Boosting Fronthaul Capacity: Global Optimization of Power Sharing for Centralized Radio Access Network

Jiankang Zhang [ID], *Senior Member, IEEE*, Sheng Chen [ID], *Fellow, IEEE*, Xinying Guo, Jia Shi [ID], and Lajos Hanzo [ID], *Fellow, IEEE*

Abstract—The limited fronthaul capacity imposes a challenge on the uplink of a centralized radio access network (C-RAN). We propose to boost the fronthaul capacity of a massive multiple-input multiple-output (MIMO)-aided C-RAN by globally optimizing the power sharing between channel estimation and data transmission both for the user devices (UDs) and the remote radio units (RRUs). Intuitively, allocating more power to the channel estimation will result in more accurate channel estimates, which increases the achievable throughput. However, increasing the power allocated to the pilot training will reduce the power assigned to data transmission, which reduces the achievable throughput. In order to optimize the powers allocated to the pilot training and to the data transmission of both the UDs and the RRUs, we assign an individual power sharing factor to each of them and derive an asymptotic closed-form expression of the signal-to-interference-plus-noise ratio for the massive MIMO-aided C-RAN consisting of both the UD-to-RRU links and the RRU-to-baseband unit (BBU) links. We then exploit the C-RAN architecture's central computing and control capability for jointly optimizing the UDs' power sharing factors and the RRUs' power sharing factors aiming at maximizing the fronthaul capacity. Our simulation results show that the fronthaul capacity is significantly boosted by the proposed global optimization of the power allocation between channel estimation and data transmission both for the UDs and for their host RRUs. As a specific example of 32 receive antennas (RAs) deployed by the RRU and 128 RAs deployed by the BBU, the sum rate of 10 UDs achieved with the optimal power sharing factors improves by 33% compared with the one attained without optimizing power sharing factors.

Index Terms—Fronthaul, small cell, massive MIMO, C-RAN, power allocation.

J. Zhang and L. Hanzo are with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: jz09v@ecs.soton.ac.uk; lh@ecs.soton.ac.uk).

S. Chen is with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K., and also with the King Abdulaziz University, Jeddah 21589, Saudi Arabia (e-mail: sqc@ecs.soton.ac.uk).

X. Guo is with the College of Information Science and Engineering, Henan University of Technology, Zhengzhou 450066, China (e-mail: guoxinying@haut.edu.cn).

J. Shi is with the State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an 710071, China (e-mail: jiashi@xidian.edu.cn).

Digital Object Identifier 10.1109/TVT.2018.2890640

I. INTRODUCTION

THE fifth-generation (5G) wireless communication system is expected to meet the ever-increasing mobile traffic, predicted to be 291.8 Exabytes by 2019 [1]. The number of mobile user devices (UDs) of 5G connections is forecast to reach a figure between 25 million and 100 million by 2021 [2]. In order to provide ubiquitous service access for such huge number of UDs, the small cell concept [3] combined with massive multiple-input multiple-output (MIMO) [4] and/or millimeter-wave (mmWave) technologies [5] has emerged as one of the most promising network structures for 5G systems. However, substantial processing power is required for jointly exploiting the small cell's spatial diversity and temporal diversity. This motivates the concept of centralized radio access network (C-RAN) [6], [7], which consists of baseband units (BBUs) and remote radio units (RRUs). In a C-RAN, multiple RRUs are connected to a BBU, which carries out all the baseband signal processing centrally, whilst the RRUs handle the radio frequency processing. Consequently, substantial amount of information is exchanged over fronthaul links between RRUs and their host BBU, which imposes a bottleneck on C-RAN [8] and prevents its large-scale practical deployment.

A. Related Works

Extensive efforts have been devoted to the enhancement of fronthaul capacity, including compression/quantization [9]–[11], quality of service (QoS) guarantee [12], [13], interference mitigation [14], [15], user/access link selection [16]–[18], as well as resource allocation and optimization [8], [19]–[21].

To elaborate, compression/quantization schemes were designed for reducing the data traffic of fronthaul links to meet their capacity constraint. Explicitly, Zhou *et al.* [9] studied a joint fronthaul compression and transmit beamforming scheme relying on noise covariance matrix quantization in the context of traditional MIMO having a small number of antennas. Lee *et al.* [10] investigated multivariate fronthaul quantization motivated by the network-information theory, which is capable of jointly optimizing the downlink precoding and quantization at a reduced-complexity. By contrast, Vu *et al.* [11] derived a block error rate metric in the context of Rayleigh fading channels for designing an adaptive compression scheme.

QoS represents a general terminology that includes packet loss ratio, bit error ratio (BER), throughput, transmission delay, etc. Since it is challenging to investigate all the QoS metrics

jointly, the existing studies mainly concentrate on just a single or a few aspects of QoS. Explicitly, by relying on content caching, Zhao *et al.* [12] improved the link-level effective capacity, while the dynamic network slicing scheme of [13] improved the QoS in terms of rate-fairness, as well as maximum and minimum rates.

Through interference reduction, the achievable fronthaul capacity can also be increased. Hence Liu *et al.* [14] focused the attention on interference mitigation by exploiting the inherent sparsity of C-RAN. Hao *et al.* [15] considered the mitigation of intra-cluster interference and inter-tier interference by exploiting coordinated multipoint transmission and by allocating distinct bandwidths to BBUs and RRUs, respectively.

Fronthaul link selection and user association, which are typically investigated together with precoding [17], [18], maximize the achievable sum-rate of a C-RAN, given a fixed fronthaul capacity constraint. Furthermore, Pan *et al.* [16] studied the joint optimization of RRU selection, user association and beamforming in the presence of imperfect channel state information (CSI). Resource allocation and optimization [8], [19]–[21] also offer an effective means of enhancing the achievable throughput, given a limited fronthaul capacity.

Upgrading the network infrastructure by replacing copper cabling with fiber cabling for fronthaul connections has also been considered as an alternative solution to provide high-capacity fronthaul [22]. However, this approach suffers from poor flexibility and high cost in large-scale deployments [23], [24]. Laying optical fiber to connect the RRUs to their host BBU is impossible in city centres of some countries [25]. Thus wireless backhauling/fronthauling [26]–[30], relying on massive MIMO [4] and mmWave [5], [30], has recently emerged as a promising solution for 5G networks due to its flexibility and cost-efficiency [23], [31], [32]. Park *et al.* [31] proposed a partially centralized C-RAN based on massive MIMO schemes, while Parsaeefard *et al.* [32] allocated resources by appropriately adjusting the parameters of the RRU, BBU and fronthaul as well as the power allocated to UDs.

B. Our Contributions

A UD's achievable throughput inherently depends on both the UD-to-RRU links and the RRU-to-BBU links. Thus all the UD-to-RRU links and RRU-to-BBU links should be considered together in order to maximize a C-RAN's capacity. By exploiting the C-RAN's superior centralized signal processing/resource control capability over both the UD-to-RRU links and the RRU-to-BBU links, we propose to boost the fronthaul capacity by globally optimizing the power sharing for both the RRUs and UDs located within a BBU's service coverage. Intuitively, allocating more power to channel estimation will result in more accurate channel estimates, which increases the achievable throughput. But increasing the power allocated to pilot training will reduce the power allocated to data transmission, which reduces the achievable data throughput. This paper addresses for the first time how to optimize the powers allocated to pilot training and to data transmission both for the UDs and RRUs. The main contributions of this paper are as follows.

1) We investigate the C-RAN's configuration, which employs large numbers of antennas both at the RRUs and their host BBU. We formulate the ultimately achievable uplink sum-rate of the C-RAN as a function of the signal-to-interference-plus-noise ratio (SINR) by considering both the UD-to-RRU links and the RRU-to-BBU links. Furthermore, we derive a closed-form expression of the asymptotic achievable uplink sum-rate in the presence of realistic channel estimation errors both for the UD-to-RRU links and the RRU-to-BBU links.

2) We propose to boost the uplink fronthaul capacity by globally optimizing the power sharing between the pilots and data transmission both for RRUs and for the UDs within a BBU's service coverage. Specifically, given the power of a UD, a UD's power sharing factor controls the specific fractions of power allocated to the UD's uplink pilot and to the UD's uplink data transmission, respectively. Similarly, given the power of a RRU, the power allocated to the RRU's uplink pilot and the power allocated to the RRU's uplink data forwarding are controlled by the RRU's power sharing factor. We formulate the uplink sum-rate as a function of the all the UDs' power sharing factors and all the RRUs' power sharing factors. We then maximize this uplink sum-rate by invoking the global optimization algorithm of the differential evolution algorithm (DEA).

C. Notations

Throughout our discussions, $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$ stand for the transpose and Hermitian transpose of vector/matrix, respectively. $\mathbf{0}_{N \times 1}$ is the $N \times 1$ zero vector, which is abbreviated as $\mathbf{0}_N$, and $\mathbf{0}_{N \times N}$ denotes the $N \times N$ zero matrix, while following the convention, the $N \times N$ identity matrix is represented by $\boldsymbol{I}_N$. The expectation operation is represented by $\mathcal{E}\{\cdot\}$ and $\boldsymbol{vec}(\boldsymbol{A})$ denotes the column stacking operation applied to matrix $\boldsymbol{A}$. The diagonal matrix with $a_1, \ldots, a_M$ at its diagonal entries is denoted by $\mathrm{diag}\{a_1, \ldots, a_M\}$ and the block diagonal matrix $\mathrm{Bdiag}\{\boldsymbol{B}_1, \ldots, \boldsymbol{B}_M\}$ has $\boldsymbol{B}_1, \ldots, \boldsymbol{B}_M$ as its block diagonal entries, while $\otimes$ denotes the Kronecker product and $\mathrm{Tr}(\cdot)$ is the trace operator. Furthermore, $\boldsymbol{A}_{[i:]}$ and $\boldsymbol{A}_{[:i]}$ denote the ith row and ith column of $\boldsymbol{A}$, respectively. Furthermore, the subscripts tx and rx indicate that the variable considered is at the transmitter and the variable considered is at the receiver, respectively. The superscript (r/u) denotes a variable between the UDs and the RRU, and the superscript (b/r) denotes a variable between RRUs and the BBU, respectively. The superscripts (u, d) and (u, p) are variables related with UD's data transmission and pilot training, and the superscripts (r, d) and (r, p) are variables related with RRU's data transmission and pilot training, respectively.

For easy reference, below we list the key mathematical symbols used in the manuscript.

K: Total number of UDs.
R: Number of RRUs.
U_r: Number of UDs served by the rth RRU.
M: Number of RAs equipped by RRU.
N: Number of RAs equipped by BBU.

P_{UD}:	Total power of an UD.
P_{RRU}:	Total power of a RRU.
ρ_r:	Receiver's efficiency factor of the rth RRU.
$P_{\text{sp},r}$:	Received signal processing power of the rth RRU.
L_{pathloss}:	Pathloss.
d_{u_r}:	Distance between the u_rth UD and the rth RRU.
d_r:	Distance between the rth RRU and its host BBU.
$\lambda_r^{(r/u)}$:	Normalization factor at the rth RRU.
$\lambda^{(b/r)}$:	Normalization factor at the BBU.
$\eta_{u_r}^{(r/u)}$:	Power sharing factor of the u_rth UD.
$\eta_{u_r}^{(b/r)}$:	Power sharing factor of the TA for forwarding the u_rth UD's data.
$P_{\text{tx},u_r}^{(u,x)}$:	Transmit power of the u_rth UD, $x = p$ for pilot and $x = d$ for data.
$P_{\text{tx},r(u_r)}^{(r,x)}$:	Transmit power of the u_rth TA at the rth RRU, $x = p$ for pilot and $x = d$ for data.
$P_{\text{rx},u_r}^{(u,x)}$:	Receive signal power of the u_rth UD at the rth RRU, $x = p$ for pilot and $x = d$ for data.
$\boldsymbol{P}_{\text{rx},r/r'}^{(u,x)}$:	Receive signal powers of the r'th RRU's UDs at the rth RRU, $x = p$ for pilot and $x = d$ for data.
$\boldsymbol{P}_{\text{rx}}^{(r,d)}$:	Receive signal powers at the BBU.
$\boldsymbol{s}_r$:	Transmit signals of the UDs served by the rth RRU.
$\boldsymbol{H}_r^{(r/u)}$:	Uplink MIMO channel between the rth RRU and its UDs.
$\boldsymbol{H}_{r/r'}^{(r/u)}$:	Interfering channel between the r'th RRU's UDs and the rth RRU.
$\widetilde{\boldsymbol{y}}_r$:	Signals received at the rth RRU.
$\widetilde{\boldsymbol{n}}_r$:	AWGN vector at the rth RRU.
$\boldsymbol{W}_r^{(r/u)}$:	Receiver combining matrix used by the rth RRU.
$\boldsymbol{y}_r$:	Uplink signals after combining at the rth RRU.
$\boldsymbol{n}_r$:	Noise output after the rth RRU's combining.
$\boldsymbol{A}_r$:	Power amplification at the rth RRU.
$\bar{\boldsymbol{y}}$:	Transmit signals consisting of all the signals transmitted by all the R RRUs.
K_{Rice}:	Rician factor.
$\boldsymbol{H}^{(b/r)}$:	MIMO Rician channel between the K forwarding TAs of all the R RRUs and their host BBU.
$\boldsymbol{H}_{\text{d}}^{(b/r)}$:	Deterministic part of the Rician channel.
$\boldsymbol{H}_{\text{r}}^{(b/r)}$:	Scattered component of the Rician channel .
$\check{\boldsymbol{y}}$:	Signals received at the BBU.
$\check{\boldsymbol{n}}$:	AWGN vector at the BBU.
$\boldsymbol{W}^{(b/r)}$:	BBU's combining matrix.
$\check{\boldsymbol{n}}$:	Noise output after BBU's combining.
$\check{\boldsymbol{y}}$:	Uplink signals after combining at BBU.

Additionally, the main abbreviations used are listed below.

AWGN:	Additive white Gaussian noise
BBU:	Baseband unit
C-RAN:	Centralized radio access network
CSI:	Channel state information
DEA:	Differential evolution algorithm
i.i.d.:	Independently and identically distributed
LoS:	Line-of-sign
MF:	Matched filter
mmWave:	Millimeter-wave
RA:	Receive antenna
RRU:	Remote radio unit
SINR:	Signal to interference plus noise ratio
TA:	Transmit antenna
UD:	User device

The rest of this paper is organized as follows. Section II describes the massive MIMO aided uplink C-RAN architecture, and Section III derives its closed-form achievable asymptotic uplink sum-rate. The global optimization metric as a function of all the power sharing factors and how to maximize it are presented in Section IV. Our simulation results are presented in Section V for demonstrating the efficiency of the proposed approach, whilst our conclusions are given in Section VI.

