User-Centric Networking for Dense C-RANs: High-SNR Capacity Analysis and Antenna Selection

Published in:
IEEE Transactions on Communications

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

User-Centric Networking for Dense C-RANs: High-SNR Capacity Analysis and Antenna Selection

Jide Yuan, *Student Member, IEEE*, Shi Jin, *Member, IEEE*,
Wei Xu, *Senior Member, IEEE*, Weiqiang Tan, *Student Member, IEEE*,
Michail Matthaiou, *Senior Member, IEEE*, and Kai-Kit Wong, *Fellow, IEEE*

Abstract—Ultra-dense cloud radio access networks (C-RANs) is one of the architectures that will be critical components of the next-generation wireless systems. In a C-RAN architecture, an amorphous cellular framework, where each user connects to a few nearby remote radio heads (RRHs) to form its own cell, appears to be promising. In this paper, we study the ergodic capacity of such amorphous cellular networks at high signal-to-noise ratios (SNRs) where we model the distribution of the RRHs by a Poisson point process. We derive tractable approximations of the ergodic capacity at high-SNRs for arbitrary antenna configurations, and tight lower bounds for the ergodic capacity when the numbers of antennas are the same at both ends of the link. In contrast to prior works on distributed antenna systems, our results are derived based on random matrix theory and involve only standard functions which can be much more easier evaluated. The impact of the system parameters on the ergodic capacity is investigated. By leveraging our analytical results, we propose two efficient scheduling algorithms for RRH selection for energy-efficient transmission. Our algorithms offer a substantial improvement in energy efficiency compared to the strategy of connecting a fixed number of RRHs to each user.

Index Terms—Dense cloud radio access network, ergodic capacity, MIMO.

I. INTRODUCTION

Mobile data traffic continues to rise with an extraordinary rate and a $1000\times$ increase is very much expected by 2020 [1]. Yet, conventional cellular networks find it difficult to catch up with these unprecedented demands and will need a serious upgrade. In 5G, one crucial technology is ultra-dense networks (UDNs), which promises to provide a massive boost to regional capacity [2–4]. Although the conventional cellular network topologies have been well suited for providing wide area coverage, they are unable to scale up regional capacity to meet users' data needs.

Dense cloud radio access network (C-RAN) is an emerging network architecture which is recognized as the enabling technology to meet these mobile traffic demands since it avails of the reduced distance between users and remote radio head (RRH) units [5]. In C-RANs, RRHs operate as soft relays by receiving the signals from mobile users and forward them to

J. Yuan, S. Jin, W. Xu and W. Tan are with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, P. R. China (e-mail: {yuanjide, jinshi, wxu, wqtan}@seu.edu.cn).

M. Matthaiou is with the Institute of Electronics, Communications and Information Technology (ECIT), Queen's University Belfast, Belfast, BT3 9DT, U.K. (e-mail: m.matthaiou@qub.ac.uk).

K. K. Wong is with the Department of Electronic and Electrical Engineering, University College London, London WC1E 7JE, United Kingdom (e-mail: kai-kit.wong@ucl.ac.uk).

a centralized baseband unit (BBU) [6, 7]. The performance of distributed antenna arrays and best base station (BS) selection schemes in C-RANs was assessed in [8], where both the users and RRHs are equipped with a single antenna. In [9], the authors investigated the sum-rate maximization problem subject to a BS backhaul constraint in a downlink C-RAN for both dynamic and static BS clustering over different time-frequency slots. However, single access point schemes are incapable of supporting enough mobile traffic due to the low power of deployed RRHs. Undoubtedly, if we intend to catch up with the traffic demands, multi-access point schemes are inevitable in C-RANs because of the substantial improvement in spectral efficiency offered by multiple-input multiple-output (MIMO) antenna topologies.

With dense RRH deployments, the BS-centric MIMO systems may not be suitable when both the users and BSs are scattered geographically due to the poor support for cell-edge users [10]. Instead, a user-centric structure, namely, amorphous cellular, is far more appropriate, and the interference coordination becomes much more feasible [11–13]. Under this structure, the user in C-RAN chooses its own serving RRH set as its amorphous cell, while the BBU schedules time-frequency recourse centrally. The authors of [14] investigated the optimal sizes of amorphous cell topologies for single-user transmission, where the locations of RRHs were modeled by a Poisson point process (PPP) to capture the irregularity of BSs [15, 16]. Then in [17], closed-form ergodic capacity expressions were presented for the N-nearest PPP distributed RRH association strategies when the path loss exponent was four. The uplink ergodic sum capacity of amorphous cellular systems was presented in [18], where the RRHs were either co-located at the cell center or uniformly distributed within each cell. As an extension of [18], the authors further investigated the downlink amorphous cellular with a large number of users randomly distributed in the system and examined the effect of cellular size on the average user rate in [10]. Note that the above results were limited to single-antenna users, and hardly provided closed-form achievable rate expressions because of the intractability of the large-scale fading (LSF) effect, especially when users are equipped with multiple antennas. However, the use of multiple antennas has been the standard practice in concurrent wireless networks, which motivates us to characterize the capacity performance of MIMO-operated C-RAN systems.

Refer to the MIMO systems, most works applied random matrix theory to pursue an analytical characterization of

MIMO systems, and insightful results were obtained under semi-correlated channels. The authors of [19] considered the outage capacity performance of a MIMO system in correlated environments and derived exact distribution functions for the capacity with a small number of antennas. In [20], a closed-form expression for the characteristic function of the MIMO system capacity with arbitrary correlation among transmit (receive) antennas was derived. The authors of [21] analyzed the capacity and corresponding optimal input density of a correlated MIMO channel, where the channel was assumed to have a (Kronecker) correlated normal structure. All the above studies provided closed-form or integral expressions for the capacity for the cases in which channel correlation is present at one of the two ends of the link. However, we note that most of prior work assumed MIMO links between a user and one multi-antenna BS, and the perspective of a network with multiple BSs is missing. A crucial difference between the two assumptions is whether LSF is taken into account in the capacity analysis, since it an intrinsic model parameter in distributed MIMO systems [22]. Dense C-RAN, a kind of distributed MIMO systems, belongs to the latter one. Utilizing random matrix theory, the effect of LSF on capacity can be formulated by regarding the LSF matrix as a correlation matrix at one end of the link. To the best of our knowledge, there is no analytical expression available for the ergodic capacity which applies for distributed MIMO systems with both arbitrary path-loss exponent and number of antennas, especially when the locations of RRHs are deployed randomly, e.g., as a PPP.

In this paper, the uplink capacity of amorphous cellular in ultra-dense C-RAN at high signal-to-noise ratio (SNRs) is investigated, where the locations of RRHs follow a PPP. We first derive an approximation of the ergodic capacity for an arbitrary number of antennas when LSF is taken into account. Moreover, a tight high-SNR lower bound of capacity is derived when the numbers of antennas are the same at both ends of the link. Based on the proposed results, the impact of the number of antennas is characterized. In contrast to prior results, our expressions involve only standard functions which can be easily and efficiently evaluated, and they illustrate the effects of path-loss exponent and the RRH intensity on the ergodic capacity. On the premise of guaranteeing the quality-of-service (QoS), two RRH scheduling algorithms on forming the amorphous cellular are proposed, namely, user-optimal scheme (UOS) and RRH-optimal scheme (ROS), to achieve energy efficient transmission for MIMO orthogonal frequency division multiple (OFDM) systems when both the transmit and receive power consumption are considered. Both algorithms are developed on the basis of our capacity analysis results, and demonstrate excellent performance for arbitrary QoS requirements.

The rest of this paper is organized as follows: Section II presents the system model of an amorphous cellular in dense C-RANs. Section III presents the capacity analysis at high SNRs. In Section IV, two scheduling algorithms developed from our analytical results are proposed. Section V summarizes the main observations and proofs are relegated to appendices.

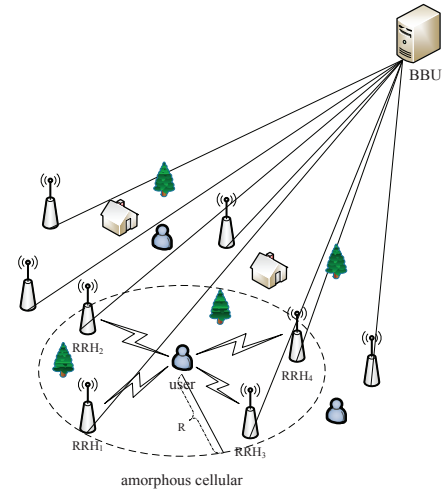*Notations*—Throughout this paper, vectors and matrices

Fig. 1. The system model of amorphous cellular in the ultra-dense C-RANs uplink, where each user connects to the n-nearest RRHs to form its own cell. All RRHs belong to one BBU.

are denoted in bold lowercase letters $\mathbf{a}$ and bold uppercase letters $\mathbf{A}$, respectively. The notation $\det\left(\mathbf{A}\right)$ stands for the determinant, and the (i,j)th entry of $\mathbf{A}$ is denoted as $\{\mathbf{A}\}_{i,j}$. The complex and real number fields are represented by $\mathcal{C}$ and $\mathcal{R}$, respectively. The superscripts $(\cdot)^{\dagger}$ denotes conjugate-transpose operation, and $\mathrm{E}\left[\cdot\right]$ evaluates the expectation of the input random entity. Additionally, $\Gamma(\cdot)$ and $\Gamma(\cdot,\cdot)$ are the gamma function and upper incomplete gamma function [23, Eq. (6.1.5), Eq. (6.5.3)] respectively, and $\psi\left(\cdot\right)$ is the digamma function [24, Eq. (8.360.1)].

