IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 9 December 2022, accepted 3 January 2023, date of publication 9 January 2023, date of current version 17 January 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3235596

RESEARCH ARTICLE

Performance Analysis of Scalable User-Centric HC-RAN Over Rician Fading Channels

HAREESH AYANAMPUDI AND RAVINDRA DHULI
School of Electronics Engineering, VIT-AP University, Amaravati, Andhra Pradesh 522237, India
Corresponding author: Ravindra Dhuli (ravindradhuli@gmail.com)

ABSTRACT In this paper, we analyze the uplink performance of a scalable user-centric heterogeneous cloud-radio access network (HC-RAN) implemented by using dynamic cooperative clustering (DCC) framework over a Rician fading channel with phase shifts. This channel describes various practical aspects, such as a deterministic line of sight (LoS) component and random non-LoS (N-LoS) components. To account for the phase shifts due to user mobility, the phase of the LoS component is modeled as a uniformly distributed random variable. We assume that phase information is available at each remote radio head (RRH). We derived the phase aware-minimum mean square error (PA-MMSE) and phase unaware Linear-MMSE estimators and obtained the channel state information (CSI). We derived a closed-form expression for the achievable spectral efficiency (SE) to evaluate the system performance with both estimators. To address the effect of coherent interference in the ultra-dense networks, we developed a two-layer decoding scheme in uplink in which maximum ratio (MR) combining is performed at the RRH and large-scale fading decoding is performed at the base-band unit (BBU) pool. Based on the obtained results, the proposed method enhanced the uplink performance in an ultra-dense scenario. It is validated by comparing it with the simulation results. Moreover, the performance loss caused by the lack of phase knowledge will depend on the pilot sequence length.

INDEX TERMS Channel estimation, HC-RAN, network scalability, Rician fading, spectral efficiency, user-centric network.

I. INTRODUCTION

With advancements in wireless technology like the Internet of Things (IoT), and massive machine-type communication (mMTC), the aggregate traffic demand over a cellular network is increasing rapidly. As per the report produced by Cisco (2017-22), the amount of global traffic will be increased by seven-fold [1]. To keep up with the rising demands of massive connectivity, data rates, capacity, and bandwidth, many researchers and industries focused on developing new radio access network (RAN) architectures. Since existing mobile network capacity is reaching the Shannon limit, many operators somehow try to satisfy these requirements by deploying more base stations (BSs). However, dense deployment of BSs will increase expenditure and operational costs and make the system more complex. Recently, the heterogeneous cloud-radio access network (HC-RAN) emerged as one of the promising network architectures which combine the features of cloud-radio access network (C-RAN) and heterogeneous BS deployment scheme of heterogeneous networks (Hetnets) with meeting the future traffic requirements. The HC-RAN has the potential to enhance the network capacity as well as energy efficiency (EE). The deployed macro base station (MBS) in HC-RAN can provide the guaranteed quality of service (QoS) to the users and reduce operational and expenditure costs by using cooperative radio resource management [2], [3], [4], [5].

In HC-RAN, the centralized signal processing feature facilitates coordination among the remote radio heads (RRHs). This can enable the coordinated multi-point (CoMP) transmission of RRHs, which can jointly serve the users to improve the ubiquitous user experience [6]. Compared to the traditional cellular transmission, the future generation wireless networks are drawing much attention towards a user-centric approach where the coverage is planned around the user than the traditional cell-centric approach where the edge user is affected by the cell edge effects. Hence, future

The associate editor coordinating the review of this manuscript and approving it for publication was Giuseppe Desolda.

wireless networks require efficient network scaling in terms of user-specific clusters, which can not be achieved by using the conventional clustering approaches. Because of their static nature and lack of user-side channel state information (CSI), they provide poor system performance. Also, they neglect the cost in terms of complexity, latency, overhead, and power consumption [7], [8], [9], [10], [11].

To address these issues, many researchers focused on developing dynamically adaptive clustering approaches [12], [13], [14], [15]. These approaches form clusters dynamically for each user and are overlapped with each other to avoid the effects of the cell edges. The performance of the user-centric network depends on the RRH selection and cluster formation. Since, the user-side CSI is continuously available at RRHs. Based on the user's mobility and QoS requirements, each user-centric cluster is updated by adding or dropping RRHs [7], [12], [16]. This will cause huge system overhead, latency, and power consumption. Also, in practical ultra-dense HC-RANs, clustering becomes more complex due to the availability of several possibilities. Despite their high complexity and cost, most of the dynamic clustering techniques ignore the key system information, such as density and available resources for RRHs, which results in significant discrepancies between the traffic loads at different RRHs. To avoid these drawbacks, in [12] and [17], the authors proposed dynamic clustering approaches that balance the traffic load between the different RRHs.

Most of the presented works unveiled the importance of dynamic user-centric clustering and computed the system performance mostly with Rayleigh fading, and a few with Rician fading [18], [19]. In practical ultra-dense systems like HC-RANs, the channel may consist of a deterministic line of sight (LoS) component and small-scale fading caused by the multi-path propagation, which can be modeled as a Rician fading [20]. A small change in the user location results in a significant difference in the phase of the channel response. Specifically, it may significantly impact the system performance in high-mobility environments. These effects are excluded in many presented works. Moreover, coherent interference due to pilot contamination is a severe issue in ultra-dense systems. Hence, mitigation of the pilot contamination and exploring the importance of the phase knowledge of the LoS components on system performance is much needed in ultra-dense systems for future wireless applications. From these observations, we will derive the phase aware-minimum mean square error (PA-MMSE) and phase unaware Linear-MMSE estimation schemes to investigate the importance of phase knowledge. Besides, we addressed the network scalability issue in dense networks by maintaining the computational complexity and fronthaul load constant irrespective of the number of users. Finally, the performance of the scalable user-centric HC-RAN over the Rician fading channel is analyzed with both estimation schemes to exploit the importance of having the phase knowledge on system performance.

The main contributions of this work are as follows:

1. We considered the user-centric HC-RAN with Rician fading channel between RRH and user pair and obtained the conditions for network scalability. Hence, the fronthaul load and computational complexity remain constant.

2. We derived the PA-MMSE channel estimator and phase unaware Linear-MMSE estimator and obtained the CSI of the users. For each coherence block, the phase of the LoS component is designed as an independent and identically distributed (i.i.d) random variable.

3. We proposed a two-layer decoding scheme to mitigate the pilot contamination effect by employing maximal ratio (MR) combining in the first layer and large-scale fading decoding in the second layer. Also, we derived the closed-form expression for the achievable spectral efficiency (SE) per user for both estimators.

4. Finally, we evaluated the uplink performance of the scalable user-centric HC-RAN over a Rician fading channel in terms of achievable SE per user (bits/s/Hz) with both estimators.

This paper is organized as follows. Related works are discussed in Section II. Section III introduces the network model for HC-RAN with the channel estimation process. Section IV describes the DCC approach for the user-centric HC-RAN and network scalability. In Section V, the uplink performance of scalable user-centric HC-RAN for both estimators is analyzed. Then, the numerical results are discussed in section VI. Finally, concluding remarks are provided in section VII.

A. NOTATIONS

The boldface upper case letters indicate the matrices, the boldface lower case letters indicate the column vectors, $\mathbf{I}_{\mathbf{M_T}}$ is an identity matrix of order $\mathbf{M_T}$. $\mathcal{N}_C$ $(\mathbf{0}, \mathbf{R})$ is a circular symmetric complex Gaussian distribution with correlation matrix $\mathbf{R}$. $(.)^H$, $(.)^T$, trace $(.)$, diag $(.)$ indicates the Hermitian, transpose, trace and diagonal. $\mathbb{E}\{.\}$ indicates the expectation of a random variable. The remaining notations are given in Table. 1.

II. RELATED WORKS

A number of studies have focused on improving the performance of user-centric HC-RANs subject to minimizing the pilot contamination and fronthaul load in an ultra-dense scenario. In general, the transmit power of the MBS in HC-RAN is much higher than the RRHs. Based on the received signal strength, users tend to associate more with the MBS than the RRHs. This will cause severe inter-user interference. Hence, to mitigate this, an optimized strategy is required for user association within the user-centric HC-RANs. User association with RRH or MBS will greatly impact the system's performance [4]. In [21], the authors proposed a joint optimization problem for FDD-based ultra-dense C-RAN by considering channel uncertainties. They optimized the user association and beam-forming vectors to minimize the transmit power subject to user rate requirements. Zhang et al. [22] proposed CoMP-based interference

TABLE 1. List of notations.

Symbol	description
K	The number of users
N	The number of RRHs
h_{lk}	The channel between the $l-$th RRH and $k-$th user
$\Phi_{\tau_{lk}}$	Pilot signal assigned to $k-$th user from $l-$th RRH
$\mathbf{h}_k$	The collective channel for $k-$th user from all the serving RRHs
$\mathbf{R}_k$	The collective spatial correlation matrix of the channel $\mathbf{h}_k$
$\hat{\mathbf{h}}_k$	The MMSE estimate of the channel $\mathbf{h}_k$
$\tilde{\mathbf{h}}_k$	Estimation error of the channel $\mathbf{h}_k$
$\mathbf{C}_k$	Estimation error correlation matrix
$\mathcal{S}_k$	Set of users that use the same pilot signal as $k-$th user
β_{lk}	The large-scale fading gain between the $l-$th RRH and $k-$th user
$\bar{g}_{lk}$	Mean value of the LoS signal
$\mathbf{\Lambda}_k$	Received signal spatial correlation matrix
$\hat{x}_k$	Estimate of the signal x_k
$\mathbf{y}_p^k$	Collective received signal for $k-$th user from all RRHs
$\mathbf{\Theta}_k$	Collective phase shifts of received $k-$th user signal from all RRHs
$\mathcal{U}_l$	Set of all users corresponding to $l-$th RRH provided with interference coordination
$\mathcal{Z}_l$	Set of users corresponding to $l-$th RRH, which are served with data along interference coordination
$\mathbf{v}_{lk}$	The combining vector from $l-$th RRH to the $k-$th user
B_k	A subset of RRHs serving $k-$th user
$\mathcal{C}_l$	The cardinality set of $l-$th RRH

mitigation in HC-RANs. The authors in [23] proposed the user-centric power allocation schemes, where the user is served by a single BS in a static user-centric cluster while the remaining BSs will adjust their power levels to mitigate the effect due to interference. Ali et al. [24] proposed joint user association, admission control, and power allocation for maximizing the throughput of the HC-RAN. The authors in [25] proposed the scalable user-centric framework for HC-RAN with dynamic clustering and evaluated the system performance over a Rayleigh fading scenario. The authors in [26] and [27] proposed dynamic cooperative clustering (DCC) based user-centric approaches for interference cooperation and evaluated the system performance. In this, each RRH has its own set of users for providing interference coordination. In [28] and [29], the authors proposed and analyzed the performance of the two-layer decoding scheme with limited BS cooperation to eliminate the effect of the interference in the uplink. This approach is more generalized for practical spatially correlated Rayleigh fading channels in [30]. In [18], the authors implemented a large-scale precoding scheme for downlink multi-cell massive multiple input multiple output (MIMO) systems under Rician fading to eliminate inter-cell interference. They have also proposed two partial two-layer precoding schemes for reducing the fronthaul signaling load. Abbas et al. [19] analyzed the performance of user-centric small-BS-based Hetnets under a hybrid Rician and Rayleigh fading environment. The Joint transmission from the RRHs requires phase synchronization, as discussed in [31].

