Complementary Networking for C-RAN: Spectrum Efficiency, Delay and System Cost

Yi Zhong, *Member, IEEE*, Tony Q. S. Quek, *Senior Member, IEEE*, and Wenyi Zhang, *Senior Member, IEEE*

***Abstract*— Cloud radio access networks (C-RANs) architecture is a cost-efficient and energy-efficient solution for increasing the capacity of the cellular network. In order to adapt the traffic and reduce the system cost, we propose complementary networking architecture for the C-RAN, which takes advantage of both the C-RAN and traditional base stations (BSs).We propose combining the Neyman–Scott cluster process and the Poisson hole process to model the architecture; then, the interference in the large-scale network with cooperation is well characterized through the Poisson point process approximation. By mixing the two multiple association mechanisms, i.e., the RRH-selection mode and the cooperation mode, we explore the tradeoffs between the area spectrum efficiency (ASE), the mean delay, and the system cost both analytically and numerically. The results reveal that the mean delay is negative correlated to the ASE for small ASE and is reversed for large ASE, while the cost is always positively correlated to the ASE. The cooperation mode increases the useful signal as well as the interference, which becomes dominant when the radius of the RRH cluster is large. The proportion of sub-frames operating in the two modes can be configured to tradeoff the ASE, the mean delay, and the cost.**

***Index Terms*— Complementary networking architecture, cooperation, C-RAN, RRH-selection, Neyman-Scott cluster process, Poisson hole process.**

I. Introduction

A. Motivation

AS THE internet content becomes data-rich, the traffic in future mobile networks is expected to be exceptionally greater. In order to meet the exponential growth of traffic [1], radio access networks need to increase the capacity and coverage. The most promising approach is to make the cells smaller, thus increasing the spatial reuse of spectrum resource. However, deploying more BSs inevitably raise the capital and operational expenses. To address this problem, the cloud-based radio access network is proposed as a cost-efficient and energy-efficient solution [2]–[7]. In the C-RAN architecture, functionality of the base band units (BBUs) is moved to a central data center. The radio units, also called remote radio heads (RRHs), are connected to the central data center via optical fibers.

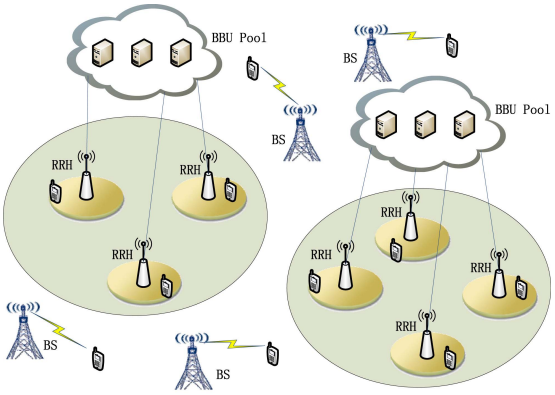

Fig. 1. Typical scenario for the complementary deployment of the C-RAN and the traditional BSs.

The benefits of C-RAN mainly include three aspects. First, integrating the BBUs into data centers greatly reduces the number of BSs, thus reducing energy consumption caused by the air conditioning and other site support equipment. Meanwhile, deploying high density of RRHs reduces the distance between the RRHs and the users, thus saving the transmit power. The energy consumption can be further reduced by selectively turning off the idle virtual BBUs. Second, since the BBUs are separated from the RRHs, only the installation of the auxiliary antenna feeder systems is required at the RRHs, which greatly reduces the cost. Third, the signalling, data and channel state information are easily shared among virtual BSs in the BBU pool, thus making the implement of joint processing quite easy in C-RAN.

Because of propagation and implementation constrains, the size of clouds is limited. Therefore, RRHs are grouped into clusters according to geographical locations. However, for regions with less traffic, deploying C-RAN is unnecessary; thus, we propose to deploy traditional BSs instead of C-RAN in those regions. Therefore, we assume a network architecture (see Figure 1) as a hybrid of C-RAN and traditional BSs, named the *complementary networking*, in which C-RAN is

Manuscript received May 9, 2016; revised December 3, 2016 and March 21, 2017; accepted April 21, 2017. Date of publication May 10, 2017; date of current version July 10, 2017. This work was supported in part by the National Natural Science Foundation of China under Grant 61379003, in part by the Fundamental Research Funds for the Central Universities under Grant WK3500000003, in part by the SUTD-ZJU Research Collaboration under Grant SUTD-ZJU/RES/01/2014, in part by the Zhejiang Provincial Public Technology Research of China under Grant 2016C31063, in part by the MOE ARF Tier 2 under Grant MOE2014-T2-2-002, in part by the NSFC Major International Joint Research Project under Grant 61210002, in part by the Fundamental Research Funds for the Central Universities under Grant 2015XJGH011, in part by the EU FP7-PEOPLE-IRSES, project acronym CROWN under Grant 610524, and in part by the China International Joint Research Center of Green Communications and Networking under Grant 2015B01008. The associate editor coordinating the review of this paper and approving it for publication was G. C. Alexandropoulos. *(Corresponding Author: Yi Zhong.)*

Y. Zhong is with the School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: yzhong@hust.edu.cn).

T. Q. S. Quek is with the Singapore University of Technology and Design, Singapore 487372 and also with the Department of Electronics Engineering, Kyung Hee University, Yongin-si 17104, South Korea (e-mail: tonyquek@sutd.edu.sg).

W. Zhang is with the Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei 230027, China (e-mail: wenyizha@ustc.edu.cn).

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

Digital Object Identifier 10.1109/TWC.2017.2701359

deployed in the hot spots to provide high rate access, while the BSs are deployed outside the RRH cluster region to reduce the cross-tier interference.

In practice, many different kinds of configurations and transmission modes can be applied in the C-RAN architecture. In this work, we consider two most often mentioned transmission modes, i.e., the RRH-selection mode and the cooperation mode. Although the BBUs are separated from the RRHs and concentrated at the data centers, each RRH is associated to one virtual BBU. If all virtual BBU-RRH pairs in the same cluster transmit different data flows to different users through orthogonal frequency bands, i.e., in the RRH-selection mode, the energy and compute resource (e.g., CPU cores, DSPs) consumption may not be optimal since all virtual BBUs should be active and generate frames, resulting in more cost in the BBU pool. Optionally, multiple RRHs can cooperatively transmit the same frame generated by one single virtual BBU to one user in the cooperation mode. This greatly reduces the number of active virtual BBUs, thus decreasing the energy and compute resource consumption. However, cooperation also leads to more interference than independent transmissions of flows through orthogonal frequency bands, since all RRHs are transmitting on the entire frequency band. Thus, it is still uncertain how the C-RAN architecture benefits from the cooperation, which will be explored in this work.

B. Contributions

The main contributions of this work are as follows

- We propose the complementary networking, which is a hybrid of C-RAN and traditional BSs, to combine the advantages of both kinds of networks. We evaluate the ASE, the mean delay and the cost of such network both analytically and numerically through the Poisson point process approximation. The effect of deployment parameters, such as the size of the C-RAN region, the density of the C-RAN and the density of the traditional BSs, are explored. We further simplify the results into closed form in special cases to gain insight. The results reveal the relationship between different system parameters and provide a useful guideline for the design of the practical C-RAN architecture.
- We propose to mix the two operation modes for the C-RAN architecture, i.e., the *RRH-selection mode* and the *cooperation mode*. The ratio of the number of sub-frames in the RRH-selection mode and that in the cooperation mode can be configured to further optimize the ASE, the mean delay and the cost of the network.
- In terms of the analytical approaches, we propose to combine the Neyman-Scott cluster process [8], [9] with the Poisson hole process [10] to model the complementary networking architecture, in which the deployment of the C-RAN and that of the traditional BSs are correlated. The nearest neighbor distribution is derived for the C-RAN users and is approximated for the users served by traditional BSs.
- Most of the existing works related to analyzing the cooperation in large scale networks, such as [11], considered a Poisson field of interference and ignored the effect of cooperation on the interference. By applying the proposed network model, the interference for both the RRH-selection mode and the cooperation mode can be well characterized.

C. Related Works

In [12], the advantages and challenges of the heterogeneous C-RAN (H-CRAN), in which the BBU pool is interfaced with BSs to mitigate the cross-tier interference between RRHs and BSs through centralized cloud computing-based cooperative processing, are summarized and analyzed. In [13] and [14], the authors considered a group of single-antenna BSs distributed in a disc and analyzed the performance of the distributed beamforming and the best BS selection while the effect of the interference is ignored. In [15], the deployment cost for the C-RAN network is modelled and analyzed; however, the cost for the energy which occupies a large proportion of the total cost has not been taken into consideration. In [16] and [17], the power consumption models for the C-RAN architecture, which reflect the dynamic allocation of the centralized BBU resources to the RRHs, are proposed and detailed. [6] elaborates that the virtual BBUs could be selectively turned off (or be taken to a lower power state) to achieve a much higher utilization of processing resources and lower energy consumption. The works related to the cooperation among clustered BSs could be found in [18]–[21].

Using tools from point process theory to model the large-scale network can be found in [22]–[27]. In [8] and [9], the Poisson cluster process was used to model and analyze the performance of the wireless networks. In [10], the Poisson hole process was used to model the active cognitive users in the Poisson cognitive networks. The analysis of delay in the Poisson networks when the interfering transmitters are full buffered can be found in [28]–[30]. The works in [21], [31], [32] analyzed the cooperation of multiple BSs based on the user-centric clustering while we focus on the location based clustering. The work in [33] considered similar topic of cooperation in RRH-selection mode and cooperation mode with the regular grid model.

The remaining part of this paper is organized as follows. Section II describes the network model and the transmission model. Section III, Section IV and Section V then establish the analytical results of the ASE, the mean delay and the system cost of the network respectively. Section III analyzed the synchronous transmission. Section VII gives the numerical results, and Section VIII concludes the paper.

II. System Model

A. Spatial Distribution Model

We use the Neyman-Scott cluster process to model the architecture of C-RAN (see Figure 2). Specially, we model the spatial distribution of data centers as the parent process of the Neyman-Scott cluster process, denoted by Φ_p, which is a Poisson point process (PPP) of intensity λ_p. For each cluster center $y \in \Phi_p$, we assume that there are N daughter points, denoted by Φ^y, which are distributed uniformly in the

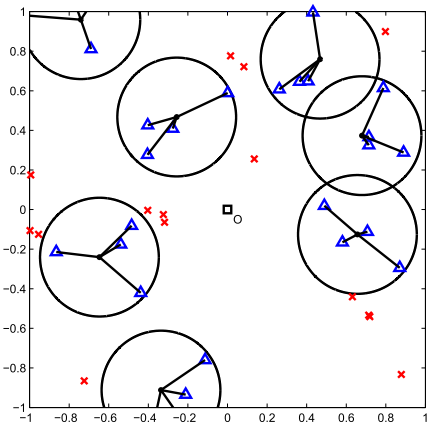

Fig. 2. The network model. The triangles are the RRHs, and the x's are the traditional BSs. The circles are the cluster regions of the C-RAN with radius R. The black straight lines denote the fronthaul between the RRHs and the BBU pools.

