Multi-User Millimeter Wave Cloud Radio Access Networks with Hybrid Precoding

Multi-User Millimeter Wave Cloud Radio Access Networks with Hybrid Precoding

Oluwatayo Y Kolawole, *Student Member, IEEE*, Satyanarayana Vuppala, *Member, IEEE*, and Tharmalingam Ratnarajah, *Senior Member, IEEE*

***Abstract*—This paper investigates the performance of cloud radio access networks (CRANs) for a downlink multi-user millimeter wave (mmWave) system, where randomly distributed remote radio heads (RRHs) supported by a central base band unit (BBU) communicate with multiple mobile equipment (MEs). The fronthaul and access link transmissions are implemented with mmWave frequency bands. The RRHs and MEs are modelled as independent poisson point processes (PPPs). We characterize the outage probability, average latency and throughput of this system under essential factors, such as blockages, RRH density and path loss. Two specific ME association scenarios are considered: best channel participation (BCP) and nearest neighbour participation (NNP). We derived for both scenarios, closed-form expressions of outage probability in the noise-limited case, and upper and lower bounds of outage probability in the interference-limited case. Our results show blockages and path loss to have a positive effect of decreasing outage probability. Larger antenna arrays are shown to compensate for communication degradation (outage performance and latency) with higher RRH deployment, which can be considered a trade-off between inter-cluster interference (ICI) and RRH density. Finally, we show that for the ME association process, BCP is the most viable for mmWave CRAN systems due to its outperforming NNP.**

***Index Terms*—cloud radio access networks, hybrid precoding, millimeter-wave networks, poison point processes**

I. Introduction

The increasing prevalence of high-speed, real-time data applications such as social networking, and high-quality (HQ) wireless video streaming is driving greater demand for more bandwidth and faster data rates from mobile users. For example, smart phone traffic, which has experienced tremendous increase over the last two decades, shows no signs of slowing and is predicted to surpass half the global traffic by 2020 [1]. Moreover, the existing frequency spectrum is congested; as a result, there is an urgent need for revolutionary approaches to meet these demands.

Millimeter wave (mmWave) frequencies have triggered increasing research attention as imperative bands for the realisation of 5G networks, because of their potential to provide large chunks of unused spectrum for growth in data rates and capacity. The prediction of utilising mmWave in communication are 10-1000 fold increase in data rates, capacities, spectrum and energy efficiencies, as well as latencies of less than 1 ms for all connected devices [2]. The authors in [3] show that wide coverage and high data rates are achievable using Constant Envelope OFDM (CE-OFDM) based mmWave systems through the use of trellis coding which effectively exploits frequency diversity and finally, the work in [4] shows how mmWave bands can be used to provide high bandwidth data transfer operations to ensure reliable connectivity especially in congested urban areas.

Propagation in mmWave communication distinctly differs from that in conventional microwave networks as it is more involved than changing the carrier frequency [5]. First, mmWave systems are characterised by massive arrays of steerable antennas for directional beamforming. This is to mitigate the effect of path losses, which are more significant in propagation at higher frequencies [6]. Second, mmWave propagation is more susceptible to blockages, such as building walls and shadowing, making line-of-sight (LOS) and non-line-of-sight (NLOS) links an integral part of mmWave cellular transmission [5]. Third, mmWave communication is affected by atmospheric conditions, although these effects are negligible for small distances as shown in [6] and [7]. The work in [8], which studied propagation at mmWave frequencies in a dense urban environment, shows that mmWave systems are viable for propagation with low gain antennas. This implies that a dense network is required for mmWave propagation. In addition, [9]–[12] show the feasibility of mmWave propagation using LOS and NLOS models.

Network architecture is critical to providing robust networks for future generations, and so cloud radio access networks (CRANs) have gained heightened interest in research and industry in recent years. CRAN is a potential architecture for dense deployment of remote radio heads (RRHs) that provides increased capacity with reduced cost and higher efficiencies [13], [14]. In the scenario where RRH nodes are fixed at specific locations, the performance of CRAN architecture for MIMO channels is studied in [13], [15], [16] showing spectral efficiency, capacity and energy efficiencies respectively. For randomly distributed RRHs in single and multi-cell scenarios, outage probability, ergodic capacity and rate analysis were studied in [17]–[19]. The authors in [20] investigated the use mobile small cells to replace RRHs in a CRAN and showed that gains can be achieved in terms of cost, spectral efficiency, and average user throughput.

These studies demonstrating different performance metrics attest to the feasibility of CRAN as a distributed antenna network architecture; however, they have only been studied

O. Kolawole, S. Vuppala, and T. Ratnarajah are with the University of Edinburgh, King's Building, Edinburgh, UK, EH9 3JL. E-mails: {o.kolawole,s.vuppala, t.ratnarajah}@ed.ac.uk. Corresponding author: S. Vuppala.

The work of S. Vuppala, and T. Ratnarajah was supported by the U.K Engineering and Physical Sciences Research Council (EPSRC) under grants EP/L025299/1 and the work of O. Kolawole was supported by the PTDF agency of the Federal government of Nigeria with the mandate to build capacities and capabilities in Nigerias Oil and Gas Industry through the development of human capacities and promotion of research and acquisition of relevant technologies.

only in the sub 3 GHz spectrum. Given that CRAN can deploy dense networks which is a requirement for mmWave propagation, this paper presents an analysis of CRAN architecture using mmWave frequencies.

The contributions of this paper are to design a system that couples two potential candidates - mmWave and CRAN - for next generation communication into multi-user mmWave CRAN systems. To the authors' best knowledge, this network setup has not been analysed previously. In addition, we build a tractable system using stochastic geometry to show the effects of outage probability, average latency and throughput against signal-to-interference-noise-ratio (SINR) threshold for mmWave CRAN systems. The performance of this system is studied for two association schemes. In the first, the RRH or BBU with the best channel participates in the transmission where the total power is given to this channel, also called best channel participation (BCP). In the second, the participating RRH is the one closest to the typical ME, also called nearest neighbour participation (NNP). We characterize this system under key factors including blockages, antenna gain, path losses, and cluster radius. Exact closed-form expressions of outage probability considering LOS and NLOS channels are derived for both scenarios in the noise-limited case (neglecting inter-cluster interference ICI). Upper and lower analytical approximations of outage probability are also developed for the interference-limited case.

The paper is organised as follows. In section II, we define the system model, channel model and blockage model. In section III, we present preliminary statistics used in the analysis of the model and then derive solutions for outage probability and throughput in section IV. We extend our analysis to the case considering ICI in section IV-B. We present Numerical results validated by simulations in section V and conclude in section VI.

The following notations are used throughout this paper: $\mathbf{B}$, $\mathbf{b}$, b represent a matrix, a vector and a scalar respectively. $\mathbf{B}^H$, $\mathbf{B}^{-1}$, and $\mathbf{B}^T$ are the Hermitian, inverse and transpose of $\mathbf{B}$ respectively. $\mathcal{CN}(m, v)$ is a circularly symmetric complex normal Gaussian random vector with mean, m and variance, v. $\mathbb{E}[.]$ denotes expectation. All other notations are defined as used.

II. System Model

We consider the downlink of a multi-user mmWave CRAN system which consists of a mmWave transmission network, a baseband unit (BBU), multiple mobile equipment (MEs) and multiple remote radio heads (RRHs). The BBU is the central intelligence unit connected to RRHs by high bandwidth links called fronthaul (FH) links whereas the links between the RRHs and MEs are referred to as access links. The interface in the FH is standardised as common public radio interface (CPRI) [21]. To meet the strict requirements imposed on the FH by CPRI, the proposed solution is to split the functionality (from baseband to packet processing and radio functions such as amplification, frequency conversion, A/D and D/A conversion) between the BBU and RRHs. Thus, the BBU performs functions such as baseband and packet processing while the RRHs perform the radio functions and transfer (send/receive) data from BBU to MEs in the system as illustrated in Fig. 1. Although dedicated fibre links are commonly used at the FH links, recent European projects including 5G PPP [22] and 5G-XHaul [23] propose that mmWave bands can be used for transmission of FH data because the centralisation in CRAN requires a large consumption of fibre cores which are scarce and expensive to deploy. Therefore, the transmission links at both FH and access in this system operate at mmWave frequency bands.

Each RRH is distributed in a randomly ordered fashion in $\mathcal{R}^2$ using independent homogeneous poisson point process (PPP), Φ_{RRH} and intensity, λ_{RRH}. Similarly, MEs are randomly distributed as PPP, Φ_{ME} with intensity λ_{ME}. The positioning of MEs are not dependent on the location of RRHs; hence, all point processes are independent of each other. The number of antennas on each RRH, ME and the BBU are denoted as $\mathrm{N_{RRH}}$, $\mathrm{N_{ME}}$ and $\mathrm{N_T}$ respectively. We assume there are K-number of MEs in the system such that $\mathrm{N_{RRH}} \geq K$ or $\mathrm{N_T} \geq K$, and focus on a circular region which we will refer to as cluster $\mathcal{G}$ illustrated in Fig. 1. The subsequent analysis will focus on the access link transmissions between RRHs and MEs.

A. Association Scenarios

When an ME associates itself with an RRH and makes a request, the RRHs help the signals from the ME to be decoded by the BBU. Thus, an ME is serviced through the access link (from ME to RRH) and fronthaul link (from tagged RRH to BBU). To analyse the performance of this system, we consider two schemes based on how the ME associates with an RRH in cluster $\mathcal{G}$ namely:

- BCP: In this scenario, BBU determines which RRH has the best channel in cluster $\mathcal{G}$ for transmission to MEs.
- NNP: This scenario is considered for its ability to reduce overhead whilst achieving acceptable performance. The metric is based on distance where the RRH closest to an ME is selected for transmission.

B. Transmission Model

It is worth noting that propagation from any RRH to each ME is via a fully connected hybrid precoder that combines radio frequency (RF) and baseband (BB) precoding. We assume that multiple MEs are served via one stream per user, thus it is sufficient that each ME employs an RF-only combiner to decode the transmitted signal as described in [24]. Hybrid precoding is also used to cancel out the unwanted signals of other MEs.

The clustered mmWave channel between the RRH located at position y and the ME located at x in $\mathcal{G}$, $\mathbf{H}_{y,x}$ is written as

$$\mathbf{H}_{y,x} = \sqrt{\frac{\mathrm{N_r N_{ME}}}{L(y,x)}} \sum_{u=1}^{\mathrm{W}_{y,x}} \eta_{u,y,x} \mathbf{a}_{\mathrm{ME}}(\theta_{u,y,x}) \mathbf{a}^{\mathrm{H}}_{\mathrm{RRH}}(\phi_{u,y,x}), \tag{1}$$

where $\eta_{y,x}$ is the complex gain assumed to be a small scale Rayleigh fading distribution [25], θ is the angle of arrival,

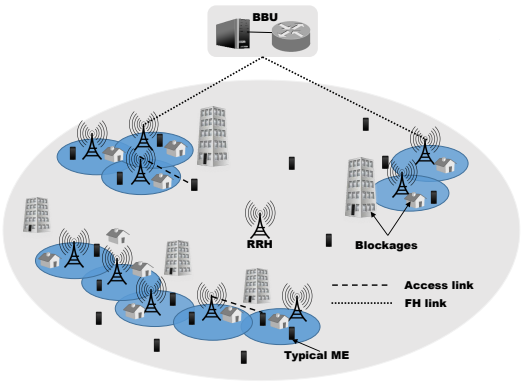

Fig. 1: An illustration of a multi-user mmWave CRAN system model.

