User-Centric Networking for Dense C-RANs: High-SNR Capacity Analysis and Remote Radio Head Selection

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

Abstract

Ultra-dense cloud radio access networks (C-RANs) is one of the well-acknowledged promising architectures in next generation wireless systems. A novel signal processing framework is desirable in response to the new architecture, which could be established based on user-centric manner. Amorphous cellular is a promising framework, in which each user connects a few neighboring remote radio heads (RRHs) to form its own cell. This paper studies the capacity of amorphous cellular in dense C-RANs at high-SNR, where the RRHs are distributed as Poisson Point Process. We derive tight upper bounds with arbitrary antenna configurations, and the exact expression of ergodic capacity when antenna numbers are the same at both ends of the link at high SNR. 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 easily evaluated. The impact of the associate RRHs number on ergodic capacity is evaluated, and the results, for links with equal number of antennas at both ends, show that the capacity increase logarithmically with increasing intensity of RRH. Further on the basis of our results, we propose two efficient scheduling algorithms for RRH selection for the multiuser frequency-division duplex system to achieve energy efficient transmission.

Index Terms

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@seu.edu.cn; jinshi@seu.edu.cn; wxu@seu.edu.cn; tanweiqiang@seu.edu.cn).

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

Amorphous cellular, ergodic capacity, multiple-input multiple-output (MIMO), dense cloud radio access network (C-RAN).

I. Introduction

The property of mobile Internet calls for new topology to meet the increasing demand for data traffic, numerically, a thousand times in the coming decade [1]. The conventional idea of cellular network is already overstretched to face the situation. In 5G, to face the challenge, ultra-dense network (UDN) has become an attractive focus thanks to its considerable potential on capacity enhancement [2–4].

The typical cellular network topology has been well-suited for providing wide-area coverage in the past, but not handling the exponential increasing requirement of data. Dense cloud radio access network (C-RAN) is an emerging network architecture which is recognized as the enabling technologies to meet such demands of emerging mobile traffics, due to the enhanced received signal strength because of the reduced distance between user and remote radio head (RRH) [5]. In C-RANs, RRHs operate as soft relays receiving signals from mobile users to a centralized base band unit (BBU) [6]. The performance of distributed antenna array and best base station selection schemes in C-RANs were presented in [7], in which both user and RRH equipped a single antenna. In [8], authors studied the sum-rate maximization problem under BS backhaul constraint in a downlink C-RAN for both dynamic and static BS clustering over different time-frequency slots. However, single access point scheme is incapable of supporting enough mobile traffic due to the low power of deployed RRH. Intuitively, if we intend to catch up with the traffic demand, multi-access point scheme is inevitable in C-RANs because of the substantial improvement in spectral efficiency of multiple-input multiple-output (MIMO) technology.

With dense RRH deployment, the BS-centric structure MIMO system cannot be justified when both users and BSs are scattered around due to the poor support for cell-edge users [9]. Instead, the user-centric structure, namely, amorphous cellular, would be more appropriate, and the interference coordination is much more feasible [10–12]. The user in C-RAN chooses its own serving RRH set as its amorphous cellular while BBU schedules time-frequency recourse centrally. The authors of [13] investigated the optimal sizes of amorphous cellular for single-user transmission, where locations of RRHs were distributed according to poisson point process (PPP) which is an effective mathematical model that can capture the irregularity of BSs [14, 15]. The authors of [16] presented closed-form ergodic capacity expressions for

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 [17], where RRHs were either co-located at cell center or uniformly distributed within each cell. As an extension of this work, the authors further studied the downlink amorphous cellular with a large number of users randomly distributed in the system and derived the effect of cellular size on the average user rate [9]. Note that the above results are obtained by assuming single-antenna users, and can hardly provide closed-formed achievable rate expression. This is mainly because of the intractable of the effect of large scale fading (LSF), especially when users equip multiple antennas. However, multi-antenna has been the standard configuration of present devices, which forces us to form a more accurate result in characterizing the capacity performance of MIMO system.

In order to measure the system behavior over amorphous cellular in ultra-dense C-RANs, we tend to evaluate the capacity by utilizing an efficient mathematic tool, namely, random matrix theory. A large amount of prior works applied random matrix theory to obtain analytical characterization of MIMO systems, and insightful results were also derived under semi-correlated channel. The authors of [18] considered the capacity outage performance of a MIMO system in correlated environments and derived exact distribution functions for the capacity with a small number of antennas. In [19], a closed-form expression for the characteristic function of MIMO system capacity with arbitrary correlation among transmit (receive) antennas was derived. The authors of [20] analyzed the capacity and corresponding optimal input density of a correlated MIMO channel, where the channel was assumed to have a (Kronecker) correlated normal distribution. All the above studies provided closed-form expressions of capacity for the cases where channel correlations present at one of the two ends of the link. However, we note that most of the researches assume MIMO links between user and one multi-antenna BS, rather than multi-BS. A crucial difference between the two assumptions is whether taking LSF into account in terms of capacity analysis, which is, however, considered as an important factor of capacity performance in distributed MIMO systems [21]. Dense C-RAN, equivalently a kind of distributed MIMO system, belongs to the latter one. Utilizing random matrix theory, we note that the effect of LSF on capacity can be formulated by regarding the LSF matrix as a correlation matrix at the end of link. To the best of our knowledge, there appears no analytical expression available for ergodic capacity which applies for distributed MIMO systems with both arbitrary path loss exponent and number of antennas, especially when locations of RRHs are deployed randomly, e.g., as

PPP instead of some predetermined regular topology.

In this paper, the uplink capacity of amorphous cellular in ultra-dense C-RAN at high-SNR is investigated, where locations of RRHs obey the PPP distribution. We derive the upper bounds of the capacity with an arbitrary number of antennas at high SNR when LSF is taken into account. Moreover, the exact expression of capacity is derived when the antenna numbers are the same at both ends of the link. Based on the proposed upper bounds, the impact of the number of antennas is characterized. The expressions, in contrast to prior results, involve only standard functions which can be easily and efficiently evaluated, and they illustrate the effects of path loss exponent and the RRH density on 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 multiuser frequency division duplex (FDD) systems when both transmission and reception power are considered. Both algorithms are heavily on the basis of our capacity analysis results, and maintain excellent performance in arbitrary QoS requirements.

The remainder of this paper is organized as follow. Section II presents the system model of an amorphous cellular in dense C-RAN under consideration. Section III presents the results of capacity analysis at high SNRs. In Section IV, two scheduling algorithms are proposed to achieve energy efficient transmission. Section V summarizes the main observations and proofs are relegated to appendices.

Throughout the paper, vectors and metrics are denoted in bold lowercase $\mathbf{a}$ and bold uppercase $\mathbf{A}$, respectively. The notation $\det(\mathbf{A})$ 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 opeartion, and $E[\cdot]$ evaluates the expectation of the input random entity. Additionally, $\Gamma(\cdot)$ and $\Gamma(\cdot,\cdot)$ are the gamma function and incomplete gamma function [22, Eq. (6.1.5), Eq. (6.5.3)] respectively, and $\psi(\cdot)$ is the digamma function [23, Eq. (8.360.1)].