II. System Model

Consider a dense C-RAN, depicted in Fig. 1, in which each RRH has a single antenna, collecting the signals from a user who has m antennas for processing in the BBU. The overall system constitutes a distributed MIMO system. We assume a user-centric network under the framework of MIMO-OFDM where the user connects to n nearest RRHs to form its own cell. In the context of interference control with efficient resource reuse, we assume that only a single user is supported in each subcarrier of the OFDM system [25, 26]. The locations of RRHs are assumed to be modeled as a two-dimensional PPP having intensity λ in a plane whose radius is R. As it is a user-centric cell, the desired user is located at the origin of the plane. Thus, the whole number of RRHs, N, in this plane is a random variable with probability distribution function (pdf) [27]

$$f_N\left(N\right) = \frac{\left(\lambda\pi R^2\right)^N}{N!}e^{-\lambda\pi R^2}. \tag{1}$$

We consider the uplink from the user to the RRHs. Assuming that user connects with the n-nearest neighboring RRHs, the n-dimensional coordinated received signal in the BBU can be written as [28, 29]

$$\mathbf{r} = \mathbf{G}\mathbf{s} + \mathbf{n}, \tag{2}$$

where $\mathbf{s}$ is the m-dimensional transmitted signal vector with the total transmission power $\mathrm{E}\left[\mathbf{s}^{\dagger}\mathbf{s}\right] = P_t$. The entries of the

additive noise $\mathbf{n}$ are modeled as zero-mean circular-symmetric complex Gaussian with variance σ_{n}^2. In this paper, the channel is mathematically modeled by the $n \times m$ random matrix $\mathbf{G}$, defined by

$$\mathbf{G} = \mathbf{B}^{\frac{1}{2}}\mathbf{H}, \tag{3}$$

where $\mathbf{H} \in \mathcal{C}^{n \times m}$ is the small-scale fast fading channel matrix with complex elements $\{h_{i,j}\} \sim \mathcal{CN}(0,1)$, and $\mathbf{B} \in \mathcal{R}^{n \times n}$ is a diagonal matrix accounting for the LSF consisted of shadowing and path loss fading, whose ith diagonal element is given by

$$\{\mathbf{B}\}_{ii} = \beta_i = 10^{-\frac{c_0 + \sigma z_i}{10}} \cdot d_i^{-\alpha}, \tag{4}$$

where c_0 is a constant value (in dB), $10^{-\frac{\sigma z_i}{10}}$ represents the shadowing with the standard deviation σ and $z_i \sim \mathcal{N}(0,1)$, d_i is the distance between the user and the ith nearest RRH, and α stands for the path-loss exponent with typical values $\alpha \in [2,4]$ [30].

In this paper, we consider the uplink scenarios in which the user (i.e., the transmitter side) has no channel state information (CSI) while the BBU (i.e., the receiver side) has perfect CSI.[1] In this case, the transmitter typically uses a uniform power allocation across all spatial subchannels, and, therefore, the ergodic capacity (in bit/s/Hz) is written as[2]

$$C_{m,n} = \mathrm{E}\left[\log_2 \det\left(\mathbf{I} + \frac{\rho}{m}\mathbf{W}\right)\right], \tag{5}$$

where $\rho = P_t/\sigma^2$ is the transmit SNR, and

$$\mathbf{W} = \begin{cases} \mathbf{B}\mathbf{H}\mathbf{H}^{\dagger}, & n \leqslant m, \\ \mathbf{H}^{\dagger}\mathbf{B}\mathbf{H}, & n > m. \end{cases} \tag{6}$$

The capacity can also be written in terms of the nonzero eigenvalues of $\mathbf{W}$. Let $\boldsymbol{\lambda}|\mathbf{B} = [\lambda_1|\mathbf{B}, \dots, \lambda_q|\mathbf{B}]^T$ denote the nonzero eigenvalues of the matrix $\mathbf{W}$, conditioned on $\mathbf{B}$, with $q = \min\{n,m\}$. Then, the ergodic capacity in (5) can alternatively be written as [28]

$$C_{m,n} = \int_{\mathbf{B}} \int_{\boldsymbol{\lambda}} \sum_{i=1}^{q} \log_2\left(1 + \frac{\rho}{m}\lambda_i\right) \\ \times f_{\lambda_i|\mathbf{B}} f_{\mathbf{B}}(\beta_1 < \dots < \beta_n)\, d\boldsymbol{\lambda}d\mathbf{B}, \tag{7}$$

where $f_{\lambda_i|\mathbf{B}}$ is the marginal pdf of the ordered eigenvalues λ_i conditioned on $\mathbf{B}$, $f_{\mathbf{B}}(\cdot)$ represents the ordered joint pdf of $\mathbf{B}$. We note that the conditional unordered eigenvalue pdf has been investigated in [32, Eq. (95)]. The ordered eigenvalue pdf can be easily derived by considering all the possible matrix patterns of $\mathbf{B}$ which consist of the same entries but in different orders, as given in

$$f_{\lambda_i|\mathbf{B}}(\lambda_i) = \frac{\Gamma(q)}{\prod_{i<j}^{p}(\beta_j - \beta_i)} \\ \times \sum_{s=p-q+1}^{p} \frac{\lambda_i^{s+q-p-1}}{\Gamma(s+q-p)} \det\left(\tilde{\boldsymbol{\Xi}}_s\right), \tag{8}$$

where $\tilde{\boldsymbol{\Xi}}_s$ is a $p \times p$ matrix whose entries are given by

$$\left\{\tilde{\boldsymbol{\Xi}}_s\right\}_{i,j} = \begin{cases} \beta_i^{j-1}, & i = 1, \dots, n, \ j \neq s, \\ \beta_i^{p-q+1}e^{-\frac{\lambda_i}{\beta_i}}, & i = 1, \dots, n, \ j = s, \end{cases} \tag{9}$$

with $p = \max\{n,m\}$. To remove the condition on $\mathbf{B}$, we need to integrate $\beta_{i=1,\dots,n}$ term-by-term. By doing so, this yields

$$f(\lambda_i) = \Gamma(q) \sum_{s=p-q+1}^{p} \frac{\lambda_i^{s+q-p-1}}{\Gamma(s+q-p)} \\ \times \int_{\mathbf{B}} \frac{\det\left(\tilde{\boldsymbol{\Xi}}_s\right) f_{\mathbf{B}}(\beta_1 < \dots < \beta_n)}{\prod_{i<j}^{p}(\beta_j - \beta_i)} d\mathbf{B}. \tag{10}$$

This integration has proved to be too complex, if not impossible, to be evaluated due to the fact that each entry of $\mathbf{B}$ has its own pdf, and with a Vandermonde determinant in the integrand. Thus, we resort to analyzing the capacity in the high-SNR regime, and obtain some interesting insights.

III. HIGH SNR ANALYSIS

In this section, we derive approximations of the ergodic capacity for amorphous cellular at high SNRs, and characterize the impact of the number of antennas on the capacity performance. We also provide tight lower bounds on the capacity when both ends of the link have the same number of antennas. According to our results, we reveal some insightful observations on the system parameters, such as the intensity of RRHs, etc.

A. High-SNR Approximation

In the high-SNR regime, we evaluate the approximation of capacity with an arbitrary number of antennas. It is obvious that the capacity is lower bounded by

$$C_{m,n} \geqslant \mathrm{E}\left[\log_2\left(\det\left(\frac{\rho}{m}\mathbf{W}\right)\right)\right]. \tag{11}$$

Note that (11) coincides with the ergodic capacity at high-SNRs, and will be more and more accurate as the SNR grows. To be more specific, we provide the constraint of transmit SNR for our results to hold in *Theorem* 1. Then, utilizing the Jensen's inequality, we now provide the following approximation to the ergodic capacity

$$C_{m,n}^{\mathrm{a}} = \log_2\left(\mathrm{E}\left[\det\left(\frac{\rho}{m}\mathbf{W}\right)\right]\right). \tag{12}$$

The tightness of the approximation, which is obtained by the adoption of the Jensen's bounds, has been discussed in [33]. According to the results in [33], the offset of (12) converges to a certain value as ρ grows, while decreases with the maximum dimension of $\mathbf{W}$, and increases with the minimum dimension of $\mathbf{W}$. In practice, the minimum dimension of the matrix typically corresponds to the number of user antennas which normally varies from 1 to 4. Thus, the approximation in (12) is considerably tight in most of the scenarios. We now investigate the expected determinant of $\mathbf{W}$ in the following lemma.

Lemma 1: The expected determinant of $\mathbf{W}$, where the entries of $\mathbf{B}$ follow the distribution given in *Lemma* 2, is given

---

[1]CSI can be obtained at RRHs through uplink training, especially when no interference is introduced by other users [7]. We consider that perfect CSI is acquired in this work and leave the channel estimation to future work.

[2]Although the $\log\det(\cdot)$ in (5) achieves capacity only under some strict assumptions [28], we call our results as "capacity" for the sake of consistency with the vast body of MIMO literature (see [19–21, 31]).

by

$$
\mathrm{E}\left[\det\left(\mathbf{W}\right)\right] = \begin{cases} \frac{\Gamma(m+1)}{\Gamma(m-n+1)} \prod_{i=1}^{n} \kappa\xi_\lambda(i), & n \leqslant m, \\ \Gamma(m+1) \sum_{B_{m,n}} \prod_{i\in B_{m,n}} \kappa\xi_\lambda(i), & n > m, \end{cases}
\tag{13}
$$

where $B_{m,n}$ is the length-m subset of $\{1,2,\ldots,n\}$, $\kappa = 10^{\frac{1}{10}\left(\frac{\sigma^2 \ln 10}{20} - c_0\right)}$, and

$$
\xi_\lambda(i) = \frac{(\lambda\pi)^{\frac{\alpha}{2}}}{\Gamma(i)} \bar{\Gamma}\left(i - \frac{\alpha}{2}\right),
\tag{14}
$$

where

$$
\bar{\Gamma}(t) = \begin{cases} \Gamma(t, d_{\mathrm{th}}), & t \text{ is nonpositive integer}, \\ \Gamma(t), & \text{otherwise}. \end{cases}
\tag{15}
$$

where d_{th} is a small constant.

Proof: See Appendix A. ∎

The results are valid for arbitrary antenna configurations. We find that the expression is easy to evaluate for the case of $n \leqslant m$, while the complexity of evaluation can be rather high when n is large for the case of $n > m$, since the number of the combinations of $B_{m,n}$ depends on m and n. Note that if the argument of $\Gamma(t)$ in (14) is a negative integer, we will run into a singularity. This is because of the inaccuracy of the LSF model when the user is close to the RRH. This can be easily avoided by introducing a small perturbation term d_{th}. Having established *Lemma* 1, we are ready to derive the approximation of the ergodic capacity.