ϕ is the angle of departure, $\mathrm{W}_{y,x}$[1] is the number of paths from RRH at y to ME at x, $L(y,x)$ is the path loss given as $L(y,x) = r_{y,x}^{\alpha_i}$, where α_i denotes the path loss exponent, and can either be LOS and NLOS depending on the link between them, $\mathbf{a}_{\mathrm{RRH}}$ and $\mathbf{a}_{\mathrm{ME}}$ are the array response vectors of each RRH and ME respectively.

Due to the sparsity of mmWave channels, we assume that all scattering happens in the azimuth plane, and model these array vectors as uniform linear arrays ULAs written as

$$\mathbf{a}_{\mathrm{RRH}}(\phi) = \frac{1}{\sqrt{\mathrm{N_r}}}\left[1, e^{\mathrm{jkd}\sin(\phi)}, \ldots, e^{(\mathrm{N_r}-1)\mathrm{jkd}\sin(\phi)}\right], \quad (2)$$

where $\mathrm{k} = \frac{2\pi}{\lambda}$, λ is the wavelength and d is the distance between antenna elements. $\mathbf{a}_{\mathrm{ME}}$ and $\mathbf{a}_{\mathrm{BBU}}$ are defined in like fashion.

C. Received Signal

Consider the links between the k^{th} RRH located at y in a circular region, $\mathcal{G}$, communicating with several MEs in an instant. Let us denote $\mathcal{U}$ as the maximum number of MEs in $\mathcal{G}$ to which a single RRH can instantaneously communicate to, and assume that the maximum value of $\mathcal{U}$ is the number of RF chains ($\mathcal{U} \leq \mathrm{N_{RF}}$) because hybrid precoding gain is constrained by $\min(\mathrm{N_{RF}}, \mathrm{K})$ [26].

The baseband and RF precoders are presented as $\mathbf{V}_{\mathrm{BB}}^k$ and $\mathbf{V}_{\mathrm{RF}}^k$ respectively. Thus,$\mathbf{V}_{\mathrm{BB}}^k = [\mathbf{v}_{\mathrm{BB}}^{k,1}, \mathbf{v}_{\mathrm{BB}}^{k,2}, \ldots, \mathbf{v}_{\mathrm{BB}}^{k,\mathcal{U}}]$ and $\mathbf{V}_{\mathrm{RF}}^k = [\mathbf{v}_{\mathrm{RF}}^{k,1}, \mathbf{v}_{\mathrm{RF}}^{k,2}, \ldots, \mathbf{v}_{\mathrm{RF}}^{k,\mathcal{U}}]$.

After passing $\mathbf{x}_k$ through $\mathbf{V}_{\mathrm{BB}}^k$, $\mathbf{V}_{\mathrm{RF}}^k$,and $\mathrm{U}_{\mathrm{RF}}^k$ respectively, the received signal, y_0 is given as

$$y_0 = \varrho_{k,1}\mathbf{h}_{k,0}^{\mathrm{eff}}\mathbf{v}_{\mathrm{BB}}^{k,0}s_1 + \sum_{\mathrm{g}\in\Phi_{\mathrm{ME}},\mathrm{g}\neq 1}\varrho_{k,\mathrm{g}}\mathbf{h}_{k,\mathrm{g}}^{\mathrm{eff}}\mathbf{v}_{\mathrm{BB}}^{k,\mathrm{g}}s_{\mathrm{g}} + z_0, \quad (3)$$

where $\mathbf{h}_{k,0}^{\mathrm{eff}} = (\mathbf{u}_{\mathrm{RF}}^0)^{\mathrm{H}}\mathbf{h}_{k,0}\mathbf{V}_{\mathrm{RF}}^k$ is called the effective channel, $\mathbf{u}_{\mathrm{RF}}^0$ is the RF combiner, $\varrho_{k,0}$ is the average received power given as $\varrho_{k,0} = \frac{P_T}{K}$, z_0 is noise, $z_0 \sim \mathcal{CN}(0, N_0)$ and P_T is the total transmit power enforced by the normalization of $\mathrm{V_{BB}}$ [27] such that $||\mathbf{V}_{\mathrm{RF}}\mathbf{V}_{\mathrm{BB}}||_F^2 = \mathrm{N_d}$.

We employ near-optimal methods proposed in [26] and [27] to design the hybrid precoder. Accordingly, $\mathbf{u}_{\mathrm{RF}}^0 = \mathbf{a}_{\mathrm{ME}}(\theta_{k,0})$ and $\mathbf{v}_{\mathrm{RF}}^{k,0} = \mathbf{a}_{\mathrm{RRH}}(\phi_{k,0})$. For interference management at baseband level, zero-forcing (ZF) is utilised at the transmitters. Thus, $\mathbf{v}_{\mathrm{BB}}^{k,0} = (\mathbf{h}_{k,0}^{\mathrm{eff}})^{\mathrm{H}}\left(\mathbf{h}_{k,0}^{\mathrm{eff}}(\mathbf{h}_{k,0}^{\mathrm{eff}})^{\mathrm{H}}\right)^{-1}$.

D. Blockage Model

To model blockages in this mmWave network, we use the probabilistic model validated in [9] which defines a link of length a as LOS with probability p_{L}[2], if a is less than or equal to radius of some region D, $a \leq \mathrm{D}$ [3]. However, when $a > \mathrm{D}$, the link is NLOS with probability of p_{N}.

E. SINR Model

To model the SINR of a mmWave network, we consider the two different schools of thoughts about propagation in mmWave networks; On one hand, the authors in [28]–[30] mention that mmWave networks tend to be noise-limited due to high blockage density which make signals from unwanted sources negligible. Whereas, authors in [31]–[33] considering high base station density with moderate blockages assume that mmWave networks tend to be interference-limited. Therefore, we model the SINR in both noise-limited and interference-limited cases:

1) Noise-limited case: In this scenario, we consider only intra cluster interference from other users associated with an RRH. The analysis is performed for a typical ME following from Slivnyak's theorem. The SINR between the k^{th} RRH and the typical ME is gotten from equation (3). Denoting SINR as ρ, we obtain

$$\rho_{k,0} = \frac{\varrho_{k,0}\|\mathbf{h}_{k,0}^{\mathrm{eff}}\mathbf{v}_{\mathrm{BB}}^{k,0}\|^2}{\sum_{\substack{\mathrm{g}\in\mathrm{n_y}\\ \mathrm{g}\neq 0}}\varrho_{k,\mathrm{g}}\|\mathbf{h}_{k,\mathrm{g}}^{\mathrm{eff}}\mathbf{v}_{\mathrm{BB}}^{k,\mathrm{g}}\|^2 + \sigma_z^2}, \quad (4)$$

where the first term of the denominator denotes intra-cluster interference to the typical ME which is zero after successful interference cancellation, and σ_z^2 is the noise.

For tractability in the statistical analysis in this paper, we transform the SNR in (4) given as

$$\rho_{k,0} \approx \frac{\mathrm{B}_{k,0}\,\eta_{k,0}^2\,\gamma(\mathrm{N_r},\mathcal{U})}{r_{k,0}^{\alpha_i}\,\sigma_z^2}, \quad (5)$$

where $\mathrm{B}_{k,0} = \varrho_{k,0}\,\mathrm{N_r}\,\mathrm{N_{ME}}$, and $\gamma(\mathrm{N_r},\mathcal{U})$ is the precoding penalty defined as

$$\gamma(\mathrm{N_r},\mathcal{U}) = \begin{cases} 1, & \text{w.p. } \left(1-\frac{1}{\mathrm{N_r}}\right)^{\mathcal{U}-1} \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

The explanation of this penalty can be found in [12, Proposition 1].

[1]Unit paths are assumed for each ME in this system.

[2]It is important to note that the values of p_L are dependent on the geography of an area i.e. a low value is assumed for dense urban areas and a higher one for semi-urban areas.

[3]D is defined in [9] as the radius of a circle from actual measurements in urban regions of New York

2) Interference-limited case: In this scenario, the typi ME experiences interference from within the cluster as wel from neighbouring clusters. It should be noted that the RR from other clusters causing interference are those minima affected by blockages [4]. Consequently, the received signal a applying the RF precoders and combiners to the transmit signal is given as

$$\mathbf{y}_0 = \varrho_{k,0}\, \mathbf{h}_{k,0}^{\text{eff}}\, \mathbf{s}_0 + \sum_{\mathrm{g}\in\Phi'_{\mathrm{ME}},\mathrm{g}\neq 0} \varrho_{k,g}\ \mathbf{h}_{k,g}^{\text{eff}}\, \mathbf{s}_g + \mathrm{ICI} + \mathbf{z}_0,$$

where ICI is the inter-cluster interference which will modelled as $\mathbf{I}_{\Phi'_{\mathrm{BS}}}$, and $\mathbf{z}_0$ is the complex Gaussian noise wit the cluster, $\mathbf{z}_0 \sim \mathcal{CN}(0, \sigma^2)$.

Thus, the SINR of the typical ME served by the k^{th} RRH is gotten from the received signal in (7) and expressed as

$$\mathrm{SINR_{k,0}} = \frac{\varrho_{k,0}\ \left|\mathbf{h}_{k,0}^{\text{eff}}\mathbf{v}_{\mathrm{BB}}^{0}\right|^2}{\sigma_z^2 + \sum\limits_{\mathrm{g}\in\Phi'_{\mathrm{ME}},\mathrm{g}\neq 0} \varrho_{k,g}\ \left|\mathbf{h}_{k,g}^{\text{eff}}\mathbf{v}_{\mathrm{BB}}^{g}\right|^2 + \mathbf{I}_{\Phi'_{\mathrm{BS}}}}, \quad (8)$$

where $\mathbf{I}_{\Phi'_{\mathrm{BS}}} = \sum_{b\in\Phi'_{\mathrm{BS}}} \varrho_b \left|\mathbf{h}_{b,v}^{\text{eff}}\mathbf{v}_{\mathrm{BB}}^{b}\right|^2$ outside cluster, $\mathcal{G}$, $\varrho_b = \frac{\mathrm{P_{BS}}}{\mathcal{U}_b}$, $\mathcal{U}_b$ is the number of users connected to a BSs in another cluster, and $\Phi'_{\mathrm{BS}} = \Phi_{\mathrm{BS}} \setminus \mathcal{G} \cap \Phi_{\mathrm{BS}}$.