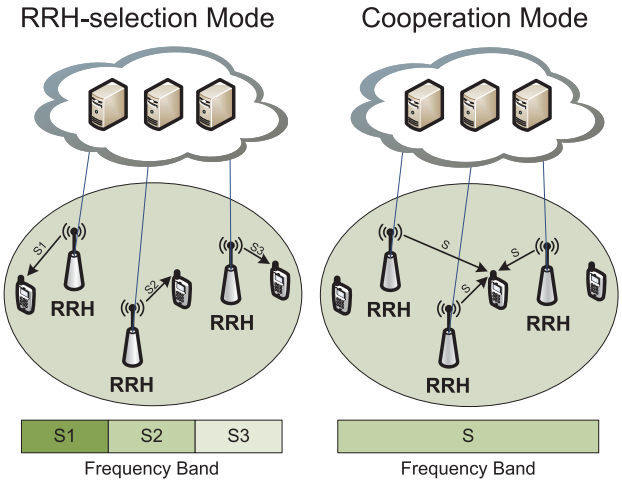

Fig. 3. Two operation modes for C-RAN: the RRH-selection mode and the cooperation mode.

circular covered area of radius R around the cluster center y. The daughter points are used to model the RRHs connected to the data center of the said cluster, which constitute a Neyman-Scott cluster process Φ_{CR}, denoted by

$$\Phi_{\mathrm{CR}} \triangleq \bigcup_{y\in\Phi_p} \Phi^y. \tag{1}$$

Therefore, the intensity of the cluster process Φ_{CR} is $N\lambda_p$. Each cluster is a unit for cooperation and resource allocation. We assume that the traditional BSs are deployed uniformly as a PPP with intensity λ_b outside the RRH cluster region of the C-RAN architecture, denoted by

$$\Xi \triangleq \bigcup_{y\in\Phi_p} B(y, R), \tag{2}$$

where $B(y, R)$ is the ball centered at y with radius R. Let Φ_b be a PPP of intensity λ_b in the whole plane; then, the locations of the traditional BSs construe a Poisson hole process as

$$\Phi_{\mathrm{BS}} \triangleq \{x \in \Phi_b : x \notin \Xi\} = \Phi_b \setminus \Xi. \tag{3}$$

Since the probability of a point of Φ_b being retained is the probability that there is no point within distance R, the intensity of the traditional BSs is $\lambda_b \exp(-\lambda_p \pi R^2)$.

B. Transmission Model

We adopt a standard path loss model, $l(x) = |x|^{-\alpha}$, with $\alpha > 2$. Regarding fading, we assume that all links experience Rayleigh fading with unit mean. Let the entire bandwidth of the spectrum to be fixed to W. The transmit power of RRHs is P_{s}, and that of traditional BSs is P_{m}. We consider the interference limited case and ignore the background thermal noise.

For C-RAN, we consider two operation modes, i.e., the *RRH-selection mode* and the *cooperation mode* (see Figure 3). In the RRH-selection mode, each RRH cluster reuses entire spectrum and divides it into N sub-bands. The frequency splitting and sub-bands allocations are also elaborated and justified in. Each RRH constructs one RRH-user link and occupies one sub-band for transmission; thus at most N users can be served simultaneously by a RRH cluster in [2] and [6]. All virtual BBUs should be active to generate different frames for different RRHs. In the RRH-selection mode, the C-RAN users will associate to the nearest RRH in the serving RRH cluster. In the cooperation mode, only one virtual BBU needs to be active to generate a common frame, which is then delivered to all N RRHs and transmitted cooperatively to the same user. The entire bandwidth is used for cooperative transmission, and all signals from the RRHs are combined by adding up the amplitudes non-coherently. Since the cooperation among RRHs can only serve one user, we assume that multiple users are served through round-robin, i.e., N users are served by time-sharing with equal time proportion. Thus, each user occupies only $1/N$ of the transmission duration, and the rate for each user will be scaled down by $1/N$.

The RRH-selection mode is capable to reduce the interference, while the cooperation mode improves the received signal and consumes less energy. Since the two modes cannot coexist, we assume that they are interwove in the time domain. For example, in TDD system, the multiplexed time scale can be one or several frames. When the multiplexed time scale is one frame which typically spans 10 ms, each frame can be divided into several sub-frames. In each sub-frame, the BBU-RRH mappings can be configured either in RRH-selection mode or in cooperation mode. The ratio of the number of sub-frames within RRH-selection mode to that within cooperation mode is an adjustable parameter determined by the traffic. The structure of a frame could be reconfigured in the granularity of one or several frames depending on the dynamic of traffic.

We assume traditional BSs use the same spectrum as C-RAN. Each BS divides entire bandwidth into N sub-bands, and each BS-user link occupies one sub-band. Thus, at most N users are served simultaneously by each BS. Users served by BSs always access their nearest BSs.

There are two types of users, i.e., the C-RAN users within the RRH cluster region Ξ and the users outside of Ξ served by traditional BSs. We consider a typical user at the origin.

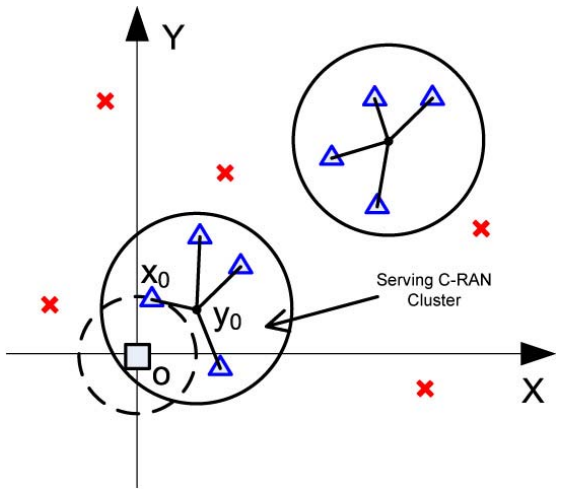

Fig. 4. Illustration of the typical C-RAN user at the origin served by the RRHs in the serving RRH cluster centered at y_0. The triangles are the RRHs, and x_0 is the nearest RRH in the serving RRH cluster.

When the typical user is served by C-RAN, there will be a serving RRH cluster, whose cluster center is within the distance R to the said user. The manner by which the typical user is served may switch between two operation modes. Without loss of generality, we consider the time scale of one frame, which is divided into several sub-frames. Each sub-frame is configured into RRH-selection mode or cooperation mode. Let ρ be the proportion of the total number of sub-frames operating in RRH-selection mode. The proportion between the number of sub-frames in RRH-selection mode and that in cooperation mode, denoted by $\rho/(1-\rho)$, could be configured as an arbitrary value in $[0,\infty]$, i.e., ρ is chosen from $[0,1]$. Otherwise, if the typical user is served by a traditional BS, the user accesses the nearest traditional BS x_b.

In order to decrease energy consumption and manage interference, we assume each RRH cluster or BS is muted independently in each sub-frame with probability $1-p$. Then, each RRH cluster or BS will be active in each sub-frame with probability p.

C. SIR and Cost Model

Without loss of generality, we consider one sub-frame. Let $\Phi_p^a \subset \Phi_p$ be the parent points of the active RRH clusters in the considered sub-frame and $\Phi_{\mathrm{BS}}^a \subset \Phi_{\mathrm{BS}}$ be the active traditional BSs; thus, the set of the active RRHs is $\Phi_{\mathrm{CR}}^a = \bigcup_{y\in\Phi_p^a} \Phi^y$.

When the typical user is a C-RAN user, see Figure 4, in the RRH-selection mode, letting $\mathcal{S}(x)$ be the index of the sub-band that the RRH $x \in \Phi_{\mathrm{CR}}^a$ is using, the received signal is

$$Y_{x_0} = \frac{P_{\mathrm{s}}^{1/2}}{|x_0|^{\alpha/2}} h_{x_0} X_{x_0} + \sum_{x\in\Phi_{\mathrm{CR}}^a\setminus x_0} \frac{P_{\mathrm{s}}^{1/2}}{|x|^{\alpha/2}} h_x X_x \mathbf{1}(\mathcal{S}(x)=\mathcal{S}(x_0)) + \sum_{x\in\Phi_{\mathrm{BS}}^a} \frac{P_{\mathrm{m}}^{1/2}}{|x|^{\alpha/2}} h_x X_x, \tag{4}$$

where x_0 denotes the location of the serving RRH, which is the nearest RRH in the serving RRH cluster to the typical user; h_x denotes the random fading coefficient between the BS located at x and the user located at the origin, and X_x denotes the transmitted symbol sent by the RRH located at x. The fading coefficients $\{h_x\} \sim \mathcal{CN}(0,1)$ are i.i.d. Gaussian random variables because of the Rayleigh fading assumption. Each RRH occupies one sub-band, and each BS occupies all sub-bands. Thus, the term for the interference from the BSs does not multiply by $\mathbf{1}(\mathcal{S}(x)=\mathcal{S}(x_0))$. Assuming that the transmitted symbols of different RRHs, $\{X_x\}$, are independent zero-mean Gaussian random variables of unit variance, the SIR is SIR $=$

$$\frac{P_{\mathrm{s}}|x_0|^{-\alpha}|h_{x_0}|^2}{\sum_{x\in\Phi_{\mathrm{CR}}^a\setminus x_0} \frac{P_{\mathrm{s}}|h_x|^2\mathbf{1}(\mathcal{S}(x)=\mathcal{S}(x_0))}{|x|^\alpha} + \sum_{x\in\Phi_{\mathrm{BS}}^a} \frac{P_{\mathrm{m}}|h_x|^2}{|x|^\alpha}}. \tag{5}$$

In the cooperation mode, letting $\Phi^{y_0} \subset \Phi_{\mathrm{CR}}^a$ be the set of RRHs in the serving RRH cluster centered at y_0, and $\Phi_{\mathrm{CR}}^a \setminus \Phi^{y_0}$ be the set of interfering RRHs, the received signal is

$$Y_{x_0} = \sum_{x\in\Phi^{y_0}} \frac{P_{\mathrm{s}}^{1/2}}{|x|^{\alpha/2}} h_x X_{y_0} + \sum_{y\in\Phi_p^a\setminus\{y_0\}} \sum_{x\in\Phi^y} \frac{P_{\mathrm{s}}^{1/2}}{|x|^{\alpha/2}} h_x X_y + \sum_{x\in\Phi_{\mathrm{BS}}^a} \frac{P_{\mathrm{m}}^{1/2}}{|x|^{\alpha/2}} h_x X_x, \tag{6}$$

where h_x is the fading coefficients; X_{y_0} is the symbol sent by the serving RRH cluster centered at y_0, and X_y is the symbol sent by the interfering RRH cluster centered at y. Then, the SIR is SIR $=$

$$\frac{P_{\mathrm{s}}|\sum_{x\in\Phi^{y_0}}|x|^{-\alpha/2}h_x|^2}{\sum_{y\in\Phi_p^a\setminus\{y_0\}} P_{\mathrm{s}}|\sum_{x\in\Phi^y}\frac{h_x}{|x|^{\alpha/2}}|^2 + \sum_{x\in\Phi_{\mathrm{BS}}^a}\frac{P_{\mathrm{m}}|h_x|^2}{|x|^\alpha}}. \tag{7}$$

When the user is served by a BS, if C-RAN works in RRH-selection mode, the SIR is SIR $=$

$$\frac{P_{\mathrm{m}}|x_b|^{-\alpha}|h_{x_b}|^2}{\sum_{x\in\Phi_{\mathrm{CR}}^a} \frac{P_{\mathrm{s}}|h_x|^2\mathbf{1}(\mathcal{S}(x)=\mathcal{S}(x_0))}{|x|^\alpha} + \sum_{x\in\Phi_{\mathrm{BS}}^a\setminus x_b} \frac{P_{\mathrm{m}}|h_x|^2}{|x|^\alpha}}, \tag{8}$$

and if C-RAN works in cooperation mode, the SIR is SIR $=$

$$\frac{P_{\mathrm{m}}|x_b|^{-\alpha}|h_{x_b}|^2}{\sum_{y\in\Phi_p^a} P_{\mathrm{s}}|\sum_{x\in\Phi^y}\frac{h_x|^2}{|x|^{\alpha/2}} + \sum_{x\in\Phi_{\mathrm{BS}}^a\setminus x_b}\frac{P_{\mathrm{m}}|h_x|^2}{|x|^\alpha}}. \tag{9}$$

The SIR expressions given by (5), (8) and (9) are for the users served by one sub-band, while the SIR expression given by (7) is for the users served by the whole band. Note that the transmit power is the same on different sub-bands. The forms of the SIR expressions are not related to the bandwidth due to the full-loaded assumption.

A transmission failure occurs if a user cannot achieve the predetermined SIR threshold θ; otherwise, the user transmits at the spectrum efficiency $\log_2(1+\theta)$ bits/s/Hz. We assume the retransmission mechanism is applied, so that when the transmission of sub-frame fails, it will be retransmitted in the next frame with probability p. In fact, the proposed re-transmission strategy is a specific instance of the automatic repeat request (ARQ), named stop-and-wait ARQ.

