The Non-coherent Ultra-Dense C-RAN is Capable of Outperforming its Coherent Counterpart at a Limited Fronthaul Capacity

Cunhua Pan, Hong Ren, Maged Elkashlan, Arumugam Nallanathan, *Fellow, IEEE*, and Lajos Hanzo, *Fellow, IEEE*

***Abstract*—The weighted sum rate maximization problem of ultra-dense cloud radio access networks (C-RANs) is considered, where realistic fronthaul capacity constraints are incorporated. To reduce the training overhead, pilot reuse is adopted and the transmit-beamforming is designed to be robust to the channel estimation errors. In contrast to the conventional C-RAN where the remote radio heads (RRHs) coherently transmit their data symbols to the user, we consider their non-coherent transmission, where no strict phase-synchronization is required. By exploiting the classic successive interference cancellation (SIC) technique, we first derive the closed-form expressions of the individual data rates from each serving RRH to the user and the overall data rate for each user that is not related to their decoding order. Then, we adopt the reweighted l_1-norm technique to approximate the l_0-norm in the fronthaul capacity constraints as the weighted power constraints. A low-complexity algorithm based on a novel sequential convex approximation (SCA) algorithm is developed to solve the resultant optimization problem with convergence guarantee. A beneficial initialization method is proposed to find the initial points of the SCA algorithm. Our simulation results show that in the high fronthaul capacity regime, the coherent transmission is superior to the non-coherent one in terms of its weighted sum rate. However, significant performance gains can be achieved by the non-coherent transmission over the coherent one in the low fronthaul capacity regime, which is the case in ultra-dense C-RANs, where mmWave fronthaul links with stringent capacity requirements are employed.**

***Index Terms*—Ultra-dense networks (UDN), C-RAN, limited fronthaul capacity, pilot reuse, non-coherent transmission.**

I. Introduction

Ultra dense networks (UDN), where a large number of base stations (BS) are installed within a given area, have been regarded as one of the most promising techniques to achieve the ambitious goal of the fifth-generation (5G) wireless system [1]. In UDNs, the received useful signal power is enhanced at the user side due to the reduced distance to its serving BSs. However, the user also receives the severe interference from its neighboring BSs, which constitutes the performance limiting factor for UDNs. Cloud radio access networks (C-RANs) have been recently proposed as the most promising solution to dealing with this issue. Under the C-RAN architecture, all the signal processing tasks are performed at the baseband unit (BBU) pool, which is located at the data center and equipped with powerful cloud computing capabilities. Then, the traditional full-functionality BSs are replaced by the low-functionality remote radio heads (RRHs), which only needs to support the simple transmission/reception. Under this centralized architecture, some advanced signal processing techniques can be realized in C-RANs, such as the coordinated multi-point (CoMP) transmission, where the interference signals can be transformed into useful signals. Furthermore, due to their low functionalities, the RRHs can be densely deployed with low operational cost. As a result, the C-RAN is an ideal platform to realize the benefits of UDNs.

To enable the centralized signal processing, the RRHs need to exchange data and channel information with the BBU pool via high-speed links named the fronthaul links. Conventionally, the fronthaul links are provisioned by the wired links such as optical fibers or high-speed Ethernet. However, in ultra-dense C-RANs, there are a large number of RRHs, and huge deployment costs will be incurred when connecting every RRH to the BBU pool using the wired links. Furthermore, some RRHs are located at places which are not accessible, and it may not be feasible to provide the dedicated wired links between each RRH to the BBU pool. Therefore, the millimeter wave (mmWave) wireless fronthaul is appealing for ultra-dense C-RANs due to its flexibility and low deployment cost, which has received extensive attention [2]–[4]. However, even at the mmWave frequency, the available bandwidth is limited compared with the wired links. Hence, the fronthaul capacity constraints should be taken into account when optimizing the system design.

Recently, transmission design has been extensively studied to deal with the fronthaul capacity constraints for C-RANs [4]–[10]. However, these contributions mainly focus on the coherent joint transmission scenario, where multiple RRHs coherently transmit the same data symbol to the served user and strict phase-synchronization among these RRHs is required. In ultra-dense C-RANs with excessive number of RRHs, it is difficult to satisfy the phase-synchronization requirement. Another alternative transmission strategy is the so-called non-coherent transmission, where multiple RRHs can transmit

Manuscript received April 16, 2018; revised September 13, 2018; accepted September 26, 2018. C. Pan, H. Ren, M. Elkashlan, and A. Nallanathan are with the Queen Mary University of London, London E1 4NS, U.K. (Email:{c.pan, h.ren, maged.elkashlan, a.nallanathan}@qmul.ac.uk). L. Hanzo is with the School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, U.K. (e-mail:lh@ecs.soton.ac.uk).

The corresponding author is Lajos Hanzo.

M. Elkashlan would like to acknowledge the financial support of the EPSRC project EP/N029666/1. A. Nallanathan would like to acknowledge the financial support of the EPSRC project EP/M016145/2. L. Hanzo gratefully acknowledges the financial support of the EPSRC projects EP/Noo4558/1, EP/PO34284/1, of the Royal Society as well as of the European Research Council's Advanced Fellow Grant QuantCom.

different data streams to the user. At the user side, successive interference cancellation technique can be adopted to decode the data streams from the corresponding RRHs. In this scheme, the strict phase-synchronization requirement is not necessary, which is more amenable to implement in ultra-dense C-RANs. The data rate expression achieved by each user for both strategies are different. For the coherent transmission, the serving RRHs for each user can be regarded as a unique antenna array with a large number of antennas. However, for the non-coherent transmission, the data rate transmitted from each RRH to the user is different, and the data rate of the user is the summation of individual data rates from all the RRHs in its serving cluster. Due to this difference, the fronthaul usage model for these two different transmission schemes are different. To the best of our knowledge, this paper is the first attempt to consider the non-coherent transmission with the fronthaul capacity constraints.

Besides the fronthaul capacity limit issue, another troublesome problem in ultra-dense C-RANs is that a large amount of channel state information (CSI) is required for the facilitation of centralized signal processing. To acquire these CSIs, an excessive amount of training overhead will be required, which significantly reduces the available resources for data transmission. One promising solution is to consider the incomplete CSI scenario, where each user only needs to estimate the CSI from its nearby RRHs, and only the large-scale channel gains are obtained for distant RRHs. Recently, we have studied the network power minimization problem under this scenario in [8], where a novel two-stage optimization framework was proposed to solve this problem. However, [8] assumed the perfect intra-cluster CSI, which is impractical due to the limited amount of pilot resources. To save the pilot consumption, pilot reuse scheme should be adopted, which incurs sizeable channel estimation error. Hence, it is imperative to design transmission schemes that are robust to the channel estimation error. Most recently, we have considered the transmit power minimization problem for time division duplex (TDD) C-RANs in [9] and frequency division duplex (FDD) C-RANs in [11]. However, both papers considered the coherent transmission in C-RANs.

In this paper, we study the weighted sum rate maximization problem for TDD ultra-dense C-RANs with non-coherent joint transmission and imperfect CSI, where the fronthaul capacity constraints are taken into account. In contrast to the transmit power minimization problem in [7], [9], [12]–[14] where the problem can be transformed into a convex second-order cone programming (SOCP) or semi-definite programming (SDP), the weighted sum rate maximization problem cannot be transformed into a convex optimization problem, which is proved to be NP-hard in [15]. In general, there are three approaches to solve this kind of optimization problems. The first one is the cyclic coordinate descent algorithm developed in [15], which enables each transmitter to update its beamforming vector locally. However, this algorithm relies on the assumption that the optimization variables are not coupled in the constraints, which is not the case in ultra-dense C-RAN where each RRH serves multiple users. The second is the weighted minimum mean square error (WMMSE) method [16] that has been successfully applied to solving diverse problems in C-RAN [6], [17]–[20]. However, since we consider the non-coherent transmission, the WMMSE used in the abovementioned papers cannot be used. The third one is the iterative SOCP-based algorithm developed in [21], which has been used in [5], [22] to deal with the transmission issues of C-RANs. However, again due to the fact that we consider the non-coherent transmission, the method developed in [21] is not applicable. Hence, the contributions of this paper can be summarized as follows:

1) We first derive the closed-form expression of the individual achievable data rates of each RRH in the serving cluster to each user by exploiting the statistical properties of the channel estimation errors, which depends on the decoding order. Then, we obtain the closed-form expression of the achievable sum-rate, which is not related to the decoding order. Based on the results derived, we formulate the weighted sum-rate maximization optimization subject to both the per-RRH power constraints and to the fronthaul capacity constraints.
2) To handle the l_0-norm in the fronthaul capacity constraints, we adopt the reweighted l_1-norm technique of compressive sensing for approximating the l_0-norm as the weighted power constraints. Then, a novel sequential convex approximation (SCA) algorithm is adopted for solving the resultant optimization problem with convergence guarantee. Furthermore, we also provide a novel method of initializing the algorithm.
3) Our simulation results show that the performance-advantage of the coherent versus non-coherent transmission depends on the fronthaul capacity limit. The weighted sum-rate of the coherent transmission is higher than that of its non-coherent counterpart, when a high fronthaul capacity is available. However, the non-coherent transmission significantly outperforms its coherent counterpart, when the fronthaul capacity is more limited, which is often the case for mmWave fronthaul links.

The rest of this paper is organized as follows. In Section II, we present the system model along with the problem formulation. A low-complexity algorithm is provided in Section III, where the novel initialization method is also proposed. Extensive simulation results are shown in Section IV. Conclusions are finally drawn in Section V. The notations are deferred to [9].

II. System Model

We consider a downlink TDD UD-CRAN, which has I RRHs and K UEs as shown in Fig. 1. Each RRH is equipped with M antennas and each UE has a single antenna. Denote $\mathcal{I} = \{1, \cdots, I\}$ and $\mathcal{U} = \{1, \cdots, K\}$ as the sets of RRHs and UEs, respectively. To reduce the complexity, the user-centric cluster method is adopted, where each UE is exclusively served by its nearby RRHs. Hence, we can define $\mathcal{I}_k \subseteq \mathcal{I}$ as the set of RRHs that potentially serve UE k, and $\mathcal{U}_i \subseteq \mathcal{U}$ as the set of UEs that can be potentially served by RRH i. These clusters are determined based on the large-scale channel gains that change very slowly, which are assumed to be fixed in this paper.