Next, the SINR in (8) is approximated for the statistical analysis [5] and expressed as:

$$\mathrm{SINR_{k,0}} \approx \frac{G_0\, |\eta_{\mathrm{k,0}}|^2\ \mathrm{r}_{k,0}^{-\alpha_i}\, \gamma(\mathrm{N_r}, \mathcal{U})}{\sigma_z^2 + \mathbf{I}_{\Phi'_{\mathrm{BS}}}}, \quad (9)$$

where $G_0 = \varrho_{k,0}\, \mathrm{N_{ME}N_r}$, $r_{k,0}$ is the distance between the typical ME, and kth RRH, α_i is the path loss exponent which could be LOS or NLOS,[6] $\gamma(\mathrm{N_r}, \mathcal{U})$ is the precoding penalty similarly defined in (6), $\mathbf{I}_{\Phi'_{\mathrm{BS}}}$ is the inter-cluster interference given as

$$\mathbf{I}_{\Phi'_{\mathrm{BS}}} = \sum_{b\in\Phi'_{\mathrm{BS}},v\neq 0} G_{\mathrm{v}}\, |\eta_{\mathrm{b,v}}|^2\, \mathrm{r}_{\mathrm{v}}^{-\alpha} \times \sum_{v\in\mathcal{U}_v} ||\mathbf{a}_{\mathrm{ME}}(\theta_{v,y,x})\mathbf{a}_{\mathrm{RRH}}^{\mathrm{H}}(\phi_{v,y,x})||^2, \quad (10)$$

where G_{v} is the gain similarly defined as G_{n} but scaled by the precoding penalty [7]. Now, we leverage on the analysis from [33] such that $\mathbf{I}_{\Phi'_{\mathrm{BS}}}$ can further be expressed as

$$\mathbf{I}_{\Phi'_{\mathrm{BS}}} = \sum_{b\in\Phi'_{\mathrm{BS}},v\neq 0} G_{\mathrm{v}}\, |\eta_{\mathrm{b,v}}|^2\, \mathrm{r}_{\mathrm{v}}^{-\alpha} \Upsilon_v, \quad (11)$$

where $\Upsilon_v = \begin{cases} 1, & \theta_{v,y,x} = \phi_{v,y,x} \\ \tau_{\mathrm{RRH}}, & \text{otherwise,} \end{cases}$ where $\tau_{\mathrm{RRH}} < 1$.

[4] Although we focus on downlink transmission, the received signal for uplink transmissions, where multiple MEs transmit independent messages to their associated RRHs can be formulated similarly. Interference is caused by the other MEs associated to the same RRH as well as outwith the cluster and these interferences can be characterised with similar stochastic geometry tools.

[5] $\eta = \mathbf{h}_{\mathrm{EFF}}^{m,n}\mathbf{v}_{\mathrm{BB}}^{n}$ and represents the dominant element in $\mathbf{h}_{\mathrm{EFF}}$.

[6] The parameters for α used in simulations are adopted from measurements validated in [8], [9], [34].

[7] For simplicity we assume the sum of penalties from BSs in interfering clusters to be constant.

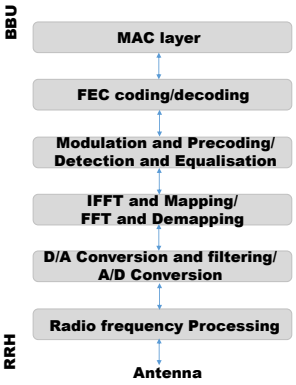

Fig. 2: CRAN Fronthaul link logical structure [36].

F. Fronthaul Model

For completeness of the paper, we include the FH model although this FH link is not the focus of our analysis. It is important to note that CPRI imposes strict requirement on the FH network [21] and to achieve functionality between BBU and RRHs, the BBU does not have to be physical. This implies that operators can dynamically maps radio signals from the BBU to any RRH using software-defined networking (SDN) concepts [35]. Fig. 2 presents a logical structure illustrating the downlink and uplink processing chain of the FH link. In the downlink, data packets for the MEs are processed at the MAC layer, there appropriate headers and schedules are added on followed by forward error correction (FEC) encoding and then, modulation and precoding. These operations are performed with channel state information available at the BBU. Next, inverse Fourier transform is performed to map ME data to physical resources in time slots and sub-carriers. After, the data is converted to the analogue domain using a digital to analogue converter (DAC) and then up-converted to the carrier frequency and transmitted to the antenna of the RRH. More details of the processing can be found in [36]. Note that for the uplink, the process is reversed as indicated by the upward flow of fig. 2.

G. Traffic Model

To measure the quality of service (QoS) and model the traffic delivery from the BBU to an ME via its associated RRH in this mmWave CRAN system, we adopt the queueing theory in [37]. Let us denote a typical ME as ME_0 and it's associated RRH as RRH_j. We assume that the traffic arrival to MEs in the network follows a PPP with parameter λ which denotes the arrival rate per unit area and the traffic load is distributed exponentially with the average traffic load of μ [8]. On the account of a constant FH rate, it is reasonable to assume that the traffic delivery time of the FH is also an exponential distribution. Thus, the traffic delivery in FH actualizes an M/M/1 queuing system [37].

[8] This assumption is based on the fact that the MEs are distributed according to a PPP, Φ_{ME}.

Similarly, in the access links, the traffic arrival to ME_0 follows a PPP with average arrival rate of λ_0. However, since the MEs at various locations associated with RRH_j have different rates depending on the channel conditions, the service time in RRH_j is generally distributed [38]. Therefore, the traffic delivery from RRH_j to ME_0 realises an M/G/1 queueing model [37]. Accordingly, the traffic delivery from BBU to ME_0 via RRH_j is modelled as a combined queue system.

III. Introductory Statistics

This section gives preliminary statistics which will be used in subsequent analysis.

Channel statistics: We present the statistical properties of mmWave channel used in the access links. Given that η is the small scale Rayleigh fading and is distributed as a complex Gaussian distribution, $\eta \sim \mathcal{CN}(0,1)$, we assume that the transmitted symbols, s_k are also Gaussian distributed i.e. $x_k \sim \mathcal{CN}(0,1)$. Thus, if $s_k \sim \mathcal{CN}(0,1)$, then the probability distribution function (PDF) of the SNR, $\rho_{k,0}$ is F-scaled [39] and expressed as

$$f_{\rho_{k,1}}(z) = \left(1-\frac{1}{\mathrm{N_r}}\right)\frac{\mathcal{U}-1}{}\frac{\mathrm{N_r}!\,(B_{k,o})^K z^{\mathrm{N_r}-K}}{(K-1)!(\mathrm{N_r}-K)!(B_{k,0}+z)^{\mathrm{N_r}+1}}. \tag{12}$$

The work in [40] shows that the F-distribution can be approximated to Chi-square distribution. Thus, we rewrite (12) in terms of Chi-squared distribution with $2(\mathrm{N_r}-K+1)$ degrees of freedom

$$f_{\rho_{k,1}}(z) \approx \left(1-\frac{1}{\mathrm{N_r}}\right)^{\mathcal{U}-1}\sum_{i\in \mathrm{L,N}}\frac{(K r_{k,0}^{\alpha_i})^{\mathrm{N_r}-K+1}}{(\mathrm{N_r}-K)!} \times \frac{1}{(B_{k,0})^{\mathrm{N_r}-K+1}} z^{\mathrm{N_r}-K}\ \mathrm{e}^{-\frac{z\,\mathrm{K}\,\mathrm{r}_{\mathrm{k,0}}^{\alpha_i}}{\mathrm{B_{k,0}}}}. \tag{13}$$

Moreover, integrating (13) for a given distance, $r_{k,0}$ yields the conditional cumulative distribution function ($F_{\rho_{k,0}|r_{k,0}}$) expressed as

$$F_{\rho_{k,0}|r_{k,0}}(z) = \left(1-\frac{1}{\mathrm{N_r}}\right)^{\mathcal{U}-1}\sum_{i\in \mathrm{L,N}} p_i\,\frac{1}{(\mathrm{N_r}-K)!} \times \gamma\left(\mathrm{N_r}-K+1, \frac{z\,K\,r_{k,0}^{\alpha_i}}{B_{k,0}}\right). \tag{14}$$

Outage probability of a generic RRH: The outage probability of any RRH in our system is given as

$$\mathbb{P}_{out}(z) = \int_0^R f_r(r) F_{\rho|r}(z)\,\mathrm{d}r, \tag{15}$$

where $f_r(r)$ represents the uniform distribution of all RRHs in $\mathcal{G}$ given as $f_r(r) = \dfrac{2\,r}{R^2}$, $F_{\rho|r}$ is the CCDF defined in (14). It is worthwhile to notice that the CDF of received SNR at k-th RRH from BBU follows from the equation (14) and is given as

$$F_{\rho_f}(\xi) = \left(1-\frac{1}{\mathrm{N_t}}\right)^{\mathcal{M}-1}\sum_{i\in \mathrm{L,N}}\frac{p_i\,\gamma\left(\mathrm{N_t}-M+1, \frac{\xi\,M\,r_f^{\alpha_i}}{B_f}\right)}{(\mathrm{N_t}-M)!}. \tag{16}$$

In the following proposition, we show the overall outage probability of the k^{th} RRH which will be used in analysis of the association schemes.

Proposition 1. *The outage probability of the k^{th}* RRH *to the typical* ME *is given as*

$$\mathbb{P}_{out}(t) = \left(1-\frac{1}{\mathrm{N_r}}\right)\sum_{i\in \mathrm{L,N}}^{\mathcal{U}-1} p_i\,\frac{2\,(t\,K)^{-\frac{2}{\alpha_i}}}{R^2} \times \left(1-\sum_{m=0}^{N_r-K}\frac{\left[\Gamma\left(\frac{2}{\alpha_i}+m\right)-\Gamma\left(\frac{2}{\alpha_i}+m, KR^{\alpha_i}t\right)\right]}{m!}\right). \tag{17}$$

Proof. See Appendix A. □

Throughput: By definition, throughput, $\mathcal{R}$ is given by

$$\mathcal{R} = \log_2(1+\rho)\,(1-\mathbb{P}_{out}), \tag{18}$$

where ρ is the received SNR at the ME defined in (5), and the expression of $\mathbb{P}_{\mathrm{out}}$ can be found in (17).

Latency ratio: This performance metric is used to measure the quality of service (QoS) of the network. It is defined as the delayed ME duration per unit of service time during the transmission process and will be discussed in section IV.

IV. Propagation Analysis and Performance Measures

In this section, we study the access link performance of our system model outlining outage probability, average latency, and throughput in the two different scenarios aforementioned in section II-A.

A. Noise-Limited Scenario

1) NNP: Outage occurs in this scenario when the channel of a RRH closest to the typical ME is in outage. It should be noted that the BBU transmits in this scenario only when there are no RRHs in $\mathcal{G}$. Thus, outage probability for the nearest neighbour is defined by the following proposition:

Proposition 2. *The outage probability of a* RRH *closest to the typical* ME *assuming it is the k^{th}* RRH *is expressed as*

$$\mathbb{P}_{out}^{\mathrm{NNP}}(\xi) = \left(1-\frac{1}{\mathrm{N_t}}\right)^{\mathcal{U}-1}\sum_{i\in \mathrm{L,N}}\frac{p_i\,\gamma\left(\mathrm{N_t}-K+1,\,\xi\,K\,R^{\alpha_i}\right)}{(\mathrm{N_t}-K)!} \times \left(1-\frac{1}{\mathrm{N_r}}\right)^{\mathcal{U}-1}\frac{e^{-\lambda_{\mathrm{RRH}}\,\pi\,R^2}}{(\mathrm{N_r}-K)!}\sum_{n=0}^{\infty}\frac{(-1)^n}{n!}\frac{(\xi\,K)^{(N_r-K+n+1)}}{(N_r-K+n+1)} \times \left[\sum_{i\in L,N} p_i\,\frac{\gamma\frac{(\alpha_i(N_r-K+n+1)}{2}+1, \pi\lambda_{\mathrm{RRH}}R^2)}{(\pi\lambda_{\mathrm{RRH}})^{\frac{(\alpha_i(N_r-K+n+1)}{2}}}\right]. \tag{19}$$

Proof. See Appendix B. □

2) BCP: In this scenario, the best channel in cluster $\mathcal{G}$ (either an RRH or the BBU) is selected to participate in transmission with the typical ME, This implies that outage occurs when this channel is in outage. Thus, the outage probability for the best channel is given as

$$\mathbb{P}_{out}^{\mathrm{BCP}}(\xi) = F_{\rho_0}(\xi)\,\mathbb{P}_{out}^{k}(\xi), \tag{20}$$

where $\mathbb{P}_{out}^{k}$ is the outage probability of the k^{th} RRH given in (17), and $F_{\rho_0}(t)$ is defined as:

$$F_{\rho_0}(\xi)=\left(1-\frac{1}{\mathrm{N_t}}\right)^{\mathcal{U}-1}\sum_{i\in\mathrm{L,N}}\frac{p_i\,\gamma\left(\mathrm{N_t}-K+1,\frac{\xi\,K\,R^{\alpha_i}}{B_0}\right)}{(\mathrm{N_t}-K)!}. \tag{21}$$

B. Interference-Limited Scenario

In the sequel, we present the analyses of this system with inter-cluster interference for both association scenarios:

1) BCP: In this scenario, the typical ME associates itself with the RRH that provides it with the best signal in the network. This can also be explained from the path loss perspective where the best channel is the one that offers the least path loss.