II. SYSTEM MODEL

Consider a dense C-RAN, sketched in Fig. 1, in which each RRH has a single antenna, helping the signals of a user who has m antennas to be processed in the BBU, which constitutes a distributed MIMO system. The user connects n nearest RRHs to form its own cell. The locations of RRHs are assumed to be modeled as a two-dimensional PPP having

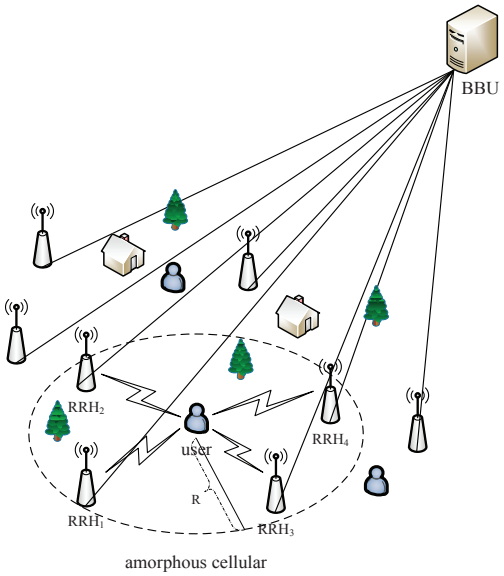

Fig. 1. System model of amorphous cellular in ultra-dense C-RANs uplink, where user connects with the n-nearest RRHs to form its own cell. All RRHs belong to one BBU.

intensity λ in a plane whose radius is R. As it is an user-centric cell, the desired user is located at the origin of the plane. Thus, the number of serving RRHs, n, in this plane is a random variable with probability distribution [24]

$$f_n(n) = \frac{(\lambda\pi R^2)^n}{n!} e^{-\lambda\pi R^2}. \tag{1}$$

We consider uplinks from the user to RRHs. The m-dimensional coordinated received signal in BBU can be written as [25, 26]

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

where $\mathbf{s}$ is n-dimensional transmitted signal vector and the total transmission power is $E\left[\mathbf{s}^{\dagger}\mathbf{s}\right] = P_t$. The entries of the additive noise $\mathbf{n}$ are modeled as zero-mean circular symmetric complex Gaussian with variance σ^2. In this paper, we ignore the shadowing effect, and the channel is mathematically written by the $m \times n$ random matrix $\mathbf{G}$, defined by

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

where $\mathbf{H} \in \mathcal{C}^{m\times n}$ is small-scale fast fading with complex elements $\{h_{i,j}\} \sim \mathcal{CN}(0,1)$, and $\mathbf{D} \in \mathcal{R}^{n\times n}$ accounts for the LSF which is given by

$$\mathbf{D} = \mathrm{diag}\left\{d_i^{-\alpha}\right\}_{i=1}^{n}, \tag{4}$$

where d_i is the distance between the user and the ith nearest RRH, and α stands for the path loss exponent with typical values $\alpha \in [2, 6]$ [21].

In this paper, we consider the scenario in which the user has no channel statement information (CSI) while BBU has perfect CSI. In what follows, the transmitter applies a uniform power allocation scheme across all spatial subchannels, and therefore, the ergodic capacity (in bits/s/Hz) is written as

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

where $\rho = P_t/\sigma^2$ is the average signal-to-noise ratio (SNR)[1], and

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

Equivalently, the capacity can be written in term of nonzero eigenvalues of $\mathbf{W}$. Let $\vec{\lambda}\Big|\,\mathbf{D} = [\lambda_1|\,\mathbf{D}, \ldots, \lambda_q|\,\mathbf{D}]^T$ denote the nonzero eigenvalues of the matrix $\mathbf{W}$, conditioned on $\mathbf{D}$, with $q = \min\{n, m\}$. Then, the ergodic capacity in (5) can alternatively be written as [25]

$$C_{m,n} = \int_{\mathbf{D}}\int_{\lambda}\sum_{i=1}^{q}\log_2\left(1 + \frac{\rho}{m}\lambda_i\right) f_{\lambda_i|\mathbf{D}} f_{\mathbf{D}}\left(d_1 < \cdots < d_n\right) d\vec{\lambda} d\mathbf{D}, \tag{7}$$

where $f_{\lambda_i|\mathbf{D}}$ is the marginal probability distribution function (pdf) of ordered eigenvalue λ_i conditioned on $\mathbf{D}$, $f_{\mathbf{D}}\left(d_1 < \cdots < d_n\right)$ represents the ordered joint pdf of $\mathbf{D}$. We note that the conditional unordered eigenvalue pdf has been investigated in [27, Eq. (95)]. The ordered eigenvalue pdf can be easily derived by considering all the possible matrix patterns of $\mathbf{D}$ which consists of the same entries but in different orders, as given in

$$f_{\lambda_i|\mathbf{D}}\left(\lambda_i\right) = \frac{\Gamma\left(q\right)}{\prod_{i<j}^{p}\left(d_j^{-\alpha} - d_i^{-\alpha}\right)} \sum_{s=p-q+1}^{p} \frac{\lambda_i^{s+q-p-1}}{\Gamma\left(s+q-p\right)} \det\left(\tilde{\mathbf{\Xi}}_s\right). \tag{8}$$

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

$$\left\{\tilde{\mathbf{\Xi}}_s\right\}_{i,j} = \begin{cases} d_i^{-\alpha(j-1)}, & i = 1, \ldots, n, \quad j \neq s, \\ d_i^{-\alpha(p-q+1)} e^{-\frac{\lambda_i}{d_i^{-\alpha}}}, & i = 1, \ldots, n, \quad j = s \end{cases} \tag{9}$$

with $p = \max\{n, m\}$. To remove the condition on $\mathbf{D}$, we need to integrate $d_{i=1,\ldots,n}$ term-

[1]Throughout the paper, we use ρ to represent the ratio of the signal power at transmitter to the noise power at receiver.

by-term. This yields

$$f\left(\lambda_i\right)=\Gamma\left(q\right)\sum_{s=p-q+1}^{p}\frac{\lambda_i^{s+q-p-1}}{\Gamma\left(s+q-p\right)}\int_{\mathbf{D}}\frac{\det\left(\tilde{\mathbf{\Xi}}_s\right)f_{\mathbf{D}}\left(d_1<\cdots<d_n\right)}{\prod_{i<j}^{p}\left(d_j^{-\alpha}-d_i^{-\alpha}\right)}\mathrm{d}\mathbf{D}. \tag{10}$$

This integration is really too complicated, if not impossible, to be evaluated due to the fact that each entry of $\mathbf{D}$ has its own pdf, and with a Vandermonde in the integrand. Thus, we try to analyze the capacity in the high-SNR regime, and provide some useful insights into implications of system behavior.

III. HIGH SNR ANALYSIS

In this section, we derive tight upper bounds of ergodic capacity for amorphous cellular at high SNRs, and characterize the impact of antenna number on the capacity performance. We also provide exact closed-form expressions of 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 RRH, etc.

We first establish the pdf of d_i which is useful in deriving the expression of capacity.