As mentioned earlier, in an ultra-dense scenario, the channel can have a deterministic LoS component and small-scale fading caused by the multi-path propagation, which can be modeled as a Rician fading [20]. A small change in user location results in a significant change in the phase shift of the LoS component, which can considerably influence the system performance, especially in high-mobility environments. These effects are neglected in the analysis of Rician fading channels in [19], and the downlink scenario is considered in [18] with phase shifts for massive MIMO systems. In [32], the authors investigated the joint optimization problem to enhance the user QoS subject to the fronthaul capacity constraints. The authors in [33] proposed a user-centric framework for enhancing the capacity of the target users by jointly optimizing the CoMP cell selection and resource allocation. Zaidi et al. [34] proposed new dynamic clustering approaches for improving the user throughput and reducing the overhead cost and signaling changes at both user and network sides compared to the existing techniques. In [35], the authors proposed a dynamic user-centric clustering scheme for millimeter-wave networks and evaluated the performance in terms of coverage probability and average SE. Most of the existing works analyzed the system performance with dynamic user-centric clustering techniques over a Rayleigh fading, and few are explored with Rician fading. However, the impact of phase information of the LoS component on system performance and pilot contamination is not yet addressed in scalable user-centric HC-RANs. These gaps are addressed in our work. In the next section, we will explain the network model for user-centric HC-RAN.

III. NETWORK MODEL

The architecture of the proposed user-centric HC-RAN is shown in Fig. 1. It consists of a pool of base-band units (BBU pool) and N RRHs, each equipped with M_T antennas which are distributed over a geographical area. Assume that there are K single antenna users within the coverage area of the HC-RAN. Each RRH in the network is connected to the BBU pool via fronthaul links. We assumed the block fading model and channels are constant within a coherence block of length τ_c channel uses. For a single antenna RRHs, the channel between the $k-$th user and $l-$th RRH is h_{lk} given by

$$h_{lk} = \bar{g}_{lk} \exp(j\theta_{lk}) + g_{lk}, \tag{1}$$

where h_{lk} is a Rician fading channel, $j = \sqrt{-1}$ is the complex notation, $g_{lk} \sim \mathcal{N}_{\mathbb{C}}(0, \beta_{lk})$ indicates the non-LoS (N-LoS) components with variance β_{lk}, which can model the large-scale fading including path loss and shadow fading, $\bar{g}_{lk} \in \mathbb{R}$ represents the mean value of the LoS components, and $\theta_{lk} \in [-\pi, \pi]$ indicates the phase shift of LoS components [20]. From equation (1), it is noted that the channel h_{lk} is Rician distributed and is independent for every user and RRH. The channel realization h_{lk} in each coherence block is independent and identically distributed (i.i.d). Each coherence block is divided into τ_p channel uses for pilot transmission, τ_u channel uses for uplink data transmission, and τ_d channel uses for downlink data transmission. Assume that the system operates in time division duplex (TDD) mode. The uplink and downlink channels are estimated by using

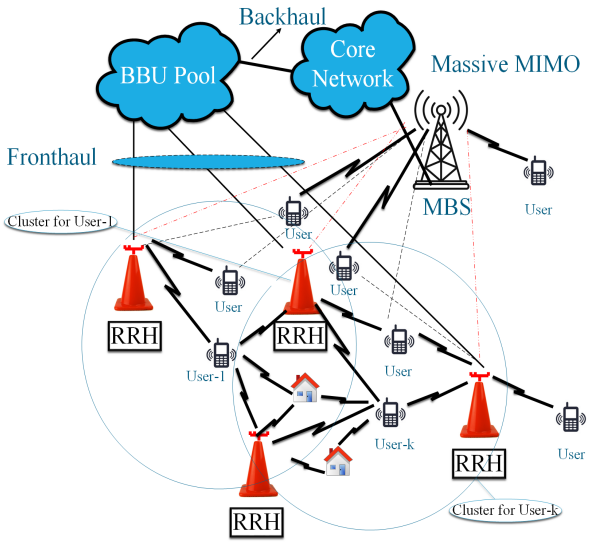

FIGURE 1. Architecture of user-centric HC-RAN.

the uplink pilot transmission phase and channel reciprocity which is discussed in the subsequent section.

A. CHANNEL MODEL

Each RRH requires the CSI for receiver processing. Therefore, let us assume that there are τ mutually orthogonal pilots signals each of length τ_p samples reserved within each coherence block of length τ_c for estimating the channel. Let the pilot signal allocated to the $k-$th user through $l-$th RRH is denoted by $\Phi_{\tau_{lk}} \in \mathbb{C}^{\tau_p \times 1}$. The pilot signals are designed to satisfy $\Phi_{lk}^H \Phi_{lk} = \tau$. Each user in the network sends a known pilot signal to the RRHs. For a single antenna RRHs, the received pilot signal at the $l-$th RRH is $\mathbf{y}_p^l \in \mathbb{C}^{\tau_p \times 1}$ given by

$$\mathbf{y}_p^l = \sum_{i=1}^{K} \sqrt{p_i} h_{li} \Phi_{\tau_{li}} + \mathbf{n}_p^l, \tag{2}$$

where p_i is the transmit power of the $i-$th user, $\mathbf{n}_p^l \sim \mathcal{NC}(0, \sigma_{ul}^2)$ is the received additive noise.

For obtaining the CSI of the desired $k-$th user, the BBU pool correlates $\mathbf{y}_p^l$ with the known pilot signal $\Phi_{\tau_{lk}}^H$. This results in

$$\begin{aligned} y_p^{lk} &= \Phi_{\tau_{lk}}^H \mathbf{y}_p^l \\ &= \sqrt{p_k}\tau h_{lk} + n^{lk}, \end{aligned} \tag{3}$$

where τ is the processing gain, $n^{lk} = \Phi_{\tau_{lk}}^H \mathbf{n}_p^l$ is the received noise, which is distributed as $n^{lk} \sim \mathcal{NC}(0, \tau\sigma_{ul}^2)$.

In practical ultra-dense networks, more than one user can share the same pilot signal when the number of users $K > \tau$. Assume that the set $\mathcal{S}_k$ represents the users that use the same pilot signal as the k-th user, including itself. Then, the received signal in (3) can be modified as

$$y_p^{lk} = \sqrt{p_k}\tau h_{lk} + \sum_{(l',i') \in \mathcal{S}_k \setminus (l,k)} \sqrt{p_{i'}}\tau h_{l'i'} + n^{lk}. \tag{4}$$

From equation (4), sharing a pilot signal with more than one user can lead to pilot contamination. The main consequence of this is that it reduces the estimation quality and causes coherent interference by correlating the channel estimates of users $i \in \mathcal{S}_k$. The effect of coherent interference increases with the number of antennas, hence gaining much attention in dense networks [36], [37]. In the subsequent sections, we will derive the PA-MMSE and phase unaware Linear-MMSE channel estimators for the Rician fading channel. The CSI is extracted to characterize the importance of having the phase information of LoS components. The minimum mean square error (MMSE) estimator in this work will give an optimal solution by considering the interference from the users of neighboring BSs and interference from the neighboring users of the same BS and noise. It exploits the statistical characteristics of the channel to obtain better channel estimates. The other alternative schemes employed for this are element wise-MMSE (EW-MMSE) and least square (LS) channel estimators. But they provide sub-optimal results. The EW-MMSE channel estimation is less complex but ignores the correlation between the antenna elements. The LS channel estimation is used when no channel statistical information is available at the BS [37], [38].

B. PHASE UNAWARE LINEAR-MMSE CHANNEL ESTIMATION

The channel statistics $\bar{g}_{lk}$, and β_{lk} are available at the RRHs. The Linear-MMSE channel estimate of the k-th user without having phase knowledge for the Rician faded channel $\hat{h}_{lk}$ corresponding to the $l-$th RRH is given as [37] and [39]

$$\hat{h}_{lk} = \sqrt{p_k}\beta'_{lk} \left(\lambda'_{lk}\right)^{-1} y_p^{lk}, \tag{5}$$

where $\beta'_{lk} = \beta_{lk} + \bar{g}_{lk}^2$, $\lambda'_{lk} = \sum_{(l,i) \in \mathcal{S}_k} p_i \tau \left(\beta_{li} + \bar{g}_{li}^2\right) + \sigma_{ul}^2$. The proof of expression in (5) is given in the Appendix. A. The estimate $\hat{h}_{lk}$ has zero mean and a variance $p_k \beta'_{lk} \left(\lambda'_{lk}\right)^{-1} \beta'_{lk}$.

The estimation error can be expressed as $\tilde{h}_{lk} = h_{lk} - \hat{h}_{lk}$. It has zero mean and a variance c'_{lk}. By using standard estimation theory, c'_{lk} can be expressed as (From Appendix. A)

$$c'_{lk} = \beta'_{lk} - p_k \beta'_{lk} \left(\lambda'_{lk}\right)^{-1} \beta'_{lk}. \tag{6}$$

The channel estimate $\hat{h}_{lk}$ and estimation error $\tilde{h}_{lk}$ both are uncorrelated random variables. For a generalized multi-antenna scenario, the collective channel estimate for the $k-$th user $\hat{\mathbf{h}}_k \in \mathbb{C}^{NM_T \times 1}$ is given by

$$\hat{\mathbf{h}}_k = \sqrt{p_k} \mathbf{R}'_k \Lambda'_k \mathbf{y}_p^k, \tag{7}$$

where $\mathbf{R}'_k = \text{diag}\left(\beta'_{1k}, \ldots, \beta'_{Nk}\right) \in \mathbb{C}^{NM_T \times NM_T}$, $\Lambda'_k = \text{diag}\left(\lambda'_{1k}, \ldots, \lambda'_{Nk}\right)^{-1} \in \mathbb{C}^{NM_T \times NM_T}$, and $\mathbf{y}_p^k = \left[y_p^{1k}, \ldots, y_p^{Nk}\right]^T \in \mathbb{C}^{NM_T \times 1}$. The Linear-MMSE estimate $\hat{\mathbf{h}}_k$

has a zero mean and co-variance $\mathbf{R}'_k - \mathbf{C}'_k$, where $\mathbf{C}'_k = \mathbf{R}'_k - p_k \mathbf{R}'_k \Lambda'_k \mathbf{R}'_k$. The mean square error (MSE) can be expressed as $\mathbb{E}\left\{\left\|\mathbf{h}_k - \hat{\mathbf{h}}_k\right\|^2\right\} = \text{trace}\left(\mathbf{C}'_k\right)$.