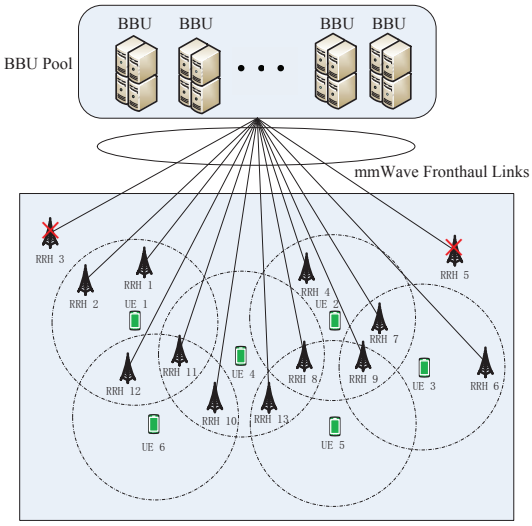


Fig. 1. Illustration of a C-RAN with thirteen RRHs and six UEs, i.e., $I = 13$, $K = 6$. To reduce the complexity, each UE is served by the RRHs within the dashed circle centered at the UE.

Let $\mathbf{x}_i \in \mathbb{C}^{M\times 1}$ be the transmit signal at RRH i, the received signal at UE k can be written as

$$y_k = \sum_{i\in\mathcal{I}} \mathbf{h}_{i,k}^{\mathrm{H}} \mathbf{x}_i + z_k, \tag{1}$$

where $\mathbf{h}_{i,k} \in \mathbb{C}^{M\times 1}$ represents the channel vector, z_k is the zero-mean additive complex white Gaussian noise with variance σ^2. The channel vector $\mathbf{h}_{i,k}$ is expressed as $\mathbf{h}_{i,k} = \sqrt{\alpha_{i,k}}\bar{\mathbf{h}}_{i,k}$, which consists of two parts: the large-scale channel gains $\alpha_{i,k}$ that includes the shadowing and path loss, and the small-scale channel fading $\bar{\mathbf{h}}_{i,k}$ following the distribution of $\mathcal{CN}(\mathbf{0}, \mathbf{I})$.

A. Channel Estimation

Similar to our previous work [9], [11], we only consider the channel estimation for the intra-cluster CSI. For the CSI out of the UE's cluster, it is assumed that only large-scale channel gains are obtained, i.e., $\{\alpha_{i,k}, \forall i \in \mathcal{I}\backslash\mathcal{I}_k, \forall k\}$. The channels are assumed to be frequency-flat within a coherence interval with T time slots, among which τ time slots are used for channel estimation, and the remaining $T - \tau$ time slots are dedicated for the data transmission. Hence, the number of orthogonal pilot sequences is equal to τ. In UD-CRANs, the number of UEs is much larger than τ. To enable the channel estimation, some UEs should reuse the same pilot.

Denote the set of available pilot sequences as $\mathcal{Q} = [\mathbf{q}_1, \cdots, \mathbf{q}_\tau] \in \mathbb{C}^{\tau\times\tau}$ that satisfies the orthogonal condition. For TDD UD-CRANs, the UEs send the pilot sequences to the RRHs. In specific, the training signals received at RRH i is given by

$$\mathbf{Y}_i = \sum\nolimits_{k\in\mathcal{U}} \sqrt{p_t}\mathbf{h}_{i,k}\mathbf{q}_{\pi_k}^{\mathrm{H}} + \mathbf{N}_i, \tag{2}$$

where p_t is the pilot power for each UE, $\mathbf{N}_i \in \mathbb{C}^{M\times\tau}$ is the Gaussian noise matrix, the elements of which follow the same distribution of $\mathcal{CN}(0, \sigma^2)$, $\mathbf{q}_{\pi_k} \in \mathbb{C}^{\tau\times 1}$ is the pilot training vector sent from UE k. In addition, to differentiate the channels from different UEs, the pilot sequences used by the UEs sharing the same RRH should also be orthogonal, i.e. $\mathbf{q}_{\pi_k}^{\mathrm{H}}\mathbf{q}_{\pi_{k'}} = 0, \text{for } k, k' \in \mathcal{U}_i, k \neq k', \forall i \in \mathcal{I}$. Furthermore, to control the estimation error, the maximum reuse time for each pilot should be below a fixed value $n_{\max}$, i.e., $n_l \leq n_{\max}, \forall l$, where n_l denotes the reuse time for pilot l. This paper aims to minimize the number of required orthogonal pilots while guaranteeing the above constraints. The Dsatur algorithm of graph theory can be used for solving the pilot allocation problem as detailed in [23]. In TDD UD-CRANs, the number of pilots required increases with the number of UEs and decreases with the number of RRHs. However, it is not related to the number of antennas at each RRH. The other key parameter is the maximum pilot reuse time $n_{\max}$. A higher $n_{\max}$ leads to more severe channel estimation errors, while a smaller $n_{\max}$ will require a larger number of pilots, hence the number of time slots remaining for transmission will be reduced. Denote $c^\star$ as the minimum number of different colors, which is equal to τ.

Denote $\mathcal{K}_{\pi_k}$ as the set of UEs that reuse the same pilot of UE k, that includes UE k. Then the minimum mean square error (MMSE) estimation of channel $\mathbf{h}_{i,k}$ is given by [24]

$$\hat{\mathbf{h}}_{i,k} = \frac{\alpha_{i,k}}{\sum_{k'\in\mathcal{K}_{\pi_k}} \alpha_{i,k'} + \hat{\sigma}^2} \frac{1}{\sqrt{p_t}} \mathbf{Y}_i \mathbf{q}_{\pi_k} \tag{3}$$

where $\hat{\sigma}^2 = \sigma^2/p_t$. According to the property of MMSE estimation [24], the channel estimation $\hat{\mathbf{h}}_{i,k}$ is independent of the channel estimation error $\mathbf{e}_{i,k} = \mathbf{h}_{i,k} - \hat{\mathbf{h}}_{i,k}$. The estimation error $\mathbf{e}_{i,k}$ follows the distribution of $\mathcal{CN}(\mathbf{0}, \delta_{i,k}\mathbf{I})$, where $\delta_{i,k}$ is given by

$$\delta_{i,k} = \frac{\alpha_{i,k}\left(\sum_{l\in\mathcal{K}_{\pi_k}\backslash\{k\}} \alpha_{i,l} + \hat{\sigma}^2\right)}{\sum_{l\in\mathcal{K}_{\pi_k}} \alpha_{i,l} + \hat{\sigma}^2}. \tag{4}$$

B. Downlink Data Transmission Model

The non-coherent joint transmission is considered, where the RRHs in $\mathcal{I}_k$ send different data streams to UE k [1]. Then, the signal received at UE k is given by

$$y_k = \sum_{i\in\mathcal{I}_k} \mathbf{h}_{i,k}^{\mathrm{H}} \mathbf{w}_{i,k} s_{i,k} + \sum_{l\neq k, l\in\mathcal{U}} \sum_{i\in\mathcal{I}_l} \mathbf{h}_{i,k}^{\mathrm{H}} \mathbf{w}_{i,l} s_{i,l} + z_k, \tag{5}$$

where $\mathbf{w}_{i,k} \in \mathbb{C}^{M\times 1}$ represents beam-vector from RRH i to UE k, respectively, $s_{i,k}$ denotes the data stream that RRH i intends to send to UE k. The data streams are assumed to be independent of each other, and have zero mean and unit variance.

It is assumed that each UE k has perfect knowledge of the effective precoded channels $\hat{\mathbf{h}}_{i,k}^{\mathrm{H}}\mathbf{w}_{i,k}, \forall i \in \mathcal{I}_k$. In general, some downlink pilot resources are required to train the effective precoded channels as detailed in [25], which will incur

[1]It should be noted that the non-coherent joint transmission is different from the coherent joint transmission where all the RRHs serving the same UE will transmit the same data symbol to the UE as seen in our previous work [8], [9], [11]. The non-coherent joint transmission is easier to implement at the transmitter side than the coherent joint transmission, since no strict phase-synchronization of the RRHs is required. For the receiver side, SIC is required in support of the non-coherent transmission, which incurs additional complexity. However, since the cluster size for each user in practical ultra-dense C-RANs is generally small, the computational complexity remains affordable at the user.

estimation error. However, Caire *et al.* showed in [25] that this error is marginal compared with that of channel estimation. Hence, this error is not considered here.

Since different RRHs send different signals to each UE, each UE needs to detect all the signals from all its serving RRHs. One low-complexity detection algorithm named successive interference cancellation (SIC) [26]–[29] can be adopted, where each UE sequentially detects its signals from different RRHs [2]. In particular, each UE first detects its signal from the first RRH in $\mathcal{I}_k$, while regarding the other desired signals as interference. For the second RRH, the UE already knows the signal from the first RRH. Hence, this UE can subtract it from the received signals and detect the signal from the second RRH. Repeat this procedure until all the desired signals are detected. As a result, the data rate transmitted on each RRH depends on the decoding order. Denote $\mathcal{I}_k = \{c_1^k, \cdots, c_{|\mathcal{I}_k|}^k\}$ as the candidate set of RRHs for UE k. Without loss of generality, we assume that the decoding order at UE k is $c_1^k, \cdots, c_{|\mathcal{I}_k|}^k$, namely, UE k first decodes the signal from RRH c_1^k, then decodes the signal of RRH c_2^k second, etc.