The ASE is defined as the mean rate per area per Hz. The queueing delay is difficult to analyze because of the interacting queueing problem [34], [35]. Thus, the delay denotes the number of frames to successfully deliver a packet. Since the network is full-loaded and the queueing delay is ignored, the delay lower bounds all other types of delays. The cost consists of the deployment cost and the energy cost. We consider the equipment cost and the fronthaul cost in the deployment cost. Denote the equipment cost C_{BP}, C_{RRH} and C_{BS} as the cost of deploying a BBU pool, a RRH and a traditional BS respectively. The fronthaul cost is the cost of connecting the RRHs to the BBU pool, which is assumed to be as the form $A_0|x|^{\beta}$ [15], where $A_0 > 0$ is the base cost and β is the exponent which determines how the base cost scales with the deployment distance $|x|$. The power consumption for a RRH is P_{RRH} and that for a traditional BS is P_{BS}. The power consumption model for a BBU pool is $P_{\mathrm{BP}}(n) = P_{\mathrm{C}} + \sum_{i=1}^{n} P_{\mathrm{BBU}}$, where n is the number of active virtual BBUs in the said BBU pool. The power consumption of the network in turn translates to the electricity bill in dollars, and we denote the translation factor as T_0 which typically takes value 787.5$/W for the use of a decade [36].

III. Area Spectrum Efficiency

In this work, we assume that the network is full-loaded, i.e., the time-frequency resources are fully occupied. However, in practical system, the traffic varies dynamically, and the statuses of the time-frequency resources change between idle and busy. Though the interference may increase due to the full-loaded assumption, the number of concurrent transmissions is maximized. Therefore, the ASE obtained under the full-loaded assumption is the maximum ASE for various loads in the practical system.

A. C-RAN Users

If the typical user is served by a RRH cluster, there will be a serving RRH cluster centered at y_0 with $|y_0| \leq R$. y_0 is uniformly distributed in the circular region centered at the origin with radius R, see Figure 4. Thus, the probability density function (pdf) of $|y_0|$ is

$$f_{|y_0|}(u) = \frac{2u}{R^2}, \quad 0 \leq u \leq R. \tag{10}$$

In RRH-selection mode, the typical user accesses the nearest RRH x_0 in the serving RRH cluster centered at y_0. The cumulative distribution function (cdf) of $|x_0|$ given $|y_0|$ is

$$\begin{aligned} &F_{|x_0||y_0|}(u) \\ &= 1 - \mathbb{P}\{\Phi^{y_0}(B(y_0, R) \cap B(o, u)) = 0 \mid |y_0|\} \\ &= \begin{cases} 1 - \left(1 - \dfrac{u^2}{R^2}\right)^N & \text{if } 0 \leq u \leq R - |y_0| \\ 1 - \left(1 - \dfrac{S_{y_0}(u)}{\pi R^2}\right)^N & \text{if } R - |y_0| < u \leq R + |y_0|, \end{cases} \end{aligned} \tag{11}$$

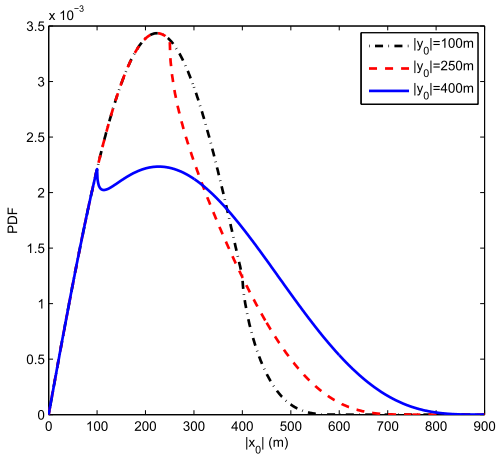

Fig. 5. The pdf of the distance between the serving RRH and the typical user, denoted by $|x_0|$, when the distance between the serving RRH cluster center and the typical user, denoted by $|y_0|$, is set as 100m, 250m and 400m respectively. The radius of each RRH cluster is 500m, and the number of RRHs in each RRH cluster is 3.

where

$$\begin{aligned} S_{y_0}(u) = {} & R^2 \arccos\left(\frac{R^2 + |y_0|^2 - u^2}{2R|y_0|}\right) \\ & + u^2 \arccos\left(\frac{u^2 + |y_0|^2 - R^2}{2u|y_0|}\right) \\ & - R|y_0| \sin\left(\arccos\left(\frac{R^2 + |y_0|^2 - u^2}{2R|y_0|}\right)\right). \end{aligned} \tag{12}$$

The probability density function (pdf) of $|x_0|$ given $|y_0|$ is

$$f_{|x_0||y_0|}(u) = F'_{|x_0||y_0|}(u). \tag{13}$$

Figure 5 gives the pdfs of $|x_0|$ for different values of $|y_0|$. It reveals that when the distance between the serving RRH cluster center and the typical user $|y_0|$ is large, the distance between the serving RRH and the typical user is more likely to be chosen from larger range of values.

In the RRH-selection mode, the interference comes from the RRHs using the same sub-band as the typical user and from all the BSs. Since only one RRH in each interfering RRH cluster causes interference, by the translation invariance of the PPP, the interfering RRHs construes a PPP whose distribution is the same as Φ_p^a. The interfering BSs construe a Poisson hole process; however, since the analysis of Poisson hole process is difficult, it is unlikely to obtain compact results. Nonetheless, when ignoring the dependence between the C-RAN and the BSs, we get an independent thinning of Φ_b with retaining probability $\exp(-\lambda_p \pi R^2)$, denoted by Φ_b', which yields a good approximation on the spatial distribution of the BSs [10].

Approximation 1: The active traditional BSs are approximately construed to an independently thinning version of Φ_b', which is a PPP $\Phi_b^{a'}$ with intensity $p\lambda_b \exp(-\lambda_p \pi R^2)$.

Comparing with the PPP, the points in Poisson hole process attract each other and show clustering property. The work [8] reveals that the parameters of Poisson hole process determine whether the clustering property brings more interference. The intensity and the radius of holes determine whether the results

obtained by the PPP approximation are lower or upper bounds. Based on these discussion, we get the following lemma.

Lemma 1: In the RRH-selection mode, the success probability of the C-RAN users is (14), as shown at the bottom of this page, where $C_0 = \pi p\theta^{2/\alpha}\Gamma(1+2/\alpha)\Gamma(1-2/\alpha)$, $\eta(x,y) = \left(x\cos y + \sqrt{R^2 - x^2\sin^2 y}\right)^2$, ${}_2F_1(a,b;c;z)$ *is the Gaussian hypergeometric function.*

Proof: See Appendix A □

Remark 1: To simplify the results, consider the performance of a C-RAN user located at the center of a RRH cluster. Then, we have $y_0 = o$ *and* $f_{|x_0|||y_0|}(u) = \frac{2uN}{R^2}\left(1 - \frac{u^2}{R^2}\right)^{N-1}$, $0 \le u \le R$. *For the special case that the density of C-RAN clusters is much larger than that of the traditional BSs, i.e.* $\lambda_b e^{-\lambda_p \pi R^2}$ *approaches* 0, *(14) will be simplified as*

$$\begin{aligned}\mathcal{P}_1 &\simeq \int_0^R e^{-\lambda_p x^2 C_0}\frac{2Nx}{R^2}\left(1-\frac{x^2}{R^2}\right)^{N-1}\mathrm{d}x \\ &= \frac{Ne^{-\lambda_p R^2 C_0}}{(-\lambda_p R^2 C_0)^N}\gamma\left(N, -\lambda_p R^2 C_0\right). \end{aligned} \tag{15}$$

where $\gamma(s,x) = \int_0^x t^{s-1}e^{-t}\mathrm{d}t$ *is the lower incomplete gamma function. The result is closed-form with a special function which could be easily evaluated, and it also reveals how the relationship between different system parameters that affect the success probability. For example, increasing* λ_p *four-fold or doubling R will result in the same success probability.*

In the cooperation mode, the typical C-RAN user is served by all RRHs in the serving RRH cluster centered at y_0 cooperatively. Therefore, the signal is the superposition of all signals from the RRHs in the serving RRH cluster. However, the interference signals from the RRHs is different from that in the RRH-selection mode since the signal from the same interfering RRH cluster can be treated as one signal due to the fact that the same symbol is transmitted.

Lemma 2: In the cooperation mode, the success probability of the C-RAN users is (16), as shown at the bottom of this page, where $f_{\Phi^y}(z) = \int_0^\infty \ldots \int_0^\infty g_y(z_N)\ldots g_y(z_2)g_y(x - \sum_{i=2}^N z_i)\mathrm{d}z_2\ldots\mathrm{d}z_N$ *is the pdf of* $\sum_{x\in\Phi^y}|x|^{-\alpha}$ *given by the N-fold convolution of* $g_y(z) = \frac{1}{\alpha\pi R^2}S_y'(z^{-\frac{1}{\alpha}})z^{-\frac{1}{\alpha}-1}$ *with* $S_y(u)$ *given by (12).*

Proof: See Appendix B. □

Remark 2: For the special case that $\lambda_p \gg \lambda_b$ *and R is small, the desired link distance approaches R, and the interfering RRHs can be considered as locating at the centers of the RRH clusters. Thus, (16) can be simplified as*

$$\begin{aligned}\mathcal{P}_2 &\simeq \exp\left(-2\pi p\lambda_p\int_0^\infty \frac{\theta r^{-\alpha}}{R^{-\alpha}+\theta r^{-\alpha}}r\mathrm{d}r\right) \\ &= \exp(-\lambda_p R^2 C_0). \end{aligned} \tag{17}$$

Remark 1 and Remark 2 show that the success probabilities of the C-RAN user in the two different modes are linearly correlated in the corresponding special case as

$$\mathcal{P}_1 \simeq \frac{N\gamma\left(N, -\lambda_p R^2 C_0\right)}{(-\lambda_p R^2 C_0)^N}\mathcal{P}_2. \tag{18}$$

B. Users Served by Traditional BSs

If the typical user is served by a traditional BS, it will access the nearest BS x_b. The interfering traditional BSs construe a Poisson hole process with no BSs within the circular region $B(o, |x_b|)$, see Figure 6. By Approximation 1, deriving the mean rate for the typical user becomes a standard problem which is well studied. Thus, the pdf of the distance between the typical user at the origin and the nearest traditional BS x_b is

$$f_{|x_b|}(u) = \exp\left(-\pi u^2\lambda_b e^{-\lambda_p\pi R^2} - \lambda_p\pi R^2\right)2\pi\lambda_b u. \tag{19}$$

For the users served by the traditional BSs, there is no RRH cluster center within the distance R of those users; otherwise, they will be served by the RRHs. Therefore, the interfering RRHs will construe a cluster process with no cluster center in the circular region $B(o,R)$. In the RRH-selection mode, only one RRH in each interfering RRH

$$\begin{aligned}\mathcal{P}_1 \simeq \int_0^R\int_0^{R+y}\exp\Bigg(&-\lambda_p x^2 C_0 - \lambda_b e^{-\lambda_p\pi R^2}x^2\frac{P_m^{2/\alpha}}{P_s^{2/\alpha}}C_0 \\ &+ p\lambda_b e^{-\lambda_p\pi R^2}\int_0^\pi \eta(y,\psi)\,{}_2F_1\left(1,\frac{2}{\alpha};1+\frac{2}{\alpha};-\frac{P_s\eta^{\alpha/2}(y,\psi)}{P_m\theta x^\alpha}\right)\mathrm{d}\psi\Bigg)\frac{2y}{R^2}f_{|x_0|||y_0|=y}(x)\mathrm{d}x\mathrm{d}y, \end{aligned} \tag{14}$$

$$\begin{aligned}\mathcal{P}_2 \simeq \int_0^R\int_0^\infty\exp\Bigg(&-2\pi p\lambda_p\int_0^\infty\int_0^\infty\frac{\theta w}{z+\theta w}f_{\Phi^{(r,0)}}(w)r\mathrm{d}w\mathrm{d}r - \frac{\lambda_b e^{-\lambda_p\pi R^2}P_m^{2/\alpha}C_0}{P_s^{2/\alpha}z^{2/\alpha}} \\ &+ p\lambda_b e^{-\lambda_p\pi R^2}\int_0^\pi\eta(t,\psi)\,{}_2F_1\left(1,\frac{2}{\alpha};1+\frac{2}{\alpha};-\frac{P_s\eta^{\alpha/2}(t,\psi)}{P_m\theta}z\right)\mathrm{d}\psi\Bigg)\frac{2t}{R^2}f_{\Phi^{(t,0)}}(z)\mathrm{d}z\mathrm{d}t, \end{aligned} \tag{16}$$

$$\begin{aligned}\mathcal{P}_3 = \pi\lambda_b\int_0^\infty\exp\Bigg(&-\frac{2\ p\lambda_p}{R^2}\int_R^\infty\int_0^{2\pi}\int_0^R\frac{rt\mathrm{d}r\mathrm{d}\beta\mathrm{d}t}{u^{-\alpha/2}(r^2+t^2+2rt\cos\beta)^{\alpha/2}P_m/(\theta P_s)+1} \\ &-\lambda_b e^{-\lambda_p\pi R^2}uC_0 + \pi p\lambda_b e^{-\lambda_p\pi R^2}u\,{}_2F_1\left(1,\frac{2}{\alpha};1+\frac{2}{\alpha};-\frac{1}{\theta}\right) - \pi u\lambda_b e^{-\lambda_p\pi R^2} - \lambda_p\pi R^2\Bigg)\mathrm{d}u. \end{aligned} \tag{20}$$