Thus, in the following proposition, we present the outage probability of SINR at the typical ME from the best RRH within the cluster for a predefined rate ξ.

Proposition 3. *The outage probability of received SINR at the typical* ME *from the best* RRH *in the cluster is given as*

$$\mathbb{P}_{out}^{\mathrm{BCP}}(\xi) = \sum_{j\in\mathrm{L,N}} p_j\left(1-\frac{1}{\mathrm{N_r}}\right)^{\mathcal{U}-1}\sum_{k=0}^{\nu}\binom{\nu}{k}(-1)^k \tag{22}$$
$$\times\int_{y>0} e^{\frac{-A\,k\,\xi\,y\sigma_z^2}{G_0}}\prod_{j\in\mathrm{L,N}}\mathbb{E}_{\mathbf{I}^j_{\Phi'_{\mathrm{BS}}}}\left[\exp\left(\frac{-A\,k\,\xi\,y\mathbf{I}^j_{\Phi'_{\mathrm{BS}}}}{G_0}\right)\right] f_\zeta(y)dy,$$

where $\nu=\mathrm{N_r}-K$ is a parameter from the tight upper bound of Gamma distribution given as $\mathbb{P}[|\eta_{\mathrm{m,n}}|^2<\gamma<(1-e^{-A\gamma})^\nu]$ with $A=\nu(\nu!)^{\frac{-1}{\nu}}$ and $y\triangleq r_n^{\alpha_j}$, and f_ζ is the distribution of the least path loss, which is given by

$$f_\xi(x) = \prod_{j\in\{L,N\}}\frac{2p_j}{\alpha_j}\pi\lambda_{\mathrm{BS}}P_{\mathrm{BS}}^{\frac{2}{\alpha_j}}\,x^{\frac{2}{\alpha_j}-1}e^{-\pi p_j\lambda_{\mathrm{BS}}P_{\mathrm{BS}}^{\frac{2}{\alpha_j}}x^{\frac{2}{\alpha_j}}}, \tag{23}$$

Proof. Refer Appendix C. □

2) NNP: We consider the RRH closest to the typical ME, which experiences outage when its channel is in less than the predefined rate ξ.

Proposition 4. *The outage probability of the* RRH *closest to the typical* ME *considering the impact of inter-cluster interference is expressed as*

$$\mathbb{P}_{out}^{\mathrm{NNP}}(\xi) = \sum_{j\in\mathrm{L,N}} p_j\left(1-\frac{1}{\mathrm{N_r}}\right)^{\mathcal{U}-1}\int_{r>0}\sum_{k=0}^{\nu}\binom{\nu}{k}(-1)^k \tag{24}$$
$$\times e^{\frac{-A\,k\,\xi\,y\sigma_z^2}{G_0}}\prod_{j\in\mathrm{L,N}}\mathbb{E}_{\mathbf{I}^j_{\Phi'_{\mathrm{BS}}}}\left[\exp\left(\frac{-A\,k\,\xi\,y\mathbf{I}^j_{\Phi'_{\mathrm{BS}}}}{G_0}\right)\right] f_r(r)\,\mathrm{d}r,$$

where y and ν are defined in (22)*, f_r is the pdf of the nearest distance to the typical* ME *given as [41]:*

$$f_r(r) \;=\; e^{-\lambda_{\mathrm{RRH}}\,\pi\,r^2}\,2\,\pi\,\lambda_{\mathrm{RRH}}\,r. \tag{25}$$

Proof. Using f_r defined in (25) and following similar steps in Appendix B, the equation (24) can be obtained. □

3) Average Rate: In order to compute the average latency, we need to characterise the average rate. Therefore in this subsection we adopt the framework developed by authors in [42] to evaluate average rate in terms of moment generating functions (MGFs). Thus, we present the average rate between the typical ME and its associated RRH in the next proposition[9]

Proposition 5. *The average rate between a typical* ME *and it's associated* RRH *is given as*

$$\bar{\mathcal{R}} = \left(1-\frac{1}{\mathrm{N_r}}\right)^{\mathcal{U}-1}\int_0^\infty(1-\mathcal{L}_S(t))\;\mathcal{L}_{\mathbf{I}_{\Phi'_{\mathrm{BS}}}}(t)\frac{e^{-t}}{t}\,\mathrm{dt}, \tag{26}$$

where S is the SNR given as $S=\dfrac{G_0\,|\eta_{\mathrm{k,0}}|^2\,\mathrm{r}_{k,0}^{-\alpha_j}}{\sigma_z^2}$ and $\mathbf{I}_{\Phi'_{\mathrm{BS}}}$ is similarly defined as in (10)*,*

$$\mathcal{L}_S(t) = \sum_{k=0}^{\nu}\binom{\nu}{k}(-1)^{k+1}\left(1+\frac{t\,G_0}{A\,k\,r_{k,0}^{\alpha_{\mathrm{L}}}\sigma_z^2}\right)^{-1}p_{\mathrm{L}} \tag{27}$$
$$+\sum_{k=0}^{\nu}\binom{\nu}{k}(-1)^{k+1}\left(1+\frac{t\,G_0}{A\,k\,r_{k,0}^{\alpha_{\mathrm{N}}}\sigma_z^2}\right)^{-1}p_{\mathrm{N}},$$

and

$$\mathcal{L}_{\mathbf{I}_{\Phi'_{\mathrm{BS}}}}(t) = \prod_{q\in\mathrm{L,N}} p_q\exp\Bigg(-2\pi\lambda_{\mathrm{RRH}} \tag{28}$$
$$\times\int_r^\infty 1-\left(1+\frac{A\,k\,t\,G_{\mathrm{v}}\,r_{k,0}^{\alpha_q}}{G_0\,r_v^{\alpha_q}}\right)^{-\nu}\mathrm{dr}\Bigg).$$

Proof. Refer Appendix D. □

Remark: *The analysis in this paper is done for access links between* RRH*s and* ME*s. However, worthy of mention is that the FH rate is constant and dependent on digitised I and Q samples and number of antennas as described in section 4.4.1 of [43] and given as*

$$\mathcal{R}_{FH} = 2\gamma f_s N_A N_q, \tag{29}$$

where the factor 2 accounts for I and Q phases of the signal, γ represents the overhead introduced by FEC and control signals, N_A is the number of antennas, and f_s, N_q are the sampling frequency and resolution of the quantiser respectively.

[9]We note that this proposition also holds for the noise-limited case when the interference is equal to zero.

C. Traffic Latency

In this subsection, we utilise queuing theory model described in II-G to analyse the delay in data delivery from BBU to ME_0. If ME_0 associated with RRH_j is assumed to have traffic load of κ_0 expressed as $\kappa_0 = \frac{\lambda_0}{\mu_0}$, where λ_0 is the arrival rate per unit area and μ_0 is the average number of requested volumes (average traffic load).

In the network, an RRH is connected to different MEs, thus we denote the coverage area of all RRHs as $\mathcal{B}$ and assume that each ME connected to an RRH is served in a round robin manner. Then, the average traffic load density at ME_0 from RRH_j is given as $\delta_j = \frac{\lambda_0\,\kappa_0}{\mathcal{R}_0}$, where $\mathcal{R}_0$ is the access link rate defined in (26) and the total traffic load in RRH_j is expressed as $D_j = \int_{\mathcal{B}} \delta(x)\,\mathrm{dx}$, where x denotes the location of ME connected to RRH_j. Next, we can compute the required service time to satisfy the demands of ME_0 is given as $\nu_0 = \frac{\kappa_0}{\mathcal{R}_0}$. Given that traffic delivery in the access link emulates the M/G/1 queueing model, the average traffic delivery time for ME_0 in RRH_j is given as [37] $T_j = \frac{\kappa_0}{\mathcal{R}_0(1-D_j)}$ In RRH_j, the average waiting time for the traffic load of ME_0 is obtained by subtracting the required service time from the average traffic delivery time and expressed as $\chi_j = T_j - \nu_0 = \frac{D_j\,\kappa_0}{\mathcal{R}_0\,(1-D_j)}$. The latency in the access link of RRH_j and ME_0 is calculated as the ratio of waiting time and service time. Thus,

$$\Psi_j = \frac{\chi_j}{\nu_0} = \frac{D_j}{1-D_j}. \tag{30}$$

In the FH link between the BBU and RRH_j, the required time to satisfy the traffic demand of ME_0 is dependent on the FH rate, $\mathcal{R}_{\mathrm{FH}}$ defined in (29) and expressed as $\hat{\nu}_0 = \frac{\kappa_0}{\mathcal{R}_{\mathrm{FH}}}$. From the M/M/1 queueing model in [37], the average wait time for ME_0's traffic load in RRH_j's FH is defined as $\hat{\chi}_j = \frac{\hat{D}_j\,\kappa_0}{\mathcal{R}_{\mathrm{FH}}\,(1-\hat{D}_j)}$, where $\hat{D}_j$ is the total load in the FH of RRH_j. Then, the latency ratio to measure how much time ME_0 waits per unit service time in the FH of RRH_j is given as

$$\hat{\Psi}_j = \frac{\hat{\chi}_j}{\hat{\nu}_0} = \frac{\hat{D}_j}{1-\hat{D}_j}. \tag{31}$$

Hence, the traffic delivery latency from BBU to ME_0 via RRH_j is given as $\Delta_j = \hat{\Psi}_j + \Psi_j$. It is worthy of mention that since the latency ratio of FH link in (31) is dependent only on the traffic load at the FH of RRH_j, it implies that all MEs associated with RRH_j have the same latency ratio and a smaller Δ_j suggests that RRH_j introduces a low latency to its associated MEs.

V. Numerical Results

In this section, we validate the system model and analytical derivations from the aforementioned propositions using Monte-Carlo simulations. The values for most of the parameters used in simulation are motivated from literature mentioned in the references. For completeness of the paper, and unless otherwise stated, Table I presents these parameters and their equivalent values.

A cluster radius of 250m is considered, with RRH intensity expressed as $\lambda_{\mathrm{RRH}} = \frac{\Lambda}{\pi R^2}$, where Λ represents the average number of RRHs in a cluster sized πR^2 with transmit power of 30 dBm and noise power of -70dB.