II. SYSTEM MODEL

Consider an uplink C-RAN architecture as illustrated in Fig. 1. Each RRU is connected to its host BBU via a wireless fronthaul link. Assume that the frequency is reused within the coverage of a BBU, while orthogonal access is adopted by the different BBUs. Thus there is no interference between the UDs located in the different BBUs' coverage areas. Hence we only have to consider a single BBU's coverage. The BBU employs N receive antennas (RAs) to serve R RRUs, while each RRU is equipped with M RAs. The rth RRU employs U_r transmit antennas (TAs) and serves U_r single-TA UDs, where $1 \leq r \leq R$. The total number of UDs within the BBU's coverage is $K = \sum_{r=1}^{R} U_r$. According to the well-known spatial multiplexing gaining or spatial degree of freedom, the number of independent data streams supported cannot be higher than the number of receive antennas. The spatial degree of freedom between the rth RRU and its supporting UDs is given by $\min\{U_r, M\}$ [33], [34]. Therefore, the number of UDs served by the rth RRU is no more than the number of the rth RRU's RAs, i.e., $U_r \leq M$. We further assume that $M \gg U_r$ for $r \in \{1, 2, \ldots, R\}$. Since the rth RRU uses U_r TAs for forwarding its serving UDs' data to the BBU, the total number of TAs for all the R RRUs is K, which equals to the number of uplink fronthaul streams. Clearly, K is no more than the number of the BBU's RAs, i.e., $K \leq N$.

A. Power Consumption Preliminaries

For a fair power allocation, all the UDs have the same total power P_{UD}, which is shared by pilot training and data transmission of each UD. Explicitly, for the link of the u_rth UD to the rth RRU, the power allocation between pilot training and data transmission is controlled by a power sharing factor $0 < \eta_{u_r}^{(r/u)} < 1$. Let the power of pilot training be $P_{\text{tx},u_r}^{(u,p)}$ and that of data transmission be $P_{\text{tx},u_r}^{(u,d)}$, respectively, for the u_rth UD served by the rth RRU. Then,

$$P_{\text{tx},u_r}^{(u,p)} = \eta_{u_r}^{(r/u)} P_{\text{UD}}, \tag{1}$$

$$P_{\text{tx},u_r}^{(u,d)} = \left(1 - \eta_{u_r}^{(r/u)}\right) P_{\text{UD}}. \tag{2}$$

For convenience, we introduce the new symbol x with $x = p$ indicating pilot training and $x = d$ indicating data transmission. At the rth RRU, the received signal power $P_{\text{rx},u_r}^{(u,x)}$ of the u_rth
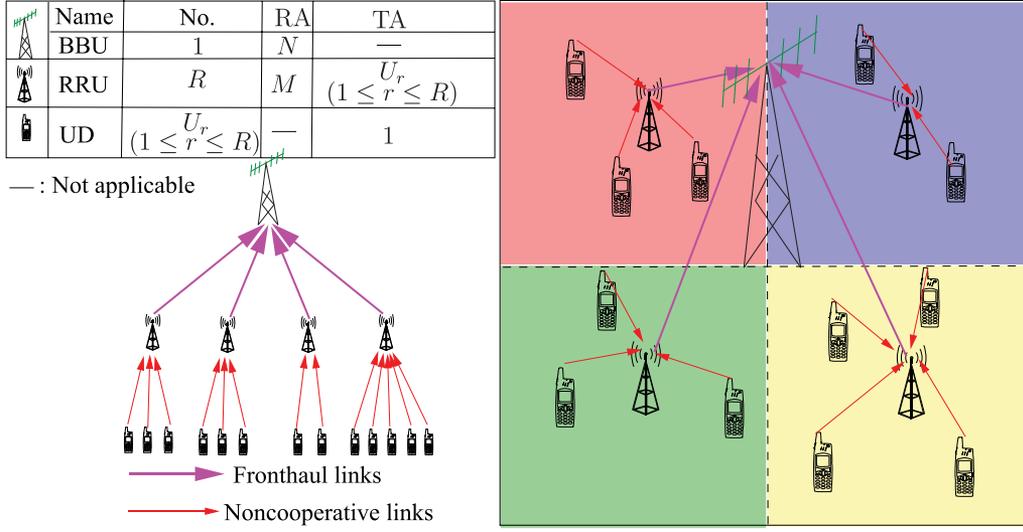
	Name	No.	RA	TA
	BBU	1	N	—
	RRU	R	M	U_r $(1 \le r \le R)$
	UD	U_r $(1 \le r \le R)$	—	1

— : Not applicable

Fig. 1.   An ultra-dense C-RAN architecture for uplink transmission.

UD is given by

$$P_{\mathrm{rx},u_r}^{(u,x)} = P_{\mathrm{tx},u_r}^{(u,x)} 10^{-L_{\mathrm{pathloss}}(u_r)/10}, \tag{3}$$

where the pathloss $L_{\mathrm{pathloss}}(u_r)$ is given by [35]

$$L_{\mathrm{pathloss}}(u_r)\,[\mathrm{dB}] = -154 + 20\log_{10}\left(f_c\right) + 20\log_{10}\left(d_{u_r}\right), \tag{4}$$

in which d_{u_r} [m] is the distance between the u_rth UD and its host, the rth RRU, and f_c [Hz] is the carrier frequency. For the 5G system, $f_c = 3.4\,\mathrm{GHz}$ has been allocated in United Kingdom (UK) [36], which will be considered in our investigations.

Each RRU is allocated with the same total power P_{RRU} for the sake of fairness, which is shared by the received signal processing of detecting the served UDs' data as well as pilots and data transmission for forwarding these UDs' data to the BBU. Intuitively, the power consumed by the rth RRU's received signal processing is related to the uplink sum-rate of its serving UDs, which is dominated by the uplink transmit power. Since the accurate modeling of this received signal processing is absent in the literature, we approximately model the rth RRU's received signal processing power consumption $P_{\mathrm{sp},r}$ as a function of its serving UDs' uplink transmit power by

$$P_{\mathrm{sp},r} = \rho_r U_r P_{\mathrm{UD}}, \tag{5}$$

where ρ_r is the receiver's efficiency factor of the rth RRU. The remaining power $(P_{\mathrm{RRU}} - P_{\mathrm{sp},r})$ of the rth RRU is shared by its U_r TAs for pilots and data transmission. Let $\eta_{u_r}^{(b/r)}$ be the power sharing factor of the TA for forwarding the u_rth UD's data. Then, the pilot training power $P_{\mathrm{tx},r(u_r)}^{(r,p)}$ and data transmission power $P_{\mathrm{tx},r(u_r)}^{(r,d)}$ for the u_rth TA are given by

$$P_{\mathrm{tx},r(u_r)}^{(r,p)} = \eta_{u_r}^{(b/r)}\left(P_{\mathrm{RRU}} - P_{\mathrm{sp},r}\right)/U_r, \tag{6}$$

$$P_{\mathrm{tx},r(u_r)}^{(r,d)} = \left(1 - \eta_{u_r}^{(b/r)}\right)\left(P_{\mathrm{RRU}} - P_{\mathrm{sp},r}\right)/U_r. \tag{7}$$

At the BBU, the received signal power $P_{\mathrm{rx},r(u_r)}^{(r,x)}$ is related to the transmit signal power $P_{\mathrm{tx},r(u_r)}^{(r,x)}$ by a similar pathloss model $P_{\mathrm{rx},r(u_r)}^{(r,x)} = P_{\mathrm{tx},r(u_r)}^{(r,x)} 10^{-L_{\mathrm{pathloss}}[r(u_r)]/10}$. Since mmWave communication is established between the RRUs and their host BBU,

the pathloss $L_{\mathrm{pathloss}}[r(u_r)]$ of the RRU-to-BBU link is given by [37], [38]

$$L_{\mathrm{pathloss}}\left[r(u_r)\right]\,[\mathrm{dB}] = 3.34 + 18.62\log_{10}(f_{mm}) + 22\log_{10}(d_r), \tag{8}$$

where d_r [m] is the distance between the rth RRU and its host BBU, while f_{mm} [GHz] is the carrier frequency. For the mmWave based 5G system in the UK, $f_{mm} = 26\,\mathrm{GHz}$ is allocated [36], which is used for our investigations.

B. Signal Model of UD-to-RRU

Obviously the pilot training and data transmission have the same signal model. Again, we introduce the symbol x, with $x = p$ indicating the pilot training and $x = d$ representing data transmission, respectively. Then, the signals received at the rth RRU $\widetilde{y}_r \in \mathbb{C}^M$ can be expressed generically as

$$\widetilde{y}_r = H_r^{(r/u)}\left(P_{\mathrm{rx},r}^{(u,x)}\right)^{\frac{1}{2}} s_r + \sum_{r'=1, r' \neq r}^{R} H_{r/r'}^{(r/u)}\left(P_{\mathrm{rx},r/r'}^{(u,x)}\right)^{\frac{1}{2}} s_{r'} + \widetilde{n}_r, \tag{9}$$

where $H_r^{(r/u)} \in \mathbb{C}^{M \times U_r}$ is the uplink MIMO channel between the rth RRU and its UDs, $s_r = [s_1^{(r)} s_2^{(r)} \cdots s_{U_r}^{(r)}]^{\mathrm{T}}$ is the transmit signal vector with $\mathcal{E}\{|s_{u_r}^{(r)}|\} = 1$ for $u_r \in \{1, 2, \ldots, U_r\}$, and $\widetilde{n}_r \in \mathbb{C}^M$ is the additive white Gaussian noise (AWGN) vector with the distribution $\mathcal{CN}(\mathbf{0}_M, \sigma_{\widetilde{n}}^2 I_M)$. Still referring to (9), and $P_{\mathrm{rx},r}^{(u,x)} = \mathrm{diag}\{P_{\mathrm{rx},r(1)}^{(u,x)}, \ldots, P_{\mathrm{rx},r(U_r)}^{(u,x)}\}$ are the received powers of the U_r UDs' signals at the rth RRU, while $H_{r/r'}^{(r/u)}$ is the interfering channel matrix between the r'th RRU's UDs and the rth RRU, and $P_{\mathrm{rx},r/r'}^{(u,x)} = \mathrm{diag}\{P_{\mathrm{rx},r/r'(1)}^{(u,x)}, \ldots, P_{\mathrm{rx},r/r'(U_{r'})}^{(u,x)}\}$ are the received powers of the r'th RRU's UDs at the rth RRU. Since all the UDs' signals suffer from the same noise at the RRU, the noise power at all the UD-to-RRU links are identical. Furthermore, the second term in (9) represents the interference imposed by the UDs of the adjacent RRUs.

Because the UDs are randomly distributed in the RRUs' coverage areas and there are many obstructions between the UDs and their host RRU, the direct line-of-sight (LoS) paths may always be blocked. Hence, the channels between the UDs and their host RRU are Rayleigh channels, and the UD-to-RRU MIMO channel matrix can be expressed as

$$\boldsymbol{H}_r^{(r/u)} = \left(\boldsymbol{R}_{\text{rx},r}^{(r/u)}\right)^{\frac{1}{2}} \boldsymbol{G}_r^{(r/u)} \left(\boldsymbol{R}_{\text{tx},u}^{(r/u)}\right)^{\frac{1}{2}}, \qquad (10)$$

where $\boldsymbol{R}_{\text{rx},r}^{(r/u)} \in \mathbb{C}^{M \times M}$ is the spatial correlation matrix of the rth RRU's M RAs and $\boldsymbol{R}_{\text{tx},u}^{(r/u)} \in \mathbb{C}^{U_r \times U_r}$ is the spatial correlation matrix of the U_r UDs, while $\boldsymbol{G}_r^{(r/u)} \in \mathbb{C}^{M \times U_r}$ has the independently identically distributed (i.i.d.) complex entries and each of them has the distribution of $\mathcal{CN}(0,1)$. Because the UDs are randomly distributed and they are independent of each other, there is no correlation between the TAs of different UDs and we have $\boldsymbol{R}_{\text{tx},u}^{(r/u)} = \boldsymbol{I}_{U_r}$.

Let the receiver combining matrix used by the rth RRU be $\boldsymbol{W}_r^{(r/u)} \in \mathbb{C}^{U_r \times M}$. The uplink signals $\boldsymbol{y}_r \in \mathbb{C}^{U_r}$ after combining at the rth RRU can be expressed as

$$\boldsymbol{y}_r = \sqrt{\lambda_r^{(r/u)}} \boldsymbol{W}_r^{(r/u)} \boldsymbol{H}_r^{(r/u)} \left(\boldsymbol{P}_{\text{rx},r}^{(u,x)}\right)^{\frac{1}{2}} \boldsymbol{s}_r$$
$$+ \sqrt{\lambda_r^{(r/u)}} \boldsymbol{W}_r^{(r/u)} \sum_{r'=1, r' \neq r}^{R} \boldsymbol{H}_{r/r'}^{(r/u)} \left(\boldsymbol{P}_{\text{rx},r/r'}^{(u,x)}\right)^{\frac{1}{2}} \boldsymbol{s}_{r'} + \boldsymbol{n}_r, \quad (11)$$

where $\lambda_r^{(r/u)} = 1/(\frac{1}{U_r} \mathcal{E}\{\text{Tr}\{\boldsymbol{W}_r^{(r/u)}(\boldsymbol{W}_r^{(r/u)})^{\text{H}}\}\})$ is the normalization factor, and $\boldsymbol{n}_r = \sqrt{\lambda_r^{(r/u)}} \boldsymbol{W}_r^{(r/u)} \tilde{\boldsymbol{n}}_r \in \mathbb{C}^{U_r}$ is the effective noise vector having the distribution of $\mathcal{CN}(\boldsymbol{0}_{U_r}, \boldsymbol{\Sigma}_{\boldsymbol{n}_r})$ with the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{n}_r} = \lambda_r^{(r/u)} \sigma_{\tilde{n}}^2 \boldsymbol{W}_r^{(r/u)} \left(\boldsymbol{W}_r^{(r/u)}\right)^{\text{H}}$. When a matched-filter (MF) is used for uplink combining, we have

$$\boldsymbol{W}_r^{(r/u)} = \left(\widehat{\boldsymbol{H}}_r^{(r/u)}\right)^{\text{H}}, \qquad (12)$$

where $\widehat{\boldsymbol{H}}_r^{(r/u)}$ is the estimate of the uplink channel $\boldsymbol{H}_r^{(r/u)}$.