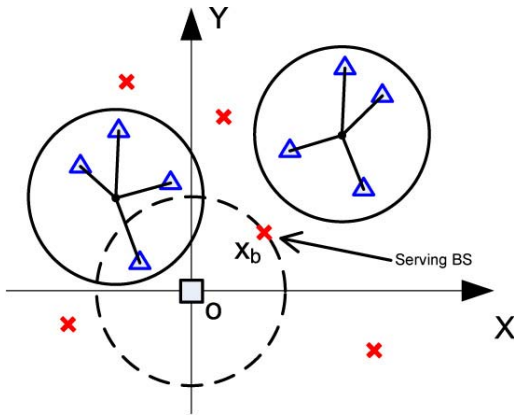

Fig. 6. Illustration of the typical user at the origin served by the nearest traditional BS at x_b. The triangles are RRHs, and the x's are traditional BSs.

cluster causes interference, while in the cooperation mode, all RRHs in the interfering RRH clusters cause interference.

Lemma 3: In RRH-selection mode, the success probability of the users served by traditional BS is (20), as shown at the bottom of the previous page. In cooperation mode, the success probability of the typical user served by traditional BS is (21), as shown at the bottom of this page.

Proof: See Appendix C □

Remark 3: When $P_m \gg P_s$ or λ_p is small, the interference from the C-RAN clusters can be ignored. Thus, (20) and (21) can be simplified as

$$\mathcal{P}_3 \simeq \mathcal{P}_4 \simeq \frac{1}{C_0/\pi - p_2 F_1(1, 2/\alpha; 1+2/\alpha; -1/\theta) + 1}, \tag{22}$$

In the RRH-selection mode, each user occupies one sub-band with bandwidth W/N. In the cooperation mode, the whole bandwidth W is used by each user; however, since each user only occupies $1/N$ of the time, the rate should be scaled down by $1/N$. Therefore, the mean rate for the typical user in both modes is $p\frac{W}{N}\log_2(1+\theta)\mathbb{P}\{\text{SIR} \geq \theta\}$. Noticing that the proportion of the total time in the RRH-selection mode is ρ, and that in the cooperation mode is $1-\rho$, we get the throughputs for the users served by the C-RAN or the BSs as

$$\tau_{\text{CR}} = (\rho\mathcal{P}_1 + (1-\rho)\mathcal{P}_2)\frac{pW}{N}\log_2(1+\theta). \tag{23}$$

$$\tau_{\text{BS}} = (\rho\mathcal{P}_3 + (1-\rho)\mathcal{P}_4)\frac{pW}{N}\log_2(1+\theta). \tag{24}$$

Since the intensity of the parent process of the RRH clusters is λ_p and the intensity of the traditional BSs is $\lambda_b \exp(-\lambda_p \pi R^2)$, we get the intensity of the users served by the C-RAN and the BSs as $N\lambda_p$ and $N\lambda_b \exp(-\lambda_p \pi R^2)$ respectively. Therefore, we get the ASE as

$$A_e = \frac{N}{W}\lambda_p \tau_{\text{CR}} + \frac{N}{W}\lambda_b e^{-\lambda_p \pi R^2}\tau_{\text{BS}}. \tag{25}$$

Theorem 1: In the proposed C-RAN architecture, the ASE of the network is

$$A_e = \big(\lambda_p(\rho\mathcal{P}_1 + (1-\rho)\mathcal{P}_2) + \lambda_b e^{-\lambda_p \pi R^2}(\rho\mathcal{P}_3 + (1-\rho)\mathcal{P}_4)\big)p\log_2(1+\theta). \tag{26}$$

IV. Mean Delay

The queueing delay is difficult to analyze because of the interacting queueing problem [34], [35]. Therefore, we analyze the mean number of frames required to successfully deliver a packet.

A. C-RAN Users

The set of active RRHs and active BSs are different in different frames due to the randomly muting and the fading. Conditioned on the realization of the point processes, the success transmission is a random event determined by the muting process and the fading. Therefore, the success probability conditioned on the realization of the point processes is the same in different frames. Then, the mean delay conditioned on the realization of the point processes is a random variable of geometric distribution. Based on the above discussion, we obtain the following lemma.

Lemma 4: In the RRH-selection mode, the mean delay of the typical C-RAN user is (27), as shown at the bottom of this page.

Proof: See Appendix D. □

Remark 4: Consider the performance of a C-RAN user located at the center of a RRH cluster. For the special case that the density of C-RAN clusters is much larger than that of the traditional BSs, (27) will be simplified as

$$\mathcal{D}_1 \simeq \frac{N\gamma\left(N, \lambda_p(1-p)^{2/\alpha-1}R^2C_0\right)}{p(\lambda_p(1-p)^{2/\alpha-1}C_0R^2)^N}e^{\lambda_p(1-p)^{2/\alpha-1}C_0R^2}. \tag{28}$$

Lemma 5: In the cooperation mode, the mean delay of the typical C-RAN user is (29), as shown at the top of the next page.

$$\mathcal{P}_4 = \pi\lambda_b\int_0^\infty \exp\Bigg(-2\pi p\lambda_p\int_R^\infty\int_0^\infty \frac{f_{\Phi^{(r,0)}}(z)r\,\mathrm{d}z\,\mathrm{d}r}{1+P_m/(P_s\theta u^{\alpha/2}z)} - \lambda_b e^{-\lambda_p\pi R^2}uC_0 + \pi pu\lambda_b e^{-\lambda_p\pi R^2}{}_2F_1\left(1, \frac{2}{\alpha}; 1+\frac{2}{\alpha}; -\frac{1}{\theta}\right) - \pi u\lambda_b e^{-\lambda_p\pi R^2} - \lambda_p\pi R^2\Bigg)\mathrm{d}u. \tag{21}$$

$$D_1 \simeq \frac{1}{p}\int_0^R\int_0^{R+y}\exp\Bigg(\lambda_p x^2(1-p)^{2/\alpha-1}C_0 + \lambda_b e^{-\lambda_p\pi R^2}\frac{P_m^{2/\alpha}}{P_s^{2/\alpha}}x^2(1-p)^{2/\alpha-1}C_0 - \frac{p\lambda_b}{1-p}e^{-\lambda_p\pi R^2}\int_0^\pi \eta(y,\psi){}_2F_1\left(1, 2/\alpha; 1+2/\alpha; -\frac{P_s\eta^{1/2/\alpha}(y,\psi)}{(1-p)\theta x^\alpha P_m}\right)\mathrm{d}\psi\Bigg)\frac{2y}{R^2}f_{|x_0||y_0=y}(x)\mathrm{d}x\,\mathrm{d}y. \tag{27}$$

$$D_2 \simeq \frac{1}{p}\int_0^R\int_0^\infty \exp\left(2\pi p\lambda_p \int_0^\infty\int_0^\infty \frac{r f_{\Phi^{(r,0)}}(w)\mathrm{d}w\mathrm{d}r}{1-p+\frac{z}{\theta w}} + \frac{\lambda_b C_0\ e^{-\lambda_p\pi R^2}(1-p)^{2/\alpha-1}P_m^{2/\alpha}}{P_s^{2/\alpha}z^{2/\alpha}}\right.$$
$$\left. - \frac{p\lambda_b e^{-\lambda_p\pi R^2}}{1-p}\int_0^\pi \eta(y,\psi)\,{}_2F_1\left(1,\frac{2}{\alpha};1+\frac{2}{\alpha};-\frac{P_s z\eta^{\alpha/2}(y,\psi)}{(1-p)\theta P_m}\right)\mathrm{d}\psi\right)\frac{2y}{R^2}f_{\Phi^{(y,0)}}(z)\mathrm{d}z\mathrm{d}y. \quad (29)$$

Proof: See Appendix E. □

Remark 5: For the special case that $\lambda_p \gg \lambda_b$ and R is small, (29) could be simplified as

$$\mathcal{D}_2 \simeq \frac{1}{p}\exp(\lambda_p R^2(1-p)^{2/\alpha-1}C_0). \quad (30)$$

Remark 4 and Remark 5 show that the mean delays of the C-RAN user in two different modes are linearly correlated in the corresponding special case as

$$\mathcal{D}_1 \simeq \frac{N\gamma\left(N,-\lambda_p(1-p)^{2/\alpha-1}C_0\ R^2\right)}{(\lambda_p(1-p)^{2/\alpha-1}C_0\ R^2)^N}\mathcal{D}_2. \quad (31)$$

B. Users Served by Traditional BSs

By Approximation 1, the spatial distribution of the traditional BSs is a PPP Φ_b' outside the circular region $B(o,|x_b|)$ with intensity $\lambda_b\exp(-\lambda_p\pi R^2)$. For the interference from the RRHs, there is no interfering RRHs cluster center within the circular region $B(o,R)$. Through the same derivations as the previous lemmas, we get the following lemma.

Lemma 6: In the RRH-selection mode, the mean delay of the typical user served by a traditional BS is (32), as shown at the bottom of this page. In the cooperation mode, the mean delay of the typical user served by a traditional BS is (33), as shown at the bottom of this page.

Remark 6: For the special case that $P_m \gg P_s$ or λ_p is small, (32) and (33) could be simplified as

$$\mathcal{D}_3 \simeq \mathcal{D}_4 \simeq \frac{(1-p)/p}{1-p-\frac{(1-p)^{2/\alpha}C_0}{\pi}+p_2F_1(1,\frac{2}{\alpha};1+\frac{2}{\alpha};-\frac{1}{\theta(1-p)})}. \quad (34)$$

The mean delay for the users served by the C-RAN and the BSs are

$$D_{\mathrm{CR}} = \rho D_1 + (1-\rho)D_2, \quad (35)$$
$$D_{\mathrm{BS}} = \rho D_3 + (1-\rho)D_4. \quad (36)$$

Theorem 2: In the proposed C-RAN architecture, the mean delay for all users in the network is

$$D = \frac{\lambda_p(\rho D_1+(1-\rho)D_2)+\lambda_b e^{-\lambda_p\pi R^2}(\rho D_3+(1-\rho)D_4)}{\lambda_p+\lambda_b e^{-\lambda_p\pi R^2}}. \quad (37)$$

V. System Cost

The cost per unit area for deploying the BBU pools, the RRHs and the traditional BSs is $\lambda_p C_{\mathrm{BP}} + N\lambda_p C_{\mathrm{RRH}} + \lambda_b e^{-\lambda_p\pi R^2}C_{\mathrm{BS}}$. The fronthaul connects the BBU pools to the RRHs, and since each RRH is uniformly distributed in the circular cluster region, the pdf of the distance of each BBU pool to RRH connection, denoted by L, is

$$f_{\mathrm{L}}(l) = \frac{2l}{R^2}, \quad 0 \le l \le R. \quad (38)$$

The average cost for each BBU pool to RRH connection is

$$C_{\mathrm{L}} = \int_0^R A_0 l^\beta f_L(l)\mathrm{d}l = \frac{2A_0R^{\beta+2}}{\beta+2}. \quad (39)$$