TABLE I: Simulation Parameters

Notation	Parameter	Values
$\mathrm{N_{ME}}$	Number of ME antennas	2x2, 4x4 UPA
B	Bandwidth	2GHz
$\mathrm{N_{RRH}}$	Number of RRH antennas	4x4,8x8 UPA
$\mathrm{N_{BBU}}$	Number of BBU antennas	4x4, 8x8 UPA
K	Number of MEs	4
p_{L}, R	Blockage Model	0.1, 250 meters
λ_{RRH}	Density of RRH nodes	$\frac{6}{1000\pi}$
α	Path loss exponent	L-2, N-3.5
$\mathrm{N_0}$	Noise power	-70 dB

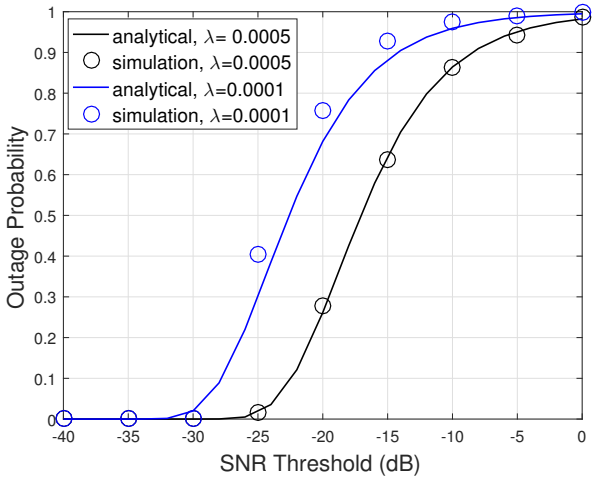

Fig. 3: BCP outage probability comparison with varying RRH densities.

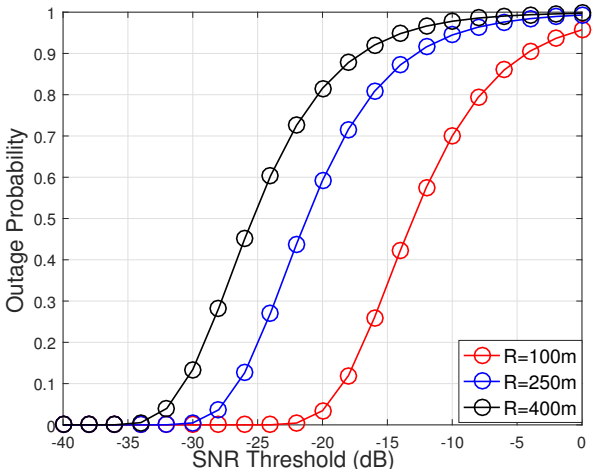

Fig. 4: Comparison of best channel participation scenario for varying radii.

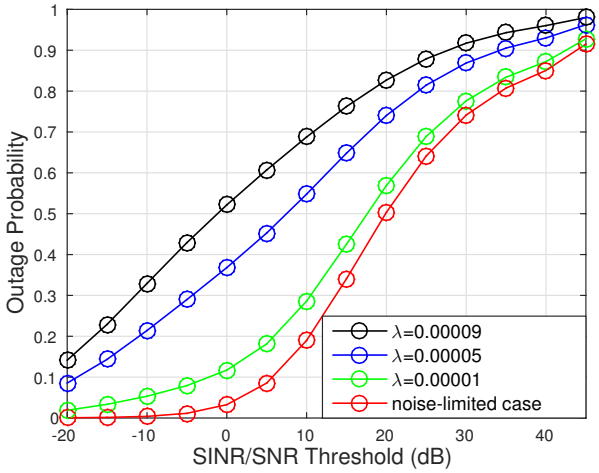

Fig. 5: NNP Comparison of outage probability in noise and interference-limited regimes.

Starting with the BCP scenario, Fig. 3 presents a plot of outage probability against SNR threshold, and we observe that as with an increase in the required threshold, the probability of outage also increases. However, as the average number of RRHs in the cluster is increased, outage probability is reduced. An interesting observation is the impact of inter-cluster interference on the outage probability. This will be explained from Fig. 6 subsequently.

Next, we consider the effect of changing the cluster radius on outage probability in Fig. 4. Here we observe that increasing the radius does not increase the outage probability for the best channel scenario. We may infer by this result that there is a trade-off between cluster radius and outage performance in mmWave CRAN as smaller cluster radii lead to better performance. This can be used in network planning by operators especially in urban areas where many small clusters can be formed.

Fig. 5 shows the comparison of outage probability in both noise-limited and interference-limited scenarios for NNP. It is evident from the figure that for a given moderate number of blockages, the outage probability is much less in the noise-limited regime when compared to the interference-limited regime. It can also be observed that as the number of interferers is decreased (by the reduction in RRH node density), the performance in the interference-limited scenario tends to that in the noise-limited regime. This outcome indicates that in systems employing mmWave links, successful transmission is largely dependent on blockage and nodal densities.

Fig. 6 shows that an increase in average number of RRHs does not lead to a decrease in outage probability. Although this result appears to be counter intuitive, it can be explained by the fact that increasing the number of RRHs also increases the probability of interfering RRHs. We establish by the small difference in outage probabilities between RRH intensities in Fig. 6, a trade-off between outage performance and interference. It is also evident from Fig. 6 that increasing the array size improves the performance leading to smaller outage probabilities.

In Fig. 7, we plot the outage probability as a function of SINR threshold. In measuring the effect of path loss on the transmission from the best channel to a typical ME, we observe that outage probability decreases with reducing α. We may thus deduce that increasing the values of path loss exponent will degrade communication.

Fig. 8 illustrates the outage probability of the BCP and NNP scenarios against SINR threshold. It is clear that BCP always outperforms NNP. For example, to achieve an SINR threshold of 20dB when RRH density is 0.0001, the outage probability for BCP and NNP are 0.25 and 0.8 respectively. Although NNP is considered in some CRAN applications for the reduction of overhead in the selection process, it is worthwhile to note that in multi-user mmWave CRAN systems, we opt for the best channel selection process for better transmission due to the big difference between NNP and BCP performances.

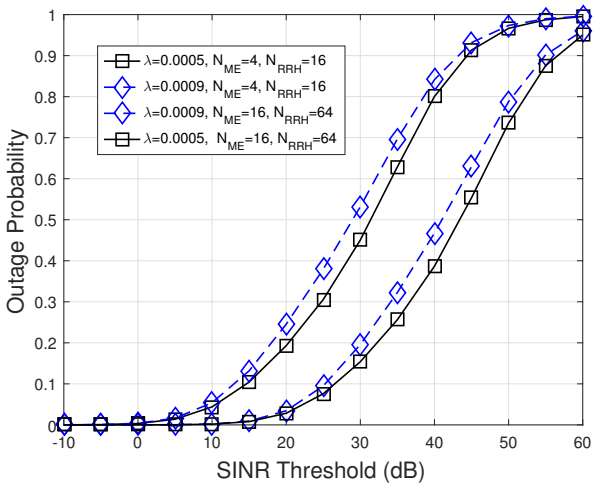

Fig. 6: Best channel partipation with varying Lambda considering interference.

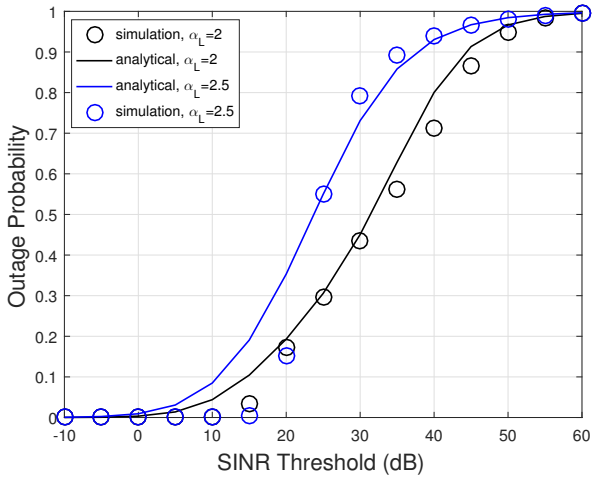

Fig. 7: BCP Outage Probability with varying path loss exponents considering interference

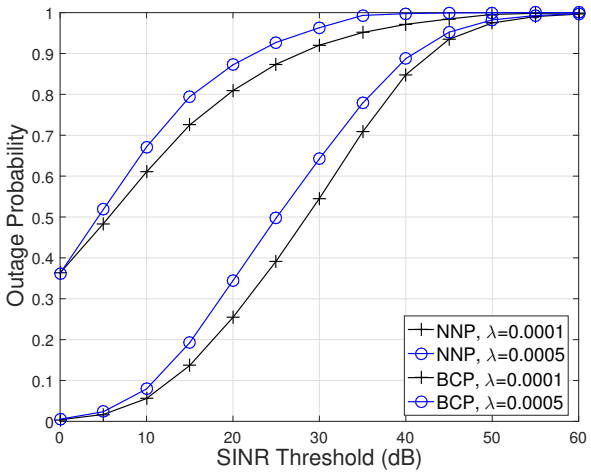

Fig. 8: Outage Probability of two ME association scenarios.

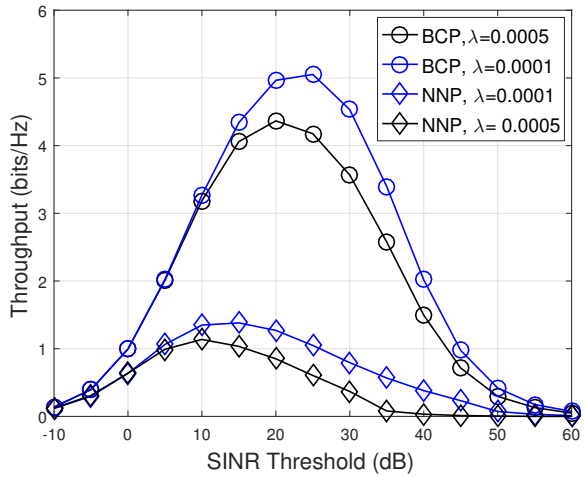

Fig. 9: Throughput comparison with varying RRH densities of two ME association scenarios.

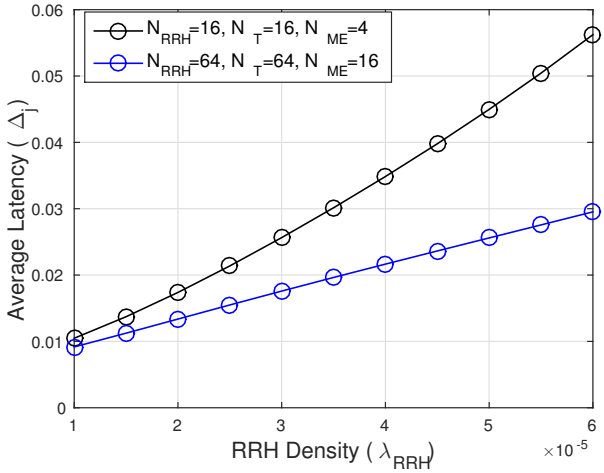

Fig. 10: Average latency against RRH density with varying fronthaul rates.

Figures 9 illustrates the throughput performance of BCP and NNP schemes against SINR threshold. In Fig. 9, we show the comparison of throughput performance for both BCP and NNP scenarios. Following from the result that increasing the number of RRHs can cause a degradation to communication between either the best or nearest RRH and the typical ME, the throughput performance for RRH intensity of 0.0001 is better than that for an intensity of 0.0005 in both cases. Furthermore, Fig. 9 shows that BCP significantly outperforms NNP. To achieve an SINR of 20dB, the BCP throughput for $\lambda_{\mathrm{RRH}} = 0.0001$ is 5bits/Hz while that of NNP is 1.3bits/Hz, this confirms the results of Fig. 8.It is worthy of mention that there exists an optimal value of throughput as shown by the shape of curves in Fig. 9, with the implication that increasing SINR does not indefinitely lead to better performance. Determination of the optimal point can be explored in future works.