C. PA-MMSE CHANNEL ESTIMATION

We assume that the phase information θ_{lk} is known at the $l-$th RRH along with channel statistics $\bar{g}_{lk}$ and β_{lk}. For single antenna RRHs, the PA-MMSE channel estimate of the Rician faded channel $\hat{h}_{lk}^{\text{PA-MMSE}}$ is given by

$$\hat{h}_{lk}^{\text{PA-MMSE}} = \bar{g}_{lk} e^{j\theta_{lk}} + \sqrt{p_k}\beta_{lk} (\lambda_{lk})^{-1} \left(y_p^{lk} - \bar{y}_p^{lk}\right), \quad (8)$$

where $\lambda_{lk} = \sum_{(l,i)\in\mathcal{S}_k} \sqrt{p_i}\tau\beta_{li} + \sigma_{ul}^2$, and $\bar{y}_p^{lk} = \sum_{(l,i)\in\mathcal{S}_k} \sqrt{p_i}\tau\bar{g}_{li} e^{j\theta_{li}}$. The equation in (8) is computed based on Appendix. A. For a given phase information, the estimate $\hat{h}_{lk}^{\text{PA-MMSE}}$ has a mean $\bar{g}_{lk} e^{j\theta_{lk}}$ and a variance $p_k\beta_{lk}\lambda_{lk}\beta_{lk}$. The estimation error can be expressed as $\tilde{h}_{lk}^{\text{PA-MMSE}} = h_{lk} - \hat{h}_{lk}^{\text{PA-MMSE}}$. It has zero mean and a variance c_{lk}, which is given by (From Appendix. A)

$$c_{lk} = \beta_{lk} - p_k\beta_{lk}\lambda_{lk}\beta_{lk}. \quad (9)$$

The terms y_p^{lk}, $\bar{y}_p^{lk}$ and θ_{lk} are constant within each coherence block and these will differ in each coherence block. For generalized multi-antenna scenario, the collective channel estimate for $k-$th user $\hat{\mathbf{h}}_k^{\text{PA-MMSE}} \in \mathbb{C}^{NM_T\times 1}$ is given by

$$\hat{\mathbf{h}}_k^{\text{PA-MMSE}} = \Theta_k \bar{\mathbf{g}}_k + \sqrt{p_k}\mathbf{R}_k \Lambda_k \left(\mathbf{y}_p^k - \bar{\mathbf{y}}_p^k\right), \quad (10)$$

where $\bar{\mathbf{g}}_k = [\bar{g}_{1k}, \ldots, \bar{g}_{Nk}]^T \in \mathbb{C}^{NM_T\times 1}$, $\Theta_k = \text{diag}\left[e^{j\theta_{1k}}, \ldots, e^{j\theta_{Nk}}\right] \in \mathbb{C}^{NM_T\times NM_T}$, $\bar{\mathbf{y}}_p^k = \left[\bar{y}_p^{1k}, \ldots, \bar{y}_p^{Nk}\right]^T \in \mathbb{C}^{NM_T\times 1}$, $\Lambda_k = \text{diag}\left(\lambda_{1k}, \ldots, \lambda_{Nk}\right)^{-1} \in \mathbb{C}^{NM_T\times NM_T}$, and $\mathbf{R}_k = \text{diag}\left(\beta_{1k}, \ldots, \beta_{Nk}\right) \in \mathbb{C}^{NM_T\times NM_T}$. The collective channel estimate $\hat{\mathbf{h}}_k^{\text{PA-MMSE}}$ has a mean $\mathbb{E}\left\{\hat{\mathbf{h}}_k^{\text{PA-MMSE}} \mid \bar{\mathbf{g}}_k\Theta_k\right\} = \bar{\mathbf{g}}_k\Theta_k$ and the co-variance $cov\left\{\hat{\mathbf{h}}_k^{\text{PA-MMSE}} \mid \bar{\mathbf{g}}_k\Theta_k\right\} = \mathbf{R}_k - \mathbf{C}_k$, where $\mathbf{C}_k = \mathbf{R}_k - p_k\mathbf{R}_k\Lambda_k\mathbf{R}_k$. The MSE of the estimate can be expressed as $\mathbb{E}\left\{\left\|\mathbf{h}_k - \hat{\mathbf{h}}_k^{\text{PA-MMSE}}\right\|^2\right\} = \text{trace}\left(\mathbf{C}_k\right)$.

In the next section, we will develop the user-centric framework for HC-RAN using the DCC approach and obtain the conditions for network scalability.

IV. DCC APPROACH FOR USER-CENTRIC HC-RAN AND NETWORK SCALABILITY

The DCC scheme was proposed in [26] and [27] to provide a unified analysis of the channels with interference in the cellular network. This approach is characterized by two sets $\mathcal{Z}_l$ and $\mathcal{U}_l$ for all $l \in N$ as shown in Fig. 2. The set of users served with data by the $l-$th RRH is denoted as $\mathcal{Z}_l$. The set $\mathcal{U}_l$ contains all the users corresponding to the $l-$th RRH, which are provided with interference coordination based on the available CSI. Let us define the set of diagonal matrices

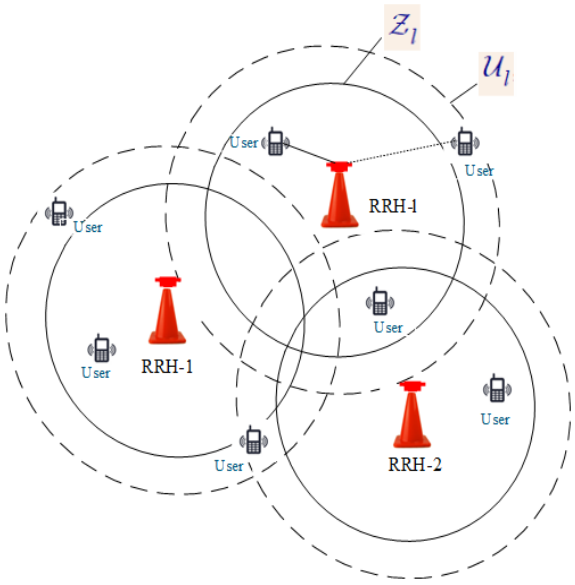

FIGURE 2. Illustration of the DCC based clustering.

$\mathbf{Z}_{li} \in \mathbb{C}^{M_T\times M_T}$ to determine which antenna of $l-$th RRH is serving data to the $i-$th user and $\mathbf{U}_{li} \in \mathbb{C}^{M_T\times M_T}$ defines the interference coordination to users in the $l-$th cell. It can be noted that $\mathbf{Z}_{li} \subseteq \mathbf{U}_{li}$.

Based on the DCC scheme, only certain channel elements will carry the information and interference. These can be selected by using the diagonal matrices $\mathbf{U}_i = \text{diag}\left[\mathbf{U}_{1i}, \ldots, \mathbf{U}_{Ni}\right] \in \mathbb{C}^{NM_T}$ and $\mathbf{Z}_i = \text{diag}\left[\mathbf{Z}_{1i}, \ldots, \mathbf{Z}_{Ni}\right] \in \mathbb{C}^{NM_T}$. The non-zero elements in the diagonal matrix will serve the $i-$th user. Specifically, the $m-$th diagonal element of the matrix $\mathbf{Z}_{li}$ is 1 only if the $m-$th antenna of the $l-$th RRH is serving the $i-$th user and similarly for $\mathbf{U}_{li}$. For each user in a cell, the sets $\mathbf{Z}_{li}$ and $\mathbf{U}_{li}$ can be expressed as

$$\mathbf{Z}_{li} = \begin{cases} \mathbf{I}_{M_T} & i \in \mathcal{Z}_l, \\ \mathbf{0}_{M_T} & \text{otherwise.} \end{cases}, \quad (11)$$

$$\mathbf{U}_{li} = \begin{cases} \mathbf{I}_{M_T} & i \in \mathcal{U}_l, \\ \mathbf{0}_{M_T} & \text{otherwise.} \end{cases} \quad (12)$$

For the desired $k-$th user, the user-centric framework for HC-RAN can be obtained based on the DCC scheme by letting B_k, a subset of RRHs serving the $k-$th user. Hence, the matrices $\mathbf{Z}_{lk}$ and $\mathbf{U}_{lk}$ can be modified as

$$\mathbf{Z}_{lk} = \begin{cases} \mathbf{I}_{M_T} & l \in B_k, \\ \mathbf{0}_{M_T} & \text{otherwise.} \end{cases}, \quad (13)$$

$$\mathbf{U}_{lk} = \begin{cases} \mathbf{I}_{M_T} & l \in B_k, \\ \mathbf{0}_{M_T} & \text{otherwise.} \end{cases} \quad (14)$$

The joint transmission and interference coordination with DCC provides more spatial degrees of freedom to separate the users.

In practice, as the number of users in the network increases, the complexity of computation and the number of scalars sent over a fronthaul link increases. The limited fronthaul capacity

and computational complexity make the system unscalable. The DCC scheme proposed in [27] addresses these issues in multi-cell coordinated systems by defining the set of users which are served by at least one of the RRH antennas and is given by

$$C_l = \{i : \mathrm{tr}\,(\mathbf{Z}_{li}) \geqslant 1, l \in \{1, \ldots, N\}\, i \in \{1, \ldots, K\}\}\,. \quad (15)$$

If the cardinality of each RRH, $\mid C_l \mid$ is constant with an increase in the number of users, then the number of scalars sent over the fronthaul for each RRH is limited to only $\mid C_l \mid$ number of users. And also, cardinality $\mid C_l \mid \leq \tau_p$ for all $l = 1, \ldots, N$ and τ_p is independent of K. Hence, the computational complexity and fronthaul load remain constant irrespective of the number of users. Hence, the system is scalable. In this proposed system, the dynamic user-centric clusters are formed based on the sub-optimal algorithms proposed in [35] and [40]. Since it will at least guarantees the service to the user by avoiding the risk of a user being dropped abruptly from service when no RRH is ready to serve it.

The previous user-centric approaches [12], [15] doesn't guarantee QoS to the users when there exists a pilot contamination effect, affecting the system performance severely in an ultra-dense network. To avoid this drawback, we developed the two-layer decoding scheme, which can have the ability to reduce the effect of pilot contamination in an ultra-dense scenario.

The uplink performance analysis of DCC based user-centric HC-RAN with network scalability for both estimators is presented in the following section.

V. UPLINK PERFORMANCE ANALYSIS

During the uplink data transmission, each user sends the data symbols to the RRHs. The received signal at the $l-$th RRH $\mathbf{y}_l^{ul} \in \mathbb{C}^{M_T}$ is given by

$$\mathbf{y}_l^{ul} = \sum_{i=1}^{K} \mathbf{h}_{li} x_i + \mathbf{n}_l^{ul}, \quad (16)$$

where $x_i \sim \mathbb{C}\,(0, p_i)$ is the transmit power of $i-$th user and $\mathbf{n}_l^{ul} \sim \mathcal{N}_{\mathcal{C}}\left(0, \sigma_{ul}^2\right)$ is the received additive noise at $l-$th RRH. With the available channel estimates, the $l-$th RRH selects the combining vector $\mathbf{v}_{li} \in \mathbb{C}^{M_T}$ to obtain the estimate of the data signal. The estimate of the received information signal at the $l-$th RRH for the $k-$th user is given by

$$\tilde{x}_{lk} = \mathbf{v}_{lk}^H \mathbf{y}_l^{ul}. \quad (17)$$

By using DCC approach, equation in (17) can be modified as

$$\begin{aligned}\tilde{x}_{lk} &= \mathbf{v}_{lk}^H \mathbf{U}_{lk} \mathbf{Z}_{lk} \mathbf{y}_l^{ul} \\ &= \mathbf{v}_{lk}^H \mathbf{U}_{lk} \mathbf{Z}_{lk} \mathbf{h}_{lk} x_k + \sum_{i=1, i\neq k}^{K} \sum_{l=1}^{N} \mathbf{v}_{lk}^H \mathbf{U}_{lk} \mathbf{Z}_{lk} \mathbf{h}_{li} x_i \\ &\quad + \mathbf{v}_{lk}^H \mathbf{U}_{lk} \mathbf{Z}_{lk} \mathbf{n}_l^{ul} \\ &= \mathbf{v}_{lk}^H \mathbf{Z}_{lk} \mathbf{h}_{lk} x_k + \sum_{i=1, i\neq k}^{K} \sum_{l=1}^{N} \mathbf{v}_{lk}^H \mathbf{Z}_{lk} \mathbf{h}_{li} x_i + \mathbf{v}_{lk}^H \mathbf{Z}_{lk} \mathbf{n}_l^{ul}\end{aligned} \quad (18)$$

since $\mathbf{Z}_{li} \subseteq \mathbf{U}_{li}$ and product of the matrix results in $\mathbf{Z}_{li}$.