$$\boldsymbol{vec}\left(\widehat{\boldsymbol{H}}_r^{(r/u)}\right) = \boldsymbol{\Psi}_{\boldsymbol{H}_r^{(r/u)}} \left(\sigma_{\tilde{n}}^2 \left(\boldsymbol{P}_{\text{rx},r}^{(u,p)}\right)^{-1} \otimes \boldsymbol{I}_M + \boldsymbol{\Psi}_{\boldsymbol{H}_r^{(r/u)}}\right)^{-1}$$
$$\times \boldsymbol{vec}(\widetilde{\boldsymbol{Y}}_r \boldsymbol{X}_r^{(p)}), \qquad (13)$$

The optimal minimum mean square error (MMSE) channel estimator [39] is given by (13),[1] where $\widetilde{\boldsymbol{Y}}_r \in \mathbb{C}^{M \times U_r}$ is the received signal matrix over the U_r pilot symbols, and $\boldsymbol{X}_r^{(p)} \in \mathbb{C}^{U_r \times U_r}$ is the pilot symbol matrix with $\boldsymbol{X}_r^{(p)}(\boldsymbol{X}_r^{(p)})^{\text{H}} = \boldsymbol{I}_{U_r}$, while $\boldsymbol{P}_{\text{rx},r}^{(u,p)} = \text{diag}\left\{P_{\text{rx},r(1)}^{(u,p)}, \ldots, P_{\text{rx},r(U_r)}^{(u,p)}\right\}$ are the received powers of the pilot symbols. Furthermore, in (13), $\boldsymbol{\Psi}_{\boldsymbol{H}_r^{(r/u)}}$ denotes the covariance matrix of $\boldsymbol{vec}(\boldsymbol{H}_r^{(r/u)})$, which is

given by

$$\boldsymbol{\Psi}_{\boldsymbol{H}_r^{(r/u)}} = \mathcal{E}\left\{ \boldsymbol{vec}\left(\boldsymbol{H}_r^{(r/u)}\right) \boldsymbol{vec}\left(\boldsymbol{H}_r^{(r/u)}\right)^{\text{H}} \right\}$$
$$= \boldsymbol{I}_{U_r} \otimes \boldsymbol{R}_{\text{rx},r}^{(r/u)} \in \mathbb{C}^{MU_r \times MU_r}. \qquad (14)$$

The MMSE estimate (13) follows the distribution [39]

$$\boldsymbol{vec}\left(\widehat{\boldsymbol{H}}_r^{(r/u)}\right) \sim \mathcal{CN}\left(\boldsymbol{0}_{MU_r}, \boldsymbol{\Psi}_{\widehat{\boldsymbol{H}}_r^{(r/u)}}\right), \qquad (15)$$

with the covariance matrix $\boldsymbol{\Psi}_{\widehat{\boldsymbol{H}}_r^{(r/u)}}$ given by

$$\boldsymbol{\Psi}_{\widehat{\boldsymbol{H}}_r^{(r/u)}} = \boldsymbol{\Psi}_{\boldsymbol{H}_r^{(r/u)}} \left(\sigma_{\tilde{n}}^2 \left(\boldsymbol{P}_{\text{rx},r}^{(u,p)}\right)^{-1} \otimes \boldsymbol{I}_M + \boldsymbol{\Psi}_{\boldsymbol{H}_r^{(r/u)}}\right)^{-1}$$
$$\times \boldsymbol{\Psi}_{\boldsymbol{H}_r^{(r/u)}} \in \mathbb{C}^{MU_r \times MU_r}. \qquad (16)$$

The relationship between the MMSE channel estimate $\boldsymbol{vec}(\widehat{\boldsymbol{H}}_r^{(r/u)})$ and the true channel $\boldsymbol{vec}(\boldsymbol{H}_r^{(r/u)})$ is given by

$$\boldsymbol{vec}\left(\boldsymbol{H}_r^{(r/u)}\right) = \boldsymbol{vec}\left(\widehat{\boldsymbol{H}}_r^{(r/u)}\right) + \boldsymbol{vec}\left(\widetilde{\boldsymbol{H}}_r^{(r/u)}\right), \qquad (17)$$

where the channel estimation error $\boldsymbol{vec}(\widetilde{\boldsymbol{H}}_r^{(r/u)})$ is statistically independent of both $\boldsymbol{vec}(\widehat{\boldsymbol{H}}_r^{(r/u)})$ and $\boldsymbol{vec}(\boldsymbol{H}_r^{(r/u)})$. Moreover, the distribution of $\boldsymbol{vec}(\widetilde{\boldsymbol{H}}_r^{(r/u)})$ is

$$\boldsymbol{vec}\left(\widetilde{\boldsymbol{H}}_r^{(r/u)}\right) \sim \mathcal{CN}\left(\boldsymbol{0}_{MU_r}, \boldsymbol{\Psi}_{\widetilde{\boldsymbol{H}}_r^{(r/u)}}\right), \qquad (18)$$

with the covariance matrix $\boldsymbol{\Psi}_{\widetilde{\boldsymbol{H}}_r^{(r/u)}}$ given by

$$\boldsymbol{\Psi}_{\widetilde{\boldsymbol{H}}_r^{(r/u)}} = \boldsymbol{\Psi}_{\boldsymbol{H}_r^{(r/u)}} - \boldsymbol{\Psi}_{\widehat{\boldsymbol{H}}_r^{(r/u)}}. \qquad (19)$$

C. Signal Model of RRU-to-BBU

The UDs' data signals received by their host RRUs are forwarded to the BBU after power amplification. Ideally, the u_rth UD's data received by the rth RRU is scaled to its transmit signal constellation by a power amplification coefficient $a_{u_r}^{(r)} = \frac{1}{P_{\text{rx},r(u_r)}^{(u,d)}}$. Hence, the power amplification at the rth RRU is represented by the diagonal matrix given by

$$\boldsymbol{A}_r = \text{diag}\{a_1^{(r)}, \cdots a_{U_r}^{(r)}\} = \left(\boldsymbol{P}_{\text{rx},r}^{(u,d)}\right)^{-1}, \qquad (20)$$

and the rth RRU forwards the amplified signal $\bar{\boldsymbol{y}}_r = \boldsymbol{A}_r^{\frac{1}{2}} \boldsymbol{y}_r \in \mathbb{C}^{U_r}$ to its host BBU via the fronthaul links. Explicitly, let $\bar{\boldsymbol{y}} \in \mathbb{C}^K$ be the transmit signal vector consisting of all the signals transmitted by the R RRUs, which is given by

$$\bar{\boldsymbol{y}} = \left[\bar{\boldsymbol{y}}_1^{\text{T}} \cdots \bar{\boldsymbol{y}}_R^{\text{T}}\right]^{\text{T}} = \boldsymbol{A}^{\frac{1}{2}} \left(\boldsymbol{\lambda}^{(r/u)}\right)^{\frac{1}{2}} \boldsymbol{W}^{(r/u)} \boldsymbol{H}^{(r/u)} \left(\boldsymbol{P}_{\text{rx}}^{(u,d)}\right)^{\frac{1}{2}} \boldsymbol{s}$$
$$+ \boldsymbol{A}^{\frac{1}{2}} \left(\boldsymbol{\lambda}^{(r/u)}\right)^{\frac{1}{2}} \boldsymbol{W}^{(r/u)} \sum_{r'=2}^{R} \boldsymbol{V}_{r'}^{(r/u)} \widetilde{\boldsymbol{s}}_{r'} + \boldsymbol{A}^{\frac{1}{2}} \boldsymbol{n}, \qquad (21)$$

---

[1]We assume that the pilot contamination imposed by the UDs served by the adjacent RRUs has been eliminated by optimal pilot design [40], [41].

where we have

$$\boldsymbol{A} = \text{Bdiag}\{\boldsymbol{A}_1, \ldots, \boldsymbol{A}_R\} \in \mathbb{C}^{K \times K}, \qquad (22)$$

$$\boldsymbol{\lambda}^{(r/u)} = \text{Bdiag}\left\{\lambda_1^{(r/u)} \boldsymbol{I}_{U_1}, \ldots, \lambda_R^{(r/u)} \boldsymbol{I}_{U_R}\right\} \in \mathbb{C}^{K \times K}, \qquad (23)$$

$$\boldsymbol{W}^{(r/u)} = \text{Bdiag}\left\{\boldsymbol{W}_1^{(r/u)}, \ldots, \boldsymbol{W}_R^{(r/u)}\right\} \in \mathbb{C}^{K \times MR}, \qquad (24)$$

$$\boldsymbol{H}^{(r/u)} = \text{Bdiag}\left\{\boldsymbol{H}_1^{(r/u)}, \ldots, \boldsymbol{H}_R^{(r/u)}\right\} \in \mathbb{C}^{MR \times K}, \qquad (25)$$

$$\boldsymbol{P}_{\text{rx}}^{(u,d)} = \text{Bdiag}\left\{\boldsymbol{P}_{\text{rx},1}^{(u,d)}, \ldots, \boldsymbol{P}_{\text{rx},R}^{(u,d)}\right\} \in \mathbb{C}^{K \times K}, \qquad (26)$$

$$\boldsymbol{s} = \left[\boldsymbol{s}_1^{\text{T}} \cdots \boldsymbol{s}_R^{\text{T}}\right]^{\text{T}} \in \mathbb{C}^K, \qquad (27)$$

$$\boldsymbol{n} = \left[\boldsymbol{n}_1^{\text{T}} \cdots \boldsymbol{n}_R^{\text{T}}\right]^{\text{T}} \in \mathbb{C}^K, \qquad (28)$$

$$\boldsymbol{V}_{r'}^{(r/u)} = \text{Bdiag}\left\{\boldsymbol{V}_{1/(r' \bmod R)}^{(r/u)}, \ldots, \boldsymbol{V}_{R/((r'+R-1) \bmod R)}^{(r/u)}\right\}$$
$$\in \mathbb{C}^{MR \times K}, \qquad (29)$$

$$\widetilde{\boldsymbol{s}}_{r'} = \left[\boldsymbol{s}_{(r' \bmod R)}^{\text{T}}, \ldots, \boldsymbol{s}_{((r'+R-1) \bmod R)}^{\text{T}}\right]^{\text{T}} \in \mathbb{C}^K. \quad (30)$$

In (29), $\boldsymbol{V}_{r/(t \bmod R)}^{(r/u)} \in \mathbb{C}^{M \times U_t}$, $t = r', r'+1, \ldots, (r'+R-1)$, depends on the value of $(t \bmod R)$, which is given by

$$\boldsymbol{V}_{r/(t \bmod R)}^{(r/u)}$$
$$= \begin{cases} \boldsymbol{H}_{r/(t \bmod R)}^{(r/u)} \left(\boldsymbol{P}_{\text{rx},r/(t \bmod R)}^{(u,d)}\right)^{\frac{1}{2}}, & t \bmod R \neq 0, \\ \boldsymbol{H}_{r/R}^{(r/u)} \left(\boldsymbol{P}_{\text{rx},r/R}^{(u,d)}\right)^{\frac{1}{2}} & t \bmod R = 0. \end{cases} \quad (31)$$

Similarly, in (30), if $t \bmod R \neq 0$, $\boldsymbol{s}_{(t \bmod R)}$ is as it is, while if $t \bmod R = 0$, $\boldsymbol{s}_{(t \bmod R)} = \boldsymbol{s}_R$.

The signals received at the BBU $\check{\boldsymbol{y}} \in \mathbb{C}^N$ are expressed as

$$\check{\boldsymbol{y}} = \boldsymbol{H}^{(b/r)} \left(\boldsymbol{P}_{\text{rx}}^{(r,d)}\right)^{\frac{1}{2}} \bar{\boldsymbol{y}} + \check{\boldsymbol{n}}, \qquad (32)$$

where $\check{\boldsymbol{n}} \sim \mathcal{CN}\left(\boldsymbol{0}_N, \sigma_{\check{n}}^2 \boldsymbol{I}_N\right)$ is the AWGN vector, $\boldsymbol{P}_{\text{rx}}^{(r,d)} = \text{diag}\{P_{\text{rx},1}^{(r,d)}, \ldots, P_{\text{rx},K}^{(r,d)}\}$ are the received signal powers at the BBU, and $\boldsymbol{H}^{(b/r)} \in \mathbb{C}^{N \times K}$ is the MIMO channel matrix between the K forwarding-TAs of the R RRUs and their host BBU. The RRUs are generally stationary and are carefully positioned, so that the direct LoS paths always exist between the RRUs and their host BBU. Thus the channels between the RRUs and their host BBU are Rician channels and, therefore, the MIMO channel matrix $\boldsymbol{H}^{(b/r)}$ is given by

$$\boldsymbol{H}^{(b/r)} = \nu \boldsymbol{H}_{\text{d}}^{(b/r)} + \zeta \boldsymbol{H}_{\text{r}}^{(b/r)}$$
$$= \nu \boldsymbol{H}_{\text{d}}^{(b/r)} + \zeta \left(\boldsymbol{R}_{\text{rx},b}^{(b/r)}\right)^{\frac{1}{2}} \boldsymbol{G}^{(b/r)} \left(\boldsymbol{R}_{\text{tx},r}^{(b/r)}\right)^{\frac{1}{2}}, \quad (33)$$

where $\boldsymbol{H}_{\text{d}}^{(b/r)}$ is the deterministic part of the Rician channel satisfying $\text{Tr}\{\boldsymbol{H}_{\text{d}}^{(b/r)}(\boldsymbol{H}_{\text{d}}^{(b/r)})^{\text{H}}\} = NK$, $\nu = \sqrt{\frac{K_{Rice}}{1+K_{Rice}}}$ and $\zeta = \sqrt{\frac{1}{1+K_{Rice}}}$ with K_{Rice} being the Rician factor, while

$\boldsymbol{H}_{\text{r}}^{(b/r)} = \left(\boldsymbol{R}_{\text{rx},b}^{(b/r)}\right)^{\frac{1}{2}} \boldsymbol{G}^{(b/r)} \left(\boldsymbol{R}_{\text{tx},r}^{(b/r)}\right)^{\frac{1}{2}}$ is the scattered component of the Rician channel in which $\boldsymbol{R}_{\text{rx},b}^{(b/r)} \in \mathbb{C}^{N \times N}$ is the spatial correlation matrix of the N RAs at the BBU, $\boldsymbol{R}_{\text{tx},r}^{(b/r)} \in \mathbb{C}^{K \times K}$ is the spatial correlation matrix of the K forwarding TAs of the R RRUs, and $\boldsymbol{G}^{(b/r)} \in \mathbb{C}^{N \times K}$ has i.i.d. complex entries and each of them has the distribution $\mathcal{CN}(0, 1)$. The rth RRU has a total of M antennas, which is much more than U_r, and it can always select its U_r forwarding TAs to be spaced sufficiently far apart. Consequently, the correlations between the U_r TAs can be assumed to be zero. Furthermore, there exists no correlation between the antennas of different RRUs. Thus we can assume that $\boldsymbol{R}_{\text{tx},r}^{(b/r)} = \boldsymbol{I}_K$.