Lemma 1: 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

$$f_{d_i}=2\lambda\pi d_i e^{-\lambda\pi d_i^2}\frac{\left(\lambda\pi d_i^2\right)^{i-1}}{(i-1)!}. \tag{11}$$

Proof: See Appendix A. ■

In contrast to prior works [28], this lemma presents the pdf of the distance between the origin to an arbitrary particle instead of the nearest one. Taking the derivative with the respect to λ in (11), we note the result indicates that, for a certain intensity of RRH, there is most likely $\langle\lambda\pi d_i^2\rangle^2$ RRHs in the area of πd_i^2. Having established *Lemma* 1, we are now ready to derive the upper bounds of the capacity at high-SNR.

A. High-SNR Upper Bound

In the high-SNR regime, without loss any of generality, we evaluate the upper bounds of capacity with an arbitrary number of antennas. Utilizing the Jensen's inequality, the capacity

[2]The operation $\langle\cdot\rangle$ is the rounding function.

with n associated RRHs in (5) is upper bounded by

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

In what follows, the $C_{m,n}$ further upper bounded as

$$C_{m,n}^{U}=\log_2\left(E\left[\det\left(\frac{\rho}{m}\mathbf{W}\right)\right]\right). \tag{13}$$

We investigate the expected determinant of $\mathbf{W}$ in the first place, given in following Lemma.

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

$$E\left[\det\left(\mathbf{W}\right)\right]=\begin{cases}\frac{\Gamma(m+1)}{\Gamma(m-n+1)}\prod\limits_{i=1}^{n}\xi_\lambda\left(i\right), & n\leqslant m,\\ \Gamma\left(m+1\right)\sum\limits_{B_{m,n}}\prod\limits_{i\in B_{m,n}}\xi_\lambda\left(i\right), & n>m,\end{cases} \tag{14}$$

where $B_{m,n}$ is length-m subset of $\{1,2,\ldots,n\}$, and

$$\xi_\lambda\left(i\right)=\frac{\left(\lambda\pi\right)^{\frac{\alpha}{2}}}{\Gamma\left(i\right)}\Gamma'\left(i-\tfrac{\alpha}{2}\right). \tag{15}$$

where

$$\Gamma'\left(t\right)=\begin{cases}\Gamma\left(t,d_{\text{th}}\right), & t\leqslant 0,\\ \Gamma\left(t\right), & t>0.\end{cases} \tag{16}$$

Proof: See Appendix B. ■

Note that 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 which 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. It has to be noted that the value of $\Gamma\left(a\right)$ is infinite (positive or negative) if a is a non-positive integer. Thus, we set a protect distance d_{th} for feasibility of the integration. Having established *Lemma* 2, we are ready to derive the upper bound of the ergodic capacity.

Theorem 1: For dense C-RAN systems where RRHs are stationary point process following PPP, the capacity with n nearest RRHs can be upper bounded at high SNRs by

$$C_{m,n}^{U}=\begin{cases}\log_2\frac{\Gamma(m+1)}{m^n\Gamma(m-n+1)}+\sum\limits_{i=1}^{n}\log_2\left(\rho\xi_\lambda\left(i\right)\right), & n\leqslant m,\\ \log_2\frac{\Gamma(m+1)}{m^m}+\log_2\left(\sum\limits_{B_{m,n}}\prod\limits_{i\in B_{m,n}}\rho\xi_\lambda\left(i\right)\right), & n>m.\end{cases} \tag{17}$$

Moreover, ρ must satisfy the SNR constraint

$$\rho \gg \frac{1}{\xi_\lambda(n)}. \tag{18}$$

Proof: The result in (17) directly follows by applying *Lemma* 2 to the upper bound in (13). However, being conscious that the result only holds for the high-SNR cases, i.e., the SNR of signals at receiver should be considerably larger than 1 under the effect of LSF. For which reason, we evaluate the ρ from the weakest subchannel under LSF, given by [30]

$$\min_i \left\{ E\Big[\frac{\rho}{m} d_i^{-\alpha} \mathbf{h}_i^\dagger \mathbf{h}_i\Big] \right\} = \min_i \left\{ \rho E\left[d_i^{-\alpha}\right] \right\} \gg 1, \tag{19}$$

where $\mathbf{h}_i$ is ith column of $\mathbf{H}$. The expected value of $d_i^{-\alpha}$ can be evaluated as the same as in (15). 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 $\min_i \left\{ E\left[d_i^{-\alpha}\right] \right\} = \xi_\lambda(n)$. ■

It is important to note that the result in (17) holds for arbitrary numbers of antennas. The result implies the joint impact of various factors on the capacity upper bound, e.g. intensity λ, path loss exponent α and number of associated RRHs n. By establishing the constraint of signal SNR, we note that more transmission power is needed for (17) to hold with increasing

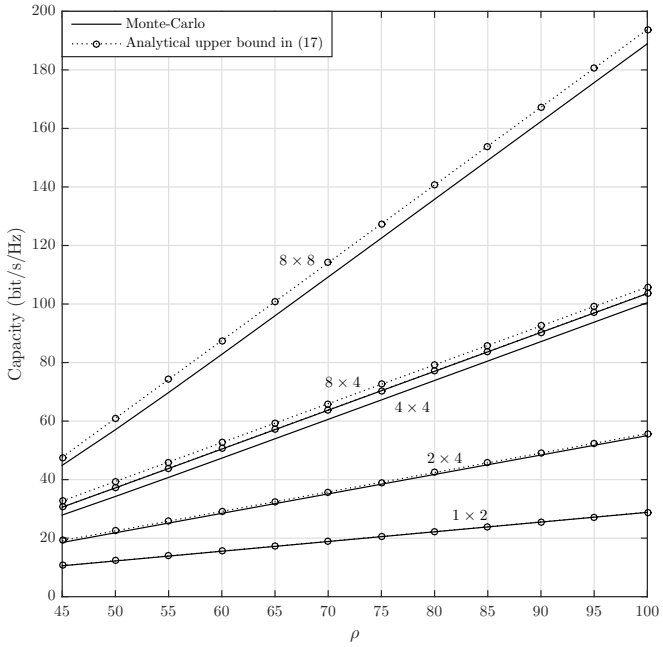

Fig. 2. Comparison of Monte Carlo simulation with upper bounds on ergodic capacity against the ρ. The different antenna configurations are denoted as $m \times n$. Results are shown for $\alpha = 2$, $\lambda = 10^{-3}$ and $\sigma^2 = -70$dBm [29].