These estimates of the received information signals at each RRH are sent to the BBU pool for obtaining the final estimate of data corresponding to the $k-$th user. The collective signal estimate corresponding to the $k-$th user at the BBU pool with large-scale fading coefficients is given by

$$\begin{aligned}\hat{x}_k &= \sum_{l=1}^{N} a_{lk} \tilde{x}_{lk} \\ &= \sum_{l=1}^{N} a_{lk} \mathbf{v}_{lk}^H \mathbf{Z}_{lk} \mathbf{h}_{lk} x_k \sum_{i=1, i\neq k}^{K} \sum_{l=1}^{N} a_{lk} \mathbf{v}_{lk}^H \mathbf{Z}_{lk} \mathbf{h}_{li} x_i \\ &\quad + \sum_{l=1}^{N} a_{lk} \mathbf{v}_{lk}^H \mathbf{Z}_{lk} \mathbf{n}_l^{ul} \\ &= \mathbf{a}_k^H \mathbf{d}_k x_k + \sum_{i=1, i\neq k}^{K} \mathbf{a}_k^H \mathbf{d}_i x_i + \mathbf{a}_k^H \mathbf{v}_k^H \mathbf{Z}_k \mathbf{n}_l^{ul},\end{aligned} \quad (19)$$

where $\mathbf{d}_k = \begin{bmatrix} \mathbf{v}_{1k}^H \mathbf{Z}_{1k} \mathbf{h}_{1k} \\ . \\ . \\ . \\ \mathbf{v}_{Nk}^H \mathbf{Z}_{Nk} \mathbf{h}_{Nk} \end{bmatrix} \in \mathbb{C}^{N\times 1}$, $\mathbf{a}_k = [a_{1k}, \ldots, a_{Nk}]^T \in \mathbb{C}^{N\times 1}$ is the large-scale fading decoding weight vector corresponding to the $k-$th user and received noise $n = \sum_{l=1}^{N} a_{lk} \mathbf{v}_{lk}^H \mathbf{Z}_{lk} \mathbf{n}_l^{ul}$. The BBU pool computes weights based on the received large-scale fading coefficients. The value of the large-scale fading coefficients depends upon the distance and shadowing. Hence, these will give the quality of the received signal. Based on the quality of the received signal, it will assign the weights to each RRH with an estimated signal of the $k-$th user. The RRH, which has a higher signal-to-interference and noise ratio (SINR) to a user, is assigned a large weight. This will reduce inter-user interference effectively. The achievable SE [bit/s/Hz] of the $k-$th user is given by [37] and [38]

$$\mathrm{SE}_{\mathrm{ul}}^k = \frac{\tau_u}{\tau_c} \log_2 \left(1 + \Upsilon_{\mathrm{ul}}^k\right), \quad (20)$$

where $\frac{\tau_u}{\tau_c}$ is the pre-log factor, Υ_{ul}^k is the effective SINR of $k-$th user in the uplink and is given by with large-scale fading coefficients, the equation in (21), as shown at the bottom of the next page, can be modified as given in equation (22), as shown at the bottom of the next page.

From equation (22), we need to select the value of the large-scale fading decoding weight vector such that it needs to maximize the SINR of the user. It is obtained as equation (23), as shown at the bottom of the next page.

And the corresponding maximized achievable SE of the k-th user is given by

$$\mathrm{SE}_{\mathrm{ul}}^k = \frac{\tau_u}{\tau_c} \times \log_2 \left(1 + p_k \left(\mathbb{E}\left\{\mathbf{d}_k\right\}\right)^H \mathbf{a}_k\right). \quad (24)$$

The achievable SE expression given in (24) can be valid for any fading scheme, combining scheme, or channel estimation scheme [37], [38]. Every choice of fading scheme and combining scheme will only affect the parameters $\mathbf{d}_i$, $\mathbf{d}_k$, and $\mathbb{E}\left\{\|\mathbf{v}_k\mathbf{Z}_k\|^2\right\}$. In the subsequent sections, we present the uplink SE analysis of the two-layer decoding scheme and computed the Υ_{ul}^k for MR combining scheme by using the PA-MMSE and phase unaware Linear-MMSE estimators.

A. UPLINK SE ANALYSIS USING PHASE UNAWARE LINEAR-MMSE ESTIMATOR

The uplink achievable SE of the proposed HC-RAN using Linear-MMSE estimator without having the phase information is given by the equation in (20). From (22), by substituting all the expectations values derived in Appendix. B, the SINR in Rayleigh coefficient form [37] can be expressed as

$$\Upsilon_{\text{ul}}^k = \frac{p_k \mathbf{a}_k^H \mathbf{b}_k \mathbf{b}_k^H \mathbf{a}_k}{\mathbf{a}_k^H \Gamma_k^{\text{ul,Linear-MMSE}} \mathbf{a}_k}, \tag{25}$$

where the expression for the $\Gamma_k^{\text{ul,Linear-MMSE}}$ is given by equation in (26), as shown at the bottom of the page, and $\mathbf{b}_k = p_k \tau \text{tr}\left(\mathbf{Z}_k \Omega_k'\right)$.

The maximized achievable SE can be expressed as

$$\text{SE}_{\text{ul}}^k = \frac{\tau_u}{\tau_c} \log_2 \left(1 + p_k \mathbf{b}_k^H \left(\Gamma_k^{\text{ul,Linear-MMSE}}\right)^{-1} \mathbf{b}_k\right). \tag{27}$$

B. UPLINK SE ANALYSIS USING PA-MMSE ESTIMATOR

The uplink achievable SE of the proposed HC-RAN by using the PA-MMSE estimator is given by (20). From equation (22), by substituting all the expectations values given in Appendix. C, the resulting SINR can be expressed as

$$\Upsilon_{\text{ul}}^k = \frac{p_k \left|\text{tr}\left(\mathbf{Z}_k \mathbf{D}_k\right)\right|^2}{\Gamma_k^{\text{ul,PA-MMSE}}}, \tag{28}$$

where, the expression for $\Gamma_k^{\text{ul,PA-MMSE}}$ is given by equation in (29), as shown at the bottom of the page, $\mathbf{D}_k = \hat{p}_k \tau_p \Omega_k + \mathbf{L}_k$, and $\mathbf{L}_k = \text{diag}\left(\bar{h}_{1k}^2, \ldots, \bar{h}_{Nk}^2\right)$.

$$\Upsilon_{\text{ul}}^k = \frac{p_k \mathbf{a}_k^H \mathbf{b}_k \mathbf{b}_k^H \mathbf{a}_k}{\mathbf{a}_k^H \Gamma_k^{\text{ul,PA-MMSE}} \mathbf{a}_k}. \tag{30}$$

The maximum value of the achievable SE can be obtained by selecting the weight vector $\mathbf{a}_k = \left(\Gamma_k^{\text{ul,PA-MMSE}}\right)^{-1} \mathbf{b}_k$. Hence the uplink achievable SE can be expressed as

$$\text{SE}_{\text{ul}}^k = \frac{\tau_u}{\tau_c} \log_2 \left(1 + p_k \mathbf{b}_k^H \left(\Gamma_k^{\text{ul,PA-MMSE}}\right)^{-1} \mathbf{b}_k\right). \tag{31}$$

$$\Upsilon_{\text{ul}}^k = \frac{p_k \left|\mathbb{E}\left\{\mathbf{a}_k^H \mathbf{d}_k\right\}\right|^2}{\sum\limits_{l \in B_k} \sum\limits_{i=1}^{K} p_i \mathbb{E}\left\{\left|\mathbf{a}_k^H \mathbf{d}_i\right|^2\right\} - p_k \left|\mathbb{E}\left\{\mathbf{a}_k^H \mathbf{d}_k\right\}\right|^2 + \sigma_{ul}^2 \mathbb{E}\left\{\left\|\mathbf{a}_k^H \mathbf{Z}_k \mathbf{v}_k\right\|^2\right\}}. \tag{21}$$

$$\Upsilon_{\text{ul}}^k = \frac{p_k \left|\mathbf{a}_k^H \mathbb{E}\left\{\mathbf{d}_k\right\}\right|^2}{\mathbf{a}_k^H \left(\sum\limits_{i=1}^{K} p_i \mathbb{E}\left\{|\mathbf{d}_i|^2\right\} - p_k \left|\mathbb{E}\left\{\mathbf{d}_k\right\}\right|^2 + \sigma_{ul}^2 \mathbb{E}\left\{\|\mathbf{v}_k \mathbf{Z}_k\|^2\right\}\right) \mathbf{a}_k}. \tag{22}$$

$$\mathbf{a}_k = \left(\sum_{i=1}^{K} p_i \mathbb{E}\left\{|\mathbf{d}_i|^2\right\} - p_k \left(\mathbb{E}\left\{\mathbf{d}_k\right\}\right)\left(\mathbb{E}\left\{\mathbf{d}_k\right\}\right)^H + + \sigma_{ul}^2 \mathbb{E}\left\{\|\mathbf{v}_k \mathbf{Z}_k\|^2\right\}\right)^{-1} \mathbb{E}\left\{\mathbf{d}_k\right\}. \tag{23}$$

$$\begin{aligned}\Gamma_k^{\text{ul,Linear-MMSE}} = {} & p_i \sum_{i=1}^{K} \hat{p}_k \tau \text{tr}\left(\mathbf{Z}_k^H \mathbf{R}_l' \mathbf{Z}_k \Omega_k'\right) + p_i \hat{p}_k \hat{p}_i \tau^2 \sum_{i=S_k \setminus \{k\}} \left(\text{tr}\left(\mathbf{Z}_k^H \mathbf{R}_l^2 \Lambda_k' \Omega_k' \mathbf{Z}_k\right) + 2\text{tr}\left(\mathbf{Z}_k^H \Omega_k' \Lambda_k' \mathbf{L}_l \mathbf{R}_l \mathbf{Z}_k\right)\right. \\ & \left. + \left|\text{tr}\left(\mathbf{Z}_k \mathbf{R}_l' \Lambda_k' \mathbf{R}_k'\right)\right|^2 - \text{tr}\left(\mathbf{Z}_k^H \left(\mathbf{R}_l' \Lambda_k' \mathbf{R}_k'\right)^2 \mathbf{Z}_k\right)\right) - p_k \mathbf{b}_k \mathbf{b}_k^H + p_k \tau \sigma_{ul}^2 \text{tr}\left(\mathbf{Z}_k^H \Omega_k' \mathbf{Z}_k\right).\end{aligned} \tag{26}$$

$$\begin{aligned}\Gamma_k^{\text{ul,PA-MMSE}} = {} & \sum_{i=1}^{K} p_i \text{tr}\left(\mathbf{Z}_k^H \left(\mathbf{R}_n + \mathbf{L}_n\right) \mathbf{D}_k \mathbf{Z}_k\right) + \sum_{i=S_k \setminus \{k\}} p_i \hat{p}_k \hat{p}_i \tau_p^2 \left|\text{tr}\left(\mathbf{Z}_k \mathbf{R}_n \Lambda_k \mathbf{R}_k\right)\right|^2 - p_k \text{tr}\left(\mathbf{Z}_k^H \mathbf{L}_k^2 \mathbf{Z}_k\right) \\ & + \sigma_{ul}^2 \text{tr}\left(\mathbf{Z}_k^H \mathbf{D}_k \mathbf{Z}_k\right)\end{aligned} \tag{29}$$

TABLE 2. Simulation parameters.

parameter	value
Number of RRHs, N	100
Number of users, K	50
Bandwidth	20MHz
Maximum uplink transmit power, p_i	100mW
Number of samples per coherence block, τ_c	200
Receiver noise power, σ^2_{ul}	-94 dBm
Decorrelation distance d_{dc}	0.1Km
Standard deviation of log-normal shadowing σ_{sf}	8

VI. RESULTS AND DISCUSSION

In this section, we evaluate the performance of the proposed system with a two-layer decoding scheme in uplink over Rician fading channels with phase information. We assumed that phase information is known at each RRH, and they are synchronized. The phase noise due to impairments is neglected in the analysis. In [19] and [35] the authors evaluated system performance by considering the Rician fading channel without taking phase shifts into account. In this work, we considered the network with N single-antenna RRHs and K single-antenna users, which are distributed over a square kilometer. The simulation parameters used in this work are presented in Table. 2.

