference, and a finite UE density exploited by the IMC. Such law has three profound implications:

- First, network densification should be *stopped* at a certain BS density for a given UE density, because the network capacity reaches a limit due to (i) the bounded signal and interference powers, and (ii) a finite spatial spectrum reuse because of a finite UE density. Such BS density can be optimized by solving *the proposed BS deployment problem*, which also answers the fundamental question of “for a given UE density, how dense an UDN should be?”.
- Second, the recent concerns about the network capacity crash in UDNs can be resolved by our discovered capacity scaling law with *proactively choosing and serving a subset of UEs* in a TDMA/FDMA manner. Such UE density can be optimized by solving *the proposed UE scheduling problem*, which also answers the fundamental question of “for a given BS density, what is the optimal user load that can maximize the ASE?”.
- Third, even if the UE density is tremendously large or infinite, and there is also a strong need to serve a huge number of UEs simultaneously, the network capacity crash can still be avoided by employing TD or FD ICIC among groups of BSs. Such BS partition can be optimized by solving *the proposed BS coordination problem*, which also answers the fundamental question of “for a given BS density serving a required UE density, what is the optimal resource reuse factor that can maximize the ASE?”.

The obtained complex theoretical results may be simplified using approximation methods, which may lead to a simpler formulation and more intuitive insights. Such research will be part of our future work.

REFERENCES

- [1] 3GPP, “TR 36.872: Small cell enhancements for E-UTRA and E-UTRAN - Physical layer aspects,” Dec. 2013.
- [2] D. López-Pérez, M. Ding, H. Claussen, and A. Jafari, “Towards 1 Gbps/UE in cellular systems: Understanding ultra-dense small cell deployments,” *IEEE Communications Surveys Tutorials*, vol. 17, no. 4, pp. 2078–2101, Jun. 2015.
- [3] J. Andrews, F. Baccelli, and R. Ganti, “A tractable approach to coverage and rate in cellular networks,” *IEEE Transactions on Communications*, vol. 59, no. 11, pp. 3122–3134, Nov. 2011.
- [4] S. Lee and K. Huang, “Coverage and economy of cellular networks with many base stations,” *IEEE Communications Letters*, vol. 16, no. 7, pp. 1038–1040, Jul. 2012.
- [5] M. Ding, P. Wang, D. López-Pérez, G. Mao, and Z. Lin, “Performance impact of LoS and NLoS transmissions in dense cellular networks,” *IEEE Transactions on Wireless Communications*, vol. 15, no. 3, pp. 2365–2380, Mar. 2016.
- [6] M. Ding and D. López-Pérez, “Performance impact of base station antenna heights in dense cellular networks,” *IEEE Transactions on Wireless Communications*, vol. 16, no. 12, pp. 8147–8161, Dec 2017.
- [7] M. Ding, D. López-Pérez, G. Mao, and Z. Lin, “Performance impact of idle mode capability on dense small cell networks with LoS and NLoS transmissions,” *arXiv:1609.07710 [cs.NI]*, Sep. 2016.
- [8] M. Ding, D. López-Pérez, G. Mao, and Z. Lin, “Ultra-dense networks: Is there a limit to spatial spectrum reuse?” in *2018 IEEE International Conference on Communications (ICC)*, May 2018, pp. 1–6.
- [9] H. S. Dhillon, R. K. Ganti, F. Baccelli, and J. G. Andrews, “Modeling and analysis of K-tier downlink heterogeneous cellular networks,” *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, pp. 550–560, Apr. 2012.
- [10] X. Zhang and J. Andrews, “Downlink cellular network analysis with multi-slope path loss models,” *IEEE Transactions on Communications*, vol. 63, no. 5, pp. 1881–1894, May 2015.
- [11] T. Bai and R. Heath, “Coverage and rate analysis for millimeter-wave cellular networks,” *IEEE Transactions on Wireless Communications*, vol. 14, no. 2, pp. 1100–1114, Feb. 2015.
- [12] A. H. Jafari, D. López-Pérez, M. Ding, and J. Zhang, “Performance analysis of dense small cell networks with practical antenna heights under rician fading,” *IEEE Access*, vol. 6, pp. 9960–9974, 2018.
- [13] M. Ding and D. López-Pérez, “Please Lower Small Cell Antenna Heights in 5G,” *IEEE Globecom 2016*, pp. 1–6, Dec. 2016.
- [14] I. Atzeni, J. Arnau, and M. Kountouris, “Downlink cellular network analysis with los/nlos propagation and elevated base stations,” *IEEE Transactions on Wireless Communications*, vol. 17, no. 1, pp. 142–156, Jan 2018.
- [15] A. AlAmmouri, J. G. Andrews, and F. Baccelli, “SINR and throughput of dense cellular networks with stretched exponential path loss,” *IEEE Transactions on Wireless Communications*, vol. 17, no. 2, pp. 1147–1160, Feb 2018.
- [16] J. Liu, M. Sheng, L. Liu, and J. Li, “How Dense is Ultra-Dense for Wireless Networks: From Far- to Near-Field Communications,” *arXiv:1606.04749 [cs.IT]*, Jun. 2016.
- [17] M. D. Renzo, “Stochastic geometry modeling and analysis of multi-tier millimeter wave cellular networks,” *IEEE Transactions on Wireless Communications*, vol. 14, no. 9, pp. 5038–5057, Sep. 2015.
- [18] M. D. Renzo, W. Lu, and P. Guan, “The intensity matching approach: A tractable stochastic geometry approximation to system-level analysis of cellular networks,” *IEEE Transactions on Wireless Communications*, vol. 15, no. 9, pp. 5963–5983, Sep. 2016.
- [19] Z. Luo, M. Ding, and H. Luo, “Dynamic small cell on/off scheduling using stackelberg game,” *IEEE Communications Letters*, vol. 18, no. 9, pp. 1615–1618, Sept 2014.
- [20] 3GPP, “TR 38.913: Study on Scenarios and Requirements for Next Generation Access Technologies,” July 2018.
- [21] —, “TR 36.828: Further enhancements to LTE Time Division Duplex for Downlink-Uplink interference management and traffic adaptation,” Jun. 2012.
- [22] M. Ding, D. López-Pérez, G. Mao, P. Wang, and Z. Lin, “Will the area spectral efficiency monotonically grow as small cells go dense?” *IEEE GLOBECOM 2015*, pp. 1–7, Dec. 2015.