The cost per unit area for the fronthaul is $\frac{2\lambda_p N A_0 R^{\beta+2}}{\beta+2}$. The equipment cost per unit area is

$$C_{\mathrm{eq}} = \lambda_p C_{\mathrm{BP}} + N\lambda_p C_{\mathrm{RRH}} + \lambda_b e^{-\lambda_p\pi R^2}C_{\mathrm{BS}} + \frac{2\lambda_p N A_0 R^{\beta+2}}{\beta+2}. \quad (40)$$

The energy cost is directly related to the active number of virtual BBUs. In RRH-selection mode, all N virtual BBUs in each BBU pool should be active to generate N independent frames. In cooperation mode, only one active virtual BBU in the BBU pool is active. Considering the intensities of the nodes, the energy cost per unit area for RRHs and for BSs are $Np\lambda_p T_0 P_{\mathrm{RRH}}$ and $p\lambda_b e^{-\lambda_b\pi R^2}T_0P_{\mathrm{BS}}$ respectively. In RRH-selection mode, the energy cost per unit area for the BBU pools is $p\lambda_p T_0(P_{\mathrm{C}}+NP_{\mathrm{BBU}})$, and the total energy cost per unit area is $Np\lambda_p T_0 P_{\mathrm{RRH}} + p\lambda_b e^{-\lambda_p\pi R^2}T_0P_{\mathrm{BS}} +$

$$D_3 = \frac{\pi\lambda_b}{p}\int_0^\infty \exp\left(\frac{2p\lambda_p}{R^2}\int_R^\infty\int_0^{2\pi}\int_0^R \frac{rt\mathrm{d}r\mathrm{d}\beta\mathrm{d}t}{1-p+\frac{P_m}{\theta u^{\frac{\alpha}{2}}P_s}(r^2+t^2+2rt\cos\beta)^{\frac{\alpha}{2}}} + \frac{\lambda_b u e^{-\lambda_p\pi R^2}C_0}{(1-p)^{1-2/\alpha}}\right.$$
$$\left. - \frac{\pi p\lambda_b u}{1-p}e^{-\lambda_p\pi R^2}{}_2F_1\left(1,\frac{2}{\alpha};1+\frac{2}{\alpha};-\frac{1}{\theta(1-p)}\right) - \pi u\lambda_b e^{-\lambda_p\pi R^2} - \lambda_p\pi R^2\right)\mathrm{d}u. \quad (32)$$

$$D_4 = \frac{\pi\lambda_b}{p}\int_0^\infty \exp\left(2\pi p\lambda_p\int_R^\infty\int_0^\infty \frac{r f_{\Phi^{(r,0)}}(w)\mathrm{d}w\mathrm{d}r}{1-p+\frac{P_m}{w\theta P_s}u^{-\alpha/2}} + \frac{\lambda_b u e^{-\lambda_p\pi R^2}C_0}{(1-p)^{1-2/\alpha}}\right.$$
$$\left. - \frac{\pi p\lambda_b u}{1-p}e^{-\lambda_p\pi R^2}{}_2F_1\left(1,\frac{2}{\alpha};1+\frac{2}{\alpha};-\frac{1}{\theta(1-p)}\right) - \pi u\lambda_b e^{-\lambda_p\pi R^2} - \lambda_p\pi R^2\right)\mathrm{d}u. \quad (33)$$

$p\lambda_p T_0(P_{\mathrm{C}} + NP_{\mathrm{BBU}})$. In cooperation mode, the energy cost per unit area for the BBU pools is $p\lambda_p T_0(P_{\mathrm{C}} + P_{\mathrm{BBU}})$, and the total energy cost per unit area is $Np\lambda_p T_0 P_{\mathrm{RRH}} + p\lambda_b e^{-\lambda_p \pi R^2} T_0 P_{\mathrm{BS}} + p\lambda_p T_0(P_{\mathrm{C}} + P_{\mathrm{BBU}})$. Thus, we get the average energy cost per unit area as

$$C_{\mathrm{en}} = Np\lambda_p T_0 P_{\mathrm{RRH}} + p\lambda_b e^{-\lambda_p \pi R^2} T_0 P_{\mathrm{BS}} + p\lambda_p T_0 P_{\mathrm{C}} + (\rho N + 1 - \rho) p\lambda_p T_0 P_{\mathrm{BBU}}. \quad (41)$$

Combining the equipment cost and the energy cost, we get the total cost per unit area as

$$C = C_{\mathrm{eq}} + C_{\mathrm{en}}. \quad (42)$$

VI. Asynchronous Transmission

In the previous discussions, we have analyzed the synchronous transmission, i.e., all active RRH clusters work synchronously either in the RRH-selection mode or in the cooperation mode. An advantage of synchronous transmission is that the frequent switching between two modes is avoided. However, when ignoring the expense for switching between different operation modes, the asynchronous transmission, where each individual RRH cluster chooses its own transmission mode based on the property of the scheduled user, appears to be a more efficient approach. For example, when a RRH cluster serves a user close to a RRH, the RRH-selection mode is preferred since the performance of the user can be guaranteed by the nearest RRH, and other RRHs in the same cluster can be released to serve other users. On the other hand, when serving a user far from the nearest RRH, the cooperation mode is preferred to enhance the desired signal.

Analytically, in order to evaluate the performance of asynchronous transmission, consider a typical user located at the center of a RRH cluster. We consider a practical and meaningful user association mechanism, i.e., the choice of the transmission mode is based on the distance between a user and its nearest RRH. In particular, if the distance between the typical user and its nearest RRH is smaller than R_0 (with probability $\rho = 1 - (1 - R_0^2/R^2)^N$), the RRH-selection mode will be chosen; otherwise, the cooperation mode will be used (with probability $1 - \rho$). Without loss of generality, we assume that each individual RRH cluster works in the RRH-selection mode with probability ρ and in the cooperation mode with probability $1 - \rho$.

When the typical user is served in the RRH-selection mode, the pdf of the distance between the typical user and its nearest RRH is

$$f_{|x_0||y_0=o}(u) = \frac{F'_{|x_0|||y_0|}(u)}{F_{|x_0|||y_0|}(R_0)}\Big|_{y_0=o} = \frac{2uN(1-u^2/R_0^2)^{N-1}}{(1-(1-R_0^2/R^2)^N)R_0^2}, \quad 0 \le u \le R_0, \quad (43)$$

where $F_{|x_0|||y_0|}(u)$ is given by (11). Similar to the derivations of Lemma 1 and Lemma 4, we get the success probability and the mean delay of asynchronous transmission in (44) and (45), as shown at the bottom of this page.

When the typical user is served in the cooperation mode, i.e., the distance between the typical user and its nearest RRH is larger than R_0, we assume that the cooperative RRHs are distributed on the circle centered at the origin with radius R to simplify the result, which is accurate for large R_0. Then, similar to the derivations of Lemma 2 and Lemma 5, we get the success probability and the mean delay in asynchronous transmission in (46) and (47), as shown at the bottom of the next page.

The results can be further simplified by the same approaches in the previous remarks. Figure 7 shows the comparison of the success probabilities for two different operation modes when varying the critical radius R_0 in the asynchronous transmission. The optimal value of R_0 changes for different values of the active probability p.

VII. Numerical Evaluation

In this section, we numerically evaluate the performance of the network. The default configurations of the system parameters are list in Table I. The simulation results are obtained by considering a rectangle region of 2000m×2000m, and the number of repetitions to deploy the nodes is set as 1000. The Rayleigh fading is considered, while the shadowing and the thermal noise are ignored. The ASE, the mean delay and the system cost in the simulation are statistically obtained by considering all users in the network.

Figure 8 plots the tradeoffs between the ASE, the mean delay and the system cost for different densities of the RRH clusters λ_p. The curves are obtained by varying the access probability p, i.e., each point of the curves corresponds to one triple evaluated at a particular p, which has three components: the ASE, the mean delay and the system cost.

$$\mathcal{P}_1 \simeq \int_0^{R_0} \exp\Bigg(-\rho\lambda_p x^2 C_0 - 2\pi p(1-\rho)\lambda_p \int_0^\infty \int_0^\infty \frac{\theta w}{x^{-\alpha} + \theta w} f_{\Phi^{(r,0)}}(w) r \mathrm{d}w \mathrm{d}r - \lambda_b e^{-\lambda_p \pi R^2} x^2 \frac{P_m^{2/\alpha}}{P_s^{2/\alpha}} C_0 + p\pi\lambda_b R^2 e^{-\lambda_p \pi R^2} {}_2F_1\left(1, \frac{2}{\alpha}; 1 + \frac{2}{\alpha}; -\frac{P_s R^\alpha}{P_m \theta x^\alpha}\right)\Bigg) f_{|x_0||y_0=o}(x)\mathrm{d}x, \quad (44)$$

$$D_1 \simeq \frac{1}{p}\int_0^{R_0} \exp\Bigg(\frac{\rho\lambda_p x^2 C_0}{(1-p)^{1-2/\alpha}} + 2\pi p(1-\rho)\lambda_p \int_0^\infty \int_0^\infty \frac{r f_{\Phi^{(r,0)}}(w)\mathrm{d}w\mathrm{d}r}{1 - p + \frac{x^{-\alpha}}{\theta w}} + \frac{\lambda_b x^2 C_0 e^{-\lambda_p \pi R^2} P_m^{2/\alpha}}{(1-p)^{1-2/\alpha} P_s^{2/\alpha}} - \frac{p\pi\lambda_b R^2}{1-p} e^{-\lambda_p \pi R^2} {}_2F_1\left(1, \frac{2}{\alpha}; 1 + \frac{2}{\alpha}; -\frac{P_s R^\alpha}{(1-p)\theta x^\alpha P_m}\right)\Bigg) f_{|x_0||y_0=o}(x)\mathrm{d}x. \quad (45)$$

TABLE I
SYSTEM PARAMETERS

Symbol	Description	Value
α	Path loss exponent	3
θ	SIR threshold	0dB
P_m, P_s	Transmit power of traditional BSs and RRHs	46dBm, 23dBm
λ_p	Density of parent process for the clustered RRHs	0.000008 center/m^2
λ_b	Density of traditional BSs	0.000005 BS/m^2
R	Radius of each RRH cluster	100m
N	Number of RRHs in each RRH cluster	10
T_0	Electricity bill per watt for a decade	787.5$
$C_{\mathrm{BP}}, C_{\mathrm{RRH}}, C_{\mathrm{BS}}$	Cost for a BBU pool, a RRH and a BS	40000$, 10000$, 50000$
A_0	Base cost for fronthaul	100000$
β	Scaling exponent of the deployment cost	1
$P_{\mathrm{RRH}}, P_{\mathrm{BS}}$	Power consumptions for a RRH and a BS	49.7W, 160.8W
$P_{\mathrm{C}}, P_{\mathrm{BBU}}$	Power consumptions for common site infrastructure and BBU	23.5W, 14.8W

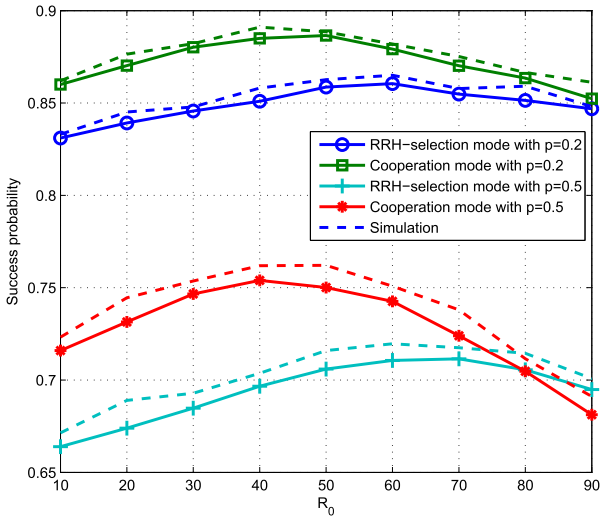

Fig. 7. Comparison of success probabilities for two different operation modes when varying R_0 in asynchronous transmission.