C. Achievable Data Rate Analysis

We first analyze the achievable data rate between RRH c_j^k and UE k. When UE k first decodes the signal from RRH c_1^k, it will not have any knowledge of the signals from all the RRHs in $\mathcal{I}_k$. Then, the received signal at UE k can be reexpressed as

$$\begin{aligned} y_{c_1^k,k} &= y_k = \hat{\mathbf{h}}_{c_1^k,k}^{\mathrm{H}} \mathbf{w}_{c_1^k,k} s_{c_1^k,k} + \mathbf{e}_{c_1^k,k}^{\mathrm{H}} \mathbf{w}_{c_1^k,k} s_{c_1^k,k} \\ &+ \sum_{m=2}^{|\mathcal{I}_k|} \mathbf{h}_{c_m^k,k}^{\mathrm{H}} \mathbf{w}_{c_m^k,k} s_{c_m^k,k} + \sum_{l\neq k, l\in\mathcal{U}} \sum_{n\in\mathcal{I}_l} \mathbf{h}_{n,k}^{\mathrm{H}} \mathbf{w}_{n,l} s_{n,l} + z_k. \end{aligned} \tag{6}$$

When UE k starts to decode the signal from RRH $c_j^k, j = 2, \cdots, |\mathcal{I}_k|$, it has already decoded the signals from the first $j-1$ RRHs, i.e., $c_m^k, m = 1, \cdots, j-1$. Then, the received signal in (5) is processed by subtracting the known signals from (5) as

$$\begin{aligned} & y_{c_j^k,k} \\ &= y_k - \sum_{m=1}^{j-1} \hat{\mathbf{h}}_{c_m^k,k}^{\mathrm{H}} \mathbf{w}_{c_m^k,k} s_{c_m^k,k} \\ &= \hat{\mathbf{h}}_{c_j^k,k}^{\mathrm{H}} \mathbf{w}_{c_j^k,k} s_{c_j^k,k} + \sum_{m=1}^{j} \mathbf{e}_{c_m^k,k}^{\mathrm{H}} \mathbf{w}_{c_m^k,k} s_{c_m^k,k} + \\ & \sum_{m=j+1}^{|\mathcal{I}_k|} \mathbf{h}_{c_m^k,k}^{\mathrm{H}} \mathbf{w}_{c_m^k,k} s_{c_m^k,k} + \sum_{l\neq k, l\in\mathcal{U}} \sum_{n\in\mathcal{I}_l} \mathbf{h}_{n,k}^{\mathrm{H}} \mathbf{w}_{n,l} s_{n,l} + z_k. \end{aligned} \tag{7}$$

Since $\hat{\mathbf{h}}_{c_j^k,k}^{\mathrm{H}} \mathbf{w}_{c_j^k,k}$ is assumed to be known at UE k, the first term in the second equality of (7) is regarded as the useful signal while the other terms are treated as uncorrelated noise. Then, the achievable data rate can be obtained by regarding the uncorrelated noise as the Gaussian noise with the same variance. The average data rate is given by

$$\bar{r}_{c_j^k,k} = \frac{T-\tau}{T} \mathbb{E}\left\{\log_2\left(1+\bar{\gamma}_{c_j^k,k}\right)\right\}, \tag{8}$$

where the expectation is taken over all uncertain terms, such as unknown channel estimation errors $\{\mathbf{e}_{i,k}, i \in \mathcal{I}_k\}$, and the small-scale inter-cluster CSI $\{\mathbf{h}_{i,k}, i \in \mathcal{I}\backslash\mathcal{I}_k\}$, $\bar{\gamma}_{c_j^k,k}$ is the effective SINR between RRH c_j^k and UE k, given by

$$\bar{\gamma}_{c_j^k,k} = \frac{\left|\hat{\mathbf{h}}_{c_j^k,k}^{\mathrm{H}} \mathbf{w}_{c_j^k,k}\right|^2}{\mathrm{SI}_{c_j^k,k} + \mathrm{MI}_k + \sigma^2}. \tag{9}$$

In (9), $\mathrm{SI}_{c_j^k,k}$ represents the self-interference incurred by channel estimation error and interference from the remaining RRHs in $\mathcal{I}_k$, and MI_k is the multi-UE cochannel interference. The expressions of $\mathrm{SI}_{c_j^k,k}$ and MI_k are respectively given by

$$\mathrm{SI}_{c_j^k,k} = \sum_{m=1}^{j} \left|\mathbf{e}_{c_m^k,k}^{\mathrm{H}} \mathbf{w}_{c_m^k,k}\right|^2 + \sum_{m=j+1}^{|\mathcal{I}_k|} \left|\mathbf{h}_{c_m^k,k}^{\mathrm{H}} \mathbf{w}_{c_m^k,k}\right|^2 \tag{10}$$

and

$$\mathrm{MI}_k = \sum_{l\neq k, l\in\mathcal{U}} \sum_{n\in\mathcal{I}_l} \left|\mathbf{h}_{n,k}^{\mathrm{H}} \mathbf{w}_{n,l}\right|^2, \tag{11}$$

Due to the incomplete CSI and the channel estimation error, it is difficult to derive the closed-form expression of the data rate. To deal with difficulty, we first derive its lower bound, which is more amenable for algorithm design.

Note that $\hat{\mathbf{h}}_{c_j^k,k}$ is the channel estimate given in (3) and $\mathbf{w}_{c_j^k,k}$ is the beamforming vector designed by the BBU pool. Hence, the numerator in (9) is available at the BBU pool, and only the denominator contains the uncertain terms. By using Jensen's inequality [30], we can derive the data rate lower bound as follows:

$$\begin{aligned} & \bar{r}_{c_j^k,k} && (12) \\ \geq\ & \log_2\left(1 + \frac{\left|\hat{\mathbf{h}}_{c_j^k,k}^{\mathrm{H}} \mathbf{w}_{c_j^k,k}\right|^2}{\mathbb{E}\left\{\mathrm{SI}_{c_j^k,k}\right\} + \mathbb{E}\left\{\mathrm{MI}_k\right\} + \sigma^2}\right) && (13) \\ \triangleq\ & r_{c_j^k,k} && (14) \end{aligned}$$

where $r_{c_j^k,k}$ is the lower bound of the average data rate $\bar{r}_{c_j^k,k}$. $\mathbb{E}\left\{\mathrm{SI}_{c_j^k,k}\right\}$ can be represented as

$$\mathbb{E}\left\{\sum_{m=1}^{j} \left|\mathbf{e}_{c_m^k,k}^{\mathrm{H}} \mathbf{w}_{c_m^k,k}\right|^2\right\} + \mathbb{E}\left\{\sum_{m=j+1}^{|\mathcal{I}_k|} \left|\mathbf{h}_{c_m^k,k}^{\mathrm{H}} \mathbf{w}_{c_m^k,k}\right|^2\right\}. \tag{15}$$

The first term in (15) can be easily calculated as $\sum_{m=1}^{j} \delta_{c_m^k,k} \left\|\mathbf{w}_{c_m^k,k}\right\|^2$. For the second term, note that channels $\mathbf{h}_{c_m^k,k}, m \in \mathcal{I}_k$ are estimated at UE k, we have

$$\begin{aligned} \mathbb{E}\left\{\left|\mathbf{h}_{c_m^k,k}^{\mathrm{H}} \mathbf{w}_{c_m^k,k}\right|^2\right\} &= \mathbb{E}\left\{\left|\left(\hat{\mathbf{h}}_{c_m^k,k}^{\mathrm{H}} + \mathbf{e}_{c_m^k,k}^{\mathrm{H}}\right) \mathbf{w}_{c_m^k,k}\right|^2\right\} \\ &= \left|\hat{\mathbf{h}}_{c_m^k,k}^{\mathrm{H}} \mathbf{w}_{c_m^k,k}\right|^2 + \delta_{c_m^k,k} \left\|\mathbf{w}_{c_m^k,k}\right\|^2 \end{aligned} \tag{16}$$

[2] In this paper, perfect SIC is assumed and the absence of error propagation is stipulated. This assumption is reasonable, since each user has to decode the signals from its serving clusters, the size of which is usually small. David Tse stated [29] that the effect of error propagation can be compensated by using a stronger code (e.g., increasing the block length), provided that the decoding size is relatively small.

where the second equality follows by using the independence of $\hat{\mathbf{h}}_{c_m^k,k}$ and $\mathbf{e}_{c_m^k,k}$, and the covariance matrix of $\mathbf{e}_{c_m^k,k}$ is $\delta_{c_m^k,k}\mathbf{I}_M$. Then, the second term is given by

$$\begin{aligned}&\mathbb{E}\left\{\sum_{m=j+1}^{|\mathcal{I}_k|}\left|\mathbf{h}_{c_m^k,k}^{\mathrm{H}}\mathbf{w}_{c_m^k,k}\right|^2\right\}\\&=\sum_{m=j+1}^{|\mathcal{I}_k|}\left(\left|\hat{\mathbf{h}}_{c_m^k,k}^{\mathrm{H}}\mathbf{w}_{c_m^k,k}\right|^2+\delta_{c_m^k,k}\left\|\mathbf{w}_{c_m^k,k}\right\|^2\right).\end{aligned}\tag{17}$$

The $\mathbb{E}\left\{\mathrm{SI}_{c_j^k,k}\right\}$ is given by

$$\mathbb{E}\left\{\mathrm{SI}_{c_j^k,k}\right\}=\sum_{m=j+1}^{|\mathcal{I}_k|}\left|\hat{\mathbf{h}}_{c_m^k,k}^{\mathrm{H}}\mathbf{w}_{c_m^k,k}\right|^2+\sum_{m\in\mathcal{I}_k}\delta_{m,k}\|\mathbf{w}_{m,k}\|^2.\tag{18}$$

To calculate $\mathbb{E}\{\mathrm{MI}_k\}$, we consider two cases: 1) $n\in\mathcal{I}_k$; 2) $n\notin\mathcal{I}_k$. For the first case, the channel $\mathbf{h}_{n,k}$ has been estimated at UE k. Hence, the expectation value can be calculated as in (16). For the latter case, UE k only knows the large-scale channel gain $\alpha_{n,k}$. Then, the expectation can be calculated as $\mathbb{E}\left\{\left|\mathbf{h}_{n,k}^{\mathrm{H}}\mathbf{w}_{n,l}\right|^2\right\}=\alpha_{n,k}\|\mathbf{w}_{n,k}\|^2$ and MI_k can be calculated as

$$\mathbb{E}\{\mathrm{MI}_k\}=\sum_{l\in\mathcal{U},l\neq k}\left(\sum_{n\in\mathcal{I}_l\cap\mathcal{I}_k}\mathbf{w}_{n,l}^{\mathrm{H}}\mathbf{J}_{n,k}\mathbf{w}_{n,l}+\sum_{n\in\mathcal{I}_l\setminus\mathcal{I}_k}\alpha_{n,k}\|\mathbf{w}_{n,l}\|^2\right),$$

where $\mathbf{J}_{n,k}=\hat{\mathbf{h}}_{n,k}\hat{\mathbf{h}}_{n,k}^{\mathrm{H}}+\delta_{n,k}\mathbf{I}_M$.