Having seen the effect on increasing the number of RRHs on outage probability, we determine the consequence of average data delivery latency on varying RRH density. Hence, Fig. 10 is plotted for different values of λ_{RRH}. We observe that increasing the RRH density leads to a rise in average latency. This outcome, although unexpected, is not so unusual given that an increased number of RRHs implies a larger traffic load at the FH queue leading to more latency. In addition, we also observe that increasing the antenna gain which in turn increases the FH and access link rates leads to reducing the latency. It is important to note that at low RRH density, the latency for both gains are almost equal due to the fact that when the traffic load is small, there is no significant gain in increasing the fronthaul or access link rate. In other words, the impact of FH and access link rates are negligible for small traffic loads.

VI. Conclusion

This paper presents the performance analysis of a system that couples mmWave and CRAN (mmWave CRANs) for future generation communication. We considered the downlink scenario of multiple distributed RRHs and a BBU in a multi-user system employing mmWave technologies including hybrid beamforming in both fronthaul and access links. Analytical expressions for outage probability, throughput were derived for two ME association scenarios, namely, BCP and NNP. We considered two cases to model practical deployment of mmWave CRANs networks namely: a) noise-limited, and b) interference-limited. In addition, we analyse the impact of fronthaul rate on the average data delivery latency in the system. These expressions were validated in the numeral section. Our results show a trade-off between density of RRHs and inter-cluster interference. It is shown that deploying larger antenna arrays can compensate for the degradation of communication in terms of delivery latencies, throughput, and outage probabilites with higher RRH deployment. This can be exploited by engineers in practice for the maintenance of high performance. In addition, there can be a positive impact of increased blockages and path loss exponents on outage probability and throughput. Moreover, simulation and analytical results show that BCP scenarios grossly outperform

NNP scenarios. Despite the attractiveness of NNP in existing literature, based on its reduction in overhead compared to other scenarios, the BCP is shown to be most viable for multi-user mmWave CRAN systems.

APPENDIX A
PROOF OF PROPOSITION 1

The outage probability of the k^{th} RRH is defined as

$$\mathbb{P}_{out}(z) = \int_0^R f_{r_{k,0}}(r) F_{\rho_{k,0}|r_{k,0}}(z)\,\mathrm{d}r. \tag{32}$$

Substituting the expressions of $f_{r_{k,0}}(r) = \frac{2r}{R^2}$ and $F_{\rho_{k,1}|r_k}$ from (14), we obtain

$$\mathbb{P}_{out}(t) = \int_0^R \frac{2r}{R^2}\left(1-\frac{1}{\mathrm{N_r}}\right)\sum_{i\in \mathrm{L,N}}^{\mathcal{U}-1} p_i \frac{1}{(\mathrm{N_r}-K)!} \times \gamma\left(\mathrm{N_r}-K+1, t\,K\,r_{k,0}^{\alpha_i}\right)\mathrm{d}r, \tag{33}$$

where $t = \dfrac{z}{B_{k,0}}$.

Starting with the LOS link, we have

$$\mathbb{P}^{\mathrm{L}}_{out}(t) = \tag{34}$$
$$\int_0^R \left(1-\frac{1}{\mathrm{N_r}}\right)^{\mathcal{U}-1}\frac{2r}{R^2}\frac{p_{\mathrm{L}}\gamma\left(\mathrm{N_r}-K+1, tKr_{k,0}^{\alpha_{\mathrm{L}}}\right)}{(\mathrm{N_r}-K)!}\mathrm{d}r,$$
$$\overset{(a)}{=}\int_0^R \frac{2r}{R^2}\left(1-\frac{1}{\mathrm{N_r}}\right)^{\mathcal{U}-1} p_{\mathrm{L}}\left[1-\sum_{m=0}^{\mathrm{N_r}-K}\frac{(tKr^{\alpha_{\mathrm{L}}})^m}{m!}\mathrm{e}^{\mathrm{tKr}^{\alpha_{\mathrm{L}}}}\right]\mathrm{d}r,$$
$$\overset{(b)}{=}\left(1-\frac{1}{\mathrm{N_r}}\right)^{\mathcal{U}-1} p_{\mathrm{L}}\left[1-\sum_{m=0}^{\mathrm{N_r}-K}\frac{2\,(t\,K)^{-\frac{2}{\alpha_{\mathrm{L}}}}}{R^2} \times\frac{\left[\Gamma\left(\frac{2}{\alpha_{\mathrm{L}}}+m\right)-\Gamma\left(\frac{2}{\alpha_{\mathrm{L}}}+m, KR^{\alpha_{\mathrm{L}}}t\right)\right]}{m!}\right],$$

where (a) follows from the series equivalent of the lower Gamma incomplete function given by $\gamma(N,t) = (N-1)!\left(1-e^{-t}\sum_{k=0}^{N-1}\frac{t^k}{k!}\right)$, and (b) follows from solving the integral with respect to r.

Following similar steps we can derive the outage probability of a NLOS link to obtain $\mathbb{P}^{\mathrm{N}}_{out}(t)$. Finally the proof of (17) can be concluded by the summation of $\mathbb{P}^{\mathrm{L}}_{out}(t)$ and $\mathbb{P}^{\mathrm{N}}_{out}(t)$.

APPENDIX B
PROOF OF PROPOSITION 2

During the nearest neighbour participation scenario, the ME is served by the RRH closest to it. The PDF of the distance between the typical ME and it's closest RRH is given as [44]

$$f_{\text{closest}} = 2\,\pi\,\lambda_{\mathrm{RRH}}\,r\,\mathrm{e}^{-\lambda_{\mathrm{RRH}}\,\pi\,\mathrm{r}^2}. \tag{35}$$

Given the closest distance, the outage probability for the closest RRH becomes

$$\mathbb{P}^{\mathrm{NNP}}_{out}(\xi) = \int F_{\rho|r}(\xi) f_{\text{closest}}(r), \tag{36}$$

where $f_{\text{closest}}(r)$ is the PDF of distance between the ME and it's closest RRH is defined in (35), and $F_{\rho|r}$ is defined in (14).

Considering that only when there is no RRH in cluster $\mathcal{G}$, does the BBU transmit to the ME, we expand the definition of nearest neighbour outage probability to the typical ME, assuming the k^{th} RRH is the closest RRH as

$$\mathbb{P}^{\mathrm{NNP}}_{out}(\xi) = \int_0^R f_{\text{closest}}(r) F_{\rho_{k,0}|r_{k,0}}(\xi)\,\mathrm{d}r + \int_R^\infty f_{\text{closest}}(r) F_{\rho_0}(\xi)\mathrm{d}r, \tag{37}$$

where

$$F_{\rho_0}(\xi) = \left(1-\frac{1}{\mathrm{N_t}}\right)^{\mathcal{U}-1}\sum_{i\in \mathrm{L,N}}^{\mathcal{U}-1}\frac{p_i\,\gamma\left(\mathrm{N_t}-K+1, \frac{\xi\,K\,R^{\alpha_i}}{B_{k,0}}\right)}{(\mathrm{N_t}-K)!}. \tag{38}$$

Substituting the expressions of $F_{\rho_0}(\xi)$, $F_{\rho_{k,0}|r_{k,0}}(\xi)$, and $f_{\text{closest}}(r)$, into (37) we obtain (39).

To integrate I in (39), we start with LOS link

$$\mathrm{I} = \left(1-\frac{1}{\mathrm{N_r}}\right)^{\mathcal{U}-1}\int_0^{\mathrm{R}} 2\pi\lambda_{\mathrm{RRH}}\,\mathrm{r}\mathrm{e}^{-\pi\lambda_{\mathrm{RRH}}\,\mathrm{r}^2} \times\sum_{i\in \mathrm{L,N}}\frac{p_{\mathrm{L}}\,\gamma\left(\mathrm{N_r}-K+1,\,\xi\,K\,r_{k,0}^{\alpha_{\mathrm{L}}}\right)}{(\mathrm{N_r}-K)!}\mathrm{d}r, \tag{40}$$

replacing the lower incomplete gamma function with its series equivalent

$$\gamma\left(\mathrm{N_r}-K+1,\,\xi\,K\,r_{k,0}^{\alpha_{\mathrm{L}}}\right) = \sum_{n=0}^{\infty}\frac{(-1)^n\left(\xi K r_{k,0}^{\alpha_{\mathrm{L}}}\right)^{Nr+n-K+1}}{n!\,(Nr+n-K+1)},$$

and integrating (40) gives

$$\mathrm{I} = \left(1-\frac{1}{\mathrm{N_r}}\right)^{\mathcal{U}-1}\frac{\mathrm{p_L}}{(\mathrm{N_r}-\mathrm{K})!}\sum_{n=0}^{\infty}\frac{(-1)^{\mathrm{n}}}{\mathrm{n}!}\frac{(\xi\,\mathrm{K})^{(\mathrm{N_r}-\mathrm{K}+\mathrm{n}+1)}}{(\mathrm{N_r}-\mathrm{K}+\mathrm{n}+1)} \times\frac{\gamma\left(\frac{\alpha_{\mathrm{L}}(N_r-K+n+1)}{2}+1,\,\pi\,\lambda_{\mathrm{RRH}}\,R^2\right)}{(\pi\,\lambda_{\mathrm{RRH}})^{\frac{(\alpha_{\mathrm{L}}\,(N_r-K+n+1)}{2}}}, \tag{41}$$

following similar steps for NLOS link, and summing them we obtain the result of integrating I given as

$$\mathrm{I} = \left(1-\frac{1}{\mathrm{N_r}}\right)^{\mathcal{U}-1}\frac{1}{(\mathrm{N_r}-\mathrm{K})!}\sum_{n=0}^{\infty}\frac{(-1)^{\mathrm{n}}}{\mathrm{n}!}\frac{(\xi\,\mathrm{K})^{(\mathrm{N_r}-\mathrm{K}+\mathrm{n}+1)}}{(\mathrm{N_r}-\mathrm{K}+\mathrm{n}+1)} \times\left[\sum_{i\in L,N} p_i\,\frac{\gamma\frac{(\alpha_i(N_r-K+n+1)}{2}+1, \pi\lambda_{\mathrm{RRH}}R^2)}{(\pi\lambda_{\mathrm{RRH}})^{\frac{(\alpha_i(N_r-K+n+1)}{2}}}\right]. \tag{42}$$

Next, integrating II, we obtain:

$$\mathrm{II} = \left(1-\frac{1}{\mathrm{N_t}}\right)^{\mathcal{U}-1}\sum_{\mathrm{i\in L,N}}\frac{\mathrm{p_i}\,\gamma\left(\mathrm{N_t}-\mathrm{K}+1,\,\xi\,\mathrm{K}\,\mathrm{R}^{\alpha_{\mathrm{i}}}\right)}{(\mathrm{N_t}-\mathrm{K})!}\,\mathrm{e}^{-\lambda_{\mathrm{RRH}}\,\pi\,\mathrm{R}^2}. \tag{43}$$