Theorem 1: For dense C-RAN systems where the RRHs are distributed according to a stationary PPP with intensity λ, the high-SNR approximation of the capacity with n nearest RRHs can be expressed as in (16) at the top of next page. Moreover, ρ must satisfy the SNR constraint

$$
\rho \gg \frac{1}{\kappa\xi_\lambda(n)}.
\tag{17}
$$

Proof: The result in (16) directly follows by applying *Lemma* 1 to the approximation of the capacity in (12). However, we should be careful that the result only holds for the high-SNR cases, i.e., the SNR of the received signals should be considerably larger than 1 under the effect of LSF. For this reason, we evaluate ρ for the weakest subchannel under LSF, given by [35]

$$
\min_i \left\{ \mathrm{E}\left[\frac{\rho}{m}\beta_i \mathbf{h}_i^\dagger \mathbf{h}_i\right] \right\} = \min_i \left\{ \rho \mathrm{E}[\beta_i] \right\} \gg 1,
\tag{18}
$$

where $\mathbf{h}_i$ is the ith column of $\mathbf{H}$. The expected value of β_i can be evaluated in the same way as in (14). Noting that $\xi_\lambda(i)$ is a decreasing function against i, we prove the relationship between ρ and the number of associated RRH by using $\min_i \{E[\beta_i]\} = \kappa\xi_\lambda(n)$. ∎

It is important to note that the result in (16) holds for arbitrary numbers of antennas. The result illustrates the combined impact of various factors on the capacity approximation, e.g., intensity λ, path-loss exponent α and number of associated RRHs n. By establishing the constraint of SNR, we note that when n grows large, a higher SNR is required for (16) to hold. Figure 2 compares the
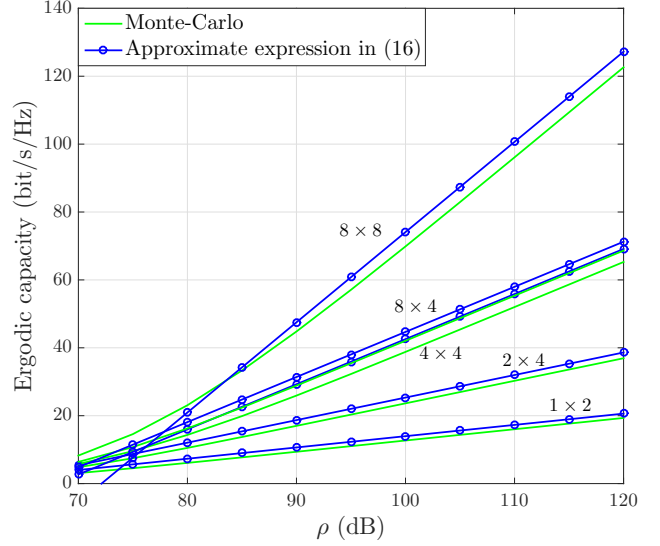

Fig. 2. Comparison of Monte-Carlo simulations with the approximate expressions of the ergodic capacity against ρ. The different antenna configurations are denoted as $m \times n$. Results are shown for $c_0 = 41.1\,\mathrm{dB}$, $\sigma^2 = 5\,\mathrm{dB}$, $\alpha = 2.09$, $\sigma_{\mathrm{n}}^2 = -70\,\mathrm{dBm}$, and $\lambda = 10^{-3}\mathrm{m}^2$ [30, 34].

approximations $C_{m,n}^{\mathrm{a}}(\lambda)$ with the Monte-Carlo simulations versus ρ. All the dotted curves are generated using (16). We confirm that the results are rather tight for both cases where the number of transmit antennas is larger or smaller than the number of associated RRHs. Moreover, we note the very limited capacity enhancement with adding antennas at the side with more antennas, indicating that the capacity performance is mainly affected by the minimum number of antennas. Additionally, we see that the tightness of bounds looses as the minimum number of antennas grows while the
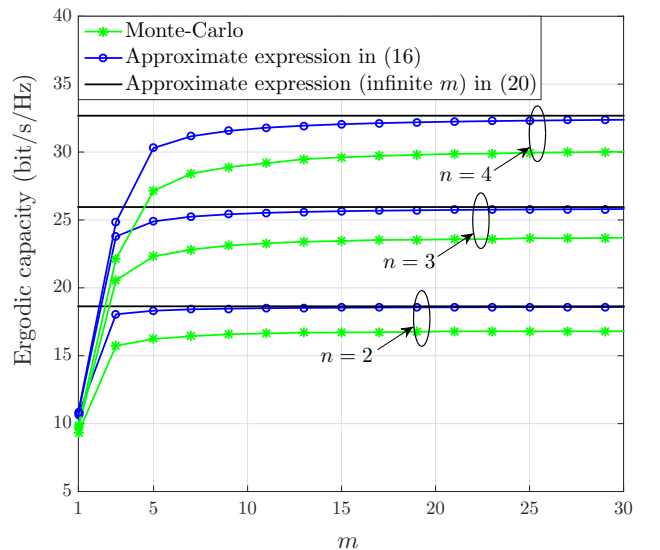

Fig. 3. Comparison of Monte-Carlo simulations with the approximate expressions of the ergodic capacity against m. Results are shown for $n = 2, 3, 4$ with $\alpha = 2.09$, $c_0 = 41.1\,\mathrm{dB}$, $\sigma^2 = 5\,\mathrm{dB}$, $\sigma_{\mathrm{n}}^2 = -70\,\mathrm{dBm}$, $\lambda = 10^{-3}\mathrm{m}^2$, and $\rho = 90\,\mathrm{dB}$.

$$C^{\mathrm{a}}_{m,n}(\lambda) = \begin{cases} \log_2 \frac{\Gamma(m+1)}{m^n \Gamma(m-n+1)} + \sum\limits_{i=1}^{n} \log_2\left(\rho\kappa\xi_\lambda(i)\right), & n \leqslant m, \\ \log_2 \frac{\Gamma(m+1)}{m^m} + \log_2\left(\sum\limits_{B_{m,n}} \prod\limits_{i\in B_{m,n}} \rho\kappa\xi_\lambda(i)\right), & n > m. \end{cases} \tag{16}$$
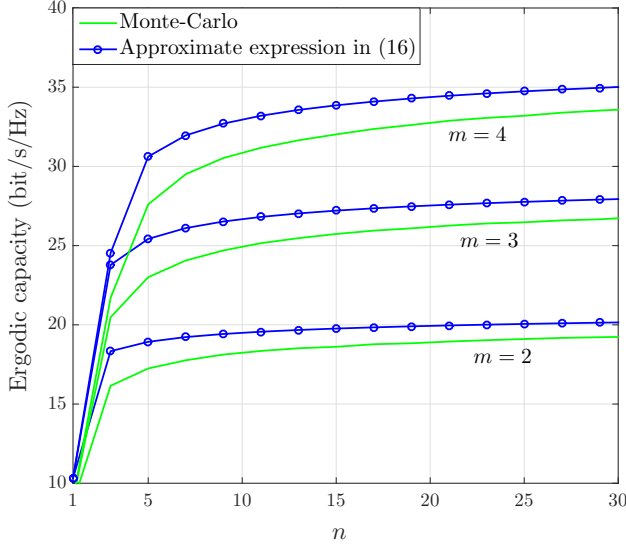

Fig. 4. Comparison of Monte-Carlo simulation with the analytical approximations on the ergodic capacity against n. Results are shown for $m = 2, 3, 4$ with $\alpha = 2.09$, $c_0 = 41.1\,\mathrm{dB}$, $\sigma^2 = 5\,\mathrm{dB}$, $\sigma_{\mathrm{n}}^2 = -70\,\mathrm{dBm}$, $\lambda = 10^{-3}\mathrm{m}^2$, and $\rho = 90\,\mathrm{dB}$.

tightness of the bound improves when the maximum number of antennas grows, which is consistent with the analysis of [33].

To gain more insights, we further examine the following cases.

- For $n < m$, adding one additional RRH, while not altering m, would give

$$C^{\mathrm{a}}_{m,n+1}(\lambda) = C^{\mathrm{a}}_{m,n}(\lambda) + \frac{m-n}{m}\log_2\left(\rho\kappa\xi_\lambda(n+1)\right), \tag{19}$$

yielding a steady improvement with growing n because of the additional power captured by every new RRH.

- For $n \leqslant m$, as $m \to \infty$, the approximation of (16) reduces to

$$\lim_{m\to\infty} C^{\mathrm{a}}_{m,n}(\lambda) \approx \sum_{i=1}^{n} \log_2\left(\rho\kappa\xi_\lambda(i)\right). \tag{20}$$

The result is obtained by using the *Stirling*'s formula approximation, and the proof is presented in Appendix B. The expression shows the upper limit of the capacity with fixed n RRHs, which also reveals the impact of each associated RRH on the capacity.

Figure 3 compares the analytical approximations based on (16) with Monte-Carlo results for the cases of $n = 2, 3, 4$ against m. We note that the tightness of approximations becomes slightly worse with more antennas as also observed in Fig. 2. Moreover, the lines of all approximations increase

sharply with increasing m when $m < n$, but remain nearly constant when $m > n$. Noting the fact that the ergodic capacity is mainly affected by $\min\{n, m\}$, this result is also manifested through a slight offset between the approximation of $m \to \infty$ and the approximation of $m = n$ in the figure.

Figure 4 gives the analytical approximations based on (16) and Monte-Carlo results for the cases of $m = 2, 3, 4$ against n, in which the locations of RRHs are modeled as a PPP. We note that the approximation of the capacity curves show a similar trend as already seen in the results in Fig. 3. Since the number of transmit antennas is generally constant, the result indicates that it is not cost-effective to associate more RRHs, i,e., expand the size of the cell, when the number of them is already greater than the number of transmit antennas.

B. Lower bound with $m = n$

We use $C^{\mathrm{L}}_{m,n}$ to represent the lower bounds of the ergodic capacity. Thus, the conditional expectation of $C^{\mathrm{L}}_{m,n}$ can be expressed as

$$\mathrm{E}\left[C^{\mathrm{L}}_{m,n}\big|\mathbf{B}\right] = \mathrm{E}\left[\log_2 \det\left(\frac{\rho}{m}\mathbf{W}\big|\mathbf{B}\right)\right]. \tag{21}$$

To derive the expression of $C^{\mathrm{L}}_{m,n}$, it is necessary to obtain the expected log-determinant of $\mathbf{W}$. Note that the conditional expected log-determinant has been investigated in [32, Lemma 4]. Utilizing this result, (21) can be rewritten as

$$\mathrm{E}\left[C^{\mathrm{L}}_{m,n}\big|\mathbf{B}\right] = m\log_2 \frac{\rho}{m}$$
$$+ \frac{1}{\ln 2}\left(\sum_{k=1}^{q} \psi(m-q+k) + \frac{\sum\limits_{k=n-q+1}^{n} \det\left(\tilde{\mathbf{\Xi}}_k\right)}{\prod_{i<j}^{n}(\beta_j - \beta_i)}\right), \tag{22}$$

where $\tilde{\mathbf{\Xi}}_k$ is an $n \times n$ matrix with entries

$$\left\{\tilde{\mathbf{\Xi}}_k\right\}_{i,j} = \begin{cases} \beta_i^{j-1}, & j \neq k, \\ \beta_i^{j-1}\ln\beta_i, & j = k. \end{cases} \tag{23}$$

Then, by using (36), the lower bound can be obtained by integration of (22) over β_i term-by-term, as given below

$$C^{\mathrm{L}}_{m,n} = m\log_2 \frac{\rho}{m} + \frac{\sum_{k=1}^{q}\psi(m-q+k)}{\ln 2}$$
$$+ \frac{\left(2(\lambda\pi)^{(n+1)/2}\right)^n}{\ln 2}\int_0^\infty \cdots \int_0^\infty \frac{\sum\limits_{k=n-q+1}^{n} \det\left(\tilde{\mathbf{\Xi}}_k\right)}{\prod_{i<j}^{n}(\beta_j - \beta_i)}$$
$$\times \prod_{i=1}^{n} \frac{\beta_i^{2i+1}e^{-\lambda\pi\beta_i^2}}{(i-1)!}d\beta_1\cdots d\beta_n. \tag{24}$$

The integrand above is extremely difficult to evaluate due to the Vandermonde determinant. In order to alleviate this difficulty, we derive the lower bound for the special case of

$m = n$, i.e., the number of RRHs is equal to the number of transmit antennas, and investigate the impact of path loss exponent α on the capacity.