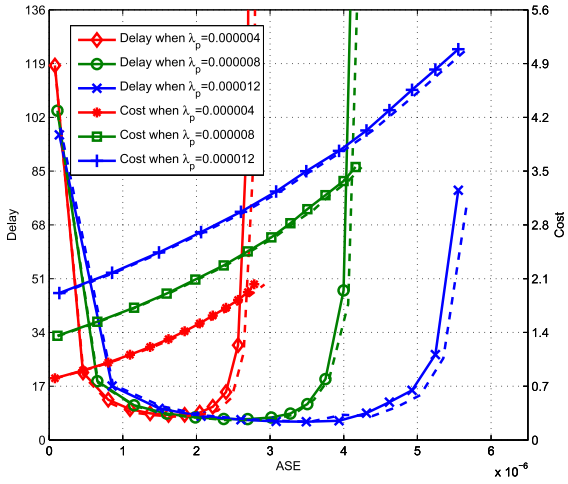

Fig. 8. Tradeoffs between the ASE, the mean delay and the system cost for different densities of the RRH clusters λ_p. The solid lines plot the numerical results, and the dashed lines plot the simulation results.

The curves show that the gap between the simulation results and the numerical results coincide well, thus verifying the accuracy of Approximation 1. In particular, the approximation is very precise for small access probability p or small densities of the RRH clusters λ_p. As the ASE increases, the mean delay first decreases then increases while the system cost always positively correlated with the ASE. This is because as p increases, there will be more simultaneous transmissions in the network, then the ASE increases; however, the mean delay of each user is limited by the muting mechanism when p is small and by the large interference as well as the large interference correlation when p is large. The system cost always increases when increasing the active probability p because the energy cost increases linearly with respective to p. Figure 8 also reveals that as λ_p increases, the curve for the delay-ASE tradeoff becomes wider and moves to the right, indicating that a better operating point could be chosen in order to increase the ASE and decrease the mean delay.

Figure 9 plots the throughputs of the users served by the C-RAN and the BSs in the two different modes as

$$\mathcal{P}_2 \simeq \exp\Bigg(-\rho\lambda_p N^{-2/\alpha} R^2 C_0 - 2\pi p(1-\rho)\lambda_p \int_0^\infty \int_0^\infty \frac{\theta w}{NR^{-\alpha}+\theta w} f_{\Phi^{(r,0)}}(w) r \mathrm{d}w \mathrm{d}r - \frac{\lambda_b R^2 C_0 e^{-\lambda_p \pi R^2} P_m^{2/\alpha}}{P_s^{2/\alpha} N^{2/\alpha}} + p\pi\lambda_b R^2 e^{-\lambda_p \pi R^2} {}_2F_1\left(1, \frac{2}{\alpha}; 1+\frac{2}{\alpha}; -\frac{P_s N}{P_m \theta}\right)\Bigg), \tag{46}$$

$$D_2 \simeq \frac{1}{p}\exp\Bigg(\frac{\rho\lambda_p R^2 N^{-2/\alpha} C_0}{(1-p)^{1-2/\alpha}} + 2\pi p(1-\rho)\lambda_p \int_0^\infty \int_0^\infty \frac{r f_{\Phi^{(r,0)}}(w) \mathrm{d}w \mathrm{d}r}{1-p+\frac{NR^{-\alpha}}{\theta w}} + \frac{\lambda_b R^2 C_0 e^{-\lambda_p \pi R^2} P_m^{2/\alpha}}{(1-p)^{1-2/\alpha} P_s^{2/\alpha} N^{2/\alpha}} - \frac{p\pi\lambda_b R^2 e^{-\lambda_p \pi R^2}}{1-p} {}_2F_1\left(1, \frac{2}{\alpha}; 1+\frac{2}{\alpha}; -\frac{P_s N}{(1-p)\theta P_m}\right)\Bigg). \tag{47}$$

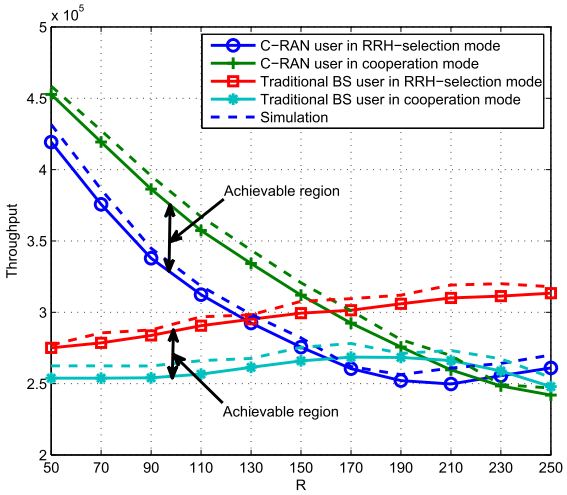

Fig. 9. Throughputs of the C-RAN users and the users served by traditional BSs in the RRH-selection mode and in the cooperation mode as functions of the radius of the RRH clusters R.

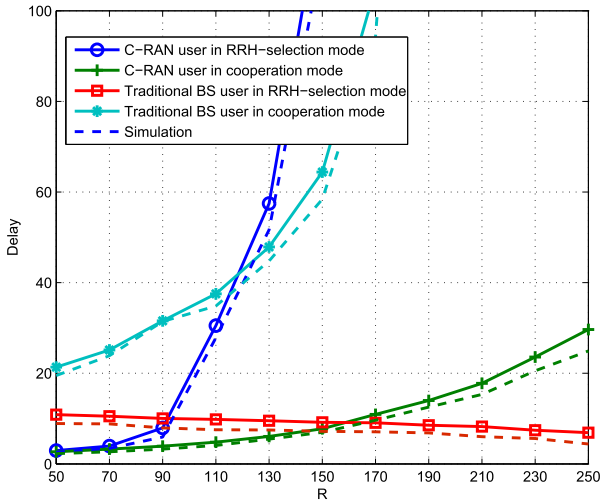

Fig. 10. Mean delays of the C-RAN users and the users served by traditional BSs in the RRH-selection mode and in the cooperation mode as functions of the radius of the RRH clusters R.

functions of the radius of the RRH clusters R. For small R, the throughput of the C-RAN users in the RRH-selection mode is smaller than that in the cooperation mode since the cooperation improves the power of the useful signal; however, the throughput of the users served by the traditional BSs in the RRH-selection mode is larger than that in the cooperation mode, since all RRHs are active in the cooperation mode which increases the interference caused by the RRHs. When R starts to grow, the throughputs of the C-RAN users in both modes decrease because when increasing R, the distances between the RRHs and the C-RAN users are increased, thus reducing the power of the useful signal. On the contrary, when R starts to grow, the throughputs of the users served by the traditional BSs in both modes increase; meanwhile, when R exceeds a certain threshold, the throughput of the C-RAN user in the RRH-selection mode starts to grow. These behaviors can be interpreted as that when increasing R the number of BSs are reduced, thus reducing the interference from the BSs. However, in the cooperation mode, when R continues to grow, although the interference from the BSs is reduced, the throughput of the C-RAN user keeps decreasing, and that of the users served by the BSs starts to decreases. This observation indicates that by increasing R, the effect of the interference caused by the RRHs becomes dominant in the cooperation mode. Figure 9 also verifies Remark 2 that the throughput decreases exponentially as the increment of R when R is small.

Figure 10 plots the mean delays of the users served by the C-RAN and the BSs in the two different modes as functions of the radius of the RRH clusters R. The trends of the curves in Figure 10 can be interpreted by the same reason for those in Figure 9 since the mean delay is negative correlated to the throughput in most of the cases. However, Figure 9 and Figure 10 reveal that the mean delay is not always inversely proportional to the throughput, which is counter-intuitive. This is because the mean delay is obtained by averaging over all transmissions in the network. Even if the mean throughput of the network is large, a small amount of transmissions in the network experience very large delay, thus resulting in large mean delay. Figure 9 and Figure 10 guide the chosen of the RRH cluster radius to optimize the throughput and the delay of the C-RAN users and the users served by the traditional BSs. Figure 10 also verifies Remark 5 that the mean delays of the C-RAN user in two different modes are linearly correlated for small R.

VIII. Conclusions

In this paper, we explored the complementary networking architecture for the C-RAN, which takes advantages of both the C-RAN and the traditional BSs. We proposed to combine the Neyman-Scott cluster process and the Poisson hole process to model the architecture of the complementary network and analyzed the two multiple association mechanisms, i.e. the RRH-selection mode and the cooperation mode, through the Poisson point process approximation.

The results reveal that when adjusting the active probability, the mean delay is negative correlated to the ASE for small ASE and is reversed for large ASE, while the cost is always positively correlated to the ASE, indicating that an appropriate active probability should be chosen to tradeoff the ASE, the mean delay and the cost. The results can be interpreted as that both the ASE and the cost are increased when increasing the number of simultaneous transmissions; however, the mean delay is limited by the muting mechanism for small p and by the large interference as well as the large interference correlation for large p. Denser deployment of C-RAN results in better ASE-delay tradeoff with more cost. The cooperation mode improves the performance of the C-RAN users by strengthening the useful signal but also increases the interference in the network. The effect of the interference caused by the cooperative RRHs becomes dominant when the radius of the RRH cluster is large. The operating points between the RRH-selection mode and the cooperation mode are achieved by time division to tradeoff the ASE, the mean delay and the system cost.

APPENDIX A
PROOF OF LEMMA 1

In the RRH-selection mode, from (5) we get the conditional success probability as

$$\mathbb{P}\{\text{SIR} \geq \theta \mid x_0, y_0\} = \mathcal{L}_{I_{\text{CR}}}\left(\frac{\theta|x_0|^\alpha}{P_s}\right)\mathcal{L}_{I_{\text{BS}}}\left(\frac{\theta|x_0|^\alpha}{P_s}\right). \quad (48)$$

where $\mathcal{L}_{I_{\text{CR}}}(s)$ and $\mathcal{L}_{I_{\text{BS}}}(s)$ are the Laplace transform of interference from the RRHs using the same sub-bands as that of the typical user and from the traditional BSs respectively. The Laplace transform of the interference from the RRHs is

$$\mathcal{L}_{I_{\text{CR}}}(s) \overset{(a)}{=} \mathbb{E}\left[\exp\left(-s\sum_{x\in\Phi_p^a} P_s|x|^{-\alpha}|h_x|^2\right)\right]$$
$$\overset{(b)}{=} \exp\left(-\pi p\lambda_p s^{\frac{2}{\alpha}} P_s^{\frac{2}{\alpha}}\Gamma\left(1+\frac{2}{\alpha}\right)\Gamma\left(1-\frac{2}{\alpha}\right)\right). \quad (49)$$

(a) holds because the interference RRHs construes a PPP whose distribution is the same as Φ_p^a. (b) follows from [23, 5.1.7].

By Approximation 1, the active traditional BSs are distributed as a PPP $\Phi_b^{a'}$ with the intensity $p\lambda_b \exp(-\lambda_p \pi R^2)$ outside the serving RRH cluster, i.e., outside the circle region $B(y_0, R)$. Thus, the Laplace transform of the interference from the traditional BSs $\mathcal{L}_{I_{\text{BS}}}(s)$ is approximated as $\mathcal{L}_{I_{\text{BS}}}(s)$

$$\simeq \mathbb{E}\left[\exp\left(-s\sum_{x\in\Phi_b^{a'}\setminus B(y_0,R)} P_m|x|^{-\alpha}|h_x|^2\right)\right]$$
$$= \exp\left(-p\lambda_b e^{-\lambda_p\pi R^2}\int_0^{2\pi}\int_{\eta(|y_0|,\psi)}^{\infty}\frac{P_m s r}{r^\alpha + P_m s}\mathrm{d}r\mathrm{d}\psi\right)$$
$$= \exp\Bigg(-p\lambda_b e^{-\lambda_p\pi R^2}\pi s^{2/\alpha}P_m^{2/\alpha}\Gamma(1+2/\alpha)\Gamma(1-2/\alpha)$$
$$+ p\lambda_b e^{-\lambda_p\pi R^2}\int_0^{\pi}\eta(|y_0|,\psi)$$
$$ {}_2F_1\left(1,\frac{2}{\alpha};1+\frac{2}{\alpha};-\frac{\eta^{\alpha/2}(|y_0|,\psi)}{P_m s}\right)\mathrm{d}\psi\Bigg). \quad (50)$$

Plugging in the conditional success probability, we get the results in the lemma.