It is shown in [8] that the gap between the data rate lower bound $r_{c_j^k,k}$ and the average data rate $\bar{r}_{c_j^k,k}$ is within 3% in ultra-dense C-RANs, which can be ignored. Hence, in the following, instead of considering the average data rate with intractable expressions, we consider its lower bound $r_{c_j^k,k}$ that is much easier to handle and more suitable for algorithm design. In the following, the data rate refers to the lower bound of data rate for simplicity. Based on the above derivations, the sum data rate of UE k is obtained as

$$r_k=\sum_{j=1}^{|\mathcal{I}_k|}r_{c_j^k,k}=\frac{T-\tau}{T}\log_2\prod_{j=1}^{|\mathcal{I}_k|}\underbrace{\left(1+\gamma_{c_j^k,k}\right)}_{J_{c_j^k,k}},\tag{19}$$

where $J_{c_j^k,k}$ is given by

$$\frac{\sum_{m=j}^{|\mathcal{I}_k|}\left|\hat{\mathbf{h}}_{c_m^k,k}^{\mathrm{H}}\mathbf{w}_{c_m^k,k}\right|^2+\sum_{m\in\mathcal{I}_k}\delta_{m,k}\|\mathbf{w}_{m,k}\|^2+\mathrm{MI}_k+\sigma^2}{\sum_{m=j+1}^{|\mathcal{I}_k|}\left|\hat{\mathbf{h}}_{c_m^k,k}^{\mathrm{H}}\mathbf{w}_{c_m^k,k}\right|^2+\sum_{m\in\mathcal{I}_k}\delta_{m,k}\|\mathbf{w}_{m,k}\|^2+\mathrm{MI}_k+\sigma^2}.\tag{20}$$

Note that the denominator of $J_{c_j^k,k}$ is exactly the numerator of $J_{c_{j+1}^k,k}$ for $j=1,\cdots,|\mathcal{I}_k|-1$. Then, with some simple manipulations, we have

$$r_k=\frac{T-\tau}{T}\log_2\left(1+\frac{\sum_{m\in\mathcal{I}_k}\left|\hat{\mathbf{h}}_{m,k}^{\mathrm{H}}\mathbf{w}_{m,k}\right|^2}{\sum_{m\in\mathcal{I}_k}\delta_{m,k}\|\mathbf{w}_{m,k}\|^2+\mathrm{MI}_k+\sigma^2}\right).\tag{21}$$

It should be emphasized that the data rate for UE k in the non-coherent transmission case is totally different from the coherent case as seen in our previous work [9], [17], where there is one term in the nominator of the SINR expression. In contrast, for the non-coherent transmission case in (21), the nominator of the SINR expression is the summation over all desired signal power towards UE k. The first term in the denominator of the SINR expression is the self-interference power due to the channel estimation error, the second term corresponds to the multiuser interference, and the final term is the noise power. From (21), we find that the decoding order is not reflected in the sum rate expression. Hence, the sum rate of UE k does not depend on the decoding order.

D. Problem Formulation

In UD-CRAN, the fronthaul links are usually deployed by using wireless transmission due to its flexibility and low cost. However, compared with the wired fronthaul links such as optical fiber, the fronthaul capacity is more stringent in wireless links. Hence, when designing the transmission strategy, the fronthaul capacity constraints in UD-CRAN should be considered, which can be expressed as

$$\sum\nolimits_{k\in\mathcal{U}_i}\mathbb{I}\left(\|\mathbf{w}_{i,k}\|^2\right)r_{i,k}\leq C_{i,\max},\forall i\in\mathcal{I}.\tag{22}$$

where $r_{i,k}$ is given in (8), $C_{i,\max}$ is the fronthaul capacity limit, and $\mathbb{I}(\cdot)$ is the indicator function, defined as

$$\mathbb{I}(x)=\begin{cases}1, \text{ if } x\neq 0,\\ 0, \text{ otherwise.}\end{cases}\tag{23}$$

Note that in our previous work [8], [9], [11], [17] for coherent joint transmission, the set of RRHs serving UE k transmit with the same data rate r_k, and the fronthaul link capacity constraint is modeled as follows

$$\sum\nolimits_{k\in\mathcal{U}_i}\mathbb{I}\left(\|\mathbf{w}_{i,k}\|^2\right)r_k\leq C_{i,\max},\forall i\in\mathcal{I}.\tag{24}$$

For the non-coherent transmission case in (22), each UE's sum rate is divided among the serving candidate of RRHs, while for the coherent transmission case in (24), each RRH has to transmit the signals with each UE's sum rate. Hence, for the case with stringent fronthaul capacity limit, the non-coherent transmission may be the better option.

In contrast to the data rate expression in (21), the fronthaul capacity constraint model in (22) is related to the decoding order and different decoding order leads to different data rates between each RRH and each UE. Hence, how to find the optimal decoding order remains a critical issue. This problem is very difficult to solve. We provide a heuristic method to solve this problem as follows. First, sort the large-scale channel gains of the serving RRHs for each UE k in the descending order, namely, $\alpha_{\pi_1^k,k}\geq\cdots\geq\alpha_{\pi_{|\mathcal{I}_k|}^k,k}$. Then, UE k decodes the data streams from the RRHs in $\mathcal{I}_k$ according to this descending order. That is, UE k decodes the signal from RRH π_1^k first, and then π_2^k in a successive interference cancellation manner until the signal from the final RRH is decoded. This decoding order is reasonable since the RRH with higher large-scale channel gain generally has a larger SINR. When decoding the signals from this RRH first, the decoding error propagation will be reduced.

In this paper, we aim to optimize the beam-vectors to maximize the weighted sum rate of UEs while guaranteeing the fronthaul capacity constraints and each RRH's power limit. Mathematically, the problem can be formally formulated as

$$\max_{\mathbf{w}} \quad \sum_{k\in\mathcal{U}} \rho_k r_k \tag{25a}$$

$$\text{s.t.} \quad \sum_{k\in\mathcal{U}_i} \|\mathbf{w}_{i,k}\|^2 \le P_{i,\max}, i\in\mathcal{I}, \tag{25b}$$

$$\sum_{k\in\mathcal{U}_i} \mathbb{I}\left(\|\mathbf{w}_{i,k}\|^2\right) r_{i,k} \le C_{i,\max}, \forall i\in\mathcal{I}, \tag{25c}$$

where $\mathbf{w}$ is the collection of all beam-vectors, ρ_k is the weightor factor assigned to UE k that is used to control the fairness among the UEs, r_k is given in (21), constraints (25b) corresponds to the per-RRH power constraints. The weighted sum rate maximization problem is an NP-hard problem. Additionally, the indicator function in the fronthaul capacity constraints complicates further the analysis of this problem. The imperfect CSI case considered here does not allow the application of the WMMSE method. In the following, we provide a low-complexity algorithm to solve the above problem.

III. Low-complexity Algorithm

In this section, we first simplify the data rate expressions, and then propose a low-complexity algorithm to solve the resultant optimization problem. Finally, we provide a novel method to initialize the algorithm.

A. Simplification of the Data Rate Expression

Before solving Problem (25), we first simplify the expressions of r_k and $r_{i,k}$. The beamforming vectors from all RRHs in $\mathcal{I}_k$ can be merged into a single large-dimension vector according to the decoding order, denoted as $\mathbf{w}_k = \left[\mathbf{w}_{\pi_1^k,k}^{\mathrm{H}}, \cdots, \mathbf{w}_{\pi_{|\mathcal{I}_k|}^k,k}^{\mathrm{H}}\right]^{\mathrm{H}} \in \mathbb{C}^{|\mathcal{I}_k|M\times 1}$. Then, the data rate r_k in (21) can be rewritten as

$$r_k = \frac{T-\tau}{T}\log_2\left(1+\frac{\mathbf{w}_k^{\mathrm{H}}\mathbf{G}_{k,k}\mathbf{w}_k}{\mathbf{w}_k^{\mathrm{H}}\mathbf{E}_{k,k}\mathbf{w}_k+\sum\limits_{l\in\mathcal{U},l\neq k}\mathbf{w}_l^{\mathrm{H}}\mathbf{G}_{l,k}\mathbf{w}_l+\sigma^2}\right). \tag{26}$$

where $\mathbf{G}_{k,k}$, $\mathbf{E}_{k,k}$ and $\mathbf{G}_{l,k}$ are respectively given by

$$\mathbf{G}_{k,k} = \mathrm{blkdiag}\left\{\hat{\mathbf{h}}_{\pi_1^k,k}\hat{\mathbf{h}}_{\pi_1^k,k}^{\mathrm{H}}, \cdots, \hat{\mathbf{h}}_{\pi_{|\mathcal{I}_k|}^k,k}\hat{\mathbf{h}}_{\pi_{|\mathcal{I}_k|}^k,k}^{\mathrm{H}}\right\}, \tag{27}$$

$$\mathbf{E}_{k,k} = \mathrm{blkdiag}\left\{\delta_{\pi_1^k,k}\mathbf{I}_M, \cdots, \delta_{\pi_{|\mathcal{I}_k|}^k,k}\mathbf{I}_M\right\}, \tag{28}$$

and

$$\mathbf{G}_{l,k} = \mathrm{blkdiag}\left\{\mathbf{A}_{\pi_1^l,k}, \cdots, \mathbf{A}_{\pi_{|\mathcal{I}_l|}^l,k}\right\}, \tag{29}$$

with $\mathbf{A}_{\pi_j^l,k}$ given by

$$\mathbf{A}_{\pi_j^l,k} = \begin{cases} \hat{\mathbf{h}}_{\pi_j^l,k}\hat{\mathbf{h}}_{\pi_j^l,k}^{\mathrm{H}} + \delta_{\pi_j^l,k}\mathbf{I}_M, & \text{if } \pi_j^l \in \mathcal{I}_l \cap \mathcal{I}_k, \\ \alpha_{\pi_j^l,k}\mathbf{I}_M, & \text{otherwise.} \end{cases} \tag{30}$$