Theorem 2: For dense C-RAN systems where the RRHs are distributed according to a stationary PPP with intensity λ, the lower bound of the capacity with $m = n$ is given by

$$C_{m,m}^{\mathrm{L}} = m\log_2 \frac{\rho}{m\varepsilon}\left(\pi\lambda\right)^{\frac{\alpha}{2}} + \frac{\left(1 - \frac{\alpha}{2}\right)}{\ln 2}\sum_{k=1}^{m}\psi\left(k\right) \qquad (25)$$

with $\varepsilon = 10^{\frac{c_0}{10}}$. When the transmit signal satisfies the power constraint given in (17), the lower bound is extremely tight to capacity.

Proof: See Appendix C. ∎

Our result in *Theorem 2* gives a very tractable mathematical relationship for the lower bound of the capacity in the arbitrary SNR regime, when $n = m$. Note the fact that since the result is derived by removing the identity matrix in (5), the lower bound in (25) will be more and more accurate as ρ increases, which is also consistent with the power constraint of *Theroem 1*. More importantly, the result shows that the capacity increases logarithmically with the intensity of RRH λ when taking LSF into account.

Corollary 1: The approximate expression of the lower bound of the capacity is given as

$$\bar{C}_{m,m}^{\mathrm{L}} \approx m\log_2\frac{\rho}{m\varepsilon}\left(\pi\lambda\right)^{\frac{\alpha}{2}} + \left(1 - \frac{\alpha}{2}\right)\log_2\Gamma\left(m + 1\right). \quad (26)$$

Proof: Utilizing $\psi\left(m\right) = \ln m + O\left(\frac{1}{m}\right)$ in (25) and after some basic operations yields the result, which completes the proof. ∎

Note that the approximation used in the proof becomes exact when m is infinity. The result involves only standard functions which can be easily evaluated. Moreover, we note
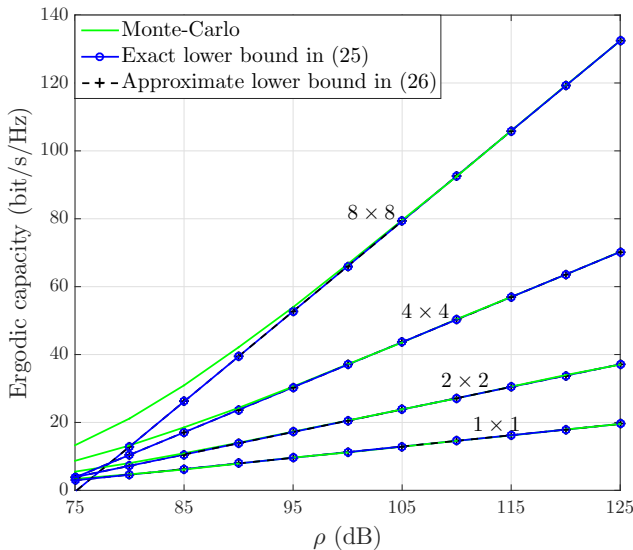

Fig. 5. Comparison of Monte-Carlo simulations, approximate analytical and exact analytical expressions of the ergodic capacity in C-RANs. The different antenna configurations are denoted as $m \times n$. Results are shown for $\alpha = 2.09$, $c_0 = 41.1\,\mathrm{dB}$, $\sigma^2 = 5\,\mathrm{dB}$, $\sigma_{\mathrm{n}}^2 = -70\,\mathrm{dBm}$, and $\lambda = 10^{-3}\mathrm{m}^2$.
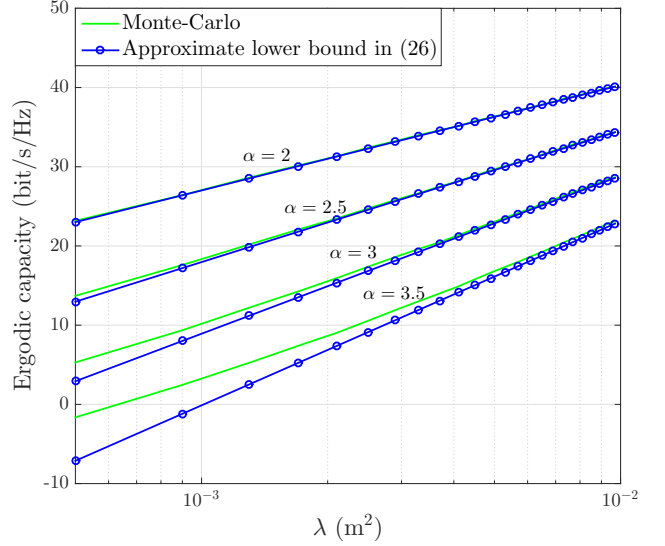

Fig. 6. Comparison of Monte-Carlo simulations and the approximate lower bound of the ergodic capacity in (26) against λ. The results are shown for 4×4, $c_0 = 41.1\,\mathrm{dB}$, $\sigma^2 = 5\,\mathrm{dB}$, $\sigma_{\mathrm{n}}^2 = -70\,\mathrm{dBm}$, and $\rho = 90\,\mathrm{dB}$.

that the capacity decreases with increasing α by taking the partial derivative of α on (26). In order to get more computational insights, we further investigate the approximate expression of capacity for the following cases:

- Adding one antenna at both ends of link, while not altering other parameters, the capacity expression is reduced to

$$\bar{C}_{m+1,m+1}^{\mathrm{L}} = \bar{C}_{m,m}^{\mathrm{L}} + m\log_2\frac{m}{m+1}$$
$$- \frac{\alpha}{2}\log_2\left(\frac{\pi\lambda}{m+1}\right) + \log_2\frac{\rho}{\varepsilon}. \quad (27)$$

The result indicates that the capacity enhancement is a decreasing function of m. Besides, for a certain ρ and m, the enhancement of capacity grows faster with larger α since $\pi\lambda$ is generally smaller than $m + 1$.

- Letting $\alpha = 2$, (25) can be reduced to

$$C_{m,m}^{\mathrm{L}} = m\log_2\frac{\rho\pi\lambda}{m\varepsilon}. \quad (28)$$

We note that the expression is really simple in the case of $\alpha = 2$, which happens in a scenario produced by little blockage and clear ground [36]. It intuitively shows the joint impact of m, ρ, and λ on the capacity, specifically, the increase with growing λ and the number of associated RRHs.

Figure 5 compares the analytical lower bound of capacity in (25), approximate lower bound in (26), and Monte-Carlo simulated curves with four different antenna configurations. As shown in Fig. 5, we note that both the exact lower bound and the approximate lower bound match precisely the simulation result in the high SNR regimes, as expected in *Theorem 2*. Moreover, the approximate expression is very close to the capacity even for the case that $m = 1$, which indicates that the approximation in (26) can be applied to arbitrary number

of antennas as long as $m = n$, and the analysis on the approximate expression in (27) and (28) is highly accurate.

The joint impact of intensity λ and path loss exponent α on the ergodic capacity is illustrated in Fig. 6. We note that the blue lines representing the approximate lower bounds in (26) converge very fast to the simulation results in the cases of $\alpha = 2, 2.5$, while are tight only in the regime of high λ in the cases of $\alpha = 3, 3.5$. The reason is that with high α in the low λ cases, meaning smaller numbers of RRHs and more intense signal attenuation, the analytical expression can hardly characterize the exact capacity due to the SNR constraint. Moreover, we find that the offsets between different α decrease with increasing λ, and the capacity grows logarithmically with increasing λ when the signal power satisfies (17), as predicted by (26).

We again emphasize that the obtained results are derived for OFDM system so that no interference is considered. However, for some extreme cases, such as the heavily-loaded scenarios, the spectrum resource may not enough to support users, the interference is hence inevitable. For practical reference, a more general system model which has considered interference should be introduced. Note that the performance of the new system model becomes even difficult to analyze. However, we are able to approximate the system behaviors by applying [37, Lemma. 1]. Since the applied lemma has employed more approximations, the tightness of the results should be comprehensive justified, and we will go through this model in our future work.

IV. RRH SELECTION AND POWER ALLOCATION

Here, the optimal number of associated RRHs is derived for achieving energy efficient transmission in the amorphous cellular. We then propose two scheduling algorithms, namely, UOS and ROS, for multiuser scenarios based on our capacity results earlier. With the help of our analytical approximation, a large amount of calculations can be avoided and replaced by simply offline searching. In the following section, we provide the analysis and the corresponding scheduling algorithms for one subcarrier. Note that the power consumed by signal transmission and reception are taken into account.

A. Energy Efficiency

We consider one subcarrier in OFDM dense C-RAN system where each single-antenna RRH is serving a single user[3] in a time resource unit. The user establishes its own cell by connecting to neighboring RRHs, and we assume full coordination among them in the cell. Without loss of generality, we suppose that the locations of users on one subcarrier are also modeled as a two-dimensional PPP having intensity λ_u with $\lambda_u \ll \lambda$.

In order to reveal the optimal number of associated RRHs, it is necessary to investigate the energy efficiency (EE) for efficient transmission. We consider the power consumption for signal transmission and reception, and the EE can be defined

---

TABLE I
SYSTEM PARAMETERS.

Path loss exponent α	2.09
Number of antenna per user m	4
Maximum number of RRH $n_{\max}$	10
RRH intensity λ	$10^{-3}\,\mathrm{m}^2$
User intensity[4] λ_u	$10^{-4}\,\mathrm{m}^2$
Protect distance d_{th}	$1\,\mathrm{m}$
Minimum transmit power $P_{\min}$	$5\,\mathrm{dBm}$
Maximum transmit power $P_{\max}$	$23\,\mathrm{dBm}$
Thermal noise	$-70\,\mathrm{dBm}$
Capacity margin Δ_r	$3\,\mathrm{bit/s/Hz}$

as the ratio of the sum spectral efficiency to the sum of the consumed power, described as

$$\eta_{\mathrm{mu}} = \frac{\sum_k C_{m_k, n_k}}{\sum_k n_k P_r + \sum_k P_{t_k}} \text{ bit/J/Hz}, \tag{29}$$

where C_{m_k, n_k} represents the total uplink spectral efficiency for the kth user; P_{t_k} denotes the signal transmission power consumed by the user k; m_k and n_k are the number of antennas of the user k and the number of RRHs which are associated to user k respectively; and P_r is defined as the power consumption at receiver including both the power consumed by signal processing and power required for supporting RRH. For simplicity, we assume that the same P_r for every activated RRHs while the P_{t_k} is adjustable for optimizing η_{mu}.