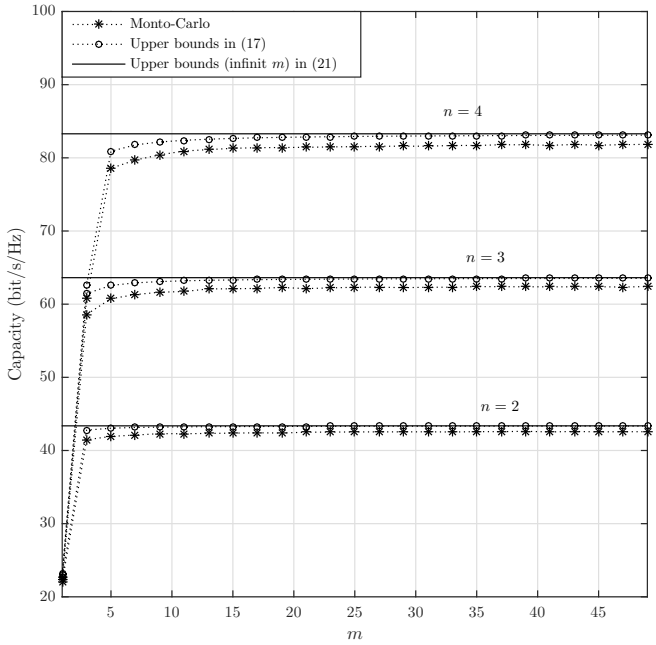

Fig. 3. Comparison of Monte Carlo simulation with upper bounds of the ergodic capacity against m. Results are shown for $n = 2, 3, 4$ with $\alpha = 2$, $\lambda = 10^{-3}$, $\rho = 85$dB and $\sigma^2 = -70$dBm.

n, and it grows non-linearly with joint effects of λ, α and n. Fig. 2 compares the upper bounds $C^U_{m,n}$ with the Monte Carlo simulations versus ρ. The curves are shown under five different antenna configurations, and all the dotted curves are generated using (17). We confirm that the results are rather tight for both cases where transmit antennas number is larger or smaller than the number of associated RRHs. Moreover, we note the very limited capacity enhancement with adding antennas at the maximum number of antenna side, which indicates the capacity performance is mainly subjected to the minimum number of antennas. Additionally, we see that the tightness of bounds looses as minimum number of antennas grows while tightens when maximum number of antennas grows, which is mainly because of the property of the Jensen's inequality.

To gain more insights, we further investigate the upper bounds by examining the following cases.

- For $n < m$, adding one additional RRH, while not altering m, would reduce the expression as

$$C^U_{m,n+1} = C^U_{m,n} + \frac{m-n}{m} \log_2 \left(\rho \xi_\lambda \left(n+1 \right) \right), \tag{20}$$

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 capacity upper bound reduces to the expression as

$$\lim_{m\to\infty} C^{U}_{m,n} \approx \sum_{i=1}^{n} \log_2 \left(\rho \xi_{\lambda}\left(i\right)\right). \tag{21}$$

The result is obtained by using the *Stirling* formula approximation, and the proof is presented in Appendix C. The expression shows the upper limit of the capacity with fix n RRHs, which implies the impact of each associated RRH on the capacity.

Fig. 3 compares the analytical upper bounds based on (17) with Monte-Carlo results for the cases of $n = 2, 3, 4$ against m. We note the fact that the tightness of upper bounds performs worse with more antennas which is confirmed in Fig. 2. Moreover, we note that the lines of upper bounds rise sharply with increasing m when $m < n$, while keep constant when $m > n$. Noting the fact that ergodic capacity is mainly subjected to $\min\{n, m\}$, which can also be recognized by being aware of a slight offset between the upper bound with $m \to \infty$ and $m = n$ in the figure.

Fig. 4 gives the analytical upper bounds based on (17) and Monte-Carlo results for the cases of $m = 2, 3, 4$ against n, in which locations of RRHs are modeled as PPP. We note

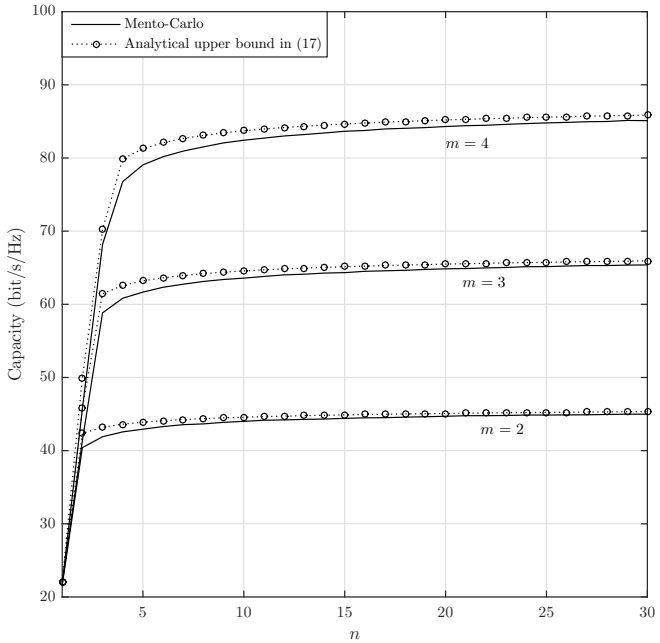

Fig. 4. Comparison of Monte Carlo simulation with upper bounds of ergodic capacity against n. Results are shown for $m = 2, 3, 4$ with $\alpha = 2$, $\lambda = 10^{-3}$, $\rho = 85$dB and $\sigma^2 = -70$dBm.

that the upper bound curves show similar trend as in Fig. 3. Moreover, simulation results show that associating one more RRH gains more but still limited capacity enhancement. Since the number of transmit antenna is generally constant, the result indicates that it is not cost-effective to associate more RRHs, i,e., enhance cellular, when the number of them is already larger than the number of transmit antennas.

B. High-SNR Capacity with $m = n$

We firstly write the conditional ergodic capacity expression in the high-SNR regime as below

$$E\left[C_{m,n}^{\mathrm{H}}\middle|\,\mathbf{D}\right] = E\left[\log_2\det\left(\frac{\rho}{m}\,\mathbf{W}\middle|\,\mathbf{D}\right)\right]. \tag{22}$$

To derive the expression of the capacity, it is necessary to obtain the expected log-determinant of a matrix as $\mathbf{W}$. Note that the conditional expected log-determinant has been investigated in [27, Lemma 4]. Utilizing this result, (22) can be rewritten as

$$E\left[C_{m,n}^{\mathrm{H}}\middle|\,\mathbf{D}\right] = m\log_2\frac{\rho}{m} + \frac{1}{\ln 2}\left(\sum_{k=1}^{q}\psi\left(m-q+k\right) + \frac{\sum\limits_{k=n-q+1}^{n}\det\left(\tilde{\mathbf{\Xi}}_k\right)}{\prod_{i<j}^{n}\left(d_j^{-\alpha}-d_i^{-\alpha}\right)}\right), \tag{23}$$

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

$$\left\{\tilde{\mathbf{\Xi}}_k\right\}_{i,j} = \begin{cases} d_i^{-\alpha(j-1)}, & j\neq k,\\ d_i^{-\alpha(j-1)}\ln d_i, & j=k. \end{cases} \tag{24}$$

Then, by using (11) and (25), the high-SNR capacity can be obtained by integrating with respect to d_i, as given below

$$\begin{aligned} C_{m,n}^{\mathrm{H}} &= m\log_2\frac{\rho}{m} + \frac{\sum_{k=1}^{q}\psi\left(m-q+k\right)}{\ln 2}\\ &+ \frac{\left(2(\lambda\pi)^{(n+1)/2}\right)^n}{\ln 2}\int_0^\infty\cdots\int_0^\infty \frac{\sum\limits_{k=n-q+1}^{n}\det\left(\tilde{\mathbf{\Xi}}_k\right)}{\prod_{i<j}^{n}\left(d_j^{-\alpha}-d_i^{-\alpha}\right)}\prod_{i=1}^{n}\frac{d_i^{2i+1}e^{-\lambda\pi d_i^2}}{(i-1)!}dd_1\cdots dd_n. \end{aligned} \tag{25}$$

The integrand above is extremely difficult due to a Vandermonde involved. In order to obtain more insights, we derive the capacity for the special case of $m = n$, i.e., the number of RRHs equal to number of transmit antennas, and investigate the impact of path loss exponent α on the capacity.