The proof of (19) is obtained from substituting (43), and (42) in (39)

$$\mathbb{P}^{\mathrm{NNP}}_{out}(\xi) = \left(1-\frac{1}{\mathrm{N_r}}\right)^{\mathcal{U}-1} \underbrace{\int_0^R 2\pi\lambda_{\mathrm{RRH}}\, r\mathrm{e}^{-\pi\lambda_{\mathrm{RRH}}\,\mathrm{r}^2} \sum_{\mathrm{i}\in\mathrm{L,N}} \frac{\mathrm{p_i}\;\gamma\left(\mathrm{N_r}-\mathrm{K}+1,\,\xi\,\mathrm{K}\,\mathrm{r}_{\mathrm{k},0}^{\alpha_\mathrm{i}}\right)}{(\mathrm{N_r}-\mathrm{K})!}\,\mathrm{dr}}_{\mathrm{I}} + \left(1-\frac{1}{\mathrm{N_t}}\right)^{\mathcal{U}-1} \underbrace{\int_R^\infty 2\pi\,\lambda_{\mathrm{RRH}}\, r\mathrm{e}^{-\pi\,\lambda_{\mathrm{RRH}}\,\mathrm{r}^2} \sum_{\mathrm{i}\in\mathrm{L,N}} \frac{\mathrm{p_i}\;\gamma\left(\mathrm{N_t}-\mathrm{K}+1,\,\xi\,\mathrm{K}\,\mathrm{R}^{\alpha_\mathrm{i}}\right)}{(\mathrm{N_t}-\mathrm{K})!}\,\mathrm{dr}}_{\mathrm{II}}. \tag{39}$$

APPENDIX C
PROOF OF PROPOSITION 3

Let $y \triangleq r_n^{\alpha_j}$ represent the path loss. The outage probability conditioned on the least path loss from the best RRH in the cluster to the typical ME averaged over the plane is defined as:

$$\mathbb{P}^{\mathrm{BCP}}_{\mathrm{out}|y}(\xi) = \mathbb{E}\left[\mathbb{P}\left[\frac{G_0\,|\eta_{\mathrm{k},0}|^2\; y^{-1}\,\gamma(\mathrm{N_r},\mathcal{U})}{\sigma_z^2+\mathbf{I}_{\Phi'_{\mathrm{BS}}}} \;<\; \xi|y\right]\right]. \tag{44}$$

For the LOS link, $y = r^{\alpha_\mathrm{L}}$, the conditional outage probability is then given as

$$\begin{aligned}\mathbb{P}^{\mathrm{L}}_{\mathrm{out}|y}(\xi) &= \mathbb{E}\left[p_\mathrm{L}\,\mathbb{P}\left[\frac{G_0\,|\eta_{\mathrm{k},0}|^2\; y^{-1}\,\gamma(\mathrm{N_r},\mathcal{U})}{\sigma_z^2+\mathbf{I}_{\Phi'_{\mathrm{BS}}}} \;<\; \xi|y\right]\right], \\ &= \int_{y>0} p_\mathrm{L}\,\mathbb{P}\left[\frac{G_0\,|\eta_{\mathrm{k},0}|^2\; y^{-1}\,\gamma(\mathrm{N_r},\mathcal{U})}{\sigma_z^2+\mathbf{I}_{\Phi'_{\mathrm{BS}}}} \;<\; \xi|y\right] f_\zeta(y)\,\mathrm{dy}.\end{aligned} \tag{45}$$

Given that the small scale fading $\eta_{k,0}$, is Rayleigh, $|\eta_{k,0}\,|^2$ follows Chi-square distribution with $2(\mathrm{N_r}-K)$ degrees of freedom and employs the upper bound of gamma distribution with parameter ν such that: $\mathbb{P}\left[\,|\eta_{k,0}\,|^2 \;<\; \gamma \;<\; (1-e^{-A\,\gamma})^\nu\right]$ with $A = \nu(\nu!)^{\frac{-1}{\nu}}$, therefore, the outage probability is expressed as

$$\begin{aligned}&\mathbb{P}\left[\frac{G_0\,|\eta_{\mathrm{k},0}|^2\; y^{-1}\,\gamma(\mathrm{N_r},\mathcal{U})}{\sigma_z^2+\mathbf{I}_{\Phi'_{\mathrm{BS}}}} \;<\; \xi|y\right] \\ &= \mathbb{E}_{\mathbf{I}_{\Phi'_{\mathrm{BS}}}}\left[\mathbb{P}\left[|\eta_{\mathrm{k},0}|^2 \;<\; \frac{\xi\,y}{G_0\,\gamma(\mathrm{N_r},\mathcal{U})}(\sigma_z^2+\mathbf{I}_{\Phi'_{\mathrm{BS}}})|y,\,\mathbf{I}_{\Phi'_{\mathrm{BS}}}\right]\right], \\ &\overset{(a)}{=} \left(1-\frac{1}{\mathrm{N_r}}\right)^{\mathcal{U}-1}\mathbb{E}_{\mathbf{I}_{\Phi'_{\mathrm{BS}}}}\left[\left(1-e^{-A\,\frac{\xi\,y}{G_0}\,(\sigma_z^2+\mathbf{I}_{\Phi'_{\mathrm{BS}}})}\right)^\nu |y,\,\mathbf{I}_{\Phi'_{\mathrm{BS}}}\right], \\ &\overset{(b)}{=} \left(1-\frac{1}{\mathrm{N_r}}\right)^{\mathcal{U}-1}\sum_{k=0}^{\nu}\binom{\nu}{k}(-1)^k e^{\frac{-A\,k\,\xi\,y\sigma_z^2}{G_0}} \\ &\qquad\qquad \times\mathbb{E}_{\mathbf{I}_{\Phi'_{\mathrm{BS}}}}\left[e^{-\frac{-A\,k\,\xi\,y\,\mathbf{I}_{\Phi'_{\mathrm{BS}}}}{G_0}}\right], \\ &\overset{(c)}{=} \left(1-\frac{1}{\mathrm{N_r}}\right)^{\mathcal{U}-1}\sum_{k=0}^{\nu}\binom{\nu}{k}(-1)^k e^{\frac{-A\,k\,\xi\,y\sigma_z^2}{G_0}} \\ &\qquad\qquad \times\prod_{j\in\mathrm{L,N}}\mathbb{E}_{\mathbf{I}^j_{\Phi'_{\mathrm{BS}}}}\left[\exp\left(\frac{-A\,k\,\xi\,y\,\mathbf{I}^j_{\Phi'_{\mathrm{BS}}}}{G_0}\right)\right],\end{aligned}$$

where (a) follows from the precoding penalty and the tight gamma approximation previously defined, (b) follows from applying binomial expansion, and (c) follows from the fact that interference links can be LOS or NLOS such that $\mathbf{I}_{\Phi'_{\mathrm{BS}}} = \mathbf{I}^{\mathrm{L}}_{\Phi'_{\mathrm{BS}}} + \mathbf{I}^{\mathrm{N}}_{\Phi'_{\mathrm{BS}}}$. Substituting (46) into (45) we obtain the LOS outage probability as

$$\begin{aligned}\mathbb{P}^{\mathrm{L}}_{\mathrm{out}}(\xi) = &\left(1-\frac{1}{\mathrm{N_r}}\right)^{\mathcal{U}-1} p_\mathrm{L}\int_{y>0}\sum_{k=0}^{\nu}\binom{\nu}{k}(-1)^k\, e^{\frac{-A\,k\,\xi\,y\sigma_z^2}{G_0}} \\ &\times\prod_{j\in\mathrm{L,N}}\mathbb{E}_{\mathbf{I}^j_{\Phi'_{\mathrm{BS}}}}\left[\exp\left(\frac{-A\,k\,\xi\,y\,\mathbf{I}^j_{\Phi'_{\mathrm{BS}}}}{G_0}\right)\right] f_\zeta(y)\,\mathrm{dy},\end{aligned} \tag{47}$$

where f_ζ is the least path loss distribution defined as

$$f_\xi(x) = \prod_{j\in\{L,N\}} \tfrac{2p_j}{\alpha_j}\pi\lambda_{\mathrm{BS}}P_{\mathrm{BS}}^{\frac{2}{\alpha_j}}\, x^{\frac{2}{\alpha_j}-1} e^{-\pi p_j\lambda_{\mathrm{BS}}P_{\mathrm{BS}}^{\frac{2}{\alpha_j}}\,x^{\frac{2}{\alpha_j}}}. \tag{48}$$

Following similar steps we can derive the NLOS outage probability, $\mathbb{P}^{\mathrm{N}}_{\mathrm{out}}(\xi)$. To obtain the expectation of the LOS interfering link, we leverage on results from [33, Lemma 6] focussing on the single path case. Thus, the expectation is given as

$$\begin{aligned}&\mathbb{E}_{\mathbf{I}^{\mathrm{L}}_{\Phi'_{\mathrm{BS}}}}\left[\exp\left(\frac{-A\,k\,\xi\,y\,\mathbf{I}^{\mathrm{L}}_{\Phi'_{\mathrm{BS}}}}{G_0}\right)\right] \\ &\qquad = \mathbb{E}_{\mathbf{I}^{\mathrm{L}}_{\Phi'_{\mathrm{BS}}}}\left[\prod_{b\in\Phi'_{\mathrm{BS}}}\exp\left(\frac{-A\,k\,\xi\,y\,G_\mathrm{v}\,x\Upsilon_v}{G_0\,r_v^{\alpha_\mathrm{L}}}\right)\right],\end{aligned} \tag{49}$$

gotten by substituting $\mathbf{I}^{\mathrm{L}}_{\Phi'_{\mathrm{BS}}} = \sum_{b\in\Phi'_{\mathrm{BS}},v\neq 0} G_\mathrm{v}\,|\eta_{\mathrm{b,v}}|^2\,\mathrm{r_v}^{-\alpha_\mathrm{L}}$ and $x = |\eta_{\mathrm{b,v}}|^2$. Applying the probability generating functional of PPP (PGFL) [41], we obtain

$$\begin{aligned}&\mathbb{E}_{\mathbf{I}^{\mathrm{L}}_{\Phi'_{\mathrm{BS}}}}\left[\exp\left(\frac{-A\,k\,\xi\,y\,\mathbf{I}^{\mathrm{L}}_{\Phi'_{\mathrm{BS}}}}{G_0}\right)\right] \\ &\exp\left(-2\,\pi\,\lambda_{\mathrm{RRH}}\int_r^\infty 1-\frac{1}{\left(1+\frac{A\,k\,G_\mathrm{v}\,\xi\,y\tau^2_{\mathrm{RRH}}}{G_0\,r_v^{\alpha_\mathrm{L}}}\right)^\nu} p_\mathrm{L}\mathrm{dr}\right).\end{aligned} \tag{50}$$

The expectation of NLOS interfering link can be obtained similarly. Finally, the proof of outage probability from the best RRH in cluster $\mathcal{G}$ to the typical ME is obtained by summation of both LOS and NLOS outage probabilities respectively.

APPENDIX D
PROOF OF PROPOSITION 6

From the generalised expression for average rate in terms of MGFs [42]

$$\bar{\mathcal{R}} = \int_0^\infty (1-\mathcal{L}_S(x))\mathcal{L}_{\mathbf{I}_{\Phi'_{\mathrm{BS}}}}(x)\frac{e^{-x}}{x}\,\mathrm{dx}, \tag{51}$$

where S is defined as $S = \dfrac{G_0 \left|\eta_{\mathrm{k},0}\right|^2 \mathrm{r}_{k,0}^{-\alpha_j}}{\sigma_z^2}$.