APPENDIX B
PROOF OF LEMMA 2

In the cooperation mode, from (7) we get the conditional success probability as

$$\mathbb{P}\{\text{SIR} \geq \theta \mid \Phi^{y_0}, y_0\}$$
$$= \mathbb{P}\left\{\frac{|\sum_{x\in\Phi^{y_0}} P_s^{1/2}|x|^{-\alpha/2}h_x|^2}{I_{\text{CR}} + I_{\text{BS}}} \geq \theta \mid \Phi^{y_0}, y_0\right\}$$
$$\overset{(a)}{=} \mathbb{P}\left\{P_s|h|^2\sum_{x\in\Phi^{y_0}}|x|^{-\alpha} \geq \theta(I_{\text{CR}} + I_{\text{BS}}) \mid \Phi^{y_0}, y_0\right\}$$
$$= \mathcal{L}_{I_{\text{CR}}}\left(\frac{\theta}{P_s\sum_{x\in\Phi^{y_0}}|x|^{-\alpha}}\right)\mathcal{L}_{I_{\text{BS}}}\left(\frac{\theta}{P_s\sum_{x\in\Phi^{y_0}}|x|^{-\alpha}}\right). \quad (51)$$

where $\mathcal{L}_{I_{\text{CR}}}(s)$ and $\mathcal{L}_{I_{\text{BS}}}(s)$ are the Laplace transforms of interference from the RRHs and the BSs respectively, and $h \sim \mathcal{CN}(0,1)$ is a complex Gaussian random variable. (a) holds because $\{h_x\} \sim \mathcal{CN}(0,1)$ are i.i.d. complex Gaussian random variables, and their weighted summation could be equivalent to one complex Gaussian random variable. The interference from the BSs in the cooperation mode is the same as that in the RRH-selection mode, whose Laplace transform is given by (50). The Laplace transform of the interference from the RRHs is $\mathcal{L}_{I_{\text{CR}}}(s)$

$$= \mathbb{E}\left[\exp\left(-s\sum_{y\in\Phi_p^a\setminus\{y_0\}}\left|\sum_{x\in\Phi^y}P_s^{1/2}|x|^{-\alpha/2}h_x\right|^2\right)\right]$$
$$\overset{(a)}{=} \mathbb{E}_{\Phi_p^a}\left[\prod_{y\in\Phi_p^a}\mathbb{E}_{\Phi^y,h_x}\left[\exp\left(-sP_s\left|\sum_{x\in\Phi^y}|x|^{-\alpha/2}h_x\right|^2\right)\right]\right]$$
$$\overset{(b)}{=} \mathbb{E}_{\Phi_p^a}\left[\prod_{y\in\Phi_p^a}\mathbb{E}_{\Phi^y}\left[\frac{1}{1+sP_s\sum_{x\in\Phi^y}|x|^{-\alpha}}\right]\right]$$
$$\overset{(c)}{=} \exp\left(-p\lambda_p\int_{\mathbb{R}^2}\left(1-\mathbb{E}_{\Phi^y}\left[\frac{1}{1+sP_s\sum_{x\in\Phi^y}|x|^{-\alpha}}\right]\right)\mathrm{d}y\right). \quad (52)$$

(a) holds from the Slivnyak-Mecke Theorem [37], which states that the reduced Palm distribution of the PPP is equal to its original distribution. (b) follows from replacing the weighted summation of multiple complex Gaussian random variables by one complex Gaussian random variable, and (c) follows from the probability generating functional (PGFL) of the PPP. Treat $\sum_{x\in\Phi^y}|x|^{-\alpha}$ with given y as a random variable, which is the summation of N i.i.d. random variables $|x|^{-\alpha}$. Given y, the cdf of $|x|^{-\alpha}$ is $G_y(z) = 1 - \frac{1}{\pi R^2}S_y(z^{-\frac{1}{\alpha}})$, and the pdf is $g_y(z) = \frac{1}{\alpha\pi R^2}S_y'(z^{-\frac{1}{\alpha}})z^{-\frac{1}{\alpha}-1}$, where $S_y(u)$ is given by (12). Plugging (50) and (52) into (51), and using the polar coordinate transformation, we get the lemma.

APPENDIX C
PROOF OF LEMMA 3

The conditional success probability of the typical user served by a traditional BS given x_b is

$$\mathbb{P}\{\text{SIR} \geq \theta \mid x_b\} = \mathcal{L}_{I_{\text{CR}}}\left(\frac{\theta|x_b|^\alpha}{P_m}\right)\mathcal{L}_{I_{\text{BS}}}\left(\frac{\theta|x_b|^\alpha}{P_m}\right). \quad (53)$$

By Approximation 1 and using the results from [27] directly, we approximate the Laplace transform of the interference from the BSs when the interference from the BSs within the distance $|x_b|$ is not included as $\mathcal{L}_{I_{\text{BS}}}(s)$

$$\simeq \exp\left(-2\pi p\lambda_b e^{-\lambda_p\pi R^2}\int_{|x_b|}^{\infty}\left(1-\frac{1}{1+P_m s v^{-\alpha}}\right)v\mathrm{d}v\right)$$
$$= \exp\Bigg(-\pi p\lambda_b e^{-\lambda_p\pi R^2}s^{2/\alpha}P_m^{2/\alpha}\Gamma(1+2/\alpha)\Gamma(1-2/\alpha)$$
$$+\pi p\lambda_b e^{-\lambda_p\pi R^2}|x_b|^2 {}_2F_1\left(1,\frac{2}{\alpha};1+\frac{2}{\alpha};-\frac{|x_b|^\alpha}{P_m s}\right)\Bigg). \quad (54)$$

For the interference from the RRHs, since there is no interfering RRHs cluster center within the circular region $B(o, R)$, the Laplace transform of the interference from the RRHs is different from (49) and (52). In the RRH-selection mode, only one RRH in each interfering RRH cluster causes

interference to the typical user. By using the PGFL of the Poisson cluster process, we get the Laplace transform of the interference from the RRHs in the RRH-selection mode as follows.

$$\mathcal{L}_{I_{\mathrm{CR}}}(s) = \exp\Bigg(-p\lambda_p \int_{\mathbb{R}^2\setminus B(o,R)} \Bigg(1 - \int_{B(o,R)} \frac{1}{1+sP_s|x+y|^{-\alpha}} \cdot \frac{1}{\pi R^2}\mathrm{d}y\Bigg)\mathrm{d}x\Bigg). \quad (55)$$

In the cooperation mode, all RRH in the interfering RRH clusters cause interference to the typical user. By the similar derivations to (52), we get the Laplace transform of the interference from the RRHs in the cooperation mode as follows.

$$\mathcal{L}_{I_{\mathrm{CR}}}(s) = \exp\Bigg(-p\lambda_p \int_{\mathbb{R}^2\setminus B(o,R)} \Bigg(1 - \mathbb{E}_{\Phi^y}\Bigg[\frac{1}{1+sP_s\sum_{x\in\Phi^y}|x|^{-\alpha}}\Bigg]\Bigg)\mathrm{d}y\Bigg). \quad (56)$$

Having derived the Laplace transform of the interference, we get the success probability for the typical user served by a traditional BS in the lemma.

Appendix D
Proof of Lemma 4

In the RRH-selection mode, from (5) we get the conditional success probability $\mathbb{P}\{\mathrm{SIR} \geq \theta \mid \Phi_{\mathrm{CR}}, \Phi_{\mathrm{BS}}\} =$

$$\mathcal{L}_{I_{\mathrm{CR}}}\left(\frac{\theta|x_0|^\alpha}{P_s} \mid \Phi_{\mathrm{CR}}\right)\mathcal{L}_{I_{\mathrm{BS}}}\left(\frac{\theta|x_0|^\alpha}{P_s} \mid \Phi_{\mathrm{BS}}\right), \quad (57)$$

where $\mathcal{L}_{I_{\mathrm{CR}}}(s \mid \Phi_{\mathrm{CR}})$ and $\mathcal{L}_{I_{\mathrm{BS}}}(s \mid \Phi_{\mathrm{BS}})$ are evaluated as

$$\mathcal{L}_{I_{\mathrm{CR}}}(s \mid \Phi_{\mathrm{CR}}) = \prod_{x\in\Phi_p}\Big(\frac{p}{1+sP_s|x|^{-\alpha}} + 1 - p\Big). \quad (58)$$

$$\mathcal{L}_{I_{\mathrm{BS}}}(s \mid \Phi_{\mathrm{BS}}) \simeq \prod_{x\in\Phi_b'\setminus B(y_0,R)}\Big(\frac{p}{1+sP_m|x|^{-\alpha}} + 1 - p\Big). \quad (59)$$

Denote the number of frames needed until a successful transmission appears as Δ. With given $\Phi_{\mathrm{CR}}, \Phi_{\mathrm{BS}}$, Δ is a random variable with geometric distribution given by

$$\mathbb{P}(\Delta = k \mid \Phi_{\mathrm{CR}}, \Phi_{\mathrm{BS}}) = (1 - p\mathbb{P}\{\mathrm{SIR} \geq \theta \mid \Phi_{\mathrm{CR}}, \Phi_{\mathrm{BS}}\})^{k-1} p\mathbb{P}\{\mathrm{SIR} \geq \theta \mid \Phi_{\mathrm{CR}}, \Phi_{\mathrm{BS}}\}. \quad (60)$$

Therefore, the mean delay for the C-RAN user in the RRH-selection mode is $D_1 = \mathbb{E}[\Delta]$

$$= \mathbb{E}_{\Phi_{\mathrm{CR}},\Phi_{\mathrm{BS}}}\Bigg[\frac{1}{p\mathbb{P}\{\mathrm{SIR} \geq \theta \mid \Phi_{\mathrm{CR}}, \Phi_{\mathrm{BS}}\}}\Bigg]$$

$$= \frac{1}{p}\mathbb{E}_{\Phi_{\mathrm{CR}},\Phi_{\mathrm{BS}}}\Bigg[\prod_{x\in\Phi_p}\Big(\frac{p}{1+\theta|x_0|^\alpha|x|^{-\alpha}} + 1 - p\Big)^{-1} \prod_{x\in\Phi_b'\setminus B(y_0,R)}\Big(\frac{p}{1+\theta|x_0|^\alpha|x|^{-\alpha}P_m/P_s} + 1 - p\Big)^{-1}\Bigg]$$

$$\simeq \frac{1}{p}\mathbb{E}_{x_0,y_0}\Bigg[\exp\Bigg(-\lambda_p\int_{\mathbb{R}^2}\Bigg(1 - \frac{1}{\frac{p}{1+\theta|x_0|^\alpha|x|^{-\alpha}} + 1 - p}\Bigg)\mathrm{d}x - \lambda_b e^{-\lambda_p\pi R^2}\int_{\mathbb{R}^2\setminus B(y_0,R)}\Bigg(1 - \frac{1}{\frac{p}{1+\theta|x_0|^\alpha|x|^{-\alpha}P_m/P_s} + 1 - p}\Bigg)\mathrm{d}x\Bigg)\Bigg]$$

$$= \frac{1}{p}\mathbb{E}_{x_0,y_0}\Bigg[\exp\Bigg(\Bigg(\lambda_p + \lambda_b e^{-\lambda_p\pi R^2}\frac{P_m^{2/\alpha}}{P_s^{2/\alpha}}\Bigg)\frac{|x_0|^2\ C_0}{(1-p)^{1-2/\alpha}} - \frac{p\lambda_b}{1-p}e^{-\lambda_p\pi R^2}\int_0^\pi \eta(|y_0|,\psi) {}_2F_1\Big(1,\frac{2}{\alpha};1+\frac{2}{\alpha};-\frac{P_s\eta^{\alpha/2}(|y_0|,\psi)}{(1-p)\theta|x_0|^\alpha P_m}\Big)\mathrm{d}\psi\Bigg)\Bigg]. \quad (61)$$

Appendix E
Proof of Lemma 5

In the cooperation mode, we get the conditional ccdf of SIR given $\Phi_{\mathrm{CR}}, \Phi_{\mathrm{BS}}$ as

$$\mathbb{P}\{\mathrm{SIR} \geq \theta \mid \Phi_{\mathrm{CR}}, \Phi_{\mathrm{BS}}\} = \mathcal{L}_{I_{\mathrm{CR}}}\left(\frac{\theta}{P_s\sum_{x\in\Phi^{y_0}}|x|^{-\alpha}} \mid \Phi_{\mathrm{CR}}\right)\mathcal{L}_{I_{\mathrm{BS}}}\left(\frac{\theta}{P_s\sum_{x\in\Phi^{y_0}}|x|^{-\alpha}} \mid \Phi_{\mathrm{BS}}\right), \quad (62)$$