Theorem 2: For dense C-RAN systems where RRHs are stationary point process following

PPP with intensity λ, the capacity at high SNR with $m = n$ is given by

$$C_{m,m}^{\mathrm{H}} = m\log_2\frac{\rho}{m}(\pi\lambda)^{\frac{\alpha}{2}} + \frac{\left(1-\frac{\alpha}{2}\right)}{\ln 2}\sum_{k=1}^{m}\psi\left(k\right) \tag{26}$$

with power constraint given in (18).

Proof: See Appendix D. ■

Our result in *Theorem* 2 gives a mathematical conclusion for the capacity in the regime of high SNR, which is valid for arbitrary antenna as long as $n = m$. By noting the power constraint, more transmission power is needed for the result to hold with more antennas. More importantly, the result shows that the capacity increases logarithmically with intensity of RRH λ with taking LSF into account, and the scaling law of which is proportional to the pass loss exponent α.

Corollary 1: Let $m = n$, the approximate expression of capacity at high SNR is given as

$$C_{m,m}^{\mathrm{H}} \approx m\log_2\frac{\rho}{m}(\pi\lambda)^{\frac{\alpha}{2}} + \left(1-\frac{\alpha}{2}\right)\log_2\Gamma\left(m+1\right). \tag{27}$$

Proof: Utilizing $\psi\left(m\right) = \ln m + O\left(\frac{1}{m}\right)$ in (26) and after some basic operations yields the result. ■

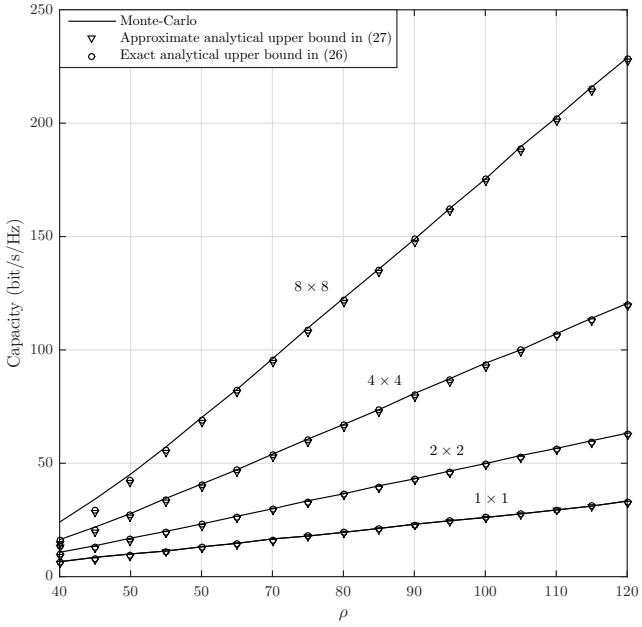

Fig. 5. Comparison of Monte Carlo simulation, approximate analytical and exact analytical expressions of the ergodic capacity in C-RANs. The different antenna configurations are denoted as $m \times n$. Results are shown for $\alpha = 2$, $\lambda = 10^{-3}$ and $\sigma^2 = -70$dBm.

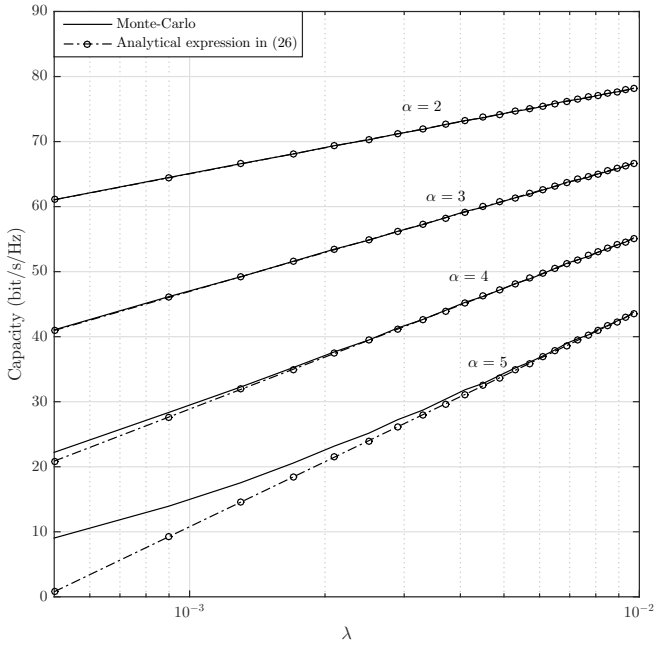

Fig. 6. Comparison of Monte Carlo simulation and exact analytical expressions of the ergodic capacity in (28) against λ. The results are shown for 4×4, SNR $= 80$dB and $\sigma^2 = -70$dBm.

Note that the approximation applied in proof can be really tight when m grows large. The result involves only standard function which can be easily evaluated. Moreover, we note that the capacity decreases with increasing α by taking the partial derivative of α on (27). In order to get more computational insights, we further investigate the approximate expression of capacity for following cases:

- Let $m = n$. Adding 1 antenna at both ends of link, while not altering other parameters, would reduce the expression of capacity to

$$C^{\mathrm{H}}_{m+1,m+1} \approx C^{\mathrm{H}}_{m,m} + m\log_2\frac{m}{m+1} - \frac{\alpha}{2}\log_2\left(\frac{\pi\lambda}{m+1}\right) + \log_2\rho. \tag{28}$$

  The result indicates that the capacity enhancement is a decreasing function of m. Besides, we note that, for a certain ρ and m, the enhancement of capacity grows faster with higher pass loss exponent.

- Let $\alpha = 2$, (26) can be reduced to

$$C^{\mathrm{H}}_{m,m} \approx m\log_2\frac{\rho\pi\lambda}{m}. \tag{29}$$

  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 [31]. It intuitively shows the joint impacts of m, ρ, and λ on the capacity, specifically, non-linearly increases with growing λ and the number of associated RRHs.

Fig. 5 compares the analytical capacity in (26), approximate expression in (27), and Monte Carlo simulated curves with four different antenna configurations. As shown in Fig. 5, we note the approximate expression is rather close to the exact expression, which indicates that the analysis on the approximate expression in (27) are highly solid. Due to the restriction on (18), the analytical formula holds for the cases of high SNR, and more power is needed with increasing number of antennas, as expected.

The joint impact of intensity λ and path loss exponent α on ergodic capacity is illustrated in Fig. 6. We note that the lines representing for the analytical expressions converge to the simulation results in the cases of $\alpha = 2, 3$, while perform tight only in the regime of high λ in the cases of $\alpha = 4, 5$. The reason is, with high α in the low λ cases, which means farther RRHs and more intense signal attenuation, the analytical expression can hardly characterize the 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 signal power satisfies with (13), as expected in (27).