The uplink signals after combining can be expressed as

$$\check{\boldsymbol{y}} = \sqrt{\lambda^{(b/r)}} \boldsymbol{W}^{(b/r)} \boldsymbol{H}^{(b/r)} \left(\boldsymbol{P}_{\text{rx}}^{(r,d)}\right)^{\frac{1}{2}} \bar{\boldsymbol{y}} + \check{\boldsymbol{n}}, \qquad (34)$$

where $\boldsymbol{W}^{(b/r)} \in \mathbb{C}^{K \times N}$ is the BBU's combining matrix, $\lambda^{(b/r)} = \frac{1}{K} \text{Tr}\{\mathcal{E}\{\boldsymbol{W}^{(b/r)}(\boldsymbol{W}^{(b/r)})^{\text{H}}\}\}$, and $\check{\boldsymbol{n}} = \sqrt{\lambda^{(b/r)}} \boldsymbol{W}^{(b/r)} \check{\boldsymbol{n}} \in \mathbb{C}^K$ is the noise output after the combining, which obeys the distribution $\mathcal{CN}(\boldsymbol{0}_K, \boldsymbol{\Sigma}_{\check{n}})$ with the covariance matrix $\boldsymbol{\Sigma}_{\check{n}} = \lambda^{(b/r)} \sigma_{\check{n}}^2 \boldsymbol{W}^{(b/r)}(\boldsymbol{W}^{(b/r)})^{\text{H}}$. Again, the MF is adopted by the BBU and $\boldsymbol{W}^{(b/r)}$ is given by

$$\boldsymbol{W}^{(b/r)} = \left(\widehat{\boldsymbol{H}}^{(b/r)}\right)^{\text{H}}, \qquad (35)$$

where $\widehat{\boldsymbol{H}}^{(b/r)}$ is the estimate of the uplink channel $\boldsymbol{H}^{(b/r)}$.

The MMSE channel estimate [39] is given by (36), shown at the bottom of this page, where $\check{\boldsymbol{Y}} \in \mathbb{C}^{N \times K}$ is the received signal matrix with removing the LoS component over the K pilot symbols before the receiver combining, and $\boldsymbol{X}_b^{(p)} \in \mathbb{C}^{K \times K}$ is the pilot symbol matrix with $\boldsymbol{X}_b^{(p)}(\boldsymbol{X}_b^{(p)})^{\text{H}} = \boldsymbol{I}_K$, while $\boldsymbol{P}_{\text{rx}}^{(r,p)} = \text{diag}\{P_{\text{rx},1}^{(r,p)}, \ldots, P_{\text{rx},K}^{(r,p)}\}$ are the received powers of the pilot symbols. In (36), $\boldsymbol{\Psi}_{\boldsymbol{H}_{\text{r}}^{(b/r)}} \in \mathbb{C}^{NK \times NK}$ is the covariance matrix of $\boldsymbol{vec}(\zeta \boldsymbol{H}_{\text{r}}^{(b/r)})$ and it is given by

$$\boldsymbol{\Psi}_{\boldsymbol{H}_{\text{r}}^{(b/r)}} = \mathcal{E}\left\{\boldsymbol{vec}\left(\zeta \boldsymbol{H}_{\text{r}}^{(b/r)}\right) \boldsymbol{vec}\left(\zeta \boldsymbol{H}_{\text{r}}^{(b/r)}\right)^{\text{H}}\right\}$$
$$= \zeta^2 \boldsymbol{I}_K \otimes \boldsymbol{R}_{\text{rx},b}^{(b/r)}. \qquad (37)$$

The distribution of the MMSE estimate $\boldsymbol{vec}(\widehat{\boldsymbol{H}}^{(b/r)})$ is [39]

$$\boldsymbol{vec}\left(\widehat{\boldsymbol{H}}^{(b/r)}\right) \sim \mathcal{CN}\left(\nu \boldsymbol{vec}(\boldsymbol{H}_{\text{d}}^{(b/r)}), \boldsymbol{\Psi}_{\widehat{\boldsymbol{H}}^{(b/r)}}\right), \qquad (38)$$

with the covariance matrix formulated as

$$\boldsymbol{\Psi}_{\widehat{\boldsymbol{H}}^{(b/r)}} = \zeta^2 \boldsymbol{\Psi}_{\boldsymbol{H}_{\text{r}}^{(b/r)}} \left(\sigma_{\check{n}}^2 \left(\boldsymbol{P}_{\text{rx}}^{(r,p)}\right)^{-1} \otimes \boldsymbol{I}_N + \zeta^2 \boldsymbol{\Psi}_{\boldsymbol{H}_{\text{r}}^{(b/r)}}\right)^{-1}$$
$$\times \zeta^2 \boldsymbol{\Psi}_{\boldsymbol{H}_{\text{r}}^{(b/r)}}. \qquad (39)$$

$$\boldsymbol{vec}\left(\widehat{\boldsymbol{H}}^{(b/r)}\right) = \nu \boldsymbol{vec}\left(\boldsymbol{H}_{\text{d}}^{(b/r)}\right) + \zeta^2 \boldsymbol{\Psi}_{\boldsymbol{H}_{\text{r}}^{(b/r)}} \left(\sigma_{\check{n}}^2 \left(\boldsymbol{P}_{\text{rx}}^{(r,p)}\right)^{-1} \otimes \boldsymbol{I}_N + \zeta^2 \boldsymbol{\Psi}_{\boldsymbol{H}_{\text{r}}^{(b/r)}}\right)^{-1} \boldsymbol{vec}\left(\check{\boldsymbol{Y}} \boldsymbol{X}_b^{(p)}\right), \qquad (36)$$

The relationship between the MMSE estimate $vec(\widehat{\boldsymbol{H}}^{(b/r)})$ and the true channel $vec(\boldsymbol{H}^{(b/r)})$ is given by

$$vec\left(\boldsymbol{H}^{(b/r)}\right) = vec\left(\widehat{\boldsymbol{H}}^{(b/r)}\right) + vec\left(\widetilde{\boldsymbol{H}}^{(b/r)}\right), \qquad (40)$$

where the channel estimation error $vec(\widetilde{\boldsymbol{H}}^{(b/r)})$ is statistically independent of both $vec(\widehat{\boldsymbol{H}}^{(b/r)})$ and $vec(\boldsymbol{H}^{(b/r)})$. Moreover, the distribution of $vec(\widetilde{\boldsymbol{H}}^{(b/r)})$ is given by

$$vec(\widetilde{\boldsymbol{H}}^{(b/r)}) \sim \mathcal{CN}\left(\boldsymbol{0}_{NK}, \boldsymbol{\Psi}_{\widetilde{\boldsymbol{H}}^{(b/r)}}\right), \qquad (41)$$

with the covariance matrix formulated as

$$\boldsymbol{\Psi}_{\widetilde{\boldsymbol{H}}^{(b/r)}} = \boldsymbol{\Psi}_{\boldsymbol{H}_{\mathrm{r}}^{(b/r)}} - \boldsymbol{\Psi}_{\widehat{\boldsymbol{H}}^{(b/r)}}. \qquad (42)$$

III. ACHIEVABLE THROUGHPUT

The achievable throughput of a UD is determined by its final SINR at the BBU. There are a total of K UDs served by the BBU. Let us map the index of the u_rth UD served by the rth RRU to k. That is, the kth UD, where $k \in \{1, 2, \ldots, K\}$, is served by the rth RRU, and we can express k as $k = \sum_{r'=1}^{r-1} U_{r'} + u_r$. The SINR of the kth UD at the BBU is expressed as

$$\gamma_k = \frac{P_{\mathrm{S},k}}{P_{\mathrm{IN},k}}, \qquad (43)$$

where $P_{\mathrm{S},k}$ is the power of the desired signal and $P_{\mathrm{IN},k}$ is the power of the interference plus noise. The ergodic achievable uplink throughput of the kth UD is defined as

$$C_k = \mathcal{E}\left\{\log_2\left(1 + \gamma_k\right)\right\}. \qquad (44)$$

Recalling Lemma 1 of Appendix A, the ergodic achievable uplink throughput of the kth UD can be approximated as

$$C_k \approx \log_2\left(1 + \frac{\bar{P}_{\mathrm{S},k}}{\bar{P}_{\mathrm{IN},k}}\right), \qquad (45)$$

where $\bar{P}_{\mathrm{S},k} = \mathcal{E}\{P_{\mathrm{S},k}\}$ and $\bar{P}_{\mathrm{IN},k} = \mathcal{E}\{P_{\mathrm{IN},k}\}$.

To calculate the ergodic signal power $\bar{P}_{\mathrm{S},k}$ and the interference plus noise power $\bar{P}_{\mathrm{IN},k}$, we substitute (21) into (34) to arrive at the k^*th UD's signal after the BBU's combining operation as given in (46) at the bottom of the page, where a_{k^*}, $P_{\mathrm{rx},k^*}^{(u,d)}$ and $P_{\mathrm{rx},k^*}^{(r,d)}$ are the k^*th diagonal elements of $\boldsymbol{A}$, $\boldsymbol{P}_{\mathrm{rx}}^{(u,d)}$ and $\boldsymbol{P}_{\mathrm{rx}}^{(r,d)}$, respectively. Furthermore, $\widetilde{s}_{r',j}$ is the jth element

of $\widetilde{\boldsymbol{s}}_{r'}$, $\check{\tilde{n}}_{k^*} = \sqrt{a_{k^*}\lambda_{k^*}^{(r/u)}P_{\mathrm{rx},k^*}^{(r,d)}}\boldsymbol{W}_{[k^*:]}^{(b/r)}\boldsymbol{H}_{[:k^*]}^{(b/r)}n_{k^*} + \check{n}_{k^*}$ is the effective noise, and the k^*th UD is served by the r^*th RRU. Moreover, $\delta(\Delta_{j,k})$ in (46) is the Dirac delta function, and the value of $\Delta_{j,k}$ depends on whether the jth and kth UDs are served by the same RRU. Specifically, if both the jth and kth UDs are served by the same RRU, we have $\Delta_{j,k} = 1$; otherwise $\Delta_{j,k} = 0$.

Then, the signal power $\bar{P}_{\mathrm{S},k^*}$ is given in (47), as shown at the bottom of the page. Intuitively, the UD-to-RRU channel links are independent of the RRU-to-BBU channel links. Hence, we can calculate their expectation separately. By denoting

$$x\Upsilon_{s,1} = \mathcal{E}\left\{\left(\widehat{\boldsymbol{H}}_{[:k^*]}^{(r/u)}\right)^{\mathrm{H}}\boldsymbol{H}_{[:k^*]}^{(r/u)}\left(\boldsymbol{H}_{[:k^*]}^{(r/u)}\right)^{\mathrm{H}}\widehat{\boldsymbol{H}}_{[:k^*]}^{(r/u)}\right\}, \quad (48)$$

$$\Upsilon_{s,2} = \mathcal{E}\left\{\left(\widehat{\boldsymbol{H}}_{[:k^*]}^{(b/r)}\right)^{\mathrm{H}}\widehat{\boldsymbol{H}}_{[:k^*]}^{(b/r)}\left(\widehat{\boldsymbol{H}}_{[:k^*]}^{(b/r)}\right)^{\mathrm{H}}\widehat{\boldsymbol{H}}_{[k^*:]}^{(b/r)}\right\}, \quad (49)$$

$\bar{P}_{\mathrm{S},k^*}$ can be expressed as

$$\bar{P}_{\mathrm{S},k^*} = a_{k^*}\lambda^{(b/r)}\lambda_{k^*}^{(r/u)}P_{\mathrm{rx},k^*}^{(r,d)}P_{\mathrm{rx},k^*}^{(u,d)}\Upsilon_{s,1}\Upsilon_{s,2}. \quad (50)$$

Recalling Lemmas 2 and 3 of Appendix A, we have

$$\Upsilon_{s,1} = \mathcal{E}\left\{\left(\widehat{\boldsymbol{H}}_{[:k^*]}^{(r/u)}\right)^{\mathrm{H}}\left(\widehat{\boldsymbol{H}}_{[:k^*]}^{(r/u)} + \widetilde{\boldsymbol{H}}_{[:k^*]}^{(r/u)}\right)\right.$$
$$\left. \times \left(\widehat{\boldsymbol{H}}_{[:k^*]}^{(r/u)} + \widetilde{\boldsymbol{H}}_{[:k^*]}^{(r/u)}\right)^{\mathrm{H}}\widehat{\boldsymbol{H}}_{[:k^*]}^{(r/u)}\right\}$$
$$= \left(\mathrm{Tr}\left\{\boldsymbol{\Psi}_{\widehat{\boldsymbol{H}}_{[:k^*]}^{(r/u)}}\right\}\right)^2 + \mathrm{Tr}\left\{\boldsymbol{\Psi}_{\widehat{\boldsymbol{H}}_{[:k^*]}^{(r/u)}}\boldsymbol{\Psi}_{\widetilde{\boldsymbol{H}}_{[:k^*]}^{(r/u)}}\right\}, \quad (51)$$

where $\boldsymbol{\Psi}_{\widehat{\boldsymbol{H}}_{[:k^*]}^{(r/u)}}$ and $\boldsymbol{\Psi}_{\widetilde{\boldsymbol{H}}_{[:k^*]}^{(r/u)}}$ denote the covariance matrices of $\widehat{\boldsymbol{H}}_{[:k^*]}^{(r/u)}$ and $\widetilde{\boldsymbol{H}}_{[:k^*]}^{(r/u)}$, respectively. Recalling Lemma 3 of Appendix A, we have

$$\Upsilon_{s,2} = \left(\mathrm{Tr}\left\{\nu^2\boldsymbol{B}_{\boldsymbol{H}_{\mathrm{d}[:,k^*]}^{(b/r)}} + \boldsymbol{\Psi}_{\widehat{\boldsymbol{H}}_{\mathrm{r}[:k^*]}^{(b/r)}}\right\}\right)^2, \quad (52)$$