To obtain the expression of $r_{\pi_j^k,k}$, we first define $\mathbf{w}_{\pi_{j+1}^k,:} = \left[\mathbf{w}_{\pi_{j+1}^k,k}^{\mathrm{H}}, \cdots, \mathbf{w}_{\pi_{|\mathcal{I}_k|}^k,k}^{\mathrm{H}}\right]^{\mathrm{H}}$. Then, the data rate $r_{\pi_j^k,k}$ in (8) can be rewritten as in (31) at the top of the next page, where $\mathbf{B}_{\pi_{j+1}^k,k}$ is given by

$$\mathbf{B}_{\pi_{j+1}^k,k} = \mathrm{blkdiag}\left\{\hat{\mathbf{h}}_{\pi_{j+1}^k,k}\hat{\mathbf{h}}_{\pi_{j+1}^k,k}^{\mathrm{H}}, \cdots, \hat{\mathbf{h}}_{\pi_{|\mathcal{I}_k|}^k,k}\hat{\mathbf{h}}_{\pi_{|\mathcal{I}_k|}^k,k}^{\mathrm{H}}\right\}.$$

B. Low-complexity Algorithm

The indicator function in the fronthaul capacity constraints can be equivalently expressed as an l_0-norm of a scalar, which represents the number of non-zero values in a vector. This reexpression allows us to approximate it as a convex reweighted l_1-norm from compressed sensing technique [6], [31]. Specifically, the indicator function $\mathbb{I}\left(\|\mathbf{w}_{i,k}\|^2\right)$ can be rewritten as

$$\mathbb{I}\left(\|\mathbf{w}_{i,k}\|^2\right) = \left\|\|\mathbf{w}_{i,k}\|^2\right\|_0. \tag{32}$$

Then, the fronthaul capacity constraints in (25c) can be reformulated as [6], [31]

$$\sum_{k\in\mathcal{U}_i} \beta_{i,k}\|\mathbf{w}_{i,k}\|^2 r_{i,k} \le C_{i,\max}, \forall i\in\mathcal{I}, \tag{33}$$

where $\beta_{i,k}$ is a constant weight factor that is related to UE k and RRH i, that is iteratively updated as

$$\beta_{i,k} = \frac{1}{\|\mathbf{w}_{i,k}\|^2+\tau}, \forall i,k, \tag{34}$$

with τ being a small constant regularization parameter [3] and $\|\mathbf{w}_{i,k}\|^2$ from the previous iteration.

However, even with the simplification of (25c) as (33), Problem (25) is still difficult to solve due to the data rate $r_{i,k}$ in (33), which is a non-convex constraint. To deal with this difficulty, we replace $r_{i,k}$ with $\hat{r}_{i,k}$ obtained from the previous iteration. With fixed $\beta_{i,k}$ and $\hat{r}_{i,k}$, Problem (25) can be transformed as

$$\max_{\mathbf{w}} \quad \sum_{k\in\mathcal{U}} \rho_k r_k \tag{35a}$$

$$\text{s.t.} \quad \sum_{k\in\mathcal{U}_i} \|\mathbf{w}_{i,k}\|^2 \le P_{i,\max}, i\in\mathcal{I}, \tag{35b}$$

$$\sum_{k\in\mathcal{U}_i} \beta_{i,k}\hat{r}_{i,k}\|\mathbf{w}_{i,k}\|^2 \le C_{i,\max}, \forall i\in\mathcal{I}, \tag{35c}$$

where the transformed fronthaul capacity constraint (35c) can be interpreted as the weighted per-RRH power constraint similar to the conventional per-RRH power constraint (35b).

Problem (35) is still difficult to solve due to the following reasons. Since we consider the non-coherent transmission, the

[3]This approximation makes the approximated problem different from the original ones. However, when the regularization parameter is chosen as a very small value, this approximation error becomes negligible. In the simulations, the regularization parameter is set to $\tau = 10^{-8}$ and the transmit power is set to zero if the calculated power to be transmitted from each RRH to any user is smaller than 10^{-8}. This incurs a negligible effect on the data rate of each user. In fact, for practical analog to digital conversion (ADC) or digital to analog conversion (DAC), there is a minimum power that can activate the ADC/DACs. This kind of value is usually higher than 10^{-8}.

$$r_{\pi_j^k,k}=\frac{T-\tau}{T}\log_2\left(1+\frac{\left|\hat{\mathbf{h}}_{\pi_j^k,k}^{\mathrm{H}}\mathbf{w}_{\pi_j^k,k}\right|^2}{\mathbf{w}_{\pi_{j+1}^k,:}^{\mathrm{H}}\mathbf{B}_{\pi_{j+1}^k,k}\mathbf{w}_{\pi_{j+1}^k,:}+\mathbf{w}_k^{\mathrm{H}}\mathbf{E}_{k,k}\mathbf{w}_k+\sum\limits_{l\in\mathcal{U},l\neq k}\mathbf{w}_l^{\mathrm{H}}\mathbf{G}_{l,k}\mathbf{w}_l+\sigma^2}\right). \tag{31}$$

WMMSE developed in [16] cannot be used since the rank of matrix $\mathbf{G}_{k,k}$ defined in (27) is generally higher than one. Hence, we cannot construct the auxiliary signal transmission model for each UE as in the coherent transmission case of [17]. Again due to that the rank of matrix $\mathbf{G}_{k,k}$ is higher than one, the nominator of the SINR of UE k cannot be transformed as in (6b) of [21] or (25) of [5]. As a result, new methods have to be developed to solve Problem (35). In the following, we develop a low-complexity algorithm based on the sequential convex approximation (SCA) algorithm.

By introducing some slack variables, Problem (35) can be rewritten as

$$\max_{\mathbf{w},\boldsymbol{\mu},\boldsymbol{\eta},\boldsymbol{\nu}}\ \sum_{k\in\mathcal{U}}\rho_k\mu_k \tag{36a}$$

$$\text{s.t. }\frac{T-\tau}{T}\log_2\left(1+\eta_k\right)\geq\mu_k,\forall k, \tag{36b}$$

$$\frac{\mathbf{w}_k^{\mathrm{H}}\mathbf{G}_{k,k}\mathbf{w}_k}{\eta_k}\geq\nu_k,\forall k, \tag{36c}$$

$$\nu_k\geq\mathbf{w}_k^{\mathrm{H}}\mathbf{E}_{k,k}\mathbf{w}_k+\sum_{l\in\mathcal{U},l\neq k}\mathbf{w}_l^{\mathrm{H}}\mathbf{G}_{l,k}\mathbf{w}_l+\sigma^2,\forall k, \tag{36d}$$

$$(35b),(35c) \tag{36e}$$

where $\boldsymbol{\mu}$, $\boldsymbol{\eta}$ and $\boldsymbol{\nu}$ denote the collections of μ_k, η_k and ν_k, respectively. The equivalence between Problem (25) and Problem (36) lies in the fact that inequalities (36b), (36c) and (36d) hold with equality at the optimal point, which can be readily proved by using the contradiction method.

Denote the left hand side (LHS) of (36c) as the following function

$$f\left(\mathbf{w}_k,\eta_k\right)=\frac{\mathbf{w}_k^{\mathrm{H}}\mathbf{G}_{k,k}\mathbf{w}_k}{\eta_k}. \tag{37}$$

It is of the form quadratic over linear, which is jointly convex over $\mathbf{w}_k$ and η_k. Hence, constraint (36c) is non-convex. We then approximate it as its first-order Taylor expansion. Specifically, by using Appendix B of [11], the following inequality holds

$$\begin{aligned}f\left(\mathbf{w}_k,\eta_k\right)&\geq\frac{\mathbf{w}_k^{(n)\mathrm{H}}\mathbf{G}_{k,k}\mathbf{w}_k^{(n)}}{\eta_k^{(n)}}-\frac{\mathbf{w}_k^{(n)\mathrm{H}}\mathbf{G}_{k,k}\mathbf{w}_k^{(n)}}{\eta_k^{(n)2}}\left(\eta_k-\eta_k^{(n)}\right)\\&\quad+\frac{2\mathrm{Re}\left\{\mathbf{w}_k^{(n)\mathrm{H}}\mathbf{G}_{k,k}\left(\mathbf{w}_k-\mathbf{w}_k^{(n)}\right)\right\}}{\eta_k^{(n)}} \qquad (38)\\&=\frac{2\mathrm{Re}\left\{\mathbf{w}_k^{(n)\mathrm{H}}\mathbf{G}_{k,k}\mathbf{w}_k\right\}}{\eta_k^{(n)}}-\frac{\mathbf{w}_k^{(n)\mathrm{H}}\mathbf{G}_{k,k}\mathbf{w}_k^{(n)}}{\eta_k^{(n)2}}\eta_k \qquad (39)\\&\triangleq\ F\left(\mathbf{w}_k,\eta_k|_{\mathbf{w}_k^{(n)},\eta_k^{(n)}}\right). \qquad (40)\end{aligned}$$

where $\mathbf{w}_k^{(n)}$ and $\eta_k^{(n)}$ are obtained from the n-th iteration. Then, we replace constraint (36c) with the following inequalities:

$$F\left(\mathbf{w}_k,\eta_k|_{\mathbf{w}_k^{(n)},\eta_k^{(n)}}\right)\geq\nu_k,\forall k. \tag{41}$$

which is a linear inequality.

Constraint (36d) can be equivalently rewritten as an SOC constraint:

$$\begin{aligned}&\left(\mathbf{w}_k^{\mathrm{H}}\mathbf{E}_{k,k}\mathbf{w}_k+\sum_{l\in\mathcal{U},l\neq k}\mathbf{w}_l^{\mathrm{H}}\mathbf{G}_{l,k}\mathbf{w}_l+\sigma^2+\tfrac{1}{4}(\nu_k-1)^2\right)^{1/2}\\&\leq\tfrac{1}{2}\left(\nu_k+1\right).\end{aligned} \tag{42}$$

The other troublesome constraint in Problem (36) is (36b). Although it is a convex constraint, it cannot be solved by the common convex solvers such as CVX package [30] due to the log-function involved. To resolve this issue, we approximate this constraint as a more tractable constraint by using the following inequality:

$$\ln\left(x\right)\geq1-\frac{1}{x}. \tag{43}$$

By substituting $x=\left(1+\eta_k\right)\Big/\left(1+\eta_k^{(n)}\right)$ into the above inequality, we have

$$\ln\left(1+\eta_k\right)\geq\ln\left(1+\eta_k^{(n)}\right)+1-\frac{1+\eta_k^{(n)}}{1+\eta_k}. \tag{44}$$

The above inequality holds with equality when $\eta_k=\eta_k^{(n)}$. In addition, the first derivatives of the both sides of (44) are the same when $\eta_k=\eta_k^{(n)}$. Hence, the right hand side (RHS) of (44) is a proper approximation for the LHS of (44). By substituting the RHS of (44) into constraint (36b), we have

$$\ln\left(1+\eta_k^{(n)}\right)+1\geq\ln 2\frac{T}{T-\tau}\mu_k+\frac{1+\eta_k^{(n)}}{1+\eta_k}, \tag{45}$$

which is convex.