IV. RRH SELECTION AND POWER ALLOCATION

In this section, the optimal number of associated RRHs is derived for the single-user energy efficient transmission in the amorphous cellular. We subsequently propose two scheduling algorithms, namely, UOS and ROS, for multiuser scenarios in the FDD system. Both algorithms are on the basis of our results on capacity. With the help of our analytical upper bound, an amount of calculations have been omitted and replaced by searching in the index derived from offline simulations. Note that the power consumed by signal transmission and reception are taken into account.

A. Energy Efficiency

We consider a FDD dense C-RAN system where each single-antenna RRH is scheduled a single user[3] with a diverse frequency band in a time resource. User establishes its own cell

[3]Actually, the assumption can be generalized to the multiuser scenarios when orthogonal multiple access approach. For convenience, we simplify the assumption to show the efficiency of our scheme and left the joint frequency and RRH scheduling to future work.

by connecting to neighboring RRHs, and we assume fully coordination among them in the cell. Without loss of generality, we suppose that the locations of users are also deployed 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 a signal transmission and reception, and the EE can be defined as the ratio of the sum rate to the sum of the consumed power, described as

$$\eta_{\text{mu}} = \frac{\sum_k C_{m_k,n_k}}{\sum_k n_k P_r + \sum_k P_{t_k}}, \tag{30}$$

where C_{m_k,n_k} represents the uplink achievable rate 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 represents the power for reception consumed by each RRH.

In contract to prior works, n_k varies according to user requirements. In general, exhaustive search for the optimal result requires extremely high complexity, which could be impossible 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, and save the results for future search.

It is straightforward to see that the expression of EE for a single user ought to be described

TABLE I
SYSTEM PARAMETERS.

Path loss exponent α	2
User antenna m	4
Maximum number of RRH $n_{\max}$	10
RRH intensity λ	10^{-3}
User intensity λ_u	10^{-4}
Protect distance d_{th}	1m
Signal processing power P_r	0.1W
Minimum transmit power $P_{\min}$	5dBm
Maximum transmit power $P_{\max}$	33dBm
Thermal noise	-70dBm
Capacity margin Δ_r	3bit/s/Hz

TABLE II
JOINT OPTIMAL CONFIGURATION OF (n, P_t)

r	$m=2$	$m=4$	r	$m=2$	$m=4$	r	$m=2$	$m=4$	r	$m=2$	$m=4$
25	$(2,9)$	$(2,7)$	34	$(3,20)$	$(3,6)$	43	$(7,31)$	$(4,7)$	52		$(4,14)$
26	$(2,10)$	$(2,8)$	35	$(3,22)$	$(3,7)$	44	$(7,33)$	$(4,8)$	54		$(4,16)$
27	$(2,12)$	$(2,10)$	36	$(3,23)$	$(3,8)$	45		$(4,9)$	56		$(4,17)$
28	$(2,13)$	$(2,11)$	37	$(4,24)$	$(3,9)$	46		$(4,10)$	58		$(4,17)$
29	$(2,15)$	$(3,5)$	38	$(4,25)$	$(3,10)$	47		$(4,11)$	60		$(4,20)$
30	$(2,16)$	$(3,5)$	39	$(4,26)$	$(3,11)$	48		$(4,11)$	62		$(4,20)$
31	$(2,18)$	$(3,5)$	40	$(5,28)$	$(3,12)$	49		$(4,12)$	64		$(5,21)$
32	$(2,19)$	$(3,5)$	41	$(5,29)$	$(3,13)$	50		$(4,13)$	66		$(5,23)$
33	$(2,21)$	$(3,5)$	42	$(7,30)$	$(3,14)$	51		$(4,14)$	68		$(5,23)$

as

$$\eta_{\text{su}} = \frac{C_{m,n}}{nP_r + P_t}. \tag{31}$$

We here use $C^U_{m,n}$ in (17) to replace $C_{m,n}$ in (31) for our offline simulations. In what follows, the approximate EE can be expressed as

$$\hat{\eta}_{\text{su}} = \frac{C^U_{m,n}}{nP_r + P_t}. \tag{32}$$

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

$$\max_{n,P_t} \ \hat{\eta}_{\text{su}} \tag{33}$$

where $1 \leqslant n \leqslant n_{\max}$ and $P_{\min} \leqslant P_t \leqslant P_{\max}$. The optimal solution of n for the above problem can be intuitively solved by 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 is shown in Table II. Based on Table II, the optimal size of single-user cell can be intuitively determined by the number of associated RRHs according to user's demand.

B. UOS and ROS

We propose two effective schemes, namely, UOS and ROS, to form user's amorphous cell on the basis of offline EE results. In UOS, we choose user to be our optimized object, specifically, the algorithm selects the required number of RRH with the strongest link for

each user from all the remaining candidates. In ROS, the scheme establishes the cell by optimizing the RRH in the system, which is arranged to the corresponding user who has the strongest link with. Note that both algorithms contain three steps and both of Step 1 and Step 3 in two algorithms are identical.

We first introduce into detail to cover the steps in UOS, and then provide the differences of ROS compared with UPS thoroughly.

In Step 1 of UOS, we identify the number of users N in the system and the optimal number of associated RRH n_k for the kth user. We use r_{th}^k (in bps/Hz) to represent the QoS of user k, and add a margin of the spectral efficiency Δ_r to each r_{th}^k to ensure the algorithm can meet the user's demand. Note that n_k can be determined by offline simulation since the density of the RRHs λ is determined. Thus, the optimal associate number of antennas of users can be obtained by indexing Table II according to their demands while avoiding a significant amount of computing to release the complexity.

In Step 2 of UOS, we use $\mathcal{M}$ to represent the set of the whole RRHs in the area; $\mathcal{N}$ denotes the set of users in the system; $\mathcal{R}_k$ is 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

Algorithm 1 UOS

Initialization:
Set $\mathcal{N} = \{1, \ldots, N\}$, $\mathcal{M}$, $r_{th}^{k=1,\ldots,N}$, $m_{k=1,\ldots,N}$, $n_{k=1,\ldots,N} = 0$,
$\mathcal{R}_{k=1,\ldots,N} = \emptyset$, $F_k = 0$, $l_k = 0$
Step 1: Determining the $n_{k=1,\ldots,N}$ and $P_{t_{k=1,\ldots,N}}$.
for $k = 1$ to N
$r^k = r_{\text{th}}^k + \Delta_r$,
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 strong uplinks for each user.
while if $F_{k=1,\ldots,N} = 0$ **do**
for $k = 1$ to N
if $F_k = 0$ **then**
$n^* = \arg\max_{i \in \mathcal{M}} \left\{ d_{ik}^{-\alpha} \right\}$, $\mathcal{R}_k = \mathcal{R}_k \cup n^*$, $\mathcal{M} = \mathcal{M} / \{n^*\}$, $l_k = l_k + 1$
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

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

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 till all n_k is satisfied. This step ensures that the 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; $C_{m_k,n_k}\left(\mathcal{R}_k,P_{t_k}\right)$ represents the capacity of 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 RRH for user, we search for the strongest link from all the possible links of the remaining user and the remaining RRH, and then schedule the RRH to its corresponding user till all the users meet n_k. In contrast to UOS, RRH is regarded as our optimized object, specifically, the selected RRHs in ROS are guaranteed to serve their optimal users. We note that Step 1 and Step 3 are extremely convenient, thus, the complexity of both algorithms is dominated by Step 2. As N and M grow large, the

TABLE III
WORST-CASE COMPLEXITY COMPARISON.