We used the COST-231 Walfisch-Ikegami model for LoS urban microcell to model the path loss (P_{Loss}) and shadowed fading. The main reason for selecting this model is it will suit well for taking the urban environment effects into account and applications like smart city and industrial IoT. The Rician $\kappa-$factor is obtained by using the expression $\kappa = 13 - 0.03d$ in dB, where d(in meters) is the distance between the RRH and the user. The correlated shadow fading (sf) coefficient is obtained by using the model sf $= \sqrt{\varepsilon}a_l + \sqrt{1-\varepsilon}b_k$ [41]. Here, ε indicates the shadow fading parameter satisfying $0 \leq \varepsilon \leq 1$, and $a_l, b_k \sim \mathcal{N}(0,1)$ are the independent random variables which will model the shadow fading effect at the RRH and user respectively. For any arbitrary RRH and user, the co-variance function of a_l and b_k is given by $\mathbb{E}\{a_l a_n\} = 2^{\frac{-d_{ln}}{d_{dc}}}$ and $\mathbb{E}\{b_i b_k\} = 2^{\frac{-d_{ik}}{d_{dc}}}$, where d_{ln} represents the distance between $l-$th RRH and $n-$th RRH and d_{dc} denotes the decorrelation distance. The fading coefficients h_{lk} for LoS and N-LoS paths can be obtained as $\bar{g}_{lk} = \sqrt{P^{lk}_{Loss}}\left(\sqrt{\frac{\kappa_{lk}}{1+\kappa_{lk}}}\right)$ and $\beta_{lk} = \sqrt{P^{lk}_{Loss}}\left(\sqrt{\frac{1}{1+\kappa_{lk}}}\right)$. In this work, $\tau_p = 5$ samples are reserved for estimating CSI, and we use $\tau_u = 195$ samples and $\tau_d = 195$ samples while analyzing the uplink and downlink performance respectively.

A. UPLINK PERFORMANCE

In uplink transmission, every user in the network transmits with maximum power. In order to reduce the effect of pilot contamination, we proposed a two-layer decoding scheme, and its performance is compared with the traditional single-layer decoding scheme. The proposed user-centric HC-RAN utilizes the distributed cooperation approach where some of the BBU pool functionalities are shifted to the RRHs, like decoding and encoding of data in uplink and downlink operation [25]. Based on the received signals, each RRH will perform MR combining in the first layer to decode the data corresponding to the desired $k-$th user. After that, the combining signals from all the RRHs are sent to the BBU pool for the final decoding of the data corresponding to the $k-$th user. In the second layer, the BBU pool computes the large-scale fading coefficients corresponding to each RRH

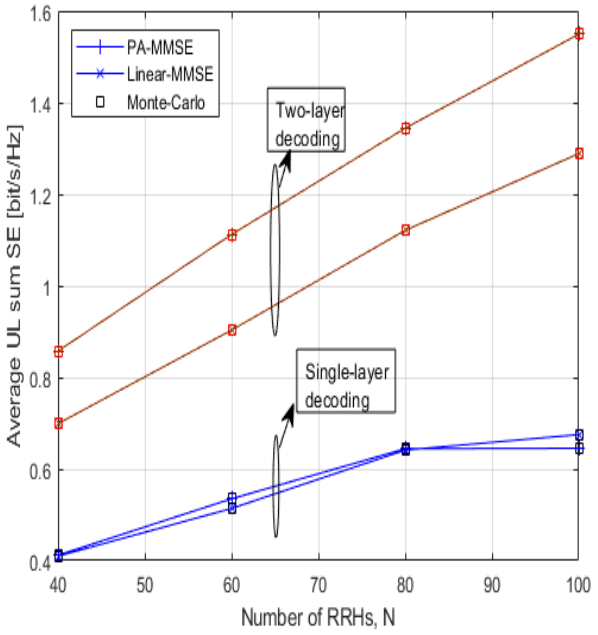

FIGURE 3. Average uplink sum SE with different values of *N* for correlated Rician fading.

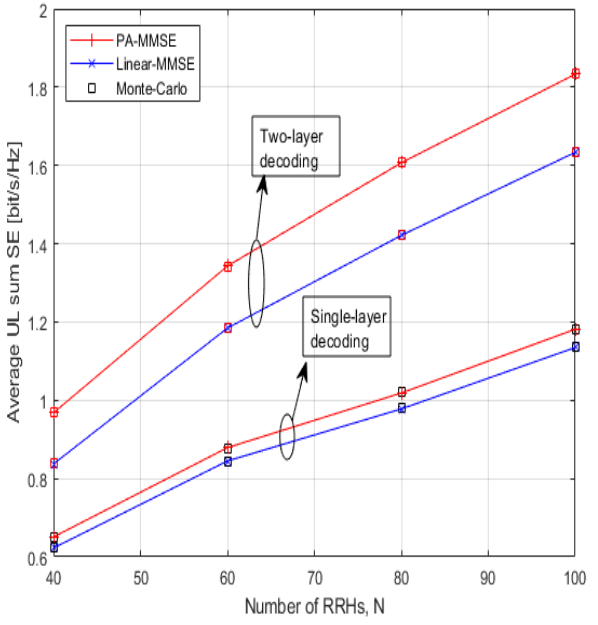

FIGURE 4. Average uplink sum SE with different values of *N* for uncorrelated Rician fading.

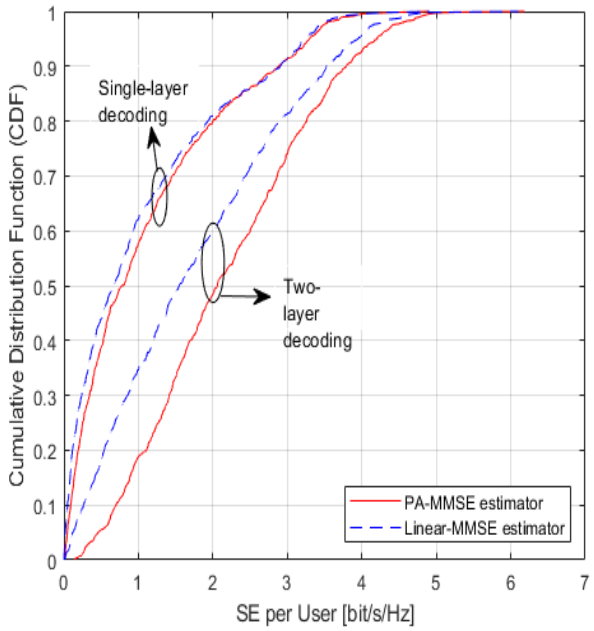

FIGURE 5. Achievable SE per user in the uplink.

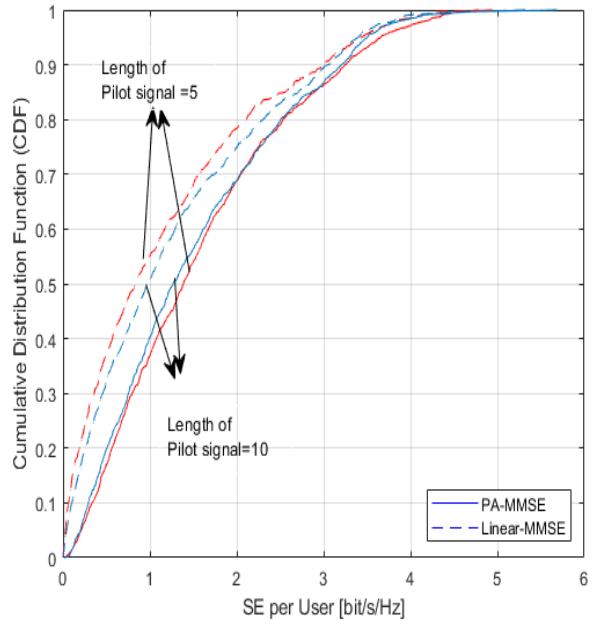

FIGURE 6. Achievable SE per user with different τ_p.

and assigns the weights to maximize the SINR of the desired user. In single-layer decoding, all the weights corresponding to each serving RRH are set to one. Here, Fig. 3 and Fig. 4 show the variation of the average achievable SE per user as a function of the number of RRHs for both the estimators over the Rician fading channel for correlated and uncorrelated scenarios. The average is taken over random user locations and the shadow fading realizations. The analytical results are validated by the simulation results.

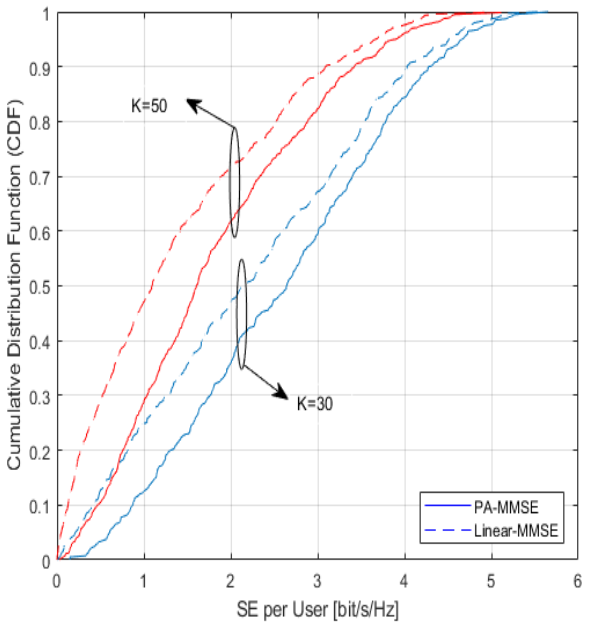

FIGURE 7. Achievable SE per user with different values of K.

Fig. 5 represents the achievable SE per user for both estimators with single and two-layer decoding schemes. From Fig. 3 and Fig. 4, it is observed that an accompanying two-layer decoding scheme in uplink can reduce inter-user interference and improve the system performance compared to single-layer decoding. Since the BBU pool assigns the weights based on the received signals from all RRHs in the second layer. These weights are computed based on the large-scale fading decoding coefficient of each RRH and user pair. Fig. 6 and Fig. 7 show the variation of the achievable SE per user with different lengths of pilot sequence and traffic loads. Fig. 6 shows that the system performance loss between two estimators depends on the phase information and degree of pilot contamination. The degree of pilot contamination decreases as the length of the pilot sequence increases. In addition, the performance of the PA-MMSE estimator with increased pilot length is reduced a bit since the reduction in the pre-log factor is more than the improved SINR. From Fig. 7, it is observed that the system performance decreases with the increase in traffic load. Increasing the pilot sequence length will compensate for this when prior phase information is unavailable. Hence, there should be a good trade-off between choosing the appropriate length of the pilot sequence and having phase knowledge of the LoS components for improving the system's performance.

VII. CONCLUSION

In this work, we analyzed the performance of the scalable user-centric HC-RAN over a Rician fading channel with phase information. The results show that the proposed system is more suitable for ultra-dense network scenarios and applications like IIoT and smart homes where an LoS path

exists between the RRH and users. The phase of the LoS component is modeled as a uniformly distributed random variable to consider the phase variations due to mobility and phase noise. We derived the closed-form expressions for the achievable SE per user using the PA-MMSE estimator and phase unaware Linear-MMSE estimator to know the importance of having phase information. We evaluated the performance of the proposed system for both estimators. Based on the obtained results, it is observed that the effect of inter-user interference can be reduced effectively by using a two-layer decoding scheme. The performance of the Linear-MMSE estimator deviates from the PA-MMSE estimator due to the lack of phase knowledge. This loss will be compensated by increasing the pilot sequence length since it reduces pilot contamination. Hence, the performance of the Linear-MMSE estimator depends on the length of the pilot sequence. Therefore, we must select an appropriate pilot length or have phase knowledge to compensate for the performance loss in high mobility regions and hardware impairment scenarios.