where we have $\boldsymbol{B}_{\boldsymbol{H}_{\mathrm{d}[:,k^*]}^{(b/r)}} = \boldsymbol{H}_{\mathrm{d}[:,k^*]}^{(b/r)}(\boldsymbol{H}_{\mathrm{d}[:,k^*]}^{(b/r)})^{\mathrm{H}}$ and $\boldsymbol{\Psi}_{\widehat{\boldsymbol{H}}_{\mathrm{r}[:k^*]}^{(b/r)}}$ is the covariance matrix of $\widehat{\boldsymbol{H}}_{\mathrm{r}[:k^*]}^{(b/r)}$. Furthermore, $\lambda^{(b/r)}$ and $\lambda_{k^*}^{(r/u)}$

$$\check{y}_{k^*} = \sqrt{a_{k^*}\lambda^{(b/r)}\lambda_{k^*}^{(r/u)}P_{\mathrm{rx},k^*}^{(r,d)}P_{\mathrm{rx},k^*}^{(u,d)}}\boldsymbol{W}_{[k^*:]}^{(b/r)}\boldsymbol{H}_{[:k^*]}^{(b/r)}\boldsymbol{W}_{[k^*:]}^{(r/u)}\boldsymbol{H}_{[:k^*]}^{(r/u)}s_{k^*}$$
$$+ \sum_{k=1,k\neq k^*}^{K}\sqrt{a_{k^*}\lambda^{(b/r)}\lambda_k^{(r/u)}P_{\mathrm{rx},k}^{(r,d)}P_{\mathrm{rx},k}^{(u,d)}}\boldsymbol{W}_{[k^*:]}^{(b/r)}\boldsymbol{H}_{[:k]}^{(b/r)}\boldsymbol{W}_{[k:]}^{(r/u)}\boldsymbol{H}_{[:k]}^{(r/u)}s_k$$
$$+ \sum_{k=1}^{K}\sum_{j=1,j\neq k}^{K}\delta\left(\Delta_{j,k}\right)\sqrt{a_{k^*}\lambda^{(b/r)}\lambda_k^{(r/u)}P_{\mathrm{rx},k}^{(r,d)}}\boldsymbol{W}_{[k^*:]}^{(b/r)}\boldsymbol{H}_{[:k]}^{(b/r)}\boldsymbol{W}_{[k:]}^{(r/u)}\sum_{r'=2}^{R}\boldsymbol{V}_{r'[:j]}^{(r/u)}\widetilde{s}_{r',j} + \check{\tilde{n}}_{k^*}, \qquad (46)$$

$$\bar{P}_{\mathrm{S},k^*} = a_{k^*}\lambda^{(b/r)}\lambda_{k^*}^{(r/u)}P_{\mathrm{rx},k^*}^{(r,d)}P_{\mathrm{rx},k^*}^{(u,d)}\mathcal{E}\left\{\left(\widehat{\boldsymbol{H}}_{[:k^*]}^{(b/r)}\right)^{\mathrm{H}}\widehat{\boldsymbol{H}}_{[:k^*]}^{(b/r)}\left(\widehat{\boldsymbol{H}}_{[:k^*]}^{(r/u)}\right)^{\mathrm{H}}\boldsymbol{H}_{[:k^*]}^{(r/u)}\left(\boldsymbol{H}_{[:k^*]}^{(r/u)}\right)^{\mathrm{H}}\widehat{\boldsymbol{H}}_{[:k^*]}^{(r/u)}\left(\widehat{\boldsymbol{H}}_{[:k^*]}^{(b/r)}\right)^{\mathrm{H}}\widehat{\boldsymbol{H}}_{[k^*:]}^{(b/r)}\right\}. \quad (47)$$

are given respectively by

$$\lambda^{(b/r)} = \left(\frac{1}{K}\mathrm{Tr}\left\{\nu^2 B_{H_{\mathrm{d}}^{(b/r)}} + \Psi_{H_{\mathrm{r}}^{(b/r)}}\right\}\right)^{-1}, \quad (53)$$

$$\lambda_{k^*}^{(r/u)} = \left(\frac{1}{U_{r^*}}\mathrm{Tr}\left\{\Psi_{H^{(r/u)}}\right\}\right)^{-1}, \quad (54)$$

where $B_{H_{\mathrm{d}}^{(b/r)}} = vec(H_{\mathrm{d}}^{(b/r)})vec(H_{\mathrm{d}}^{(b/r)})^{\mathrm{H}}$. By substituting (51) to (54) into (50), we can calculate $\bar{P}_{\mathrm{S},k^*}$.

The interference plus noise power $\bar{P}_{\mathrm{IN},k}$ is given in (55) in which $\Upsilon_{\mathrm{IN},1}$, $\Upsilon_{\mathrm{IN},2,k}$, $\Upsilon_{\mathrm{IN},3,k,j}$ and $\Upsilon_{\mathrm{IN},4}$ are defined respectively by

$$\bar{P}_{\mathrm{IN},k^*} = a_{k^*}\lambda^{(b/r)}\lambda_{k^*}^{(r/u)} P_{\mathrm{rx},k^*}^{(r,d)} P_{\mathrm{rx},k^*}^{(u,d)}\Upsilon_{\mathrm{IN},1}$$

$$+ \sum_{k=1,j\neq k^*}^{K} a_{k^*}\lambda^{(b/r)}\lambda_k^{(r/u)} P_{\mathrm{rx},k}^{(r,d)} P_{\mathrm{rx},k}^{(u,d)}\Upsilon_{\mathrm{IN},2,k}$$

$$+ \sum_{k=1}^{K}\sum_{j=1,j\neq k}^{K} \delta(\Delta_{j,k})\lambda^{(b/r)}\lambda_k^{(r/u)} P_{\mathrm{rx},k}^{(r,d)}\Upsilon_{\mathrm{IN},3,k,j}$$

$$+ \Upsilon_{\mathrm{IN},4}, \quad (55)$$

$$\Upsilon_{\mathrm{IN},1} = \mathcal{E}\left\{\left|W_{[k^*:]}^{(b/r)}\widetilde{H}_{[:k^*]}^{(b/r)}W_{[k^*:]}^{(r/u)}H_{[:k^*]}^{(r/u)}\right|^2\right\}, \quad (56)$$

$$\Upsilon_{\mathrm{IN},2,k} = \mathcal{E}\left\{\left|W_{[k^*:]}^{(b/r)}H_{[:k]}^{(b/r)}W_{[k:]}^{(r/u)}H_{[:k]}^{(r/u)}\right|^2\right\}, \quad (57)$$

$$\Upsilon_{\mathrm{IN},3,k,j} = \mathcal{E}\left\{\left|W_{[k^*:]}^{(b/r)}H_{[:k]}^{(b/r)}W_{[k:]}^{(r/u)}\sum_{r'=2}^{R}V_{r'[:j]}^{(r/u)}\right|^2\right\}, \quad (58)$$

$$\Upsilon_{\mathrm{IN},4} = \mathcal{E}\left\{\left|\breve{\tilde{n}}_{k^*}\right|^2\right\}. \quad (59)$$

Recalling Lemma 2 of Appendix A, we have $\Upsilon_{\mathrm{IN},1}$ given by (60). Similarly, recalling Lemmas 2 and 3, we can express $\Upsilon_{\mathrm{IN},2,k}$ as given in (61). Furthermore, as shown in Appendix B, $\Upsilon_{\mathrm{IN},3,k,j}$ is given in (62), in which $\Psi_{\Sigma V_{r'[:j]}^{(r/u)}}$ is expressed in (76) of Appendix B, while $\Upsilon_{\mathrm{IN},4}$ can be calculated according to (63). The equations (60) to (63) are all listed at the top of the next page. By substituting (60) to (63) into (55), we can calculate $\bar{P}_{\mathrm{IN},k^*}$.

Finally, substituting (50) and (55) into (45) leads to the ergodic achievable uplink throughput C_{k^*}.

$$\Upsilon_{\mathrm{IN},1} = \left(\left(\mathrm{Tr}\left\{\Psi_{\widehat{H}_{[:k^*]}^{(r/u)}}\right\}\right)^2 + \mathrm{Tr}\left\{\Psi_{\widehat{H}_{[:k^*]}^{(r/u)}}\Psi_{\widetilde{H}_{[:k^*]}^{(r/u)}}\right\}\right)$$

$$\times \mathrm{Tr}\left\{\Psi_{\widetilde{H}_{[:k^*]}^{(b/r)}}\left(\nu^2 B_{H_{\mathrm{d}[:k^*]}^{(b/r)}} + \Psi_{\widehat{H}_{\mathrm{r}[:k^*]}^{(b/r)}}\right)\right\}. \quad (60)$$

$$\Upsilon_{\mathrm{IN},2,k} = \left(\left(\mathrm{Tr}\left\{\Psi_{\widehat{H}_{[:k]}^{(r/u)}}\right\}\right)^2 + \mathrm{Tr}\left\{\Psi_{\widehat{H}_{[:k]}^{(r/u)}}\Psi_{\widetilde{H}_{[:k]}^{(r/u)}}\right\}\right)$$

$$\times \left(\mathrm{Tr}\left\{\left(\nu^2 B_{H_{\mathrm{d}[:k^*]}^{(b/r)}} + \Psi_{\widehat{H}_{\mathrm{r}[:k^*]}^{(b/r)}}\right)\right.\right.$$

$$\times \left.\left.\left(\nu^2 B_{H_{\mathrm{d}[:k]}^{(b/r)}} + \Psi_{H_{\mathrm{r}[:k]}^{(b/r)}}\right)\right\}\right). \quad (61)$$

$$\Upsilon_{\mathrm{IN},3,k,j} = \mathrm{Tr}\left\{\Psi_{\widehat{H}_{[:k]}^{(r/u)}}\Psi_{\Sigma V_{r'[:j]}^{(r/u)}}\right\}$$

$$\times Tr\left\{\left(\nu^2 B_{H_{\mathrm{d}[:k^*]}^{(b/r)}} + \Psi_{\widehat{H}_{\mathrm{r}[:k^*]}^{(b/r)}}\right)\right.$$

$$\times \left.\left(\nu^2 B_{H_{\mathrm{d}[:k]}^{(b/r)}} + \Psi_{H_{\mathrm{r}[:k]}^{(b/r)}}\right)\right\}. \quad (62)$$

$$\Upsilon_{\mathrm{IN},4} = a_{k^*}\lambda^{(b/r)}\lambda_{k^*}^{(r/u)} P_{\mathrm{rx},k^*}^{(r,d)}\sigma_{\widetilde{n}}^2 \mathrm{Tr}\left\{\Psi_{\widehat{H}_{[:k^*]}^{(r/u)}}\right\}$$

$$\times \mathrm{Tr}\left\{\left(\nu^2 B_{H_{\mathrm{d}[:,k^*]}^{(b/r)}} + \Psi_{H_{\mathrm{r}[:k^*]}^{(b/r)}}\right)\left(\nu^2 B_{H_{\mathrm{d}[:,k^*]}^{(b/r)}} + \Psi_{\widehat{H}_{\mathrm{r}[:k^*]}^{(b/r)}}\right)\right\}$$

$$+ \lambda^{(b/r)}\mathrm{Tr}\left\{\nu^2 B_{H_{\mathrm{d}[:,k^*]}^{(b/r)}} + \Psi_{\widehat{H}_{\mathrm{r}[:k^*]}^{(b/r)}}\right\}\sigma_{\breve{n}}^2. \quad (63)$$

IV. OPTIMAL POWER SHARING FOR UDs AND RRUs

The achievable uplink throughput depends on the accuracy of channel estimate, which in turn is dominated by the power allocated to pilots. Again, more power allocated to pilots will result in more accurate channel estimates, which will enhance the achievable throughput. However, the achievable uplink throughput also depends on the power allocated to data transmission. Allocating more power to pilots will in turn reduce the power allocated to data transmission, which will reduce the achievable throughput. Therefore, intuitively, for each UD, there exists an 'optimal' power sharing between its pilots and data transmission. Likewise, for every RRU, there exists an 'optimal' power sharing between its forwarding-TAs' pilots and data forwarding transmission. Hence, in order to maximize the system's uplink throughput, we have to jointly optimize the power sharing between pilot training and data transmission for all the UDs and their host RRUs.

Mathematically, for the k^*th UD, its achievable uplink throughput C_{k^*} is a function of the power sharing factors for all the UDs and all the RRUs in the system, that is, $C_{k^*} = C_{k^*}(\boldsymbol{\eta})$, where

$$\boldsymbol{\eta} = \left[\eta_1^{(r/u)}\eta_2^{(r/u)}\cdots\eta_K^{(r/u)}\eta_1^{(b/r)}\eta_2^{(b/r)}\cdots\eta_K^{(b/r)}\right]^{\mathrm{T}}. \quad (64)$$

The sum-rate of all the K UDs served by the host BBU, which is given by

$$C_{\mathrm{sum}}(\boldsymbol{\eta}) = \sum_{k^*=1}^{K} C_{k^*}(\boldsymbol{\eta}), \quad (65)$$

is a function of the power sharing factor set $\boldsymbol{\eta}$ for all the UDs and all the RRUs. In order to boost the achievable sum-rate, we

can optimize the power sharing factor set $\boldsymbol{\eta}$ by maximizing C_{sum} subject to the constraints of $0 < \eta_k^{(r/u)} < 1$ and $0 < \eta_k^{(b/r)} < 1$ for $1 \leq k \leq K$. Thus, we can formulate the global optimization of the power sharing for the C-RAN uplink as the following optimization problem

$$\begin{aligned} \boldsymbol{\eta}_{\text{opt}} = \;& \max_{\boldsymbol{\eta}} C_{\text{sum}}(\boldsymbol{\eta}), \\ \text{s.t.:} \;\;& 0 < \eta_k^{(r/u)} < 1, 1 \leq k \leq K, \\ & 0 < \eta_k^{(b/r)} < 1, 1 \leq k \leq K. \end{aligned} \quad (66)$$

The above problem is a multivariate optimization problem, where the cost function $C_{\text{sum}}(\boldsymbol{\eta})$ is the summation of K log-functions having $2K$ multivariate factors. Furthermore, the underlying system is a two-layer network, consisting the multiple UDs-to-RRUs links at the first layer and the multiple RRUs-to-BBC links at the second layer. In particular, as observed in (46), there is residual inter-UD interference between the K UDs. Consequently, the K decision variables $\eta_k^{(b/r)}$ for $1 \leq k \leq K$ are the functions of the other K decision variables $\eta_k^{(r/u)}$ for $1 \leq k \leq K$. Hence, the $2K$ optimization decision variables are dependent on each other. It is therefore impossible to derive a closed-form solution to this complex multivariate optimization problem, and a numerical solution must be sought. More specifically, the Karush-Kuhn-Tucker (KKT) conditions associated with the constrained optimization problem (66) are highly complex, from which no closed-form solution can be derived. Additionally, at the time of writing it remains an open question, whether (66) is convex or not.