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

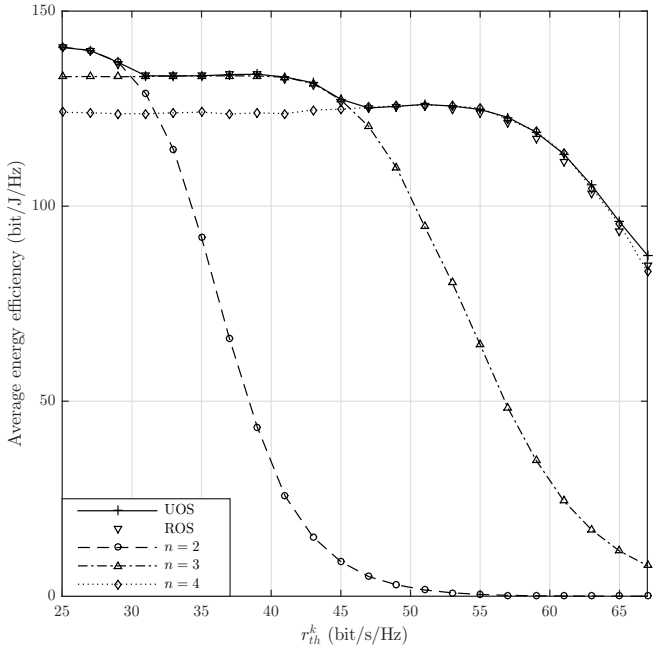

Fig. 7. Energy efficiency for UPS and RPS versus the different QoS with consideration of both power consumption by signal transmission and reception. We choose $\alpha = 2$, $\lambda_R = 10^{-3}$ and $\lambda_u = 10^{-4}$.

worst-case complexity of Step 2 in UOS and ROS are compared in Table III. As seen, the complexity of ROS is higher than RPS, 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 simulation parameters shown in Table I is illustrated in Fig. 7. We here, for convenience, set $m_{k=1,\ldots,N} = 4$ and identical user requirement. We also provide the results of fixed number of associated RRHs with $n_{k=1,\ldots,N} = 2, 3, 4$ as control groups. We first note that the EE with fewer number of RRHs performs better in low user requirements (i.e., $n = 2$ below 30bit/s/Hz and $n = 3$ below 45bit/s/Hz) since the less consumption of transmission power and fewer associated RRHs. Furthermore, the lines representing the results for the fixed n_k decrease sharply with r growing large due to the exponential increase of transmission power. In contrast to them, UOS and ROS, applying the results of our capacity analysis in Section III, show the significant advantages in EE performance, i.e., both of which remain the best average EE among all lines. There is no significant difference between UOS and ROS in the regime of low user requirement. However, when r_{th} grows large, the line of UOS wins by a narrow margin over ROS, which implies that selecting the optimal RRHs for user is the better strategy compared with selecting the

optimal users for RRH for user-centric system when take the complexity into account.

V. CONCLUSION

In this paper, we investigate amorphous cellular which is based on user-centric manner for ultra-dense C-RANs. The upper bound expression of uplink capacity at high SNR has been presented, as well as exact expression when the number of user's antennas equal to the number of associated RRHs. Based on the analytical results, we have characterized the impact of path loss exponent, RRH intensity and antenna configuration on ergodic capacity respectively. Moreover, the optimal number of associate RRHs for single user is offline derived when transmission and reception power are taken into account. These results are subsequently used in two scheduling algorithms, namely UOS and ROS, to approach energy efficient transmission for multiuser FDD systems, which shows the significant advantage compared with the fixed number of associate RRHs strategy.

APPENDIX A
PROOF OF LEMMA 1

As shown in (4), d_i represents the distance between the user and the ith nearest RRH. The cumulative distribution function of d_i is described as

$$F_{d_i}(d) = \Pr(d_i \leq d) = 1 - \Pr(d_i > d). \tag{34}$$

Since there exists exactly $(i-1)$ RRHs in the circle plane whose radius is d_i, utilizing (1), (34) can be derived as

$$F_{d_i}(d) = 1 - \frac{(\lambda\pi d_i^2)^{i-1}}{(i-1)!} e^{-\lambda\pi d_i^2}. \tag{35}$$

The pdf of d_i can be directly obtained by taking the derivative of F_{d_i}.

APPENDIX B
PROOF OF LEMMA 2

To proof this lemma, it is convenient to give separate treatments for two cases: $n \leq m$ and $n > m$.

1) $n \leq m$ Case: For this case, the expected determinant of $\mathbf{W}$, conditioned on $\mathbf{D}$, is given

by

$$\begin{aligned} E\left[\det\left(\mathbf{W}\right)|\,\mathbf{D}\right] &= E\left[\det\left(\mathbf{D}\mathbf{H}^{\dagger}\mathbf{H}\right)\big|\,\mathbf{D}\right] \\ &= E\left[\det\left(\mathbf{D}\right)\det\left(\mathbf{H}^{\dagger}\mathbf{H}\right)\big|\,\mathbf{D}\right] \\ &\overset{(a)}{=} \det\left(\mathbf{D}\right) E\left[\det\left(\mathbf{H}^{\dagger}\mathbf{H}\right)\right] \end{aligned} \tag{36}$$

where (a) is obtained due to the independence of $\mathbf{D}$ and $\mathbf{H}$. Applying the result in [32, Eq. (A.7.1)], (37) can be further represented as

$$E\left[\det\left(\mathbf{W}\right)|\,\mathbf{D}\right] = \det\left(\mathbf{D}\right)\frac{\Gamma\left(m+1\right)}{\Gamma\left(m-n+1\right)}. \tag{37}$$

Evaluate the integral of $\mathbf{D}$ utilizing *Lemma* 1, and the closed-form expression of expected determinant of $\mathbf{W}$ can be further simplified as

$$\begin{aligned} E\left[\det\left(\mathbf{W}\right)\right] &\overset{(b)}{=} \frac{\Gamma\left(m+1\right)}{\Gamma\left(m-n+1\right)}\int_{\mathbf{D}}\prod_{i=1}^{n} d_i f_{\mathbf{D}}\mathrm{d}\mathbf{D} \\ &\overset{(c)}{=} \frac{\Gamma\left(m+1\right)}{\Gamma\left(m-n+1\right)}\prod_{i=1}^{n}\frac{\left(\lambda\pi\right)^{\frac{\alpha}{2}}}{\Gamma\left(i\right)}\Gamma\left(i-\frac{\alpha}{2}\right), \end{aligned} \tag{38}$$

where (b) is derived by noting that $\mathbf{D}$ is diagonal matrix whose entries are independent to each other, and (c) is obtained from [22, Eq. (6.1.5)].