APPENDIX A
LINEAR-MMSE DERIVATION

The channel estimate $\hat{h}_{lk}$ is obtained based on the received signal (4), in which h_{lk} is the desired parameter. From [38], the expression for the Linear-MMSE estimator is given as

$$\hat{h}_{lk} = \frac{\mathbb{E}\left\{h_{lk}\left(y_p^{lk}\right)^H\right\}}{\mathbb{E}\left\{\left|y_p^{lk}\right|^2\right\}}\left(y_p^{lk}\right). \tag{32}$$

By using equations (1) and (4), the expectations can be obtained as $\mathbb{E}\left\{h_{lk}\left(y_p^{lk}\right)^H\right\} = \sqrt{p_k}\tau\left(\beta_{lk} + \bar{g}_{lk}^2\right)$, and $\mathbb{E}\left\{\left|y_p^{lk}\right|^2\right\} = \sum_{(l,i)\in\mathcal{S}_k} p_i h_{lk}\tau^2\left(\beta_{li} + \bar{g}_{li}^2\right) + \tau\sigma_{ul}^2$.

The estimation error can be expressed as $\tilde{h}_{lk}$ has zero mean and a variance c'_{lk}. By using standard estimation theory, c'_{lk} can be expressed as

$$c'_{lk} = \mathbb{E}\left\{|h_{lk}|^2\right\} - \frac{\left|\mathbb{E}\left\{h_{lk}\left(y_p^{lk}\right)^H\right\}\right|^2}{\mathbb{E}\left\{\left|y_p^{lk}\right|^2\right\}}. \tag{33}$$

By substituting the expectation values, c'_{lk} can be expressed as

$$c'_{lk} = \beta'_{lk} - p_k\beta'_{lk}\left(\lambda'_{lk}\right)^{-1}\beta'_{lk}. \tag{34}$$

where $\beta'_{lk} = \beta_{lk} + \bar{g}_{lk}^2$.

APPENDIX B
PHASE UNAWARE LINEAR-MMSE DERIVATION

The expectations of the SINR for Linear-MMSE using MR combining can be obtained as follows. In this analysis, we considered the single antenna RRHs. The term $\mathbb{E}\left\{\mathbf{v}_k^H\mathbf{h}_k\right\}$ can be calculated as [18], [37], and [38]

$$\begin{aligned}\mathbb{E}\left\{\mathbf{v}_k^H\mathbf{h}_k\right\} &= \sum_{l=1}^{N}\mathbb{E}\left\{v_{lk}^H h_{lk}\right\}\\ \mathbb{E}\left\{v_{lk}^H h_{lk}\right\} &= \mathbb{E}\left\{\hat{h}_{lk}^H\left(\hat{h}_{lk} + \tilde{h}_{lk}\right)\right\}\\ &= \mathbb{E}\left\{\hat{h}_{lk}^H\hat{h}_{lk} + \hat{h}_{lk}^H\tilde{h}_{lk}\right\}\\ &= \sqrt{p_k}\tau\beta'_{lk}\left(\lambda'_{lk}\right)^{-1}\beta'_{lk},\end{aligned} \tag{35}$$

where estimate and the estimation error both are independent random variable with zero mean. For all the RRHs, $\mathbb{E}\left\{\mathbf{v}_k^H\mathbf{h}_k\right\}$ can be expressed as

$$\mathbb{E}\left\{\mathbf{v}_k^H\mathbf{h}_k\right\} = \sqrt{p_k}\tau\,\mathrm{tr}\left(\Omega_k\right), \tag{36}$$

where $\Omega_k = \beta'_{lk}\left(\lambda'_{lk}\right)^{-1}\beta'_{lk}$. The first term can be computed as

$$\begin{aligned}\mathbb{E}\left\{\mathbf{v}_k^H\mathbf{Z}_k\mathbf{h}_k\right\} &= \mathrm{tr}\left(\mathbf{Z}_k\mathbb{E}\left\{\mathbf{v}_k^H\mathbf{h}_k\right\}\right)\\ &= \sqrt{p_k}\tau\,\mathrm{tr}\left(\mathbf{Z}_k\Omega_k\right).\end{aligned} \tag{37}$$

The second term can be computed as

$$\begin{aligned}\mathbb{E}\left\{\|\mathbf{Z}_k\mathbf{v}_k\|^2\right\} &= \mathrm{tr}\left(\mathbb{E}\left\{\|\mathbf{Z}_k\mathbf{v}_k\|^2\right\}\right)\\ &= \sqrt{p_k}\tau\,\mathrm{tr}\left(\mathbf{Z}_k\Omega_k\mathbf{Z}_k^H\right).\end{aligned} \tag{38}$$

The final term can be computed as

$$\mathbb{E}\left\{\left|\mathbf{Z}_k\mathbf{v}_k^H\mathbf{h}_i\right|^2\right\} = \mathrm{tr}\left(|\mathbf{Z}_k|^2\,\mathbb{E}\left\{\left|\mathbf{v}_k^H\mathbf{h}_i\right|^2\right\}\right)$$

By considering the effect of all possible RRH and user combinations, it can be calculated as follows

$$\mathbb{E}\left\{\left|\mathbf{v}_k^H\mathbf{h}_i\right|^2\right\} = \sum_{l=1}^{N}\sum_{n=1}^{N}\mathbb{E}\left\{\left(v_{lk}^H h_{li}\right)^H\left(v_{nk}^H h_{ni}\right)\right\}$$

For $l \neq n$ and $i \notin S_k$,

$$\begin{aligned}\mathbb{E}\left\{\left(v_{lk}^H h_{li}\right)^H\left(v_{nk}^H h_{ni}\right)\right\} &= \mathbb{E}\left\{\left(v_{lk}^H h_{li}\right)^H\right\}\mathbb{E}\left\{\left(v_{nk}^H h_{ni}\right)\right\}\\ &= 0\end{aligned}$$

since the channels are independent with zero mean values.

For $l = n$ and $i \notin S_k$,

$$\begin{aligned}\mathbb{E}\left\{\left(v_{lk}^H h_{li}\right)^H\left(v_{nk}^H h_{ni}\right)\right\} &= \mathbb{E}\left\{\left|v_{lk}^H h_{li}\right|^2\right\}\\ &= \mathbb{E}\left\{\left|\hat{h}_{lk}^H\left(\hat{h}_{li} + \tilde{h}_{li}\right)\right|^2\right\}\end{aligned}$$

using channel independence, it can be calculated as $\mathbb{E}\left\{\left|\hat{h}_{lk}^H\hat{h}_{li}\right|^2\right\} = p_k p_i\tau^2\left(\beta'_{lk}\right)^2\left(\lambda'_{lk}\right)^{-1}\left(\beta'_{li}\right)^2\left(\lambda'_{li}\right)^{-1}$, and $\mathbb{E}\left\{\left|\hat{h}_{lk}^H\tilde{h}_{li}\right|^2\right\} = p_k\tau\left(\beta'_{lk}\right)^2\left(\lambda'_{lk}\right)^{-1}c'_{li}$. Finally, $\mathbb{E}\left\{\left(v_{lk}^H h_{li}\right)^H\left(v_{nk}^H h_{ni}\right)\right\} = p_k\tau\beta'_{li}\left(\beta'_{lk}\right)^2\left(\lambda'_{lk}\right)^{-1}$.

For $l \neq n$ and $i \in S_k$,

$$\mathbb{E}\left\{\left(v_{lk}^H h_{li}\right)^H \left(v_{nk}^H h_{ni}\right)\right\} = \mathbb{E}\left\{\left(v_{lk}^H h_{li}\right)^H\right\} \mathbb{E}\left\{\left(v_{nk}^H h_{ni}\right)\right\},$$

where

$$\begin{aligned}\mathbb{E}\left\{\left(v_{lk}^H h_{li}\right)^H\right\} &= \mathbb{E}\left\{\left(\hat{h}_{lk}^H\left(\hat{h}_{li}+\tilde{h}_{li}\right)\right)^H\right\} \\ &= \mathbb{E}\left\{\hat{h}_{lk}^H \hat{h}_{li} + \hat{h}_{lk}\tilde{h}_{li}^H\right\} \\ &= \mathbb{E}\left\{\left(\sqrt{p_k}\beta'_{lk}\left(\lambda'_{lk}\right)^{-1} y_p^{lk}\right) \right. \\ &\quad \left. \times \left(\sqrt{p_i}\beta'_{li}\left(\lambda'_{li}\right)^{-1} y_p^{li}\right)^H\right\} \\ &= \sqrt{p_k p_i}\beta'_{lk}\left(\lambda'_{lk}\right)^{-2}\beta'_{li}\mathbb{E}\left\{\left(y_p^{lk}\right)\left(y_p^{li}\right)^H\right\} \\ &= \sqrt{p_k p_i}\tau\left(\beta'_{lk}\right)\left(\lambda'_{lk}\right)^{-1}\left(\beta'_{li}\right)\end{aligned}$$

and $\mathbb{E}\left\{\left(v_{nk}^H h_{ni}\right)\right\} = \sqrt{p_k p_i}\tau\left(\beta'_{nk}\right)\left(\lambda'_{nk}\right)^{-1}\left(\beta'_{ni}\right)$.

Finally,

$$\begin{aligned}\mathbb{E}\left\{\left(v_{lk}^H h_{li}\right)^H \left(v_{nk}^H h_{ni}\right)\right\} &= p_k p_i \tau^2\left(\left(\beta'_{lk}\right)\left(\lambda'_{lk}\right)^{-1}\left(\beta'_{li}\right)\right. \\ &\quad \left. \times \left(\beta'_{nk}\right)\left(\lambda'_{nk}\right)^{-1}\left(\beta'_{ni}\right)\right).\end{aligned}$$

For $l = n$ and $i \in S_k$, the term $\mathbb{E}\left\{\left(v_{lk}^H h_{li}\right)^H \left(v_{nk}^H h_{ni}\right)\right\}$ can be calculated as

$$\begin{aligned}\mathbb{E}\left\{\left(v_{lk}^H h_{li}\right)^H \left(v_{nk}^H h_{ni}\right)\right\} &= \mathbb{E}\left\{\left|v_{lk}^H h_{li}\right|^2\right\} \\ &= \mathbb{E}\left\{\left|\hat{h}_{lk}^H h_{li}\right|^2\right\}.\end{aligned}$$

this can be expressed as

$$\begin{aligned}\mathbb{E}\left\{\left|\hat{h}_{lk}^H h_{li}\right|^2\right\} &= \mathbb{E}\left\{\left|\left(\sqrt{p_k}\beta'_{lk}\left(\lambda'_{lk}\right)^{-1} y_p^{lk}\right) h_{li}\right|^2\right\} \\ &= p_k\left(\beta'_{lk}\right)^2\left(\lambda'_{lk}\right)^{-2}\mathbb{E}\left\{\left|y_p^{lk} h_{li}\right|^2\right\} \\ &= p_k p_i \tau^2\left(\beta'_{lk}\right)^2\left(\lambda'_{lk}\right)^{-2} \\ &\quad \times\left(\left(\beta_{li}\right)^2 + 2\beta_{li}\bar{h}_{li}^2\right) + \tau\beta'_{li}\lambda'_{li}\Big),\end{aligned}$$

where

$$\begin{aligned}\mathbb{E}\left\{\left|y_p^{lk} h_{li}\right|^2\right\} &= p_i\tau^2\mathbb{E}\left\{|h_{li}|^4\right\} + \mathbb{E}\left\{\left|\left(\mathbf{n}_p^l \Phi_{\tau_{lk}}^H\right)^H h_{li}\right|^2\right\} \\ &\quad + \mathbb{E}\left\{\left|\sum_{(l,q)\in\mathcal{S}_k\setminus\{i\}}\sqrt{p_q}\tau h_{lq} h_{li}\right|^2\right\} \\ &= p_i\tau^2\left(\left(\beta_{li}\right)^2 + 2\beta_{li}\bar{h}_{li}^2\right) + \tau\beta'_{li}\lambda'_{li}.\end{aligned}$$

and $\mathbb{E}\left\{|h_{li}|^4\right\} = \mathbb{E}\left\{|\bar{g}_{lk}\exp\left(j\theta_{lk}\right) + g_{lk}|^4\right\} = 2\beta_{li}^2 + 4\beta_{li}\bar{h}_{li}^2 + \bar{h}_{li}^4$. By combining all the equations, $\mathbb{E}\left\{\left|\mathbf{Z}_k \mathbf{v}_k^H \mathbf{h}_i\right|^2\right\}$ can be expressed as

$$\begin{aligned}&\mathbb{E}\left\{\left|\mathbf{Z}_k \mathbf{v}_k^H \mathbf{h}_i\right|^2\right\} \\ &= p_k\tau\,\mathrm{tr}\left(\mathbf{Z}_k^H \mathbf{R}_i \mathbf{Z}_k \Omega_k\right) \\ &\quad + p_k p_i \tau^2 \begin{cases} \mathrm{tr}\left(\mathbf{Z}_k^H \mathbf{R}_i^2 \Lambda_k \Omega_k \mathbf{Z}_k\right) + \\ 2\mathrm{tr}\left(\mathbf{Z}_k^H \Omega_k \Lambda_k \mathbf{L}_i \mathbf{R}_i \mathbf{Z}_k\right) + \\ \left|\mathrm{tr}\left(\mathbf{Z}_k \mathbf{R}_i \Lambda_k \mathbf{R}_k\right)\right| - & (39) \\ \mathrm{tr}\left(\mathbf{Z}_k^H \left(\mathbf{R}_i \Lambda_k \mathbf{R}_k\right)^2 \mathbf{Z}_k\right) & i \in S_k \\ 0 & i \notin S_k. \end{cases}\end{aligned}$$