In contrast to prior works, n_k is allowed to vary according to user requirements. In general, an exhaustive search for the optimal result requires an extremely high complexity and a large amount of overhead, which will be prohibitive for practical applications. Thus, to reduce the complexity, we choose to evaluate the optimal numbers of associated RRHs for a single user according to different demands r_{th}, and save the simulation results for future indexing of the multiuser scenarios.

We see that the expression of EE for a single user can be described as

$$\eta_{\mathrm{su}} = \frac{C_{m,n}}{n P_r + P_t}. \tag{30}$$

We here use the approximation in (16) to represent capacity for offline simulations. In what follows, the approximate EE can be expressed as

$$\hat{\eta}_{\mathrm{su},\lambda} = \frac{C_{m,n}^{\mathrm{a}}(\lambda)}{n P_r + P_t}. \tag{31}$$

We are interested in finding the optimal n and P_t that maximize $\hat{\eta}_{\mathrm{su},\lambda}$, i.e.,

$$\max_{n, P_t} \hat{\eta}_{\mathrm{su},\lambda} \tag{32}$$

---

[3]The assumption can be relaxed to cope with multiuser scenarios when orthogonal multiple access approach is adopted. We aim to show the efficiency of our scheme and leave the RRH scheduling to future work.

[4]The user intensity is reasonable since, in OFDM systems, users are arranged to hundreds of orthogonal subcarriers, i.e., users on one subcarrier are separated to the others [38]. As a consequence, the co-channel users are geographically separated.

TABLE II
JOINT OPTIMAL SOLUTIONS OF (n, P_t)

r_{th}	$m = 2$	$m = 4$	r_{th}	$m = 2$	$m = 4$	r_{th}	$m = 2$	$m = 4$	r_{th}	$m = 2$	$m = 4$
5	$(1,5)$	$(1,5)$	13	$(2,14)$	$(2,13)$	21		$(3,17)$	29		$(4,20)$
6	$(1,7)$	$(1,7)$	14	$(2,15)$	$(2,14)$	22		$(4,15)$	30		$(5,20)$
7	$(1,10)$	$(1,10)$	15	$(2,17)$	$(3,11)$	23		$(4,16)$	31		$(5,21)$
8	$(1,13)$	$(1,13)$	16	$(2,18)$	$(3,12)$	24		$(4,17)$	32		$(5,21)$
9	$(1,17)$	$(2,7)$	17	$(2,20)$	$(3,13)$	25		$(4,17)$	33		$(5,22)$
10	$(2,9)$	$(2,8)$	18	$(2,21)$	$(3,14)$	26		$(4,18)$	34		$(6,22)$
11	$(2,11)$	$(2,10)$	19	$(3,21)$	$(3,15)$	27		$(4,19)$	35		$(6,23)$
12	$(2,12)$	$(2,11)$	20	$(3,23)$	$(3,16)$	28		$(4,20)$	36		$(6,24)$

where $n \in [1, n_{\max}]$ with $n_{\max}$ being a tunable parameter representing the maximum number of associating RRHs. The parameter $P_t \in [P_{\min}, P_{\max}]$ where $P_{\min}$ and $P_{\max}$ stand for minimum transmit power and maximum transmit power respectively. The optimal solution of n_k for the above problem can be obtained through exhaustive search. The parameters of simulations can be found in Table I, and the results of joint configurations of n and P_t (in dBm) for a single user are shown in Table II. Based on Table II, the optimal size of the cell can be determined by the number of associated RRHs according to user's demands. Therefore, with the help of Table II, we are able to propose RRH scheduling algorithms with much lower complexity for multiuser scenarios.

Algorithm 1 UOS

Initialization:
  Set $\mathcal{N} = \{1, \ldots, N\}$, $\mathcal{M} = \{1, \ldots, M\}$, $r_{\text{th}}^{k=1,\ldots,N}$, $m_{k=1,\ldots,N}$, $n_{k=1,\ldots,N} = 0$,
    $\mathcal{R}_{k=1,\ldots,N} = \emptyset$, $F_k = 0$, $l_k = 0$
Step 1: Determining the $n_{k=1,\ldots,N}$ and $P_{t_{k=1,\ldots,N}}$.
  for $k = 1$ to N **do**
    $r^k = r_{\text{th}}^k + \Delta_r$,
    Search $n_{k=1,\ldots,N}$ and $P_{t_{k=1,\ldots,N}}$ in Table II with known r^k and m_k.
  end for
Step 2: Select first n_k RRHs with strongest uplinks for each user.
  while if $F_{k=1,\ldots,N} = 0$ **do**
    for $k = 1$ to N **do**
      if $F_k = 0$ **then**
        $n^* = \underset{i \in \mathcal{M}}{\arg\max} \left\{ d_{ik}^{-\alpha} \right\}$, $\mathcal{R}_k = \mathcal{R}_k \cup n^*$, $\mathcal{M} = \mathcal{M} / \{n^*\}$, $l_k = l_k + 1$
        if $l_k = n_k$ **then**
          $F_k = 1$
        end if
      end if
    end for
  end while
Step 3: Power control for achieving the optimal EE.

Algorithm 2 ROS

Initialization: : Same as initialization in UOS
Step 1: Same as Step 1 in UOS
Step 2: Select first n_k RRHs with strongest uplinks for each user.
  while if $F_{k=1,\ldots,N} = 0$ **do**
    $n^* = \underset{i \in \mathcal{M}, k \in N}{\arg\max} \left\{ d_{ik}^{-\alpha} \right\}$, $\mathcal{R}_k = \mathcal{R}_k \cup n^*$, $\mathcal{M} = \mathcal{M} / \{n^*\}$,
    $l_k = l_k + 1$
    if $l_k = n_k$ **then**
      $\mathcal{N} / \{k\}$, $F_k = 1$
    end if
  end while
Step 3: Same as Step 1 in UOS.

B. UOS and ROS

We propose two effective schemes, namely, UOS and ROS, to form user's amorphous cells on the basis of offline EE results. In UOS, the algorithm selects the required RRHs with the strongest links for each user from all the remaining candidates. On the other hand, the ROS scheme establishes the cell by optimizing the links in the system, i.e., selecting the strongest link from all the remaining users to RRHs and scheduling the RRH to its corresponding user. Note that both algorithms contain three steps and both Step 1 and Step 3 in two algorithms are identical. We first introduce into detail the steps in UOS, and then provide the differences of ROS compared with UOS.

In the initialization stage, we identify the number of users N in the system and the QoS r_{th}^k (in bit/s/Hz) of user k. We use $\mathcal{M}$ to represent the set of the whole RRHs M in the area; $\mathcal{N}$ denotes the set of users N in the system.

In Step 1 of UOS, we evaluate the optimal number of associated RRH n_k for the kth user by referring to Table II. Since our approximate capacity expression is slightly larger than real capacity, we add a margin of the spectral efficiency Δ_r to each r_{th}^k to ensure that the algorithm can always meet the user's demand. Note that n_k can be determined by offline simulation since the intensity of the RRHs λ is determined. Thus, the optimal number of associated RRHs can be selected via Table II according to their demands while avoiding a significant amount of computations.
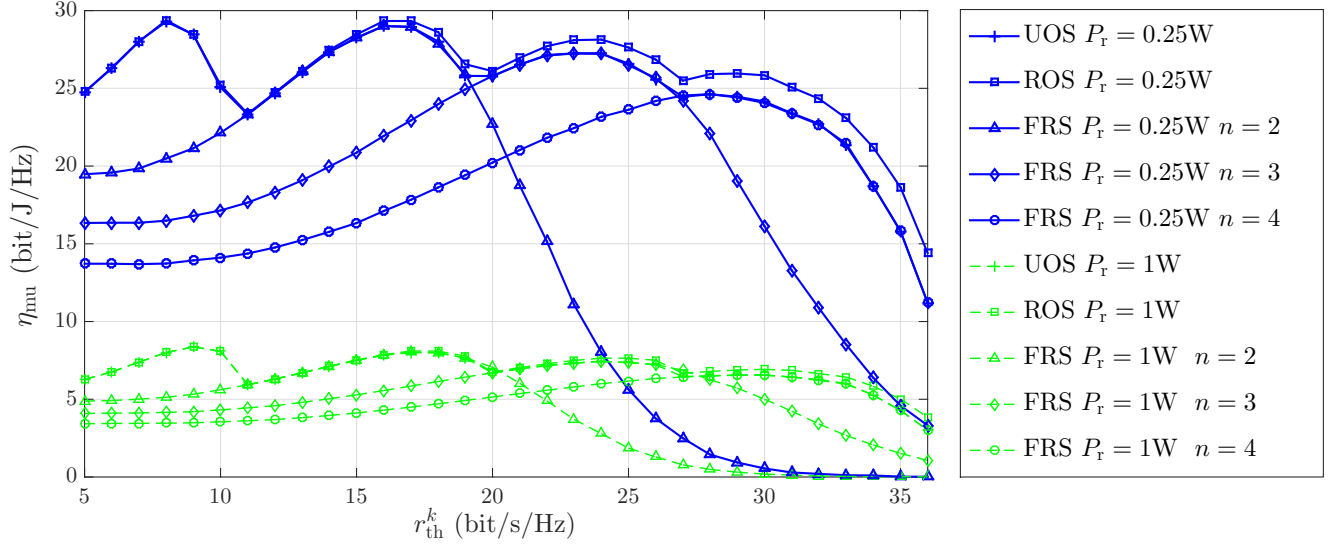
Fig. 7. Energy efficiency for UOS, ROS and FRS versus the different QoS with consideration of both power consumption by signal transmission and reception. We choose $c_0 = 41.1\,\mathrm{dB}$, $\sigma^2 = 5\,\mathrm{dB}$, $\lambda_R = 10^{-3}\mathrm{m}^2$ and $\lambda_u = 10^{-4}\mathrm{m}^2$.