Solving the optimization problem (66) using numerical optimization algorithms, such as the expectation maximization algorithm [42], the repeated weighted boosting search algorithm [43]–[45] and evolutionary algorithms [46]–[48], will impose considerable computational complexity. Fortunately, in the C-RAN architecture, the BBUs cooperate across the BBU pool, which is an aggregated collective resource shared among a large number of virtual BSs. As a benefit, the system becomes capable of achieving a much improved exploitation of the processing resources, while imposing a reduced power consumption, based on statistical computing multiplexing [49]. Therefore, the C-RAN possesses a powerful centralized signal processing capability for dynamic shared resource allocation [50]. In other words, the C-RAN has the necessary computing power for numerically solving the optimization problem (66).

Since it is unknown, whether the optimization problem (66) is convex or not, we invoke the continuous-variable DEA [51], [52] for finding the globally optimal solution $\boldsymbol{\eta}_{\text{opt}}$. The pseudo-code of the DEA is presented in Algorithm 1, where P_s is the population size, $G_{\max}$ is the pre-set maximum number of generations, $\tilde{\boldsymbol{\eta}}^{(g,p_s)}$ represents a mutated vector, $\breve{\boldsymbol{\eta}}^{(g,p_s)}$ is a trial vector, $\check{\boldsymbol{\eta}}^{(g,p_s)}$ is selected from an elite-archive having high-fitness individuals and finally $\lambda_{p_s} \in (0, 1]$ is a randomly generated scaling factor. Furthermore, A is the total number of elements in vector $\boldsymbol{\eta}^{(g,p_s)}$, $\text{rand}_\alpha(0, 1)$ denotes the random number drawn from the uniform distribution in $[0, 1]$ for the α-th element, while $C_{r_{p_s}} \in [0, 1]$ is the crossover probability.

Algorithm 1: Differential Evolutionary Algorithm.

1:     Set generation index $g = 0$ and randomly generate the initial population of P_s individuals $\boldsymbol{\eta}^{(g,p_s)}, p_s \in \{1, 2, \ldots, P_s\}$.   ▷ Initialization.

2:     **while** $g < G_{\max}$ **do**

3:       **for** $p_s = 1$ to P_s **do**   ▷ Mutation

4:         Create a mutated vector $\tilde{\boldsymbol{\eta}}^{(g,p_s)} = \boldsymbol{\eta}^{(g,p_s)} + \lambda_{p_s}$ $(\check{\boldsymbol{\eta}}^{(g,p_s)} - \boldsymbol{\eta}^{(g,p_s)}) + \lambda_{p_s} (\boldsymbol{\eta}^{(g,p_s')} - \boldsymbol{\eta}^{(g,p_s'')})$.

5:       **for** $p_s = 1$ to P_s **do**   ▷ Crossover

6:         **for** $\alpha = 1$ to A **do**

7:           Crossover the αth element of $\boldsymbol{\eta}^{(g,p_s)}$ by

$$\breve{\eta}_\alpha^{(g,p_s)} = \begin{cases} \tilde{\eta}_\alpha^{(g,p_s)}, & \text{rand}_\alpha(0, 1) \leq C_{r_{p_s}}, \\ \eta_\alpha^{(g,p_s)}, & \text{otherwise}, \end{cases}$$

8:       **for** $p = 1$ to P_s **do**   ▷ Selection

9:         Select $\boldsymbol{\eta}^{(g,p_s)}$ or $\breve{\boldsymbol{\eta}}^{(g,p_s)}$ surviving to the next generation

$$\boldsymbol{\eta}^{(g+1,p_s)} = \begin{cases} \breve{\boldsymbol{\eta}}^{(g,p_s)}, & C_{\text{sum}}(\boldsymbol{\eta}^{(g,p_s)}) \leq C_{\text{sum}}(\breve{\boldsymbol{\eta}}^{(g,p_s)}), \\ \boldsymbol{\eta}^{(g,p_s)}, & \text{otherwise}. \end{cases}$$

Additionally, the symbol ▷ denotes a comment. The detailed characteristics of this DEA can be found in [46]–[48].

V. SIMULATION STUDY

In the simulated massive MIMO aided C-RAN, the coverage of the BBU is a square area of $[-500\sqrt{2}, 500\sqrt{2}] \times [-500\sqrt{2}, 500\sqrt{2}]$ square meters (m²), and the BBU is located at the center of the square. $R = 4$ RRUs are deployed at the coordinates of $(250\sqrt{2}, 250\sqrt{2})$ m, $(250\sqrt{2}, -250\sqrt{2})$ m, $(-250\sqrt{2}, -250\sqrt{2})$ m, and $(-250\sqrt{2}, 250\sqrt{2})$ m, respectively, by default. Thus, the distance between a RRU and the BBU is 500 m by default. K UDs are independently and uniformly distributed in this square area. Each UD is equipped with a single antenna, and each RRU is equipped with $M = 32$ RAs, whilst the BBU employs $N = 128$ RAs. Naturally, the number of TAs at each RRU is an integer. For reasons of fairness, U_r for $1 \leq r \leq R$ should be similar and they are all close to K/R subject to the constraint $\sum_{r=1}^{R} U_r = K$. Consider for example the default system associated with a total number of $K = 10$ UDs, which are served by $R = 4$ RRUs. A possible choice is $(U_1, U_2, U_3, U_4) = (2, 2, 3, 3)$, which is used in the simulations. The mth-row and nth-column element of the spatial correlation matrix $\boldsymbol{R}_{\text{rx},r}^{(r/u)}$ of the rth RRU's RAs is given by $\rho^{|m-n|}$, where ρ is the spatial correlation factor of the RAs. Similarly, the mth-row and nth-column element of the spatial correlation matrix $\boldsymbol{R}_{\text{rx},b}^{(b/r)}$ of the BBU's RAs is also specified by $\rho^{|m-n|}$. The system's bandwidth is 10 MHz, and the noise power spectral density is -174 dBm/Hz. The power allocated to a single UD is $P_{\text{UD}} = 0.2$ Watts, and the power allocated to a single RRU is $P_{\text{RRU}} = 10$ Watts. The default system parameters of this simulated massive MIMO aided C-RAN are listed in

TABLE I
DEFAULT SYSTEM PARAMETERS OF THE MASSIVE MIMO AIDED C-RAN

Bandwidth	10 MHz
Number of UDs K	10
Number of RRUs R	4
Number of RAs at BBU N	128
Number of RAs at RRU M	32
Spatial correlation factor of RAs ρ	0.1
Rician factor K_{Rice}	10 dB
Receiver efficiency factor ρ_r	0.1
Noise power spectral density	-174 dBm/Hz
Total power of a single UD P_{UD}	0.2 Watts
Total power of a single RRU P_{RRU}	10 Watts

Table I. Without optimization, the power sharing factors $\eta_k^{(r/u)}$ and $\eta_k^{(b/r)}$ are all set to 0.5, for $k \in \{1, 2, \ldots, K\}$, i.e., half power for pilot training and half for data transmission/forwarding.

It is well-known that as an efficient global optimization algorithm, the DEA is capable of converging fast to an optimal solution of a complex optimization problem. The efficiency, reliability, and convergence characteristics of the DEA have been extensively investigated and documented in [46]–[48]. Hence, we focus on investigating the influence of the system's key parameters on the achievable sum-rate. These key parameters are the number of RAs at a RRU M, the number of RAs at the BBU N, the number of UDs K being served simultaneously, the number of RRUs R, the correlation factor of RAs ρ, and the Rician factor K_{Rice}. In the following, 'Without power optimization' indicates the sum-rate calculated using (65) with the non-optimized power sharing factors for UDs and RRUs, i.e., $\eta_k^{(r/u)} = 0.5$ and $\eta_k^{(b/r)} = 0.5$ for $1 \leq k \leq K$, whilst 'With power optimization' is the sum-rate using (65) with the globally optimized power sharing factors obtained by solving the optimization problem (66) using the DEA. Furthermore, 'Theoretical' indicates the sum-rate calculated using (65), while 'Simulation' is the Monte-Carlo simulation result obtained by averaging over 400 random realizations.

Fig. 2 depicts the achievable sum-rate as the function of the number of RAs M deployed at each RRU. It is widely recognized that for 'single-layer' massive MIMO systems associated with Rayleigh channel matrices, the sum rate is a smoothly increasing function of the number of receiver antennas. This smooth shape obeys the logarithmic function of the system's SINR improvement upon increasing receiver antennas. Our massive MIMO aided C-RAN is a 'two-layer' network with the first layer being Rayleigh and the second layer being Rician distributed. In the scenario of Fig. 2, the second layer is fixed, hence we can express the sum rate of (65) equivalently as

$$C(M) = \sum_{k=1}^{K} \log_2 \left(1 + \text{SINR}_k(M)\right), \qquad (67)$$

where $\text{SINR}_k(M)$ is the kth UD's SINR for this network, while M is the number of RAs at each RRU. Changing M changes the Rayleigh channel matrices of the first layer. Since this 'equivalent' system is Rayleigh, $C(M)$ is an increasing function of M
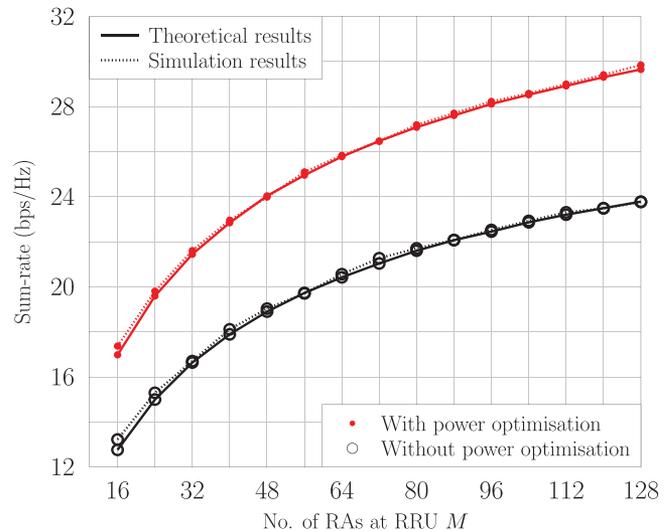

Fig. 2. The sum-rate as a function of the number of RRU's RAs M. The rest of the system parameters are as given in Table I.
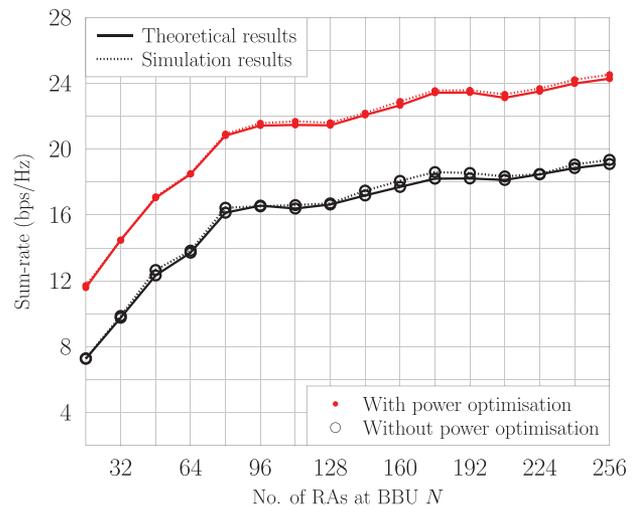

Fig. 3. The sum-rate as a function of the number of BBU's RAs N. The rest of the system parameters are as given in Table I.

having a smooth shape determined by the logarithmic function of the system's SINR. This is confirmed by the results of Fig. 2. It can also be seen from Fig. 2 that the achievable sum-rate associated with power sharing optimization is much higher than that without power sharing optimization. Specifically, the sum-rate gain achieved by the proposed globally optimal power sharing is about 33%, which increases from 4.2 bps/Hz to 5.9 bps/Hz, when M is increased from 16 to 128.

Fig. 3 portrays the achievable sum-rate as a function of the number of RAs N deployed at the BBU. It can be seen that the massive MIMO aided C-RAN associated with the proposed global optimization of power sharing is capable of offering a sum-rate gain from 4.3 bps/Hz to 5.2 bps/Hz, when N increases from 16 to 256, compared to the same system operating without power sharing optimization. Observe from Fig. 3 that the sum rate is an increasing function of N, but its shape is not very
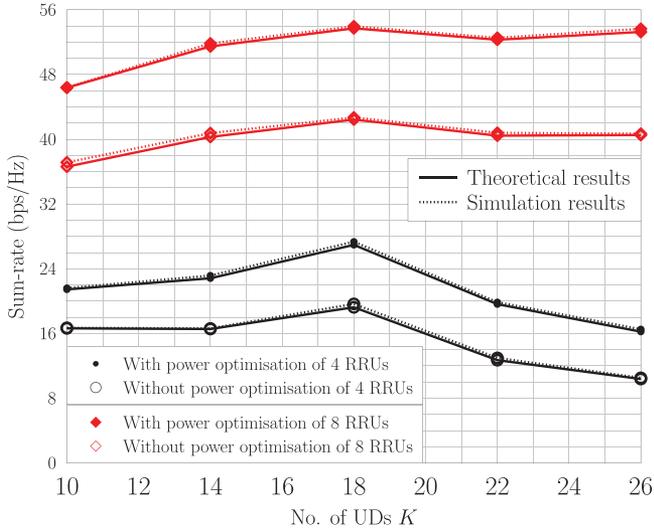
Fig. 4. The sum-rates of the 4-RRU and 8-RRU systems as a functions of the number of UDs K. The rest of the system parameters are as given in Table I.

Fig. 5. The sum-rate as a function of the spatial correlation factor ρ of RAs. The rest of the system parameters are as given in Table I.

smooth. This phenomenon may be attributed to the LoS path of the Rician channel. In the scenario of Fig. 3, the first layer is fixed, and the sum rate of (65) can be equivalently expressed as

$$C(N) = \sum_{k=1}^{K} \log_2 \left(1 + \text{SINR}_k(N)\right), \qquad (68)$$

where $\text{SINR}_k(N)$ represents the kth UD's SINR for this network, while N is the number of RAs at the BBC. Note that in contrast to the scenario of Fig. 2, this 'equivalent' system is Rician. Our experience with 'single-layer' massive MIMO scenarios associated with Rician channel matrices [33], [34] shows that the sum rate is an increasing function of the number of RAs but it exhibits slight undulations, similar to those seen in Fig. 3.