Based on the above derivations, the optimization problem to be solved in the $n+1$-th iteration is given by

$$\max_{\mathbf{w},\boldsymbol{\mu},\boldsymbol{\eta},\boldsymbol{\nu}}\quad\sum\nolimits_{k\in\mathcal{U}}\rho_k\mu_k \tag{46a}$$

$$\text{s.t.}\quad(35b),(35c),(41),(42),(45), \tag{46b}$$

which is an SOCP problem that can be effectively solved by using the CVX package [30].

Based on the above derivations, a straightforward way of solving Problem (25) involves two layers of iterations: the inner layer to solve Problem (35) with given $\beta_{i,k}$ and $\hat{r}_{i,k}$, and an outer layer to update $\beta_{i,k}$ and $\hat{r}_{i,k}$. Although the inner layer algorithm can be guaranteed to converge due to the property of SCA algorithm, the overall algorithm may have high computational complexity since two-layer iterations are involved. Here, we merge these two-layer iterations into only one layer and update $\beta_{i,k}$ and $\hat{r}_{i,k}$ inside the SCA algorithm, as shown in Algorithm 1. Although the algorithm cannot

be strictly proved to converge, the simulation section always shows the convergence of the algorithm and it converges rapidly.

Algorithm 1 SCA algorithm for solving Problem (35)

1: Initialize iteration number $n = 1$, error tolerance ε, feasible $\mathbf{w}^{(0)}$, calculate $\{\beta_{i,k}^{(0)}, \hat{r}_{i,k}^{(0)}, \forall i, k\}$ and $\boldsymbol{\eta}^{(0)}$, calculate the objective value of Problem (46), denoted as $\mathrm{Obj}^{(0)}$.
2: With fixed $\{\beta_{i,k}^{(n-1)}, \hat{r}_{i,k}^{(n-1)}, \forall i, k\}$, $\boldsymbol{\eta}^{(n-1)}$ and $\mathbf{w}^{(n-1)}$, solve Problem (46) by using the CVX package to obtain $\mathbf{w}^{(n)}$ and $\boldsymbol{\eta}^{(n)}$;
3: Update $\{\beta_{i,k}^{(n)}, \hat{r}_{i,k}^{(n)}, \forall i, k\}$;
4: Calculate new objective value $\mathrm{Obj}^{(n)}$. If $\left|\mathrm{Obj}^{(n)} - \mathrm{Obj}^{(n-1)}\right| / \mathrm{Obj}^{(n)} < \varepsilon$, terminate. Otherwise, set $n \leftarrow n + 1$, go to step 2.

C. Initialization of the algorithm

There is one critical issue that remains to be solved: how to find initial feasible beam-vectors $\mathbf{w}$ in Algorithm 1. Note that randomly choosing a set of beam-vectors that satisfy the per-RRH power constraints may not satisfy per-link capacity constraints. However, if one sets the initial beam-vectors to zero vectors, $\hat{r}_{i,k}^{(0)}$ will be zero for all UEs, which makes the constraints (35c) meaningless. Hence, one has to carefully choose the initial feasible beam-vectors. In the following, we provide an alternative to finding the initial beam-vectors.

First, we assume that each RRH is serving all the UEs in its candidate set $\mathcal{U}_i$, and the power is equally split among these UEs. Furthermore, the beam direction is chosen to align with its channel direction. Hence, one set of beam-vectors that satisfy the per-RRH power constraints is given by:

$$\tilde{\mathbf{w}}_{i,k} = \sqrt{\frac{P_{i,\max}}{|\mathcal{U}_i|}} \frac{\hat{\mathbf{h}}_{i,k}}{\left\|\hat{\mathbf{h}}_{i,k}\right\|}, \forall i \in \mathcal{I}, k \in \mathcal{U}_i. \tag{47}$$

By substituting the above beam-vectors into the data rate expression (31), we can obtain the data rates $\tilde{r}_{\pi_j^k,k}, \forall j = 1, \cdots, |\mathcal{I}_k|, k \in \mathcal{U}$. Then, check whether the beam-vectors in (47) satisfy the per-link capacity constraints or not. If yes, then set the initial beam-vectors as $\mathbf{w}_{i,k}^{(0)} = \tilde{\mathbf{w}}_{i,k}, \forall i \in \mathcal{I}, k \in \mathcal{U}_i$. Otherwise, we have to go one step further to find a set of feasible beam-vectors that satisfy the per-link capacity constraints.

To this end, we first split the fronthaul capacity limit among its serving UEs according to the following formulation:

$$C_{i,k} = \frac{\alpha_{i,k}}{\sum\limits_{k \in \mathcal{U}} \alpha_{i,k}} C_{i,\max}, \forall i \in \mathcal{I}, k \in \mathcal{U}_i. \tag{48}$$

The above rate assignment is reasonable since in general the UE with better channel condition can transmit with higher data rate. Then, we construct the following power minimization problem:

$$\min_{\mathbf{w}} \quad \sum\nolimits_{k \in \mathcal{U}} \|\mathbf{w}_k\|^2 \tag{49a}$$

$$\text{s.t.} \quad r_{\pi_j^k,k} \geq C_{\pi_j^k,k}, \forall j = 1, \cdots, |\mathcal{I}_k|, k \in \mathcal{U}, \tag{49b}$$

$$\sum_{k \in \mathcal{U}_i} \|\mathbf{w}_{i,k}\|^2 \leq P_{i,\max}, i \in \mathcal{I}. \tag{49c}$$

Since the rotation of the beam-vectors will not affect the SINR value, constraint (49b) can be equivalently cast as the SOC constraints in (50) and (51) at the top of the next page, where $\tilde{\eta}_{\pi_j^k,k} = 2^{\frac{T}{T-\tau} C_{\pi_j^k,k}} - 1$. By replacing (49b) with (50), Problem (49) is an SOCP problem that can be effectively solved.

If Problem (49) is feasible, denote the optimal solution as $\mathbf{w}^*$. Then, the rate constraint (49b) must hold with equality at the optimal solution, which can be readily proved by using the contradiction method. Hence, the obtained solution $\mathbf{w}^*$ must be feasible for Problem (25), which can be used as the initial beam-vectors of Algorithm 1, i.e., $\mathbf{w}_{i,k}^{(0)} = \mathbf{w}_{i,k}^*, \forall i \in \mathcal{I}, k \in \mathcal{U}_i$.

Otherwise, we construct an alternative optimization problem by introducing a series of auxiliary variables $\{\xi_{i,k}, \forall i, k\}$:

$$\min_{\mathbf{w}, \{\xi_{i,k} \geq 0, \forall i,k\}} \quad \sum\nolimits_{i \in \mathcal{I}} \sum\nolimits_{k \in \mathcal{U}} \xi_{i,k} + \sum\nolimits_{k \in \mathcal{U}} \|\mathbf{w}_k\|^2 \tag{52a}$$

$$\text{s.t. } \hat{\mathbf{h}}_{\pi_j^k,k}^{\mathrm{H}} \mathbf{w}_{\pi_j^k,k} + \xi_{\pi_j^k,k} \geq g_{\pi_j^k,k}(\mathbf{w}) \quad (51), (49c), \tag{52b}$$

where $g_{\pi_j^k,k}(\mathbf{w})$ is defined in (50) at the top of this page. Obviously, the above optimization problem is always feasible. Let us denote the optimal solution of Problem (52) by $\mathbf{w}^\star, \{\xi_{i,k}^\star, \forall i, k\}$. Since Problem (49) is assumed to be infeasible here, there must exist at least one $\xi_{i,k}^\star$ that is strictly larger than zero. Denote the set of indices of these $\xi_{i,k}^\star$s as $\mathcal{A} = \left\{(i,k) | \xi_{i,k}^\star > 0, \forall i, k\right\}$ and the set of indices of the $\xi_{i,k}^\star$s that are equal to zero as $\mathcal{B} = \left\{(i,k) | \xi_{i,k}^\star = 0, \forall i, k\right\}$. Note that the first set of constraints in Problem (52) hold with equality at the optimal point, which can be proved by using contradiction method. Hence, the obtained data rates corresponding to $\mathcal{A}$ are strictly smaller than $C_{i,k}$, i.e., $r_{i,k}^\star < C_{i,k}$, for $(i,k) \in \mathcal{A}$, while those corresponding to $\mathcal{B}$ are equal to $C_{i,k}$, i.e., $r_{i,k}^\star = C_{i,k}$, for $(i,k) \in \mathcal{B}$. As a result, the obtained solution $\mathbf{w}^\star$ satisfies the per-link capacity constraints, which can be used as the initial input of Algorithm 1, i.e., $\mathbf{w}_{i,k}^{(0)} = \mathbf{w}_{i,k}^\star, \forall i \in \mathcal{I}, k \in \mathcal{U}_i$.

In summary, the algorithm to find the initial feasible point is given in Fig. 2 at the top of the next page.