In Step 2 of UOS, we use $\mathcal{R}_k$ to represent the RRHs set selected for the kth user; $d_{ik}^{-\alpha}$ is the LSF between the ith RRH and the kth user; the variable l_k is used as a counter to determine if n_k RRHs have been selected for the kth user; the variable F_k is a flag and $F_k = 1$ in Step 2 means that RRH selection for kth user is completed. In the algorithm, for each round of selection, each user is scheduled one RRH with the strongest link from all the remaining candidates. The algorithm ends until all n_k are satisfied. This step ensures that each selected RRH is the optimal choice for its corresponding user from the remaining candidates, which guarantees the fairness of the scheme. In Step 3 of UOS, we adjust the transmit power to achieve optimized transmission, which is simple and straightforward; $C_{m_k,n_k}(\mathcal{R}_k, P_{t_k})$ represents the capacity of the kth user with associate RRHs $\mathcal{R}_k$ and transmission power P_{t_k}. Here, we provide one alternative scheme, ROS, as a control group, and introduce the difference between the two schemes in Step 2.

In Step 2 of ROS, instead of searching RRHs for the user, we search for the strongest link from all the possible links of the remaining users and the remaining RRHs, and then schedule the RRH to its corresponding user until all the users meet n_k. In contrast to UOS, the selected RRHs in ROS are guaranteed to serve their optimal users. We note that Step 1 and Step 3 are rather simple, and therefore, the complexity of both algorithms is dominated by Step 2. As N and M grow large, the worst-case complexity of Step 2 in UOS and ROS are compared in Table III. As seen, the complexity of ROS is higher than UOS, because searching for the strongest link from all possible links requires more calculations than just searching the strongest link for each user.

The EE against the user requirement with the simulation parameters shown in Table I is illustrated in Fig. 7. For convenience, we assume that all users have four antennas and equal QoS demand (i.e., $m_{k=1,...,N} = 4$ and $r_{\mathrm{th}}^1 = \cdots = r_{\mathrm{th}}^N$). We also provide the results for a fixed number of associated

TABLE III
WORST-CASE COMPLEXITY COMPARISON.

Proposed Algorithms	UOS	ROS
Worst Case Complexity	$O\left(M \sum_k n_k\right)$	$O\left(MN \sum_k n_k\right)$

RRHs scheme (FRS). In FRS, the scheme does not change the number of the associated RRH according to the user's demands. Here, we set $n_{k=1,...,N} = 2, 3, 4$. We first note that the EE with fewer number of RRHs performs better for low user requirements (i.e., $n = 2$ below $19\,\mathrm{bit/s/Hz}$) because of less consumption of transmission power and fewer associated RRHs. Moveover, we observe the exact same trend from the results with $P_r = 0.25\,\mathrm{W}$ and that with $P_r = 1\,\mathrm{W}$, which indicates that the optimal configurations for various requirements are non-changing with different P_r. Furthermore, the lines representing the results for the fixed n_k first increase and then reduce with r_{th}^k due to the exponential increase of transmission power. In contrast to them, UOS and ROS, by applying the results of our capacity analysis in Section III, show significant advantages in EE performance, i.e., both of which yield the best average EE among all lines. Note that there is no significant difference between UOS and ROS in the regime of low user requirement. However, when r_{th} grows large, the ROS scheme is superior by a narrow margin over the UOS scheme, which implies that selecting the users for the RRH is a better strategy compared with selecting the RRHs for the user for user-centric system. Although the worst-case complexity of the ROS is N-times that UOS has, ROS still has an advantage over UOS since N is not a big number considering that $\lambda_u \ll \lambda$.

C. λ/λ_u

For a certain number of users, the throughput of the OFDM dense C-RAN system always grows with an increasing number

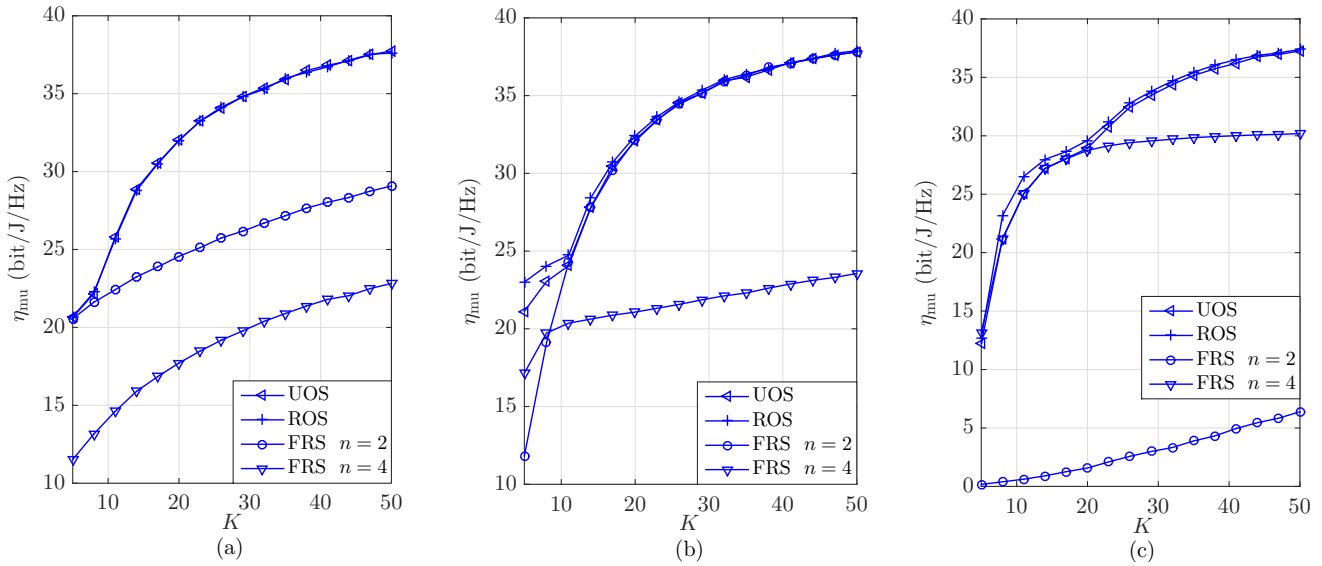Fig. 8. Energy efficiency for UOS, ROS and FRS versus K for low/medium/high QoS requirements with fixed λ_u. (a) $r_{\text{th}}^k = 10$ bit/s/Hz. (b) $r_{\text{th}}^k = 20$ bit/s/Hz. (c) $r_{\text{th}}^k = 30$ bit/s/Hz. Results are shown for $P_r = 0.25$ W, $\alpha = 3$ and $\lambda_u = 10^{-4}$ m^2.

of RRHs due to the reduced TX-RX distance and the multiuser gain. When the intensity of the RRHs grows, the EE always increases correspondingly because of the reduced distance between the user and the RRH. However, considering the cost of the RRH, deploying an excessive number of RRHs without bound is meaningless and not cost-effective.

Proposition 1: Consider the optimal configuration (n^*, P_t^*) for UOS and ROS having approximate high-SNR capacity $C_{m,n^*}^{\text{a}}(\lambda)$. When $K = \lambda/\lambda_u$ is large, to achieve the same EE performance by utilizing a FRS associated with n' RRHs, the RRH intensity λ' can be given as the solution of the equation

$$C_{m,n'}^{\text{a}}(\lambda') = \frac{n' P_r + P_t^*}{n^* P_r + P_t^*} C_{m,n^*}^{\text{a}}(\lambda). \quad (33)$$

Proof: The result can be simply obtained by solving the equation $\hat{\eta}_{\text{su},\lambda'} = \hat{\eta}_{\text{su},\lambda}$. ∎

Note that the result holds for $\lambda \gg \lambda_u$. The reason is that our algorithms schedule RRHs by avoiding any of them been selected multiple times, i.e., some of users may not be serviced by their n-nearest RRHs for small K. Moreover, the result varies according to the system parameters, such as, the intensity of the RRHs λ, rate requirement r_{th} (i.e., $C_{m,n^*}^{\text{a}}(\lambda)$), number of the antenna per user m and the power consumption for supporting RRHs P_r, and the RRH deployment can then be optimized based on the result.

In general, the intensity of the RRHs, λ, is fixed in a deployment, while the intensity of the users, λ_u, is the actual variable. With the goal of optimizing the deployment of the RRH, it is important to reveal the trend of the EE with respect to the ratio of λ and λ_u. In this subsection, we consider a system in the plane whose radius $R = 500$ m. The parameters of simulations can also be found in Table I except λ and λ_u. For convenience, we suppose all users have same user QoS demand (i.e., $r_{\text{th}}^1 = \cdots = r_{\text{th}}^N$). The simulations are shown for both fixed λ_u and λ respectively in the low/medium/high QoS requirement scenarios.

The EE against the ratio K with fixed λ_u is shown in Fig. 8. Intuitively, the EE always increases with growing K among three user QoS demand scenarios but with different speed. We find that the FRS with $n = 4$ performs poorly in the low QoS and medium demand cases as well as the FRS with $n = 2$ in the high QoS demand cases. Moreover, we see that the curves for the UOS, ROS and FRS with $n = 4$ increase sharply when $K < 25$, but grow smoothly when $K > 25$ in both medium and high QoS demand scenarios. The reason is that the distance between the associated RRH and the user can be hardly reduced when K is large, which means that the signal strength cannot be effectively enhanced, i.e., the deployment of RRH is already too "crowded" for the users. More importantly, the figures show that UOS and ROS can work effectively and keep a significant advantage in various scenarios. This is important since by applying an appropriate scheduling algorithm, the number of the deployed RRH can be sharply reduced (e.g., the number of RRH can be reduced by half in Fig. 8a and even more in the medium and high QoS demand scenarios).

Fig. 9 shows the EE against the ratio K with fixed intensity of RRH λ. This simulation corresponds to the scenarios with a given number of RRHs and very few users. The figures show that even with just a small number of users in the system, the EE can be substantially improved by applying UOS and ROS (e.g., about 10 bit/J/Hz and 15 bit/J/Hz in low and medium QoS demand cases). In Fig. 9c, the users with high QoS demands experience very poor EE if they only connect with the two nearest RRHs. We also note that the curves remain constant when K grows large, which implies that to achieve an energy-efficient transmission, many of the RRHs should not activated based on the scheduling algorithms when the number of the users is small.