APPENDIX C
PA-MMSE DERIVATION

The expectations of SINR expression for PA-MMSE channel estimator with known phase information can be obtained as follows. The term $\mathbb{E}\left\{\mathbf{v}_k^H \mathbf{h}_k\right\}$ can be calculated as [18], [37], and [38]

$$\mathbb{E}\left\{\mathbf{v}_k^H \mathbf{h}_k\right\} = \sum_{l=1}^{N}\mathbb{E}\left\{\mathbf{v}_{lk}^H \mathbf{h}_{lk}\right\}.$$

For single antenna RRHs, these will become scalars and it can be expressed as

$$\begin{aligned}\mathbb{E}\left\{\mathbf{v}_{lk}^H \mathbf{h}_{lk}\right\} &= \sum_{l=1}^{N}\mathbb{E}\left\{v_{lk}^H h_{lk}\right\} \\ &= \mathbb{E}\left\{\left(\hat{h}_{lk}^{\mathrm{PA-MMSE}}\right)^H\left(\hat{h}_{lk}^{\mathrm{PA-MMSE}}\right)\right\} \\ &= p_k\tau\beta_{lk}^2\lambda_{lk}^{-1} + \bar{h}_{lk}^2. \qquad (40)\end{aligned}$$

The first term can be calculated as

$$\begin{aligned}\mathbb{E}\left\{\mathbf{Z}_k \mathbf{v}_k^H \mathbf{h}_k\right\} &= \mathrm{tr}\left(\mathbf{Z}_k \mathbb{E}\left\{\mathbf{v}_k^H \mathbf{h}_k\right\}\right) \\ &= \mathrm{tr}\left(p_k\tau\mathbf{A}_k\Omega_k + \mathbf{A}_k L_k\right). \qquad (41)\end{aligned}$$

The second term can be obtained as

$$\begin{aligned}\mathbb{E}\left\{\|\mathbf{Z}_k \mathbf{v}_k\|^2\right\} &= \mathrm{tr}\left(|\mathbf{Z}_k|^2\,\mathbb{E}\left\{\|\mathbf{v}_k\|^2\right\}\right) \\ &= \mathrm{tr}\left(p_k\tau\mathbf{A}_k\Omega_k\mathbf{A}_k^H + \mathbf{A}_k L_k \mathbf{A}_k^H\right). \qquad (42)\end{aligned}$$

The final term

$$\mathbb{E}\left\{\left|\mathbf{Z}_k \mathbf{v}_k^H \mathbf{h}_i\right|^2\right\} = \mathrm{tr}\left(|\mathbf{Z}_k|^2\,\mathbb{E}\left\{\left|\mathbf{v}_k^H \mathbf{h}_i\right|^2\right\}\right).$$

can be computed by considering all the possible RRH and user combinations. It is as follows

$$\mathbb{E}\left\{\left|\mathbf{v}_k^H \mathbf{h}_i\right|^2\right\} = \sum_{l=1}^{N}\sum_{n=1}^{N} a_{lk} a_{lk}^H \mathbb{E}\left\{\left(v_{lk}^H h_{li}\right)^H\left(v_{nk}^H h_{ni}\right)\right\}.$$

For $l \neq n$ and $i \notin S_k$,

$$\begin{aligned}\mathbb{E}\left\{\left(v_{lk}^H h_{li}\right)^H\left(v_{nk}^H h_{ni}\right)\right\} &= \mathbb{E}\left\{\left(v_{lk}^H h_{li}\right)^H\right\}\mathbb{E}\left\{\left(v_{nk}^H h_{ni}\right)\right\} \\ &= 0.\end{aligned}$$

since channels are independent with zero means.

For $l = n$ and $i \notin S_k$,

$$\begin{aligned}\mathbb{E}\left\{\left(v_{lk}^H h_{li}\right)^H\left(v_{nk}^H h_{ni}\right)\right\} &= \mathbb{E}\left\{\left|v_{lk}^H h_{li}\right|^2\right\} \\ &= \left(p_k\tau\beta_{lk}^2\lambda_{lk}^{-1} + \bar{h}_{lk}^2\right)\left(\beta_{li} + \bar{h}_{li}^2\right).\end{aligned}$$

For $l \neq n$ and $i \in S_k$,

$$\mathbb{E}\left\{\left(v_{lk}^H h_{li}\right)^H \left(v_{nk}^H h_{ni}\right)\right\} = \mathbb{E}\left\{\left(v_{lk}^H h_{li}\right)^H\right\} \mathbb{E}\left\{\left(v_{nk}^H h_{ni}\right)\right\} = \mathbb{E}\left\{\left(\left(\hat{h}_{lk}^{\text{PA}-\text{MMSE}}\right)^H \hat{h}_{li}^{\text{PA}-\text{MMSE}}\right)^H\right\} \times \mathbb{E}\left\{\left(\left(\hat{h}_{nk}^{\text{PA}-\text{MMSE}}\right)^H \hat{h}_{ni}^{\text{PA}-\text{MMSE}}\right)\right\},$$

where

$$\mathbb{E}\left\{\left(\left(\hat{h}_{lk}^{\text{PA}-\text{MMSE}}\right)^H \hat{h}_{li}^{\text{PA}-\text{MMSE}}\right)^H\right\} = \mathbb{E}\left\{\left(\bar{g}_{lk} e^{j\theta_{lk}} + \sqrt{p_k}\beta_{lk}\lambda_{lk}^{-1}\left(y_p^{lk} - \bar{y}_p^{lk}\right)\right)^H \times \left(\bar{g}_{li} e^{j\theta_{li}} + \sqrt{p_i}\beta_{li}\lambda_{lk}^{-1}\left(y_p^{lk} - \bar{y}_p^{lk}\right)\right)\right\} = \sqrt{p_i p_k}\tau\beta_{lk}\beta_{li}\lambda_{lk}^{-1},$$

similarly

$$\mathbb{E}\left\{\left(\left(\hat{h}_{nk}^{\text{PA}-\text{MMSE}}\right)^H \hat{h}_{ni}^{\text{PA}-\text{MMSE}}\right)^H\right\} = \sqrt{p_i p_k}\tau\beta_{nk}\beta_{ni}\lambda_{nk}^{-1}$$

finally,

$$\mathbb{E}\left\{\left(v_{lk}^H h_{li}\right)^H \left(v_{nk}^H h_{ni}\right)\right\} = p_i p_k \tau^2 \beta_{lk}\beta_{li}\lambda_{lk}^{-1}\beta_{nk}\beta_{ni}\lambda_{nk}^{-1}.$$

For $l \neq n$, $i = k$, and $i \in S_k$,

$$\mathbb{E}\left\{\left(v_{lk}^H h_{li}\right)^H \left(v_{nk}^H h_{ni}\right)\right\} = \mathbb{E}\left\{\left(v_{lk}^H h_{lk}\right)^H\right\} \mathbb{E}\left\{\left(v_{nk}^H h_{nk}\right)\right\} = \left(p_k\tau\beta_{lk}^2\lambda_{lk}^{-1} + \bar{h}_{lk}^2\right) \times \left(p_k\tau\beta_{nk}^2\lambda_{nk}^{-1} + \bar{h}_{nk}^2\right)$$

For $l = n$, $i = k$, and $i \in S_k$,

$$\mathbb{E}\left\{\left(v_{lk}^H h_{li}\right)^H \left(v_{nk}^H h_{ni}\right)\right\} = \mathbb{E}\left\{\left|v_{lk}^H h_{lk}\right|^2\right\} = \mathbb{E}\left\{\left|\hat{h}_{lk}^{\text{PA}-\text{MMSE}}\right|^4\right\} + \mathbb{E}\left\{\left|\hat{h}_{lk}^{\text{PA}-\text{MMSE}}\tilde{h}_{lk}^{\text{PA}-\text{MMSE}}\right|^2\right\},$$

where

$$\mathbb{E}\left\{\left|\hat{h}_{lk}^{\text{PA}-\text{MMSE}}\tilde{h}_{lk}^{\text{PA}-\text{MMSE}}\right|^2\right\} = \left(p_k\tau\beta_{lk}^2\lambda_{lk}^{-1} + \bar{h}_{lk}^2\right) c_{lk}.$$

$$\mathbb{E}\left\{\left|\hat{h}_{lk}^{\text{PA}-\text{MMSE}}\right|^4\right\} = \mathbb{E}\left\{\left|\left(\bar{g}_{lk} e^{-j\theta_{lk}} + \sqrt{p_k}\tau\beta_{lk}\lambda_{lk}^{-1/2} w^H\right) \times \left(\bar{g}_{lk} e^{j\theta_{lk}} + \sqrt{p_k}\tau\beta_{lk}\lambda_{lk}^{-1/2} w\right)\right|^2\right\} = 2p_k^2\tau^2\beta_{lk}^4\lambda_{lk}^{-2} + 4p_k\tau\beta_{lk}^2\lambda_{lk}^{-1}\bar{h}_{lk}^2 + \bar{h}_{lk}^4$$

and let us assume $w \sim \mathcal{N}_C(0, 1)$.