The MGF of S is obtained after following similar steps from Appendix C, and expressed as

$$\mathcal{L}_S(t) = \sum_{k=0}^{\nu} \binom{\nu}{k} (-1)^{k+1} \left(1 + \frac{t\, G_0}{A\, k\, r_{k,0}^{\alpha_{\mathrm{L}}} \sigma_z^2}\right)^{-1} p_{\mathrm{L}} + \sum_{k=0}^{\nu} \binom{\nu}{k} (-1)^{k+1} \left(1 + \frac{t\, G_0}{A\, k\, r_{k,0}^{\alpha_{\mathrm{N}}} \sigma_z^2}\right)^{-1} p_{\mathrm{N}}. \quad (52)$$

In like manner, the MGF of interference to noise ratio, $\mathcal{L}_{\mathbf{I}_{\Phi'_{\mathrm{BS}}}}(t)$ in (28) is obtained using steps similar to (49)− (50). This concludes the proof.

REFERENCES

[1] C. V. Mobile, "Cisco visual networking index: global mobile data traffic forecast update, 2011–2016," *San Jose, CA*, vol. 1, 2016.

[2] T. S. Rappaport, R. W. Heath Jr, R. C. Daniels, and J. N. Murdock, *Millimeter wave wireless communications*. Pearson Education, 2014.

[3] C. Sacchi, T. F. Rahman, N. Bartolomei, S. Morosi, A. Mazzinghi, and F. Ciabini, "Design and assessment of a ce-ofdm-based mm-wave 5g communication system," in *2016 IEEE Globecom Workshops (GC Wkshps)*, Dec 2016, pp. 1–7.

[4] K. A. S. Cacciapuoti, R. Subramanian and M. Caleffi and K. C. Chowdhury, "Software-defined network controlled switching between millimeter wave and terahertz small cells," *arXiv:1702.02775*, Feb., 2017.

[5] R. W. Heath, N. Gonzlez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave mimo systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 436–453, April 2016.

[6] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5g cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.

[7] A. Ghosh, T. A. Thomas, M. C. Cudak, R. Ratasuk, P. Moorut, F. W. Vook, T. S. Rappaport, G. R. MacCartney, S. Sun, and S. Nie, "Millimeter-wave enhanced local area systems: A high-data-rate approach for future wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1152–1163, June 2014.

[8] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proceedings of the IEEE*, vol. 102, no. 3, pp. 366–385, 2014.

[9] S. Singh, M. N. Kulkarni, A. Ghosh, and J. G. Andrews, "Tractable model for rate in self-backhauled millimeter wave cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 10, pp. 2196–2211, 2015.

[10] M. Di Renzo, "Stochastic geometry modeling and analysis of multi-tier millimeter wave cellular networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 9, pp. 5038–5057, 2015.

[11] M. N. Kulkarni, S. Singh, and J. G. Andrews, "Coverage and rate trends in dense urban mmwave cellular networks," in *2014 IEEE Global Communications Conference*. IEEE, 2014, pp. 3809–3814.

[12] M. N. Kulkarni, A. Alkhateeb, and J. G. Andrews, "A tractable model for per user rate in multiuser millimeter wave cellular networks," in *2015 49th Asilomar Conference on Signals, Systems and Computers*, Nov 2015, pp. 328–332.

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

[14] P. Marsch, B. Raaf, A. Szufarska, P. Mogensen, H. Guan, M. Farber, S. Redana, K. Pedersen, and T. Kolding, "Future mobile communication networks: Challenges in the design and operation," *IEEE Vehicular Technology Magazine*, vol. 7, no. 1, pp. 16–23, March 2012.

[15] S.-R. Lee, S.-H. Moon, J.-S. Kim, and I. Lee, "Capacity analysis of distributed antenna systems in a composite fading channel," *IEEE Transactions on Wireless Communications*, vol. 11, no. 3, pp. 1076–1086, 2012.

[16] H. Zhu, "Performance comparison between distributed antenna and microcellular systems," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 6, pp. 1151–1163, 2011.

[17] J. Zhang and J. G. Andrews, "Distributed antenna systems with randomness," *IEEE Transactions on Wireless Communications*, vol. 7, no. 9, pp. 3636–3646, 2008.

[18] H. He, J. Xue, T. Ratnarajah, F. Khan, and C. B. Papadias, "Modeling and analysis of cloud radio access networks using mat'ern hard-core point processes," *IEEE Transactions on Wireless Communications*, vol. PP, no. 99, pp. 1–1, 2016.

[19] S. Jin, M. R. McKay, C. Zhong, and K.-K. Wong, "Ergodic capacity analysis of amplify-and-forward mimo dual-hop systems," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2204–2224, 2010.

[20] A. Radwan, K. M. S. Huq, S. Mumtaz, K. F. Tsang, and J. Rodriguez, "Low-cost on-demand c-ran based mobile small-cells," *IEEE Access*, vol. 4, pp. 2331–2339, 2016.

[21] "Common Public Radio Interface Specification V6. 0," http://cpri.info/spec.html, accessed: 2017-04-07.

[22] "5G PPP Architecture Working Group View on 5G Architecture," https://5g-ppp.eu/wp-content/uploads/2014/02/5G-PPP-5G-Architecture-WP-July-2016.pdf, accessed: 2017-04-03.

[23] "5G XHaul Project Deliverable D2.1 Requirements Specification and KPIs Document," http://5g-xhaul-project.eu/download/5G-XHaul_D_21.pdf, accessed: 2017-04-03.

[24] A. Alkhateeb, G. Leus, and R. W. Heath, "Limited feedback hybrid precoding for multi-user millimeter wave systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 11, pp. 6481–6494, 2015.

[25] M. N. Kulkarni, A. Alkhateeb, and J. G. Andrews, "A tractable model for per user rate in multiuser millimeter wave cellular networks," in *2015 49th Asilomar Conference on Signals, Systems and Computers*, Nov 2015, pp. 328–332.

[26] A. Alkhateeb, R. W. Heath, and G. Leus, "Achievable rates of multi-user millimeter wave systems with hybrid precoding," in *IEEE International Conference on Communication Workshop (ICCW)*, 2015, pp. 1232–1237.

[27] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave mimo systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1499–1513, 2014.

[28] A. Ghosh, T. N. Thomas, M. C. Cudak, R. Ratasuk, P. Moorut, F. W. Vook, T. S. Rappaport, G. R. MacCartney, S. Shun, and S. Nie, "Millimeter-wave enhanced local area systems: A high data-rate approach for future wireless networks," *IEEE J. Select. Areas Commun.*, vol. 32, no. 6, pp. 1153–1163, June 2014.

[29] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Shun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave channel modeling and cellular capacity evaluation," *IEEE J. Select. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, June 2014.

[30] S. Singh, M. N. Kulkarni, A. Ghosh, and J. G. Andrews, "Tractable model for rate in self-backhauled millimeter wave cellular networks," *IEEE J. Select. Areas Commun.*, vol. 33, no. 1, pp. 2196–2211, Jan. 2015.

[31] T. Bai, A. Alkhateeb, and R. W. Heath, "Coverage and capacity of millimeter-wave cellular networks," *IEEE Communications Magazine*, vol. 52, no. 9, pp. 70–77, September 2014.

[32] T. Bai and R. W. Heath, "Coverage and rate analysis for millimeter-wave cellular networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 1100–1114, Feb. 2015.

[33] M. N. Kulkarni, A. Ghosh, and J. G. Andrews, "A comparison of mimo techniques in downlink millimeter wave cellular networks with hybrid beamforming," *IEEE Transactions on Communications*, vol. 64, no. 5, pp. 1952–1967, May 2016.

[34] J. Singh and S. Ramakrishna, "On the feasibility of beamforming in millimeter wave communication systems with multiple antenna arrays," in *IEEE Global Communications Conference (GLOBECOM)*, 2014, pp. 3802–3808.

[35] J. Wu, Z. Zhang, Y. Hong, and Y. Wen, "Cloud radio access network (c-ran): a primer," *IEEE Network*, vol. 29, no. 1, pp. 35–41, Jan 2015.

[36] J. Bartelt, P. Rost, D. Wubben, J. Lessmann, B. Melis, and G. Fettweis, "Fronthaul and backhaul requirements of flexibly centralized radio access networks," *IEEE Wireless Communications*, vol. 22, no. 5, pp. 105–111, October 2015.

[37] L. Kleinrock, *Queuing systems*. wiley, 1975.

[38] T. Han and N. Ansari, "Network utility aware traffic load balancing in backhaul-constrained cache-enabled small cell networks with hybrid power supplies," *IEEE Transactions on Mobile Computing*, vol. PP, no. 99, pp. 1–1, 2017.

[39] D. B. Cheikh, J. M. Kelif, M. Coupechoux, and P. Godlewski, "Multicellular zero forcing precoding performance in rayleigh and shadow fading," in *IEEE 73rd Vehicular Technology Conference (VTC Spring)*, May 2011, pp. 1–5.

[40] N. L. Johnson, S. Kotz, and N. Balakrishnan, "Continuous univariate distributions, vol. 1-2," 1994.
[41] M. Haenggi, *Stochastic geometry for wireless networks*. Cambridge University Press, 2012.
[42] M. D. Renzo, A. Guidotti, and G. E. Corazza, "Average rate of downlink heterogeneous cellular networks over generalized fading channels: A stochastic geometry approach," *IEEE Transactions on Communications*, vol. 61, no. 7, pp. 3050–3071, July 2013.
[43] J. Bartelt, D. Wübben, P. Rost, J. Lessmann, and G. Fettweis, "Fronthaul for a flexible centralization in cloud radio access networks," *Backhauling/Fronthauling for Future Wireless Systems*, pp. 55–84, 2016.
[44] A. M. Ibrahim, T. ElBatt, and A. El-Keyi, "Coverage probability analysis for wireless networks using repulsive point processes," in *IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Sept 2013, pp. 1002–1007.

Oluwatayo Kolawole (S'15) received her B.Eng. (Hons.) degree in Electrical Electronics Engineering from Abubakar Tafawa Balewa Univeristy, Nigeria in 2010 and an M.Sc. in Signal Processing and Communications from the University of Edinburgh in 2014. She is currently a Ph.D. candidate at the University of Edinburgh's Institute for Digital Communications. Her main area of research is wireless communications, with particular focus on millimeter wave and stochastic geometry.

Satyanarayana Vuppala (S'12-M'17) received the B.Tech. degree with distinction in Computer Science and Engineering from JNTU Kakinada, India, in 2009, and the M.Tech. degree in Information Technology from the National Institute of Technology, Durgapur, India, in 2011. He received the Ph.D. degree in Electrical Engineering from Jacobs University Bremen in 2014. He is currently a post-doctoral researcher at IDCOM in University of Edinburgh. His main research interests are physical, access, and network layer aspects of wireless security. He also works on performance evaluation of mmWave systems. He is a recipient of MHRD, India scholarship during the period of 2009-2011.

Tharmalingam Ratnarajah (A'96-M'05-SM'05) is currently with the Institute for Digital Communications (IDCOM), University of Edinburgh, Edinburgh, UK, as a head of IDCOM and Professor in Digital Communications and Signal Processing. His research interests include signal processing and information theoretic aspects of 5G wireless networks, full-duplex radio, mmWave communications, random matrices theory, interference alignment, statistical and array signal processing and quantum information theory. He has published over 300 publications in these areas and holds four U.S. patents. He is currently the coordinator of the FP7 projects HARP (3.2M€) in the area of highly distributed MIMO and ADEL (3.7M€) in the area of licensed shared access. Previously, he was the coordinator of FP7 Future and Emerging Technologies project CROWN (2.3M€) in the area of cognitive radio networks and HIATUS (2.7M€) in the area of interference alignment. Dr Ratnarajah is a Fellow of Higher Education Academy (FHEA), U.K., and an associate editor of the IEEE Transactions on Signal Processing.