In summary, both the proposed schemes offer a substantial improvement in EE performance. By applying appropriate
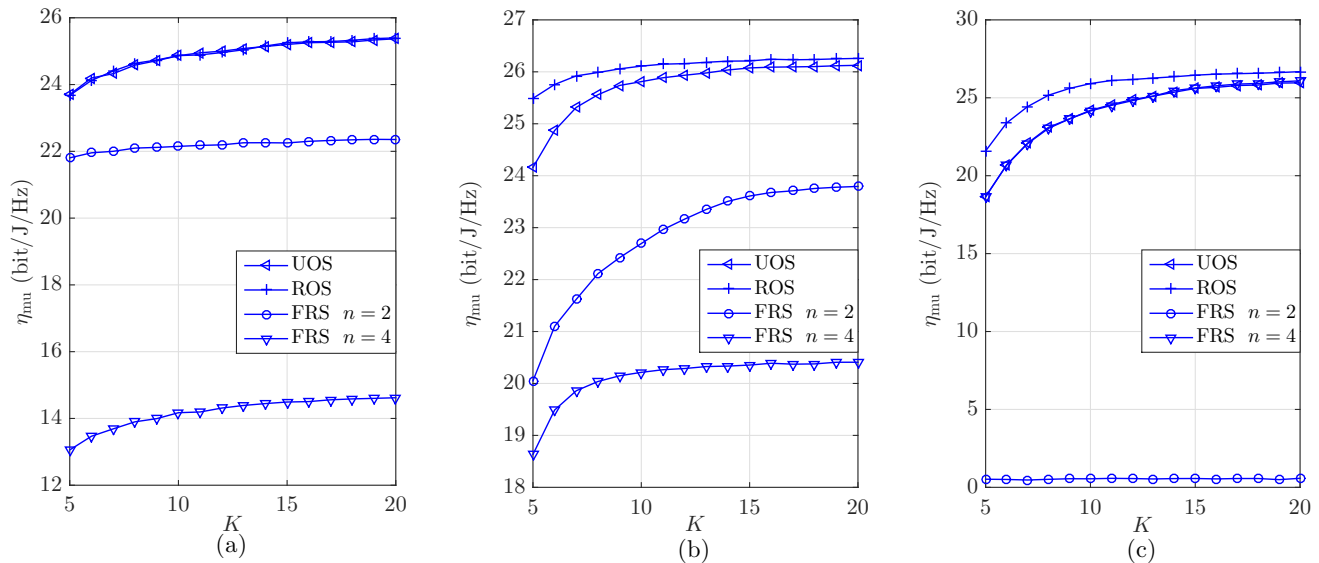
Fig. 9. Energy efficiency for UOS, ROS and FRS versus K for low/medium/high QoS requirements with fixed λ. (a) $r_{\text{th}}^k = 10$ bit/s/Hz. (b) $r_{\text{th}}^k = 20$ bit/s/Hz. (c) $r_{\text{th}}^k = 30$ bit/s/Hz. Results are shown for $P_r = 0.25$ W, $\alpha = 2.09$ and $\lambda = 10^{-3} \text{m}^2$.

scheduling schemes, we can not only achieve energy-efficient communication but also optimize the RRH deployment.

V. CONCLUSION

In this paper, we investigated amorphous cellular which is based on user-centric topologies for ultra-dense C-RANs. An approximation of the uplink capacity at high SNRs has been presented, as well as, a very tight lower bound when the number of user antennas is equal to that of associated RRHs. Based on results, we characterized the impact of path loss exponent, RRH intensity and antenna configuration on the ergodic capacity. Moreover, the optimal number of associate RRHs for single user was derived when the transmission and reception power consumption are taken into account. These results were subsequently used in two scheduling algorithms, namely UOS and ROS, to obtain energy efficient transmission. Both of the scheduling algorithms induce very low-complexity and offer a substantial improvement in EE for arbitrary QoS requirements. More importantly, according to the analytical and simulation results, a large number of deployed RRHs can be waived while still maintaining a desirable EE level by applying our scheduling algorithms.

APPENDIX A
PROOF OF LEMMA 1

To prove this lemma, it is convenient to give separate treatments for two cases.

1) $n \leqslant m$ Case: In this case, the expected determinant of $\mathbf{W}$, conditioned on $\mathbf{B}$, is given by

$$
\begin{aligned}
\text{E}\left[\det\left(\mathbf{W}\right)\middle|\mathbf{B}\right] &= \text{E}\left[\det\left(\mathbf{B}\mathbf{H}\mathbf{H}^\dagger\right)\middle|\mathbf{B}\right] \\
&= \text{E}\left[\det\left(\mathbf{B}\right)\det\left(\mathbf{H}\mathbf{H}^\dagger\right)\middle|\mathbf{B}\right] \\
&\overset{(a)}{=} \text{E}\left[\det\left(\mathbf{B}\right)\right]\text{E}\left[\det\left(\mathbf{H}\mathbf{H}^\dagger\right)\right],
\end{aligned} \quad (34)
$$

where (a) is obtained due to the independence of $\mathbf{B}$ and $\mathbf{H}$. Applying the result in [31, Eq. (A.7.1)], (34) can be further represented as

$$
\text{E}\left[\det\left(\mathbf{W}\right)\middle|\mathbf{B}\right] = \text{E}\left[\det\left(\mathbf{B}\right)\right]\frac{\Gamma\left(m+1\right)}{\Gamma\left(m-n+1\right)}. \quad (35)
$$

To evaluate the integral of $\mathbf{B}$, we quote the marginal distribution of d_i in [39, Eq. (21)], given in following lemma.

Lemma 2: For a two-dimensional PPP of particles in the plane with intensity λ, the pdf of d_i, which is the distance between the origin to the ith nearest particle, is given by

$$
f_{d_i}\left(x\right) = \frac{2(\lambda\pi)^i}{(i-1)!}x^{2i-1}e^{-\lambda\pi x^2}. \quad (36)
$$

Note that the expression (36) is the marginal distribution of d_i which has removed the correlation across the distances between user and RRHs by considering all the possibility of $(i-1)$ RRHs are nearer than the ith nearest RRHs, while the rest of them are farther than ith nearest RRHs. Therefore, the $f_{d_{i=1,\ldots,N}}\left(x\right)$ are mutual independent.

Thus, utilizing *Lemma* 2, the closed-form expression of the expected determinant of $\mathbf{B}$ can be further evaluated as

$$
\begin{aligned}
\text{E}\left[\det\left(\mathbf{B}\right)\right] &\overset{(b)}{=} \prod_{i=1}^{n}\text{E}\left[\beta_i\right] \\
&= \prod_{i=1}^{n}10^{-\frac{c_0}{10}}\text{E}_{z_i}\left[10^{-\frac{\sigma z_i}{10}}\right]\text{E}_{d_i}\left[d_i^{-\alpha}\right] \\
&\overset{(c)}{=} \prod_{i=1}^{n}\frac{\kappa(\lambda\pi)^{\frac{\alpha}{2}}}{\Gamma\left(i\right)}\Gamma\left(i-\frac{\alpha}{2}\right),
\end{aligned} \quad (37)
$$

where (b) is derived by noting that $\mathbf{B}$ is a diagonal matrix whose entries are mutual independent, and (c) is obtained from [24, Eq. (3.323.2)] and [23, Eq. (6.1.5)]. We note that $\Gamma\left(i-\frac{\alpha}{2}\right)$ will run to infinity when $i-\frac{\alpha}{2}$ is a non-positive integer. This is because of the inaccuracy of the LSF model

in (4) at point of origin. Thus, we introduce a small constant d_{th} as a guard distance, and the integrand will be evaluated as

$$\int_{d_{\text{th}}}^{\infty} d_i^{2i-1-\alpha} e^{-\lambda \pi d_i^2} \mathrm{d}d = \frac{(\lambda \pi)^{\frac{\alpha}{2}-i}}{2} \Gamma\left(i - \frac{\alpha}{2}, d_{\text{th}}\right). \quad (38)$$

Combining (37) and (38), we obtain the results.

2) $n > m$ Case: For this case, utilizing the Cauchy-Binet formula [40], the expected determinant of $\mathbf{W}$, conditioned on $\mathbf{B}$, is given by

$$\mathrm{E}\left[\det(\mathbf{W})|\mathbf{B}\right] = \mathrm{E}\left[\sum_{B_{m,n}} \det\left(\mathbf{H}^\dagger \mathbf{B}\right)_{B_{m,n}} \det(\mathbf{H})_{B_{m,n}}\middle|\mathbf{B}\right]. \quad (39)$$

Note that $(\cdot)_{B_{m,n}}$ is an $m \times m$ matrix. Thus, by utilizing the commutativity of matrix multiplication and changing the order of the matrices in (39), we can further obtain

$$\mathrm{E}\left[\det(\mathbf{W})|\mathbf{B}\right]$$
$$= \mathrm{E}\left[\det\left(\mathbf{H}_{B_{m,n}}^\dagger \mathbf{H}_{B_{m,n}}\right)\right] \sum_{B_{m,n}} \mathrm{E}\left[\det(\mathbf{B})_{B_{m,n}}\right]$$
$$= \Gamma(m+1) \sum_{B_{m,n}} \mathrm{E}\left[\det(\mathbf{B})_{B_{m,n}}\right]. \quad (40)$$

We then evaluate the integral of (40) as in (37) to yield the result.

APPENDIX B
PROOF OF (20)

Utilizing the Stirling's formula

$$\Gamma(k+1) \approx \sqrt{2k\pi}\left(\frac{k}{e}\right)^k, \quad (41)$$

the approximation in (16) with $n \leqslant m$ can be approximately rewritten as

$$C_{m,k}^{\mathrm{a}} \approx \log_2 \frac{\sqrt{2m\pi}\left(\frac{m}{e}\right)^m}{m^k \sqrt{2(m-k)\pi}\left(\frac{m-k}{e}\right)^{m-k}} + \sum_{i=1}^{k} \log_2\left(\rho\kappa\xi_\lambda(i)\right)$$
$$= \log_2 \frac{1}{e^k}\left(\frac{m}{m-k}\right)^{m-k+\frac{1}{2}} + \sum_{i=1}^{k} \log_2\left(\rho\kappa\xi_\lambda(i)\right). \quad (42)$$

According to the definition of natural constant e, we have

$$\lim_{m \to \infty}\left(\frac{m}{m-k}\right)^{m-k+\frac{1}{2}} = e^k. \quad (43)$$

Substituting (43) into (42) and after some basic operations, we conclude the proof.