For $l = n$, and $i \in S_k$,

$$\mathbb{E}\left\{\left(v_{lk}^H h_{li}\right)^H \left(v_{nk}^H h_{ni}\right)\right\} = \mathbb{E}\left\{\left|v_{lk}^H h_{li}\right|^2\right\} = \mathbb{E}\left\{\left|\hat{h}_{lk}^{\text{PA}-\text{MMSE}}\hat{h}_{li}^{\text{PA}-\text{MMSE}}\right|^2\right\} + \mathbb{E}\left\{\left|\hat{h}_{lk}^{\text{PA}-\text{MMSE}}\tilde{h}_{li}^{\text{PA}-\text{MMSE}}\right|^2\right\},$$

where

$$\mathbb{E}\left\{\left|\hat{h}_{lk}^{\text{PA}-\text{MMSE}}\tilde{h}_{li}^{\text{PA}-\text{MMSE}}\right|^2\right\} = \mathbb{E}\left\{\left|\hat{h}_{lk}^{\text{PA}-\text{MMSE}}\right|^2\right\} \times \mathbb{E}\left\{\left|\tilde{h}_{li}^{\text{PA}-\text{MMSE}}\right|^2\right\} = \left(p_k\tau\beta_{lk}^2\lambda_{lk}^{-1} + \bar{h}_{lk}^2\right) c_{li}$$

$$\mathbb{E}\left\{\left|\hat{h}_{lk}^{\text{PA}-\text{MMSE}}\hat{h}_{li}^{\text{PA}-\text{MMSE}}\right|^2\right\} = \mathbb{E}\left\{\left|\left(\bar{g}_{lk} e^{-j\theta_{lk}} + \sqrt{p_k}\tau\beta_{lk}\lambda_{lk}^{-1/2} w^H\right) \times \left(\bar{g}_{li} e^{j\theta_{li}} + \sqrt{p_i}\tau\beta_{li}\lambda_{lk}^{-1/2} w\right)\right|^2\right\} = 2p_k p_i\tau^2\beta_{lk}^2\beta_{li}^2\lambda_{lk}^{-2} + p_k\tau\beta_{lk}^2\lambda_{lk}^{-1}\bar{h}_{li}^2 + p_i\tau\beta_{li}^2\lambda_{lk}^{-1}\bar{h}_{lk}^2 + \bar{h}_{li}^2\bar{h}_{lk}^2.$$

Finally, combining all the expressions, it results in

$$\begin{aligned}\mathbb{E}\left\{\left|\mathbf{Z}_k \mathbf{v}_k^H \mathbf{h}_i\right|^2\right\} &= p_k\tau\,\text{tr}\left(\mathbf{Z}_k^H \mathbf{R}_l \mathbf{Z}_k \Omega_k\right) + p_k\tau\,\text{tr}\left(\mathbf{Z}_k^H \Omega_k \mathbf{Z}_k \mathbf{L}_l\right) + \text{tr}\left(\mathbf{Z}_k^H \mathbf{R}_l \mathbf{Z}_k \mathbf{L}_l\right) + \text{tr}\left(\mathbf{Z}_k^H \mathbf{L}_k \mathbf{L}_l \mathbf{Z}_k\right) \\ &+ \begin{cases} p_k p_i \tau^2 \left|\text{tr}\left(\mathbf{Z}_k^H \mathbf{R}_l \Lambda_k \mathbf{R}_k\right)\right|^2 & i \in S_k \\ p_k\tau^2 \left|\text{tr}\left(\mathbf{Z}_k \Omega_k\right)\right|^2 + \text{tr}\left(\mathbf{Z}_k \mathbf{L}_k\right)^2 & \\ +2p_k\tau\,\text{tr}\left(\mathbf{Z}_k \Omega_k\right)\text{tr}\left(\mathbf{Z}_k \mathbf{L}_k\right) & i = k \\ 0 & \text{otherwise.} \end{cases}\end{aligned} \tag{43}$$

REFERENCES

[1] G. Forecast, "Cisco visual networking index: Global mobile data traffic forecast update, 2017–2022," *Update*, vol. 2017, p. 2022, Feb. 2019.
[2] O. Simeone, A. Maeder, M. Peng, O. Sahin, and W. Yu, "Cloud radio access network: Virtualizing wireless access for dense heterogeneous systems," *J. Commun. Netw.*, vol. 18, no. 2, pp. 135–149, Apr. 2016.
[3] Q. Shen, Z. Ma, and S. Wang, "Deploying C-RAN in cellular radio networks: An efficient way to meet future traffic demands," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7887–7891, Aug. 2018.
[4] 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.
[5] 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.
[6] G. Nigam, P. Minero, and M. Haenggi, "Coordinated multipoint joint transmission in heterogeneous networks," *IEEE Trans. Commun.*, vol. 62, no. 11, pp. 4134–4146, Oct. 2014.

[7] S. Bassoy, H. Farooq, M. A. Imran, and A. Imran, "Coordinated multi-point clustering schemes: A survey," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 743–764, 2nd Quart., 2017.

[8] A. Davydov, G. Morozov, I. Bolotin, and A. Papathanassiou, "Evaluation of joint transmission CoMP in C-RAN based LTE—A HetNets with large coordination areas," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2013, pp. 801–806.

[9] J. Zhang, R. Chen, J. G. Andrews, A. Ghosh, and R. W. Heath Jr., "Networked MIMO with clustered linear precoding," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1910–1921, Apr. 2009.

[10] C. Pan, M. Elkashlan, J. Wang, J. Yuan, and L. Hanzo, "User-centric C-RAN architecture for ultra-dense 5G networks: Challenges and methodologies," *IEEE Commun. Mag.*, vol. 56, no. 6, pp. 14–20, Jun. 2018.

[11] U. S. Hashmi, S. A. R. Zaidi, and A. Imran, "User-centric cloud RAN: An analytical framework for optimizing area spectral and energy efficiency," *IEEE Access*, vol. 6, pp. 19859–19875, 2018.

[12] K. Zarifi, H. Baligh, J. Ma, M. Salem, and A. Maaref, "Radio access virtualization: Cell follows user," in *Proc. IEEE 25th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Sep. 2014, pp. 1381–1385.

[13] A. Papadogiannis, D. Gesbert, and E. Hardouin, "A dynamic clustering approach in wireless networks with multi-cell cooperative processing," in *Proc. IEEE Int. Conf. Commun.*, Oct. 2008, pp. 4033–4037.

[14] Q. Liu, S. Sun, and H. Gao, "Joint user-centric clustering and frequency allocation in ultra-dense C-RAN," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, May 2020, pp. 1–6.

[15] V. Garcia, Y. Zhou, and J. Shi, "Coordinated multipoint transmission in dense cellular networks with user-centric adaptive clustering," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4297–4308, Aug. 2014.

[16] Z. Zhang, N. Wang, J. Zhang, and X. Mu, "Dynamic user-centric clustering for uplink cooperation in multi-cell wireless networks," *IEEE Access*, vol. 6, pp. 8526–8538, 2018.

[17] S. Bassoy, M. Jaber, M. A. Imran, and P. Xiao, "Load aware self-organising user-centric dynamic comp clustering for 5G networks," *IEEE Access*, vol. 4, pp. 2895–2906, 2016.

[18] Ö. T. Demir and E. Björnson, "Large-scale fading precoding for spatially correlated Rician fading with phase shifts," 2020, *arXiv:2006.14267*.

[19] Z. H. Abbas, A. Ullah, G. Abbas, F. Muhammad, and F. Y. Li, "Outage probability analysis of user-centric SBS-based HCNets under hybrid Rician/Rayleigh fading," *IEEE Commun. Lett.*, vol. 24, no. 2, pp. 297–301, Feb. 2020.

[20] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[21] C. Pan, H. Ren, M. Elkashlan, A. Nallanathan, and L. Hanzo, "Robust beamforming design for ultra-dense user-centric C-RAN in the face of realistic pilot contamination and limited feedback," *IEEE Trans. Wireless Commun.*, vol. 18, no. 2, pp. 780–795, Feb. 2019.

[22] H. Zhang, C. Jiang, and J. Cheng, "Cooperative interference mitigation and handover management for heterogeneous cloud small cell networks," *IEEE Wireless Commun.*, vol. 22, no. 3, pp. 92–99, Jun. 2015.

[23] H. Zhang, Z. Yang, Y. Liu, and X. Zhang, "Power control for 5G user-centric network: Performance analysis and design insight," *IEEE Access*, vol. 4, pp. 7347–7355, 2016.

[24] M. Ali, Q. Rabbani, M. Naeem, S. Qaisar, and F. Qamar, "Joint user association, power allocation, and throughput maximization in 5G H-CRAN networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 10, pp. 9254–9262, Oct. 2017.

[25] H. Ayanampudi and R. Dhuli, "Performance analysis of heterogeneous cloud-radio access networks: A user-centric approach with network scalability," *Comput. Commun.*, vol. 194, pp. 202–212, Oct. 2022.

[26] E. Björnson and E. Jorswieck, *Optimal Resource Allocation in Coordinated Multi-Cell Systems*. Delft, The Netherlands: Now Publishers Inc, 2013.

[27] E. Bjornson, N. Jalden, M. Bengtsson, and B. Ottersten, "Optimality properties, distributed strategies, and measurement-based evaluation of coordinated multicell OFDMA transmission," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 6086–6101, Dec. 2011.

[28] A. Adhikary, A. Ashikhmin, and T. L. Marzetta, "Uplink interference reduction in large-scale antenna systems," *IEEE Trans. Commun.*, vol. 65, no. 5, pp. 2194–2206, May 2017.

[29] A. Ashikhmin and T. Marzetta, "Pilot contamination precoding in multi-cell large scale antenna systems," in *Proc. IEEE Int. Symp. Inf. Theory Proc.*, Jul. 2012, pp. 1137–1141.

[30] T. Van Chien, C. Mollën, and E. Björnson, "Large-scale-fading decoding in cellular massive MIMO systems with spatially correlated channels," *IEEE Trans. Commun.*, vol. 67, no. 4, pp. 2746–2762, Apr. 2019.

[31] R. Rogalin, O. Y. Bursalioglu, H. Papadopoulos, G. Caire, A. F. Molisch, A. Michaloliakos, V. Balan, and K. Psounis, "Scalable synchronization and reciprocity calibration for distributed multiuser MIMO," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 1815–1831, Mar. 2014.

[32] L. You and D. Yuan, "Joint CoMP-cell selection and resource allocation in fronthaul-constrained C-RAN," in *Proc. 15th Int. Symp. Modeling Optim. Mobile, Ad Hoc, Wireless Netw. (WiOpt)*, May 2017, pp. 1–6.

[33] L. You and D. Yuan, "User-centric performance optimization with remote radio head cooperation in C-RAN," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 340–353, Jan. 2020.

[34] S. Zaidi, O. B. Smida, S. Affes, U. Vilaipornsawai, L. Zhang, and P. Zhu, "User-centric base-station wireless access virtualization for future 5G networks," *IEEE Trans. Commun.*, vol. 67, no. 7, pp. 5190–5202, Jul. 2019.

[35] K. Humadi, I. Trigui, W.-P. Zhu, and W. Ajib, "Dynamic base station clustering in user-centric mmWave networks: Performance analysis and optimization," *IEEE Trans. Commun.*, vol. 69, no. 7, pp. 4847–4861, Jul. 2021.

[36] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[37] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Found. Trends Signal Process.*, vol. 11, nos. 3–4, pp. 154–655, 2017.

[38] T. L. Marzetta, *Fundamentals of Massive MIMO*. Cambridge, U.K.: Cambridge Univ. Press, 2016.

[39] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.

[40] S. Chen, F. Qin, B. Hu, X. Li, and Z. Chen, "User-centric ultra-dense networks for 5G: Challenges, methodologies, and directions," *IEEE Wireless Commun.*, vol. 23, no. 2, pp. 78–85, Apr. 2016.

[41] Z. Wang, E. K. Tameh, and A. R. Nix, "Joint shadowing process in urban peer-to-peer radio channels," *IEEE Trans. Veh. Technol.*, vol. 57, no. 1, pp. 52–64, Jan. 2008.

HAREESH AYANAMPUDI received the B.Tech. degree in electronics and communication engineering from JNTU University, Kakinada, India, in 2011, and the M.Tech. degree in communication engineering and signal processing from Acharya Nagarjuna University, India, in 2014. He is currently pursuing the Ph.D. degree with the School of Electronics Engineering, VIT-AP University, Andhra Pradesh. His research interests include wireless communication and signal processing.

RAVINDRA DHULI received the Ph.D. degree in signal processing from the Department of Electrical Engineering, IIT Delhi, in 2010. He is currently working as a Professor with the School of Electronics Engineering, VIT-AP University, India. His research interests include multi-rate signal processing, statistical signal processing, image processing, and mathematical modeling.