The joint impact of both the number of RRUs and the number of UDs is investigated next. We consider the two systems, the default system having $R = 4$ RRUs and the system associated with $R = 8$ RRUs. In the latter case, we divide the coverage area of the BBU into 8 equal-size subareas and place a RRU at the center of each subarea. Fig. 4 compares the achievable sum-rates of the 4-RRU and 8-RRU systems as the functions of K. As expected, the sum-rate of the 8-RRU system is considerably higher than that of the 4-RRU system. Furthermore, higher sum-rate gain is attained by the proposed global optimization for the 8-RRU system, than for the 4-RRU system Additionally, it is seen from Fig. 4 that the sum-rate increases as the number of UDs increases for small K but for large K, the sum-rate becomes decreasing as K increases further. This is because for large K, user interference becomes dominant, which reduces the achievable throughput.

Fig. 5 depicts the impact of the spatial correlation factor ρ of RAs on the achievable sum-rate. Note that we assume in the simulations that both the RAs of the RRU and the RAs of the BBU have the same spatial correlation factor ρ. Not surprisingly, the achievable sum-rate associated with the proposed
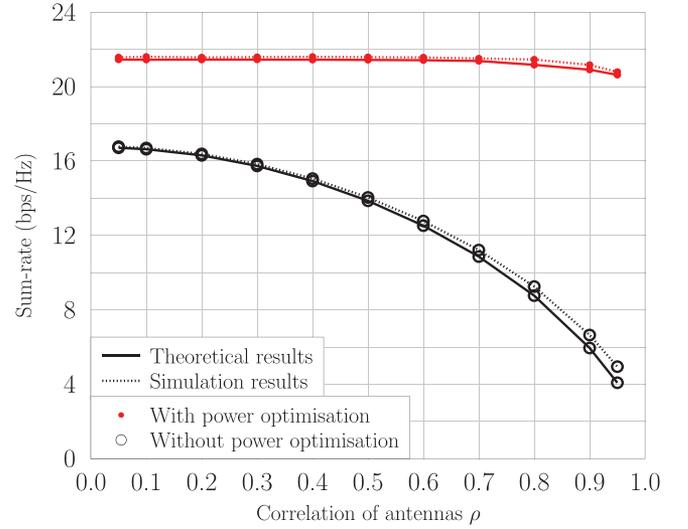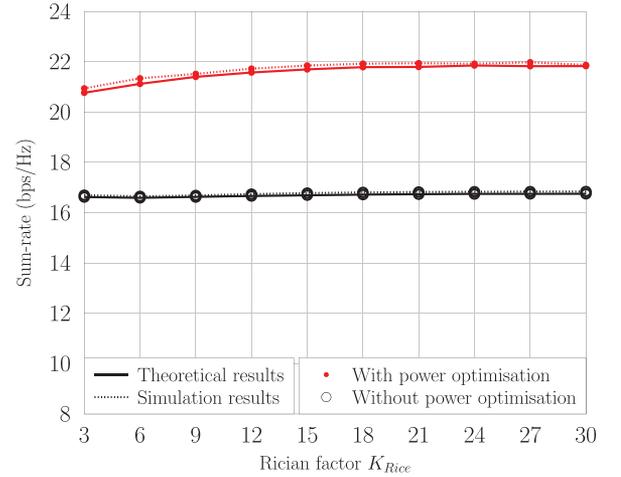

Fig. 6. The sum-rate as a function of the Rician factor K_{Rice}. The rest of the system parameters are as given in Table I.

global optimization of power sharing is significantly higher than that without power sharing optimization. Most strikingly, the spatial correlation factor ρ has a considerable negative impact on the achievable sum-rate of the system operating without power sharing optimization. By contrast, the spatial correlation factor ρ hardly impacts on the system that implements the proposed global optimization of power sharing, except for very large $\rho \geq 0.8$. This demonstrates that the optimization of power sharing is capable of mitigating the detrimental impact of the spatial correlation.

The impact of the Rician factor K_{Rice} on the achievable sum rate is investigated in Fig. 6. It can be seen from Fig. 6 that the system relying on globally optimal power sharing is capable of achieving around 5.1 bps/Hz of extra sum-rate, compared to the system dispensing with power sharing optimization. Changing K_{Rice} changes slightly the RRUs-BBU Rician channel links, which however only constitute a certain part of the

overall system. It is therefore expected that the impact of K_{Rice} on the achievable sum rate will not be dramatic. Indeed, it can be seen from Fig. 6 that K_{Rice} hardly has an impact on the sum rate of the system operating without power sharing optimization. By contrast, for the system relying on power allocation optimization, the sum rate increases slightly, as K_{Rice} increases.

VI. Conclusions

In order to boost the fronthaul capacity of massive MIMO aided C-RAN, we have proposed to globally optimize the powers shared between the pilots and data transmission both for the RRUs and UDs. In particular, we have derived an asymptotic closed-form expression of the achievable sum-rate as a function of the UDs' power sharing factors and the RRUs' power sharing factors. Based on this closed-form sum-rate expression, we have formulated the global optimal power sharing problem. Furthermore, by exploiting the central computing and control capacity of the C-RAN architecture, we have proposed to use the powerful DEA to solve this challenging optimization. Our extensive simulation results have demonstrated that the proposed global optimization of power sharing is capable of dramatically boosting the fronthaul capacity of massive MIMO aided C-RAN. Specifically, at the configuration of 128 RAs at the BBU, the sum-rate of 10 UDs achieved with the optimal power sharing factors improves about 4.2 bps/Hz to 5.9 bps/Hz, when the number of RAs M deployed in RRUs changes from 16 to 128. This represents the improvement of 33% to 25%, compared to the sum-rate obtained without optimizing power sharing factors. The influence of the key system parameters on the system's achievable sum-rate has also been extensively investigated. This study therefore has offered valuable insight and guidance concerning the practical deployment of massive MIMO aided C-RAN, particularly, on how to boost its fronthaul capacity.

Appendix

A. Gallery of Lemmas

Lemma 1: Let X_i and Y_j be nonnegative random variables for $1 \leq i \leq M$ and $1 \leq j \leq N$. Define $X = \sum_{i=1}^{M} X_i$ and $Y = \sum_{i=1}^{N} Y_i$. Then, we have the following approximation of the expectation of $\log_2(1 + \frac{X}{Y})$ [53]:

$$\mathcal{E}\left\{\log_2\left(1 + \frac{X}{Y}\right)\right\} \xrightarrow[N,M\to\infty]{\text{a.s}} \log_2\left(1 + \frac{\mathcal{E}\{X\}}{\mathcal{E}\{Y\}}\right). \quad (69)$$

Lemma 2 ([54, Lemma 1]): Let $\boldsymbol{x} \in \mathbb{C}^N$ and $\boldsymbol{y} \in \mathbb{C}^N$ be two independent random vectors, both having the distribution $\mathcal{CN}(\boldsymbol{0}_N, c\boldsymbol{I}_N)$. Then, we have

$$\frac{\boldsymbol{x}^{\mathrm{H}}\boldsymbol{y}}{N} \xrightarrow[N\to\infty]{\text{a.s}} 0, \quad (70)$$

$$\frac{\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}}{N} \xrightarrow[N\to\infty]{\text{a.s}} c. \quad (71)$$

Lemma 3 (Lemma 12 in [55]): Let $\boldsymbol{A} \in \mathbb{C}^{N \times N}$ and $\boldsymbol{x} \sim \mathcal{CN}(\boldsymbol{0}_N, \frac{1}{N}\boldsymbol{I}_N)$. $\boldsymbol{A}$ has uniformly bounded spectral norm with respect to N and it is independent of $\boldsymbol{x}$. Then, we have

$$\mathcal{E}\left\{\left|(\boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x})^2 - \left(\frac{1}{N}\mathrm{Tr}\{\boldsymbol{A}\}\right)^2\right|\right\} \xrightarrow[N\to\infty]{\text{a.s}} 0. \quad (72)$$

B. Derivation of $\Upsilon_{\mathrm{IN},3,k,j}$

We can expand $\Upsilon_{\mathrm{IN},3,k,j}$ as given in (73). Recalling Lemma 2 of Appendix A, we have (74). Furthermore, we have (75). Assuming that the jth UD is served by the r'th RRU while it is interfering the kth UD served by the r^*th RRU, then $\boldsymbol{\Psi}_{\Sigma\boldsymbol{V}_{r'[:j]}^{(r/u)}}$ is given in (76). The equations (73) to (76) are given at the bottom of the this page.

$$\Upsilon_{\mathrm{IN},3,k,j} = \mathrm{Tr}\left\{\mathcal{E}\left\{\sum_{r'=2}^{R}\boldsymbol{V}_{r'[:j]}^{(r/u)}\left(\sum_{r'=2}^{R}\boldsymbol{V}_{r'[:j]}^{(r/u)}\right)^{\mathrm{H}}\left(\boldsymbol{W}_{[k:]}^{(r/u)}\right)^{\mathrm{H}}\boldsymbol{W}_{[k:]}^{(r/u)}\right\}\right\}\mathrm{Tr}\left\{\mathcal{E}\left\{\left(\boldsymbol{W}_{[k^*:]}^{(b/r)}\right)^{\mathrm{H}}\boldsymbol{W}_{[k^*:]}^{(b/r)}\boldsymbol{H}_{[:k]}^{(b/r)}\left(\boldsymbol{H}_{[:k]}^{(b/r)}\right)^{\mathrm{H}}\right\}\right\}. \quad (73)$$

$$\mathrm{Tr}\left\{\mathcal{E}\left\{\left(\boldsymbol{W}_{[k^*:]}^{(b/r)}\right)^{\mathrm{H}}\boldsymbol{W}_{[k^*:]}^{(b/r)}\boldsymbol{H}_{[:k]}^{(b/r)}\left(\boldsymbol{H}_{[:k]}^{(b/r)}\right)^{\mathrm{H}}\right\}\right\} = \mathrm{Tr}\left\{\left(\nu^2\boldsymbol{B}_{\boldsymbol{H}_{\mathrm{d}[:k^*]}^{(b/r)}} + \boldsymbol{\Psi}_{\widehat{\boldsymbol{H}}_{\mathrm{r}[:k^*]}^{(b/r)}}\right)\left(\nu^2\boldsymbol{B}_{\boldsymbol{H}_{\mathrm{d}[:k]}^{(b/r)}} + \boldsymbol{\Psi}_{\boldsymbol{H}_{\mathrm{r}[:k]}^{(b/r)}}\right)\right\}. \quad (74)$$

$$\mathrm{Tr}\left\{\mathcal{E}\left\{\sum_{r'=2}^{R}\boldsymbol{V}_{r'[:j]}^{(r/u)}\left(\sum_{r'=2}^{R}\boldsymbol{V}_{r'[:j]}^{(r/u)}\right)^{\mathrm{H}}\left(\boldsymbol{W}_{[k:]}^{(r/u)}\right)^{\mathrm{H}}\boldsymbol{W}_{[k:]}^{(r/u)}\right\}\right\} = \mathrm{Tr}\left\{\boldsymbol{\Psi}_{\widehat{\boldsymbol{H}}_{[:k]}^{(r/u)}}\boldsymbol{\Psi}_{\Sigma\boldsymbol{V}_{r'[:j]}^{(r/u)}}\right\}. \quad (75)$$

$$\boldsymbol{\Psi}_{\Sigma\boldsymbol{V}_{r'[:j]}^{(r/u)}} = \mathcal{E}\left\{\sum_{r'=2}^{R}\boldsymbol{V}_{r'[:j]}^{(r/u)}\left(\sum_{r'=2}^{R}\boldsymbol{V}_{r'[:j]}^{(r/u)}\right)^{\mathrm{H}}\right\}$$

$$= \mathrm{Bdiag}\left\{\boldsymbol{0}_{M\times M}, \ldots, \boldsymbol{0}_{M\times M}, \underbrace{\sum_{r'=2}^{R}P_{\mathrm{rx},r^*/\left((r'+r^*-1)\bmod R\right)(j)}^{(u,d)}\boldsymbol{R}_{\mathrm{rx},r^*}^{(r/u)}}_{r^*\text{th block matrix}}, \boldsymbol{0}_{M\times M}, \ldots, \boldsymbol{0}_{M\times M}\right\}. \quad (76)$$

REFERENCES

[1] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun, "Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3029–3056, Sep. 2015.

[2] Statista, "5G—Statistics & Facts." [Online]. Available: https://www.statista.com/topics/3447/5g/. Accessed on: 2018.

[3] V. Jungnickel *et al.*, "The role of small cells, coordinated multipoint, and massive MIMO in 5G," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 44–51, May 2014.

[4] 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.

[5] M. Xiao *et al.*, "Millimeter wave communications for future mobile networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1909–1935, Sep. 2017.

[6] P. Rost *et al.*, "Cloud technologies for flexible 5G radio access networks," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 68–76, May 2014.

[7] J. Bartelt *et al.*, "Fronthaul and backhaul requirements of flexibly centralized radio access networks," *IEEE Wireless Commun.*, vol. 22, no. 5, pp. 105–111, Oct. 2015.

[8] M. Peng, Y. Wang, T. Dang, and Z. Yan, "Cost-efficient resource allocation in cloud radio access networks with heterogeneous fronthaul expenditures," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4626–4638, Jul. 2017.

[9] Y. Zhou and W. Yu, "Fronthaul compression and transmit beamforming optimization for multi-antenna uplink C-RAN," *IEEE Trans. Signal Process.*, vol. 64, no. 16, pp. 4138–4151, Aug. 2016.

[10] W. Lee, O. Simeone, J. Kang, and S. Shamai, "Multivariate fronthaul quantization for downlink C-RAN," *IEEE Trans. Signal Process.*, vol. 64, no. 19, pp. 5025–5037, Oct. 2016.

[11] T. X. Vu, H. D. Nguyen, T. Q. S. Quek, and S. Sun, "Adaptive cloud radio access networks: Compression and optimization," *IEEE Trans. Signal Process.*, vol. 65, no. 1, pp. 228–241, Jan. 2017.

[12] Z. Zhao *et al.*, "Cluster content caching: An energy-efficient approach to improve quality of service in cloud radio access networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 5, pp. 1207–1221, May 2016.

[13] Y. L. Lee, J. Loo, T. C. Chuah, and L.-C. Wang, "Dynamic network slicing for multitenant heterogeneous cloud radio access networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 4, pp. 2146–2161, Apr. 2018.