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

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

Note that $\left(\cdot\right)_{B_{m,n}}$ is a $n\times n$ matrix. Thus, utilize the commutativity of matrix multiplication and change the order of the matrices in (39), we can further obtain

$$\begin{aligned} E\left[\det\left(\mathbf{W}\right)|\,\mathbf{D}\right] &= E\left[\det\left(\mathbf{H}_{B_{m,n}}\mathbf{H}_{B_{m,n}}^{\dagger}\right)\right]\sum_{B_{m,n}}\det\left(\mathbf{D}\right)_{B_{m,n}} \\ &= \Gamma\left(m+1\right)\sum_{B_{m,n}}\det\left(\mathbf{D}\right)_{B_{m,n}}, \end{aligned} \tag{40}$$

We then evaluate the integral of (40) as in (38) to yield the result.

APPENDIX C
PROOF OF (21)

Utilizing the *Stirling* formula

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

the upper bound in (17) with $n \leqslant m$ can be approximately rewritten as

$$\begin{aligned} C_{m,k}^{U} &\approx \log_2 \frac{\sqrt{2m\pi}\left(\frac{m}{e}\right)^{m}}{m^{k}\sqrt{2\left(m-k\right)\pi}\left(\frac{m-k}{e}\right)^{m-k}} + \sum_{i=1}^{k}\log_2\left(\rho\xi_{\lambda}\left(i\right)\right) \\ &= \log_2 \frac{1}{e^{k}}\left(\frac{m}{m-k}\right)^{m-k+\frac{1}{2}} + \sum_{i=1}^{k}\log_2\left(\rho\xi_{\lambda}\left(i\right)\right). \end{aligned} \tag{42}$$

According to the definition of natural constant e, we have

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

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

APPENDIX D
PROOF OF THEOREM 2

We firstly rewrite the conditional ergodic capacity expression in the high-SNR regime

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

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

$$\begin{aligned} \sum_{k=1}^{p}\det\left(\tilde{\boldsymbol{\Xi}}_{k}\right) &\overset{(d)}{=} \sum_{k=1}^{p}\sum_{\sigma}\mathrm{sgn}\left(\sigma\right)\prod_{i=1}^{p}d_{\sigma(i)}^{-\alpha(i-1)}\ln d_{\sigma(k)}^{-\alpha} \\ &= \sum_{\sigma}\mathrm{sgn}\left(\sigma\right)\prod_{i=1}^{p}d_{\sigma(i)}^{-\alpha(i-1)}\sum_{k=1}^{p}\ln d_{\sigma(k)}^{-\alpha} \\ &= -\alpha\sum_{i=1}^{p}\ln d_i\prod\nolimits_{i<j}^{p}\left(d_j^{-\alpha}-d_i^{-\alpha}\right). \end{aligned} \tag{45}$$

where (d) is derived by the Leibniz formula [33], and the second summation is over all permutations $\sigma=\{\sigma\left(1\right),\ldots,\sigma\left(p\right)\}$ of the set $\{1,\ldots,p\}$, with $\mathrm{sgn}\left(\sigma\right)$ representing the sign

of the permutation. By substituting (45) into (44), we can rewrite the expression as

$$E\left[C_{m,n}^{\mathrm{H}}\middle|\mathbf{D}\right]=m\log_2\frac{\rho}{m}+\frac{1}{\ln 2}\left(\sum_{k=1}^{m}\psi\left(k\right)-\alpha\sum_{i=1}^{m}\ln d_i\right). \tag{46}$$

Using *Lemma* 1 and (46), the ergodic capacity in high SNR regime can be simplified as

$$\begin{aligned}C_{m,m}^{\mathrm{H}}&=\int_{\mathbf{D}}E\left[\log_2\det\left(\frac{\rho}{m}\mathbf{W}\right)\middle|\mathbf{D}\right]f_{\mathbf{D}}\mathrm{d}\mathbf{D}\\&=m\log_2\frac{\rho}{m}+\frac{\sum_{k=1}^{m}\psi\left(k\right)}{\ln 2}-\frac{\alpha}{\ln 2}\int_{\mathbf{D}}\sum_{i=1}^{m}\ln d_i f_{\mathbf{D}}\mathrm{d}\mathbf{D}.\end{aligned} \tag{47}$$

The integral in (47) can be further evaluated as

$$\begin{aligned}\int_{\mathbf{D}}\sum_{i=1}^{m}\ln d_i f_{\mathbf{D}}\mathrm{d}\mathbf{D}&=\sum_{i=1}^{m}\int_0^{\infty}\ln d_i f_{\mathbf{D}}(d_i)\mathrm{d}d_i\\&=\sum_{i=1}^{m}\frac{2(\pi\lambda)^i}{\Gamma\left(i\right)}\int_0^{\infty}d_i^{2i-1}e^{-\lambda\pi d_i^2}\ln d_i\mathrm{d}d_i\\&\overset{(e)}{=}\frac{1}{2}\left(\kappa\left(m\right)-mc-m\ln\pi\lambda\right),\end{aligned} \tag{48}$$

where (e) is obtained from [23, Eq. (4.331.1)] and [23, Eq. (4.352.2)], c is the Euler constant, and

$$\kappa\left(m\right)=\begin{cases}0, & m=1,\\ \sum\limits_{i=2}^{m}\sum\limits_{j=1}^{i-1}j^{-1}, & m>1.\end{cases} \tag{49}$$

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

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

Note that a property of digamma function is presented as [30, Eq. (2.14)]

$$\psi\left(k\right)=\psi\left(1\right)+\sum_{j=1}^{k-1}j^{-1}. \tag{51}$$

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

[3] A. G. Gotsis and A. Alexiou, "On coordinating ultra-dense wireless access networks: Optimization modeling, algorithms and insights." [Online]. Available: http://arxiv.org/abs/1312.1577

[4] T. Ihalainen, et al., "Flexible scalable solutions for dense small cell networks," in *Wireless World Research Forum (WWRF), Oulu, Finland*, 2013.

[5] P. Demestichas, A. Georgakopoulos, D. Karvounas, K. Tsagkaris, V. Stavroulaki, J. Lu, C. Xiong, and J. Yao, "5G on the horizon: Key challenges for the radio-access network," *IEEE Vehicular Technology Magazine*, 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] Z. Ding and H. V. Poor, "The use of spatially random base stations in cloud radio access networks," *IEEE Signal Processing Letters*, vol. 20, no. 11, pp. 1138–1141, Nov. 2013.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] Y. Zhang and Y. Zhang, "User-centric virtual cell design for cloud radio access networks," in *IEEE Workshop on Sign. Proc. Adv. in Wireless Comm.*, Jun. 2014, pp. 249–253.

[14] J. G. Andrews, F. Baccelliand, 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.

[15] 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.

[16] 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.

[17] 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, 2014.

[18] 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.

[19] 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.

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

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

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

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

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

[25] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European transactions on telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[26] 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.

[27] 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.

[28] G. Geraci, H. S. Dhillon, J.G. Andrews, J. Yuan, and I. B. Collings, "A new model for physical layer security in cellular networks," in *IEEE ICC 2014*, Jun. 2014, pp. 2147–2152.

[29] 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. PP, no. 99, pp. 1–1, Mar. 2016.

[30] A. M. Tulino and V. Sergio, *Random matrix theory and wireless communications*. Now Publishers, 2004.

[31] 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.

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

[33] L. N. Trefethen and D. Bau III, *Numerical linear algebra*. Siam, 1997, vol. 50.