APPENDIX C
PROOF OF THEOREM 2

We first rewrite the conditional lower bound of ergodic capacity

$$\mathrm{E}\left[C_{m,n}^{\mathrm{L}}|\mathbf{D}\right] = m\log_2 \frac{\rho}{m}$$
$$+ \frac{1}{\ln 2}\left(\sum_{k=1}^{q} \psi(m-q+k) + \frac{\sum_{k=n-q+1}^{n} \det\left(\tilde{\mathbf{\Xi}}_k\right)}{\prod_{i<j}^{n}\left(d_j^{-\alpha} - d_i^{-\alpha}\right)}\right), \quad (44)$$

When $m = n$, the sum of the determinants is reduced to

$$\sum_{k=1}^{p} \det\left(\tilde{\mathbf{\Xi}}_k\right) \stackrel{(d)}{=} \sum_{k=1}^{p}\sum_{\sigma} \mathrm{sgn}(\sigma)\prod_{i=1}^{p} \beta_{\sigma(i)}^{i-1} \ln \beta_{\sigma(k)}$$
$$= \sum_{\sigma} \mathrm{sgn}(\sigma)\prod_{i=1}^{p} \beta_{\sigma(i)}^{i-1} \sum_{k=1}^{p} \ln \beta_{\sigma(k)}$$
$$= \sum_{i=1}^{p} \ln \beta_i \prod_{i<j}^{p}(\beta_j - \beta_i), \quad (45)$$

where (d) is derived by the Leibniz formula [41], and the second summation is over all permutations $\sigma = \{\sigma(1), \ldots, \sigma(p)\}$ of the set $\{1, \ldots, p\}$, with $\mathrm{sgn}(\sigma)$ representing the sign of the permutation. By substituting (45) into (44), we can rewrite the expression as

$$\mathrm{E}\left[C_{m,n}^{\mathrm{L}}|\mathbf{B}\right] = m\log_2 \frac{\rho}{m} + \frac{1}{\ln 2}\left(\sum_{k=1}^{m} \psi(k) + \sum_{i=1}^{m} \ln \beta_i\right). \quad (46)$$

Using *Lemma* 2 and (46), the lower bound of the ergodic capacity can be simplified as

$$C_{m,m}^{\mathrm{L}} = \int_{\mathbf{B}} \mathrm{E}\left[\log_2 \det\left(\frac{\rho}{m}\mathbf{W}\right)\middle|\mathbf{B}\right] f_{\mathbf{B}}\mathrm{d}\mathbf{B}$$
$$= m\log_2 \frac{\rho}{m} + \frac{\sum_{k=1}^{m} \psi(k)}{\ln 2} + \frac{1}{\ln 2}\sum_{i=1}^{m} \mathrm{E}\left[\ln \beta_i\right]. \quad (47)$$

The integral in (47) can be further evaluated as

$$\sum_{i=1}^{m} \mathrm{E}\left[\ln \beta_i\right]$$
$$= -\left(\frac{mc_0 \ln 10}{10} + \frac{\sigma \ln 10}{10}\sum_{i=1}^{m} \mathrm{E}_{z_i}\left[z_i\right] + \alpha\sum_{i=1}^{m} \mathrm{E}_{d_i}\left[\ln d_i\right]\right)$$
$$\stackrel{(e)}{=} m\left(\frac{\alpha}{2}\left(c + \ln \pi\lambda + -\frac{\kappa(m)}{m}\right) - \frac{c_0 \ln 10}{10}\right), \quad (48)$$

where (e) is obtained from [24, Eq. (4.331.1)] and [24, Eq. (4.352.2)], c is the Euler constant, and

$$\kappa(m) = \begin{cases} 0, & m = 1, \\ \sum_{i=2}^{m}\sum_{j=1}^{i-1} j^{-1}, & m > 1. \end{cases} \quad (49)$$

Thus, by substituting (49) into (48), the expression can be rewritten as

$$C_{m,m}^{\mathrm{L}} = m\log_2 \frac{\rho}{m} + \frac{\sum_{k=1}^{m} \psi(k)}{\ln 2}$$
$$+ \frac{\alpha m}{2\ln 2}\left(c + \ln \pi\lambda - \frac{\kappa(m)}{m}\right) + \frac{mc_0}{10}\log_2 10. \quad (50)$$

Note that we are using a property of digamma function by [35, Eq. (2.14)]

$$\psi(k) = -c + \sum_{j=1}^{k-1} j^{-1}. \quad (51)$$

Substituting (51) into (50) and after some basic operations, we complete the proof.

REFERENCES

[1] Cisco, "Global mobile data traffic forecast update, 2013–2018," pp. 1–40, Feb. 2014.

[2] I. Hwang, B. Song, and S. S. Soliman, "A holistic view on hyper-dense heterogeneous and small cell networks," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 20–27, Jun. 2013.

[3] W. H. M. Kamel and A. Youssef, "Ultra-dense networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 4, pp. 2522–2545, Fourth quarter 2016.

[4] T. Ihalainen, *et al.*, "Flexible scalable solutions for dense small cell networks," in *Proc. IEEE WWRF*, Apr. 2013.

[5] P. Demestichas, *et al.*, "5G on the horizon: Key challenges for the radio-access network," *IEEE Veh. Technol. Mag.*, vol. 8, no. 3, pp. 47–53, Sept. 2013.

[6] S. H. Park, O. Simeone, O. Sahin, and S. Shamai, "Robust and efficient distributed compression for cloud radio access networks," *IEEE Trans. Veh. Technol.*, vol. 62, no. 2, pp. 692–703, Feb. 2013.

[7] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1834–1850, Mar. 2017.

[8] Z. Ding and H. V. Poor, "The use of spatially random base stations in cloud radio access networks," *IEEE Signal Process. Lett.*, vol. 20, no. 11, pp. 1138–1141, Nov. 2013.

[9] B. Dai and W. Yu, "Sparse beamforming and user-centric clustering for downlink cloud radio access network," *IEEE Access*, vol. 2, pp. 1326–1339, Oct. 2014.

[10] J. Wang and L. Dai, "Downlink rate analysis for virtual-cell based large-scale distributed antenna systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 1998–2011, Mar. 2016.

[11] M. Peng, Y. Li, J. Jiang, J. Li, and C. Wang, "Heterogeneous cloud radio access networks: A new perspective for enhancing spectral and energy efficiencies," *IEEE Wireless Commun.*, vol. 21, no. 6, pp. 126–135, Dec. 2014.

[12] C. Li, J. Zhang, M. Haenggi, and K. B. Letaief, "User-centric intercell interference nulling for downlink small cell networks," *IEEE Trans. Commun.*, vol. 63, no. 4, pp. 1419–1431, Apr. 2015.

[13] Y. Shi, J. Zhang, and K. B. Letaief, "Group sparse beamforming for green cloud-RAN," *IEEE Trans. Wireless Commun*, vol. 13, no. 5, pp. 2809–2823, May 2014.

[14] Y. Zhang and Y. J. Zhang, "User-centric virtual cell design for cloud radio access networks," in *Proc. IEEE SPAWC*, Jun. 2014, pp. 249–253.

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

[16] A. Guo and H. Martin, "Spatial stochastic models and metrics for the structure of base stations in cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 11, pp. 5800–5812, Nov. 2013.

[17] M. Peng, S. Yan, and H. V. Poor, "Ergodic capacity analysis of remote radio head associations in cloud radio access networks," *IEEE Wireless Commun. Lett.*, vol. 3, no. 4, pp. 365–368, Aug. 2014.

[18] L. Dai, "An uplink capacity analysis of the distributed antenna system (DAS): From cellular DAS to DAS with virtual cells," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2717–2731, May 2014.

[19] P. J. Smith, L. M. Garth, and S. Loyka, "Exact capacity distributions for MIMO systems with small numbers of antennas," *IEEE Commun. Lett.*, vol. 7, no. 10, pp. 481–483, Oct. 2003.

[20] M. Chiani, M. Z. Win, and A. Zanella, "On the capacity of spatially correlated MIMO Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2363–2371, Oct. 2003.

[21] L. Hanlen and A. Grant, "Capacity analysis of correlated MIMO channels," *IEEE Trans. Inf. Theory*, vol. 58, no. 11, pp. 6773–6787, Nov. 2012.

[22] M. Ahmadi, M. Ni, and J. Pan, "A geometrical probability-based approach towards the analysis of uplink inter-cell interference," in *Proc. IEEE GLOBECOM*, Dec. 2013, pp. 4952–4957.

[23] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical functions with Formulas, Graphs, and Mathematical Tables*. Courier Corporation, Jun. 1964, no. 55.

[24] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products, 7th ed*. New York: Academic Press, 2007.

[25] F. Yang and X. Zhang, "Efficient packet detection for D2D power-saving communications over mobile wireless cellular networks," in *Proc. IEEE GLOBECOM*, Dec. 2014, pp. 2492–2497.

[26] T. D. Hoang, L. B. Le and T. Le-Ngoc, "Joint subchannel and power allocation for D2D communications in cellular networks," in *Proc. IEEE WCNC*, Apr. 2014, pp. 1338–1343.

[27] S. Chiu, D. Stoyan, W. Kendall, and J. Mecke, *Stochastic Geometry and Its Applications, 3rd ed*. John Wiley and Sons, 2013.

[28] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommun*, vol. 10, no. 6, pp. 585–595, 1999.

[29] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, no. 3, pp. 311–335, 1998.

[30] J. Hoydis, K. Hosseini, S. Ten Brink, and M. Debbah, "Making smart use of excess antennas: Massive MIMO, small cells, and TDD," *Bell Labs Technical Journal*, vol. 6, no. 2, pp. 5–21, Sep. 2013.

[31] A. Grant, "Rayleigh fading multi-antenna channels," *EURASIP J. Appl. Signal Process.*, vol. 3, pp. 316–329, Mar. 2002.

[32] S. Jin, M. R. McKay, C. Zhong, and K.-K. Wong, "Ergodic capacity analysis of amplify-and-forward MIMO dual-hop systems," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2204–2224, May 2010.

[33] J. Yuan, M. Matthaiou, S. Jin, and F. Gao, "Tightness of Jensen's bounds and applications to MIMO communications," *IEEE Trans. Commun.*, vol. 65, no. 2, pp. 579–593, Feb. 2017.

[34] S. Samarakoon, M. Bennis, W. Saad, M. Debbah, and M. Latva-aho, "Ultra dense small cell networks: Turning density into energy efficiency," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 5, pp. 1267–1280, May 2016.

[35] A. M. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications*. Now Publishers, 2004.

[36] V. Erceg, L. J. Greenstein, S. Y. Tjandra, S. R. Parkoff, A. Gupta, B. Kulic, A. A. Julius, and R. Bianchi, "An empirically based path loss model for wireless channels in suburban environments," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 7, pp. 1205–1211, Jul. 1999.

[37] Q. Zhang, S. Jin, K. K. Wong, H. Zhu and M. Matthaiou, "Power scaling of uplink massive MIMO systems with arbitrary-rank channel means," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 966–981, Oct. 2014.

[38] M. Gast, *802.11 ac: A Survival Guide*. O'Reilly Media, Inc., 2013.

[39] D. Moltchanov, "Distance distributions in random networks ," *Ad Hoc Networks*, vol. 10, no. 6, pp. 1146–1166, 2012.

[40] J. G. Broida and S. G. Williamson, *A Comprehensive Introduction to Linear Algebra*. Addison-Wesley Reading, Mass., 1989.

[41] L. N. Trefethen and D. Bau III, *Numerical Linear Algebra*. Siam, 1997, vol. 50.