IV. Simulation Results

In this section, we provide simulation results to evaluate the performance of our proposed algorithm. We consider an C-RAN network covering a square area of 600 m $\times$ 600 m. Both the RRHs and UEs are randomly generated in this area. The number of RRHs and UEs are respectively given by 14 and 10, respectively. The corresponding densities of RRHs and UEs are then given by 39 RRHs/km^2 and 28 UEs/km^2, respectively. This is consistent with the requirement of 5G

$$\hat{\mathbf{h}}_{\pi_j^k,k}^{\mathrm{H}}\mathbf{w}_{\pi_j^k,k} \geq \sqrt{\tilde{\eta}_{\pi_j^k,k}}\sqrt{\mathbf{w}_{\pi_{j+1}^k,:}^{\mathrm{H}}\mathbf{B}_{\pi_{j+1}^k,k}\mathbf{w}_{\pi_{j+1}^k,:} + \mathbf{w}_k^{\mathrm{H}}\mathbf{E}_{k,k}\mathbf{w}_k + \sum_{l\in\mathcal{U},l\neq k}\mathbf{w}_l^{\mathrm{H}}\mathbf{G}_{l,k}\mathbf{w}_l + \sigma^2} \triangleq g_{\pi_j^k,k}(\mathbf{w}), \tag{50}$$

$$\mathrm{Im}\left(\hat{\mathbf{h}}_{\pi_j^k,k}^{\mathrm{H}}\mathbf{w}_{\pi_j^k,k}\right) = 0, \tag{51}$$

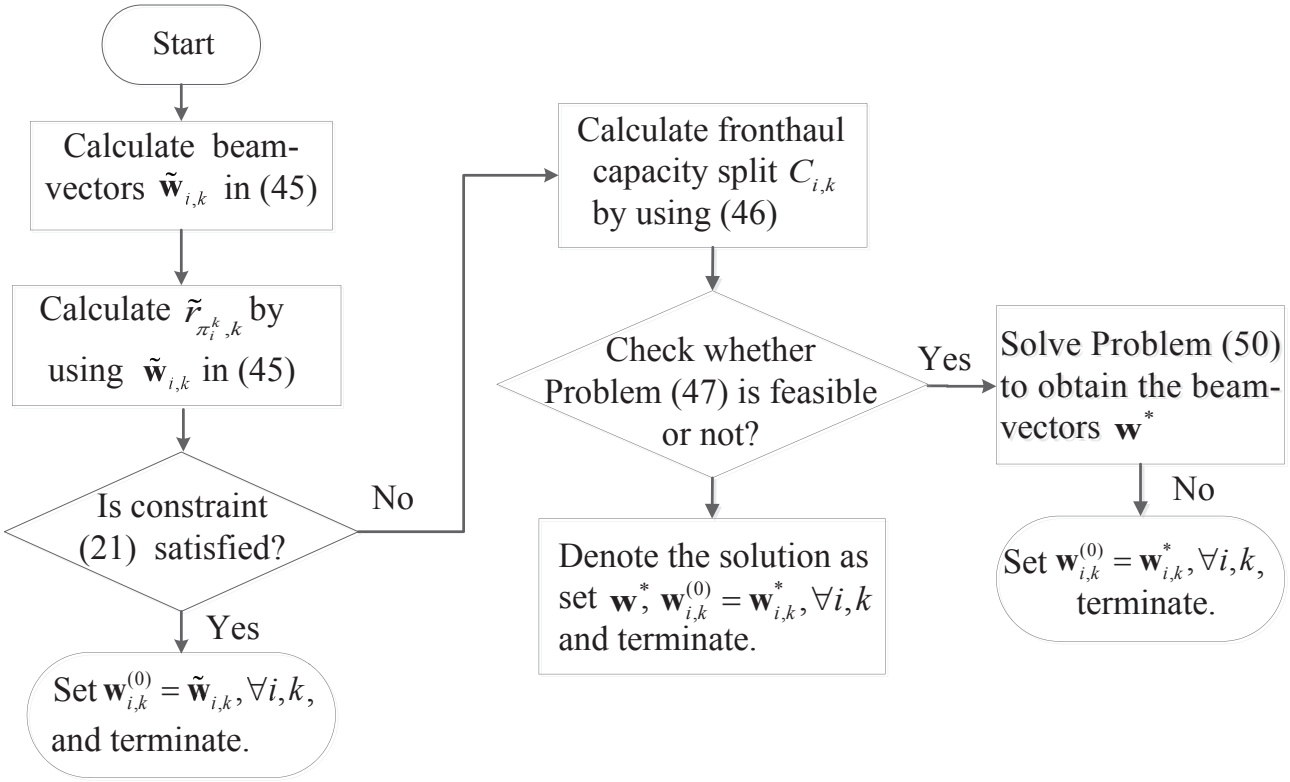


Fig. 2. Flow chart of the algorithm to find the initial feasible point for Algorithm 1.

ultra-dense network [32], where the density of the base stations (BS) is up to 40-50 BS/km^2. The channel model is set as: 1) Path loss $PL = 35.3 + 37.6\log_{10}d$ (dB) [33]; 2) log-normal shadow fading with zero mean and 8 dB standard derivation; 3) Rayleigh channel fading with zero mean and unit variance. Unless otherwise stated, the other simulation parameters are set as: Channel bandwidth of $B = 20$ MHz, the number of transmit antennas of $M = 2$, the noise power density of -174 dBm/Hz, the pilot power of $p_t = 2$ W, the maximum power of RRHs of $P_{i,\max} = 1$ W, $\forall i$, weight factor of $\rho_k = 1, \forall k$, regularization parameter $\tau = 10^{-8}$, pilot maximum reuse time of $n_{\max} = 2$. Each UE is assumed to select the RRHs with the L highest large-scale channel gains as its candidate serving set, i.e., $|\mathcal{I}_k| = L, \forall k$. The fronthaul capacity constraint for each UE is assumed to be the same, i.e., $C_{\max} = C_{i,\max}, \forall i$.

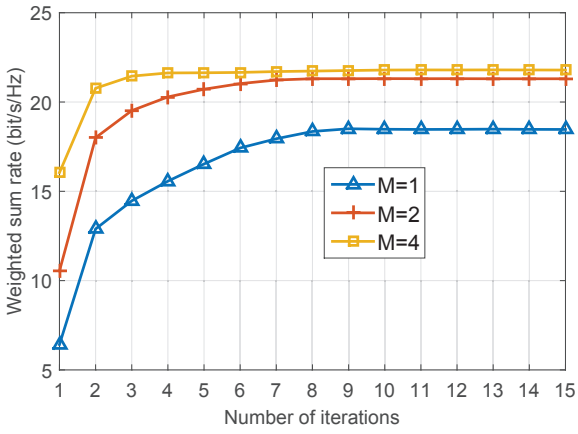


Fig. 3. Convergence behaviour of Algorithm 1 for different number of transmit antennas, where fronthaul capacity is set as $C_{\max} = 2$ bit/s/Hz.

Fig. 3 illustrates the convergence behaviour Algorithm 1 for different number of transmit antennas. It can be observed from this figure that the weighted sum rate monotonically increases with the number of iterations and converges rapidly for all numbers of transmit antennas considered. In general, all curves converge within ten iterations. It is interesting to observe that the convergence speed increases with the number of antennas. In particular, nine iterations are required for the case of $M = 1$ to converge, while four iterations are sufficient for the case of $M = 4$. As expected, the algorithm with larger number of transmit antennas will converge to a larger weighted sum rate since higher spatial degrees of freedom can be exploited. However, we find that significant gains are achieved by $M = 2$

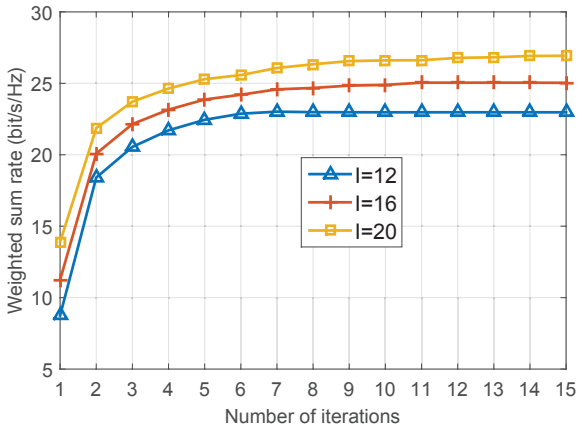


Fig. 4. Convergence behaviour of Algorithm 1 for different number of RRHs, where fronthaul capacity is set as $C_{\max} = 2$ bit/s/Hz.

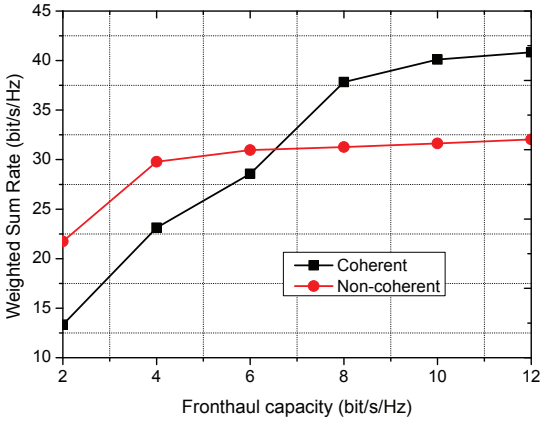


Fig. 5. Weighted sum rate versus the fronthaul capacity $C_{\max}$ for different transmission schemes.

over that of $M = 1$, while only marginal performance gains can be obtained by the case of $M = 4$ over that of $M = 2$, which comes at a higher computational complexity. Hence, this means that in ultra-dense C-RANs, the number of transmit antennas at the RRH does not have to be high.

Fig. 4 investigates the impact of the number of RRHs on the convergence behavior of the proposed algorithm, where the number of UEs is set to $K = 12$. It is observed again that the algorithm converges for $I = 12, 16, 20$, and ten iterations are sufficient for the algorithm to converge. As expected, a larger number of RRHs leads to a higher weighted sum rate due to the multi-RRH diversity. In contrast to the case seen in Fig. 3, the impact of the number of RRHs on the convergence speed is not obvious in the examples considered. The similar trend has observed for the study of the impacts of the number of UEs on the convergence speed, which is omitted here for brevity.