$\mathcal{L}_{I_{\mathrm{BS}}}(s \mid \Phi_{\mathrm{BS}})$ is given by (59). The conditional Laplace transform of interference from RRHs is

$$\mathcal{L}_{I_{\mathrm{CR}}}(s \mid \Phi_{\mathrm{CR}}) = \prod_{y\in\Phi_p}\Big(\frac{p}{1+sP_s\sum_{x\in\Phi^y}|x|^{-\alpha}} + 1 - p\Big). \quad (63)$$

Similar to the proof of Lemma 4, the mean delay for the C-RAN user in the RRH-selection mode is $D_2 =$

$$\frac{1}{p}\mathbb{E}_{\Phi^{y_0},y_0}\Bigg[\exp\Bigg(\int_0^\infty \mathbb{E}_{\Phi^{(r,0)}}\Bigg[\frac{2\pi p\lambda_p r}{1-p+\frac{\sum_{x\in\Phi^{y_0}}|x|^{-\alpha}}{\theta\sum_{x\in\Phi^{(r,0)}}|x|^{-\alpha}}}\Bigg]\mathrm{d}r + \frac{\lambda_b C_0(1-p)^{\frac{2}{\alpha}-1}P_m^{\frac{2}{\alpha}}}{(P_s\sum_{x\in\Phi^{y_0}}\frac{1}{|x|^\alpha})^{\frac{2}{\alpha}}e^{\lambda_p\pi R^2}} - \frac{p\lambda_b e^{-\lambda_p\pi R^2}}{1-p}\int_0^\pi \eta(|y_0|,\psi) {}_2F_1\Big(1,\frac{2}{\alpha};1+\frac{2}{\alpha};-\frac{P_s\eta^{\frac{\alpha}{2}}(|y_0|,\psi)}{(1-p)\theta P_m}\sum_{x\in\Phi^{y_0}}\frac{1}{|x|^\alpha}\Big)\mathrm{d}\psi\Bigg)\Bigg]. \quad (64)$$

References

[1] R. Li, Z. Zhao, X. Zhou, J. Palicot, and H. Zhang, "The prediction analysis of cellular radio access network traffic: From entropy theory to networking practice," *IEEE Commun. Mag.*, vol. 52, no. 6, pp. 234–240, Jun. 2014.

[2] M. Peng, Y. Li, Z. Zhao, and C. Wang, "System architecture and key technologies for 5G heterogeneous cloud radio access networks," *IEEE Netw.*, vol. 29, no. 2, pp. 6–14, Mar./Apr. 2015.

[3] 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.
[4] M. Tao, E. Chen, H. Zhou, and W. Yu, "Content-centric sparse multicast beamforming for cache-enabled cloud RAN," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6118–6131, Sep. 2016.
[5] S. T. Veetil, K. Kuchi, and R. K. Ganti. (2015). "Performance of cloud radio networks." [Online]. Available: https://arxiv.org/abs/1512.05904
[6] K. Chen, "C-RAN: The road towards green RAN," version 2, China Mobile Res. Inst., Beijing, China, 2011.
[7] J. Tang, W. P. Tay, and T. Q. S. Quek, "Cross-layer resource allocation with elastic service scaling in cloud radio access network," *IEEE Trans. Wireless Commun.*, vol. 14, no. 9, pp. 5068–5081, Sep. 2015.
[8] R. K. Ganti and M. Haenggi, "Interference and outage in clustered wireless ad hoc networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4067–4086, Sep. 2009.
[9] Y. Zhong and W. Zhang, "Multi-channel hybrid access femtocells: A stochastic geometric analysis," *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 3016–3026, Jul. 2013.
[10] C.-H. Lee and M. Haenggi, "Interference and outage in Poisson cognitive networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1392–1401, Apr. 2012.
[11] R. Tanbourgi, S. Singh, J. G. Andrews, and F. K. Jondral, "A tractable model for noncoherent joint-transmission base station cooperation," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4959–4973, Sep. 2014.
[12] 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.
[13] 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.
[14] Z. Yang, Z. Ding, and P. Fan, "Performance analysis of cloud radio access networks with uniformly distributed base stations," *IEEE Trans. Veh. Technol.*, vol. 65, no. 1, pp. 472–477, Jan. 2016.
[15] V. Suryaprakash, P. Rost, and G. Fettweis, "Are heterogeneous cloud-based radio access networks cost effective?" *IEEE J. Sel. Areas Commun.*, vol. 33, no. 10, pp. 2239–2251, Oct. 2015.
[16] M. Khan, R. S. Alhumaima, and H. Al-Raweshidy, "Reducing energy consumption by dynamic resource allocation in C-RAN," in *Proc. IEEE Eur. Conf. Netw. Commun. (EuCNC)*, Jun. 2015, pp. 169–174.
[17] L. Chen, H. Jin, H. Li, J.-B. Seo, Q. Guo, and V. Leung, "An energy efficient implementation of C-RAN in HetNet," in *Proc. IEEE 80th Veh. Technol. Conf. (VTC Fall)*, Sep. 2014, pp. 1–5.
[18] A. Lozano, R. W. Heath, and J. G. Andrews, "Fundamental limits of cooperation," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5213–5226, Sep. 2013.
[19] G. C. Alexandropoulos, P. Ferrand, J.-M. Gorce, and C. B. Papadias, "Advanced coordinated beamforming for the downlink of future lte cellular networks," *IEEE Commun. Mag.*, vol. 54, no. 7, pp. 54–60, Jul. 2016.
[20] A. Papadogiannis and G. C. Alexandropoulos, "The value of dynamic clustering of base stations for future wireless networks," in *Proc. IEEE Int. Conf. Fuzzy Syst.*, Jul. 2010, pp. 1–6.
[21] J. Park, N. Lee, and R. W. Heath, "Cooperative base station coloring for pair-wise multi-cell coordination," *IEEE Trans. Commun.*, vol. 64, no. 1, pp. 402–415, Jan. 2016.
[22] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 7, pp. 1029–1046, Sep. 2012.
[23] M. Haenggi, *Stochastic Geometry for Wireless Networks*. Cambridge, U.K: Cambridge Univ. Press, 2012.
[24] F. Baccelli and B. Blaszczyszyn, *Stochastic Geometry and Wireless Networks*, vol. 1. Norwell, MA, USA: Now Publishers, 2009.
[25] W. C. Cheung, T. Q. S. Quek, and M. Kountouris, "Throughput optimization, spectrum allocation, and access control in two-tier femtocell networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 561–574, Apr. 2012.
[26] S. B. Lowen and M. C. Teich, "Power-law shot noise," *IEEE Trans. Inf. Theory*, vol. 36, no. 6, pp. 1302–1318, Nov. 1990.
[27] 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.
[28] M. Haenggi, "The Local Delay in Poisson Networks," *IEEE Trans. Inf. Theory*, vol. 59, no. 3, pp. 1788–1802, Mar. 2013.
[29] F. Baccelli and B. Błaszczyszyn, "A new phase transitions for local delays in MANETs," in *Proc. IEEE INFOCOM*, San Diego, CA, USA, Mar. 2010, pp. 1–9.
[30] Y. Zhong, W. Zhang, and M. Haenggi, "Managing interference correlation through random medium access," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 928–941, Feb. 2014.
[31] N. Lee, D. Morales-Jimenez, A. Lozano, and R. W. Heath, "Spectral efficiency of dynamic coordinated beamforming: A stochastic geometry approach," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 230–241, Jan. 2015.
[32] F. Baccelli and A. Giovanidis, "A stochastic geometry framework for analyzing pairwise-cooperative cellular networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 794–808, Feb. 2015.
[33] W. Choi and J. G. Andrews, "Downlink performance and capacity of distributed antenna systems in a multicell environment," *IEEE Trans. Wireless Commun.*, vol. 6, no. 1, pp. 69–73, Jan. 2007.
[34] Y. Zhong, M. Haenggi, T. Q. S. Quek, and W. Zhang, "On the stability of static poisson networks under random access," *IEEE Trans. Commun.*, vol. 64, no. 7, pp. 2985–2998, Jul. 2016.
[35] Y. Zhong, T. Q. S. Quek, and X. Ge, "Heterogeneous cellular networks with spatio-temporal traffic: Delay analysis and scheduling," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 8, pp. 1–14, Aug. 2017.
[36] E. Oh, B. Krishnamachari, X. Liu, and Z. Niu, "Toward dynamic energy-efficient operation of cellular network infrastructure," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 56–61, Jun. 2011.
[37] D. Stoyan, W. Kendall, and J. Mecke, *Stochastic Geometry and its Applications*, 2nd ed. Hoboken, NJ, USA: Wiley, 1996.

Yi Zhong (S'12–M'15) received the B.S. and Ph.D. degrees in electronic engineering from the University of Science and Technology of China, in 2010 and 2015, respectively. In 2012, he was a visiting student with Prof. M. Haenggi's group at the University of Notre Dame, Notre Dame, IN, USA. In 2013, he was a Research Intern with Qualcomm Inc., Corporate Research and Development, Beijing. From 2015 to 2016, he was a Post-Doctoral Research Fellow with the Wireless Networks and Decision Systems Group, Singapore University of Technology and Design, led by Prof. T. Q. S. Quek. He is currently an Assistant Professor with the School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan, China. His main research interests include heterogeneous and femtocell-overlaid cellular networks, wireless ad hoc networks, stochastic geometry, and point process theory.

Tony Q. S. Quek (S'98–M'08–SM'12) received the B.E. and M.E. degrees in electrical and electronics engineering from the Tokyo Institute of Technology, and the Ph.D. degree in electrical engineering and computer science from the Massachusetts Institute of Technology. He is a tenured Associate Professor with the Singapore University of Technology and Design (SUTD). He also serves as the Associate Head of the ISTD Pillar and the Deputy Director of the SUTD-ZJU IDEA. He is a co-author of the book *Small Cell Networks: Deployment, PHY Techniques, and Resource Allocation* (Cambridge University Press, 2013) and the book *Cloud Radio Access Networks: Principles, Technologies, and Applications* (Cambridge University Press, 2017). His main research interests are the application of mathematical, optimization, and statistical theories to communication, networking, signal processing, and resource allocation problems. Specific current research topics include heterogeneous networks, wireless security, Internet of Things, and big data processing.

He has been actively involved in organizing and chairing sessions, and has served as a technical program committee member and symposium chair of a number of international conferences. He is serving as the Workshop Chair at the IEEE Globecom in 2017, the Tutorial Chair at the IEEE ICCC in 2017, and the Special Session Chair at the IEEE SPAWC in 2017. He is currently an Elected Member of the IEEE Signal Processing Society SPCOM Technical Committee. He was an Executive Editorial Committee Member of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, an Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS, and an Editor of the IEEE WIRELESS COMMUNICATIONS LETTERS.

Dr. Quek was honored with the 2008 Philip Yeo Prize for Outstanding Achievement in Research, the IEEE Globecom 2010 Best Paper Award, the 2012 IEEE William R. Bennett Prize, the IEEE SPAWC 2013 Best Student Paper Award, the IEEE WCSP 2014 Best Paper Award, the 2015 SUTD Outstanding Education Awards–Excellence in Research, the 2016 Thomson Reuters Highly Cited Researcher, and the 2016 IEEE Signal Processing Society Young Author Best Paper Award.

Wenyi Zhang (S'00–M'07–SM'11) received the bachelor's degree in automation from Tsinghua University, in 2001, and the master's and Ph.D. degrees in electrical engineering from the University of Notre Dame, IN, USA, in 2003 and 2006, respectively. He was affiliated with the Communication Science Institute, University of Southern California, as a Post-Doctoral Research Associate, and with Qualcomm Inc., Corporate Research and Development. He is currently a faculty member of the Department of Electronic Engineering and Information Science, University of Science and Technology of China. His current research interests include wireless communications and networking, information theory, and statistical signal processing.