[14] J. Liu, A. Liu, and V. K. N. Lau, "Compressive interference mitigation and data recovery in cloud radio access networks with limited fronthaul," *IEEE Trans. Signal Process.*, vol. 65, no. 6, pp. 1437–1446, Mar. 2017.

[15] W. Hao, O. Muta, and H. Gacanin, "Price-based resource allocation in massive MIMO H-CRANs with limited fronthaul capacity," *IEEE Trans. Wireless Commun.*, vol. 17, no. 11, pp. 7691–7703, Nov. 2018.

[16] C. Pan, H. Zhu, N. J. Gomes, and J. Wang, "Joint user selection and energy minimization for ultra-dense multi-channel C-RAN with incomplete CSI," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 8, pp. 1809–1824, Aug. 2017.

[17] H. Ren, N. Liu, C. Pan, and L. Hanzo, "Joint fronthaul link selection and transmit precoding for energy efficiency maximization of multiuser MIMO-aided distributed antenna systems," *IEEE Trans. Commun.*, vol. 65, no. 12, pp. 5180–5196, Dec. 2017.

[18] P. Luong, F. Gagnon, C. Despins, and L.-N. Tran, "Optimal joint remote radio head selection and beamforming design for limited fronthaul C-RAN," *IEEE Trans. Signal Process.*, vol. 65, no. 21, pp. 5605–5620, Nov. 2017.

[19] L. Liu, S. Bi, and R. Zhang, "Joint power control and fronthaul rate allocation for throughput maximization in OFDMA-based cloud radio access network," *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 4097–4110, Nov. 2015.

[20] M. Peng, C. Wang, L. Vincent, and H. V. Poor, "Fronthaul-constrained cloud radio access networks: Insights and challenges," *IEEE Wireless Commun.*, vol. 22, no. 2, pp. 152–160, Apr. 2015.

[21] W. Hao and S. Yang, "Small cell cluster-based resource allocation for wireless backhaul in two-tier heterogeneous networks with massive MIMO," *IEEE Trans. Veh. Technol.*, vol. 67, no. 1, pp. 509–523, Jan. 2018.

[22] A. Macho, M. Morant, and R. Llorente, "Next-generation optical fronthaul systems using multicore fiber media," *J. Lightw. Technol.*, vol. 34, no. 20, pp. 4819–4827, Oct. 2016.

[23] H. Zhang *et al.*, "Fronthauling for 5G LTE-U ultra dense cloud small cell networks," *IEEE Wireless Commun.*, vol. 23, no. 6, pp. 48–53, Dec. 2016.

[24] I. A. Alimi, A. L. Teixeira, and P. P. Monteiro, "Toward an efficient C-RAN optical fronthaul for the future networks: A tutorial on technologies, requirements, challenges, and solutions," *IEEE Commun. Surv. Tut.*, vol. 20, no. 1, pp. 708–769, First Quarter 2018.

[25] M. Jaber, M. A. Imran, R. Tafazolli, and A. Tukmanov, "5G backhaul challenges and emerging research directions: A survey," *IEEE Access*, vol. 4, pp. 1743–1766, 2016.

[26] U. Siddique, H. Tabassum, E. Hossain, and D. I. Kim, "Wireless backhauling of 5G small cells: Challenges and solution approaches," *IEEE Wireless Commun.*, vol. 22, no. 5, pp. 22–31, Oct. 2015.

[27] U. Habiba, H. Tabassum, and E. Hossain, "Backhauling 5G small cells with massive-MIMO-enabled mmWave communication," in *Backhauling/Fronthauling for Future Wireless Systems*, K. M. Saidul Huq and J. Rodriguez, Eds. Hoboken, NJ, USA: Wiley, 2016, pp. 29–53.

[28] M. Jaber *et al.*, "Wireless backhaul: Performance modelling and impact on user association for 5G," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 3095–3110, May 2018.

[29] B. Hu, C. Hua, C. Chen, and X. Guan, "Joint beamformer design for wireless fronthaul and access links in C-RANs," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 2869–2881, May 2018.

[30] M. A. Imran, S. A. R. Zaidi, and M. Z. Shakir, Eds., *Access, Fronthaul and Backhaul Networks for 5G & Beyond* (IET Telecommunications Series). London, U.K.: Inst. Eng. Technol., 2017.

[31] S. Park, C.-B. Chae, and S. Bahk, "Large-scale antenna operation in heterogeneous cloud radio access networks: A partial centralization approach," *IEEE Wireless Commun.*, vol. 22, no. 3, pp. 32–40, Jun. 2015.

[32] S. Parsaeefard *et al.*, "Dynamic resource allocation for virtualized wireless networks in massive-MIMO-aided and fronthaul-limited C-RAN," *IEEE Trans. Veh. Technol.*, vol. 66, no. 10, pp. 9512–9520, Oct. 2017.

[33] J. Zhang *et al.*, "Adaptive coding and modulation for large-scale antenna array-based aeronautical communications in the presence of co-channel interference," *IEEE Trans. Wireless Commun.*, vol. 17, no. 2, pp. 1343–1357, Feb. 2018.

[34] J. Zhang *et al.*, "Regularized zero-forcing precoding-aided adaptive coding and modulation for large-scale antenna array-based air-to-air communications," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 2087–2103, Sep. 2018.

[35] J. D. Parsons, *The Mobile Radio Propagation Channel*, 2nd ed. Chichester, U.K.: Wiley, 2000.

[36] Ofcom, "Update on 5G spectrum in the UK." Feb. 2017. [Online]. Available: https://www.ofcom.org.uk/__data/assets/pdf_file/0021/97023/5G-update-08022017.pdf

[37] METIS2020, "METIS Channel Model," Deliverable D1.4 v3, Jul. 2015. [Online]. Available: https://www.metis2020.com/wp-content/uploads/deliverables/METIS D1.4 v3.0.pdf

[38] T. S. Rappaport *et al.*, "Overview of millimeter wave communications for fifth-generation (5G) wireless networks with a focus on propagation models," *IEEE Trans. Antennas Propag.*, vol. 65, no. 12, pp. 6213–6230, Dec. 2017.

[39] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 2003.

[40] J. Zhang, *et al.*, "Pilot contamination elimination for large-scale multiple antenna aided OFDM systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 759–772, Oct. 2014.

[41] X. Guo, *et al.*, "Optimal pilot design for pilot contamination elimination/reduction in large-scale multiple-antenna aided OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7229–7243, Nov. 2016.

[42] J. Zhang, L. Hanzo, and X. Mu, "Joint decision-directed channel and noise-variance estimation for MIMO OFDM/SDMA systems based on expectation-conditional maximization," *IEEE Trans. Commun.*, vol. 60, no. 5, pp. 2139–2151, Jun. 2011.

[43] S. Chen, X. Wang, and C. J. Harris, "Experiments with repeating weighted boosting search for optimization signal processing applications," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 35, no. 4, pp. 682–693, Aug. 2005.

[44] J. Zhang, S. Chen, X. Mu, and L. Hanzo, "Joint channel estimation and multiuser detection for SDMA/OFDM based on dual repeated weighted boosting search," *IEEE Trans. Veh. Technol.*, vol. 60, no. 7, pp. 3265–3275, Jul. 2011.

[45] S F. Page, S. Chen, C. J. Harris, and N. M. White, "Repeated weighted boosting search for discrete or mixed search space and multiple-objective optimisation," *Appl. Soft Comput.*, vol. 12, no. 9, pp. 2740–2755, Sep. 2012.

[46] J. Zhang, S. Chen, X. Mu, and L. Hanzo, "Benchmarking capabilities of evolutionary algorithms in joint channel estimation and turbo multiuser detection/decoding," in *Proc. IEEE Congr. Evol. Comput.*, Cancun, Mexico, Jun. 20–23, 2013, pp. 3354–3362.

[47] J. Zhang, S. Chen, X. Mu, and L. Hanzo, "Evolutionary algorithm assisted joint channel estimation and turbo multi-user detection/decoding for OFDM/SDMA," *IEEE Trans. Veh. Technol.*, vol. 63, no. 3, pp. 1204–1222, Mar. 2014.

[48] J. Zhang *et al.*, "Differential evolutionary algorithm aided turbo channel estimation and multi-user detection for G.Fast systems in the presence of FEXT," *IEEE Access*, vol. 6, pp. 33111–33128, 2018.

[49] D. Pompili, A. Hajisami, and H. Viswanathan, "Dynamic provisioning and allocation in cloud radio access networks (C-RANs)," *Ad Hoc Netw.*, vol. 30, pp. 128–143, Jul. 2015.

[50] C. Pan, H. Zhu, N. J. Gomes, and J. Wang, "Joint precoding and RRH selection for user-centric green MIMO C-RAN," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2891–2906, May 2017.

[51] K. Price, R. M. Storn, and J. A. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization*. Berlin, Germany: Springer-Verlag, 2005.

[52] A. Qin, V. L. Huang, and P. N. Suganthan, "Differential evolution algorithm with strategy adaptation for global numerical optimization," *IEEE Trans. Evol. Comput.*, vol. 13, no. 2, pp. 398–417, Apr. 2009.

[53] Q. Zhang *et al.*, "Power scaling of uplink massive MIMO systems with arbitrary-rank channel means," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 966–981, Oct. 2014.

[54] F. Fernandes, A. Ashikhmin, and T. L. Marzetta, "Inter-cell interference in noncooperative TDD large scale antenna systems" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 192–201, Feb. 2013.

[55] J. Hoydis, "Random matrix theory for advanced communication systems," Ph.D. dissertation, Dept. Supélec, Gif-sur-Yvette, France, 2012.

Jiankang Zhang (S'08–M'12–SM'18) received the B.Sc. degree in mathematics and applied mathematics from the Beijing University of Posts and Telecommunications, Beijing, China, in 2006, and the Ph.D. degree in communication and information systems from Zhengzhou University, Zhengzhou, China, in 2012. He was a Lecturer from 2012 to 2013 and an Associate Professor from 2013 to 2014 with the School of Information Engineering, Zhengzhou University. From 2009 to 2011, he was a Visiting Ph.D. student with the School of Electronics and Computer Science, the University of Southampton, Southampton, U.K. From 2013 to 2014, he was a Postdoctoral Researcher with McGill University, Montreal, QC, Canada. He is currently a Senior Research Fellow with the University of Southampton. His research interests are in the areas of wireless communications and signal processing, aeronautical communications, and broadband communications. He is a recipient of a number of academic awards, including the Excellent Doctoral Dissertation of Henan Province, China, the Youth Science and Technology Award of Henan Province, China, and the Innovative Talent of Colleges and Universities of Henan Province, China. He is an Associate Editor for the IEEE ACCESS.
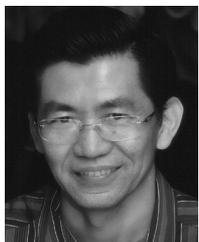
Sheng Chen (M'90–SM'97–F'08) received the B.Eng. degree from East China Petroleum Institute, Dongying, China, in 1982, and the Ph.D. degree from the City University of London, London, U.K., in 1986, both in control engineering, and the Doctor of Sciences (D.Sc.) degree from the University of Southampton, Southampton, U.K., in 2005. From 1986 to 1999, he held research and academic appointments with the University of Sheffield, University of Edinburgh, and University of Portsmouth, all in the U.K. Since 1999, he has been with the School of Electronics and Computer Science, the University of Southampton, where he is a Professor of Intelligent Systems and Signal Processing. He is a Distinguished Adjunct Professor with King Abdulaziz University, Jeddah, Saudi Arabia, and an ISI highly cited researcher in engineering (March 2004). He has 11 400 Web of Science citations and more than 24 000 Google Scholar citations. He has authored or coauthored more than 600 research papers. His research interests include adaptive signal processing, wireless communications, modeling and identification of nonlinear systems, neural network and machine learning, intelligent control system design, evolutionary computation methods, and optimization. He is a fellow of the United Kingdom Royal Academy of Engineering and the Institution of Engineering and Technology.

Xinying Guo received the B.Sc. degree in electronic and information engineering from the North University of China, Taiyuan, China, in 2011, and the Ph.D. degree in information and communication engineering from Zhengzhou University, Zhengzhou, China, in 2017. She is currently a Lecturer with the School of Information Science and Engineering, Henan University of Technology, Zhengzhou. From 2014 to 2016, she was a Visiting Researcher in electronics and computer science with the University of Southampton, Southampton, U.K. Her research interests are in the areas of wireless communications and signal processing, including channel estimation, multiuser detection, beamforming/precoding, and pilot contamination elimination.

Jia Shi received the M.Sc. and Ph.D. degrees from the University of Southampton, Southampton, U.K., in 2010 and 2015, respectively. He was a Research Associate with Lancaster University, Lancaster, U.K., during 2015–2017. He was a Research Fellow with the 5G Innovation Centre, University of Surrey, Guildford, U.K., from 2017 to 2018. In 2018, he joined Xidian University, Xi'an, China, where he is currently an Associate Professor with the State Key Laboratory of Integrated Services Networks. His current research interests include millimeter-wave communications, resource allocation in wireless systems, multicarrier communications, high secure communications, physical layer security, and cooperative communication.

Lajos Hanzo (M'91–SM'92–F'04) received the degree in electronics in 1976 and the doctorate degree in 1983. In 2009, he was awarded an honorary doctorate by the Technical University of Budapest, Budapest, Hungary, and in 2015 by the University of Edinburgh, Edinburgh, U.K. In 2016, he joined the Hungarian Academy of Science. During his 40-year career in telecommunications, he has held various research and academic posts in Hungary, Germany, and the U.K. Since 1986, he has been with the School of Electronics and Computer Science, University of Southampton, Southampton, U.K., where he holds the Chair in telecommunications. He has successfully supervised 112 Ph.D. students, co-authored 18 Wiley/IEEE Press books on mobile radio communications totalling in excess of 10 000 pages, published more than 1800 research contributions at IEEE Xplore, acted both as the Technical Program Committee Chair and the General Chair of IEEE conferences, presented keynote lectures, and has been awarded a number of distinctions. He is directing a 60-strong academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry, the Engineering and Physical Sciences Research Council, U.K., the European Research Council's Advanced Fellow Grant, and the Royal Society's Wolfson Research Merit Award. He is an enthusiastic supporter of industrial and academic liaison, and he offers a range of industrial courses. He is also a Governor of the IEEE Communications Society and the IEEE Vehicular Technology Society. During 2008–2012, he was the Editor-in-Chief of the IEEE Press and a Chaired Professor with Tsinghua University, Beijing, China. He is a Fellow of the Royal Academy of Engineering, the Institution of Engineering and Technology, and the European Association for Signal Processing.