- [23] M. Ding and D. López-Pérez, "On the performance of practical ultra-dense networks: The major and minor factors," *The IEEE Workshop on Spatial Stochastic Models for Wireless Networks (SpaSWiN) 2017*, pp. 1–8, May 2017.
- [24] Spatial Channel Model AHG, "Subsection 3.5.3, Spatial Channel Model Text Description V6.0," Apr. 2003.
- [25] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [26] R. L. Burden and J. D. Faires, *Numerical Analysis (3rd Ed.)*. PWS Publishers, 1985.
- [27] Qualcomm, "1000x: More small cells – hyper-dense small cell deployments," June 2014.

Ming Ding received the B.S. and M.S. degrees (with first class Hons.) in electronics engineering from Shanghai Jiao Tong University (SJTU), Shanghai, China, and the Doctor of Philosophy (Ph.D.) degree in signal and information processing from SJTU, in 2004, 2007, and 2011, respectively. In Sep. 2014, he joined National Information and Communications Technology Australia (NICTA), and since 2015 he worked as a Senior Research Scientist in this new R&D centre located in Sydney, N.S.W., Australia. He has authored about 40 papers in IEEE journals and conferences, all in recognized venues, and about 20 3GPP standardization contributions. Also, he holds fifteen CN, seven JP, three US, two KR patents and co-authored another 100+ patent applications on 4G/5G technologies.

David López-Pérez is a Member of Technical Staff at Nokia Bell Laboratories. He earned his PhD in Wireless Networking from the University of Bedfordshire, UK in Apr. '11, and obtained his BSc and MSc degrees in Telecommunication from the Miguel Hernandez University, Spain in Sept. '03 & Sept. '06, respectively. For his publications and patent contributions, David is a recipient of both the Bell Labs Alcatel-Lucent Award of Excellence (2013) and Certificate of Outstanding Achievement (2013, 2014, 2015). He was also finalist for the Scientist of the Year prize in The Irish Laboratory Awards (2013, 2015). David is author of the book "Heterogeneous Cellular Networks: Theory, Simulation and Deployment" Cambridge University Press, 2012. Moreover, he has published more than 90 book chapters, journal and conference papers, all in recognized venues, and filed more than 30 patents applications.

Youjia Chen received the B.S. and M.S degrees in communication engineering from Nanjing University, Nanjing, China, and Ph.D. degree in wireless engineering from the University of Sydney, Australia, in 2005, 2008 and 2017, respectively. From 2008-2009, she worked in Alcatel-Lucent Shanghai Bell. And she worked at the College of Photonic and Electrical Engineering, Fujian Normal University, China, from Aug. 2009. In 2018, she joined the College of Physics and Information Engineering, Fuzhou University, China. Her current research interests include ultra-dense networks, wireless caching and machine learning.

Guoqiang Mao received PhD in telecommunications engineering in 2002 from Edith Cowan University. He was with the School of Electrical and Information Engineering, the University of Sydney between 2002 and 2014. He joined the University of Technology Sydney in February 2014 as Professor of Wireless Networking and Director of Center for Real-time Information Networks. He has published about

200 papers in international conferences and journals, which have been cited more than 4000 times. His research interest includes intelligent transport systems, applied graph theory and its applications in telecommunications, Internet of Things, wireless sensor networks, wireless localization techniques and network performance analysis.

Zihuai Lin received the Ph.D. degree in Electrical Engineering from Chalmers University of Technology, Sweden, in 2006. Prior to this he has held positions at Ericsson Research, Stockholm, Sweden. Following Ph.D. graduation, he worked as a Research Associate Professor at Aalborg University, Denmark and currently at the School of Electrical and Information Engineering, the University of Sydney, Australia. His research interests include source/channel/network coding, coded modulation, MIMO, OFDMA, SC-FDMA, radio resource management, cooperative communications, small-cell networks, 5G cellular systems, etc.

Albert Y. Zomaya is the Chair Professor of High Performance Computing and Networking in School of Information Technologies, The University of Sydney, and he also serves as the Director of Centre for Distributed and High Performance Computing. Professor Zomaya published more than 600 scientific papers and articles and is author, co-author or editor of more than 20 books. He is the Founding Editor in Chief of the IEEE Transactions on Sustainable Computing and serves as an associate editor for more than 20 leading journals. Professor Zomaya served as an Editor in Chief for the IEEE Transactions on Computers (2011-2014). Professor Zomaya is the recipient of the IEEE Technical Committee on Parallel Processing Outstanding Service Award (2011), the IEEE Technical Committee on Scalable Computing Medal for Excellence in Scalable Computing (2011), and the IEEE Computer Society Technical Achievement Award (2014). He is a Chartered Engineer, a Fellow of AAAS, IEEE, and IET. Professor Zomaya's research interests are in the areas of parallel and distributed computing and complex systems.