Next, we compare our proposed non-coherent joint transmission with the coherent transmission, where the RRHs in each UE's serving cluster are transmitting the same signal to the UE. In Fig. 5, we plot the weighted sum rate versus the fronthaul capacity $C_{\max}$ for these two transmission schemes. As expected, the weighted sum rate achieved by both schemes

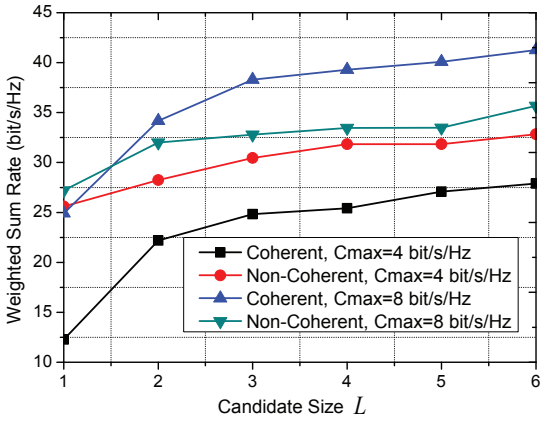


Fig. 6. Weighted sum rate versus the candidate size L for different transmission schemes.

increases with $C_{\max}$. We can observe from Fig. 5 that it is difficult to judge which transmission scheme is superior over the other one over the overall $C_{\max}$ regime. For the example in Fig. 5, in the low $C_{\max}$ regime when $C_{\max} < 6$bit/s/Hz, the non-coherent transmission significantly outperforms the coherent one. The reason can be explained as follows. In the low $C_{\max}$ regime, only a limited number of UEs can be supported by the coherent transmission scheme since the RRHs for the serving cluster have to transmit the same signal to the same UE, and the maximum data rate achieved by the UE is limited to $C_{\max}$. However, for the non-coherent transmission scheme, since different RRHs in the UE's serving cluster can transmit different signals to the same UE, and the data rate achieved by each UE is the summation of the data rates from all the RRHs in its serving cluster. Hence, the maximum data rate achieved by each UE can be up to $LC_{\max}$. However, in the high $C_{\max}$ regime, the coherent transmission scheme produces higher weighted sum rate than the non-coherent one. This is due to the fact that in the high $C_{\max}$ regime, the fronthaul capacity is abundant such that the benefits of coherent transmission can be exploited, where all RRHs in the serving cluster can be regarded as a large-dimensional base station. However, for the non-coherent transmission, the data rate from each RRH is very low due to the additional interference from other RRHs in the same cluster. Hence, the summation of these data rates will be smaller than that of the coherent one. Another interesting phenomenon is that the non-coherent transmission scheme saturates rapidly with the fronthaul capacity constraint, and increases slowly when $C_{\max} > 6$ bit/s/Hz.

Fig. 6 shows the weighted sum rate versus the candidate size L for different transmission schemes with different $C_{\max}$. It is observed from Fig. 6 that the weighted sum rate achieved by all schemes increase with L due to the fact that more spatial degrees of freedom are available for each UE. As expected, the larger $C_{\max}$ leads to higher weighted sum rate for both schemes. We can also find from Fig. 6 that the non-coherent transmission scheme converges rapidly than the the coherent one, and keeps steady when $L > 3$. When $C_{\max} = 4$ bit/s/Hz, the non-coherent transmission scheme has superior performance over the coherent one, and vice

versa.

V. Conclusions

This paper studied the weighted sum-rate maximization problem of non-coherent ultra-dense C-RANs in the face of imperfect CSI, where fronthaul capacity constraints are imposed. We first derived the closed-form expression of the achievable individual data rate by exploiting the statistical properties of the channel estimation error. To solve this optimization problem, the l_0-norm in the fronthaul capacity constraints was approximated as the weighted power constraints. Then, a low-complexity SCA algorithm was proposed for solving the resultant optimization problem, along with a novel initialization method. Our simulation results illustrated that the non-coherent scheme outperforms its coherent counterpart, when the fronthaul capacity limit is low. Hence, the non-coherent regime is an appealing scheme in ultra-dense C-RANs, when the mmWave fronthaul has a limited capacity.

References

[1] J. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. Soong, and J. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[2] Z. Gao, L. Dai, D. Mi, Z. Wang, M. A. Imran, and M. Z. Shakir, "Mmwave massive-MIMO-based wireless backhaul for the 5G ultra-dense network," *IEEE Wireless Communications*, vol. 22, no. 5, pp. 13–21, October 2015.

[3] N. Wang, E. Hossain, and V. K. Bhargava, "Joint downlink cell association and bandwidth allocation for wireless backhauling in two-tier HetNets with large-scale antenna arrays," *IEEE Transactions on Wireless Communications*, vol. 15, no. 5, pp. 3251–3268, May 2016.

[4] R. G. Stephen and R. Zhang, "Joint millimeter-wave fronthaul and OFDMA resource allocation in ultra-dense CRAN," *IEEE Transactions on Communications*, vol. 65, no. 3, pp. 1411–1423, March 2017.

[5] A. Abdelnasser and E. Hossain, "Resource allocation for an OFDMA Cloud-RAN of small cells underlaying a macrocell," *IEEE Transactions on Mobile Computing*, vol. 15, no. 11, pp. 2837–2850, Nov 2016.

[6] B. Dai and W. Yu, "Sparse beamforming and user-centric clustering for downlink cloud radio access network," *IEEE Access*, vol. 2, pp. 1326–1339, 2014.

[7] V. N. Ha, L. B. Le, and N. D. Dao, "Coordinated multipoint transmission design for Cloud-RANs with limited fronthaul capacity constraints," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 9, pp. 7432–7447, Sept 2016.

[8] 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 Journal on Selected Areas in Communications*, vol. 35, no. 8, pp. 1809–1824, Aug. 2017.

[9] C. Pan, H. Mehrpouyan, Y. Liu, M. Elkashlan, and A. Nallanathan, "Joint pilot allocation and robust transmission design for ultra-dense user-centric TDD C-RAN with imperfect CSI," *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 2038–2053, Mar. 2018.

[10] 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.

[11] 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," *arXiv preprint arXiv:1804.03990*, 2018.

[12] J. Zhao, T. Q. S. Quek, and Z. Lei, "Coordinated multipoint transmission with limited backhaul data transfer," *IEEE Transactions on Wireless Communications*, vol. 12, no. 6, pp. 2762–2775, June 2013.

[13] B. Dai and W. Yu, "Energy efficiency of downlink transmission strategies for cloud radio access networks," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 1037–1050, April 2016.

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

[15] Y. Liu, Y. Dai, and Z. Luo, "Coordinated beamforming for MISO interference channel: Complexity analysis and efficient algorithms," *IEEE Transactions on Signal Processing*, vol. 59, no. 3, pp. 1142–1157, March 2011.

[16] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.

[17] C. Pan, H. Ren, M. Elkashlan, A. Nallanathan, and L. Hanzo, "Weighted sum-rate maximization for ultra-dense user-centric TDD C-RAN with imperfect CSI," *To appear in arXiv*, 2018.

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

[19] M. Peng, Y. Yu, H. Xiang, and H. V. Poor, "Energy-efficient resource allocation optimization for multimedia heterogeneous cloud radio access networks," *IEEE Transactions on Multimedia*, vol. 18, no. 5, pp. 879–892, May 2016.

[20] J. Li, J. Wu, M. Peng, and P. Zhang, "Queue-aware energy-efficient joint remote radio head activation and beamforming in cloud radio access networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 6, pp. 3880–3894, June 2016.

[21] L. N. Tran, M. F. Hanif, A. Tolli, and M. Juntti, "Fast converging algorithm for weighted sum rate maximization in multicell MISO downlink," *IEEE Signal Processing Letters*, vol. 19, no. 12, pp. 872–875, Dec 2012.

[22] T. R. Lakshmana, A. Tlli, R. Devassy, and T. Svensson, "Precoder design with incomplete feedback for joint transmission," *IEEE Transactions on Wireless Communications*, vol. 15, no. 3, pp. 1923–1936, March 2016.

[23] Z. Chen, X. Hou, and C. Yang, "Training resource allocation for user-centric base station cooperation networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 4, pp. 2729–2735, Apr. 2016.

[24] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear estimation*. Prentice Hall Upper Saddle River, NJ, 2000, vol. 1.

[25] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2845–2866, Jun. 2010.

[26] M. Xu, D. Guo, and M. L. Honig, "Downlink noncoherent cooperation without transmitter phase alignment," *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 3920–3931, Aug. 2013.

[27] J. Li, E. Bjornson, T. Svensson, T. Eriksson, and M. Debbah, "Joint precoding and load balancing optimization for energy-efficient heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, pp. 5810–5822, Oct. 2015.

[28] T. V. Chien, E. Bjornson, and E. G. Larsson, "Joint power allocation and user association optimization for massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6384–6399, Sep. 2016.

[29] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.

[30] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

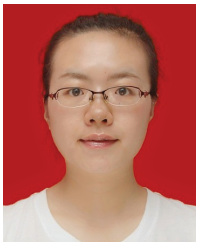
[31] B. K. Sriperumbudur, D. A. Torres, and G. R. Lanckriet, "A majorization-minimization approach to the sparse generalized eigenvalue problem," *Machine learning*, vol. 85, no. 1-2, pp. 3–39, 2011.

[32] X. Ge, S. Tu, G. Mao, C. X. Wang, and T. Han, "5G ultra-dense cellular networks," *IEEE Wireless Commun. Mag.*, vol. 23, no. 1, pp. 72–79, Feb. 2016.

[33] E. U. T. R. Access, "Further advancements for E-UTRA physical layer aspects," *3GPP TR 36.814, Tech. Rep.*, 2010.

Cunhua Pan received his B.S. and Ph.D. degrees in school of Information Science and Engineering, Southeast University, Nanjing, China, in 2010 and 2015, respectively. From 2015 to 2016, he worked as a research associate in University of Kent, UK. He is currently a postdoc with Queen Mary University of London, UK.

His research interests mainly include Ultra-dense C-RAN, machine learning, UAV, internet of things (IoT), and mobile edge computing. He serves as Editor for IEEE ACCESS, Student Travel Grant Chair for ICC 2019, and TPC Member for many conferences, such as ICC and Globecom.



Hong Ren received the B.S. degree from Southwest Jiaotong University, in Electrical Engineering in 2011, Chengdu, China, and the M.S. and Ph.D. degrees from Southeast University, Nanjing, China, in Electrical Engineering in 2014 and 2018, respectively. From October 2016 to Jan 2018, she was a visiting student in the School of Electronics and Computer Science, University of Southampton, UK. She is currently a postdoctoral scholar with the School of Electronic Engineering and Computer Science, Queen Mary University of London, UK. Her research interests lie in the areas of communication and signal processing, including green communication systems, cooperative transmission and cross layer transmission optimization.

Maged Elkashlan received the Ph.D. degree in Electrical Engineering from the University of British Columbia, Canada, 2006. From 2007 to 2011, he was with the Commonwealth Scientific and Industrial Research Organization (CSIRO), Australia. During this time, he held visiting appointments at University of New South Wales and University of Technology Sydney. In 2011, he joined the School of Electronic Engineering and Computer Science at Queen Mary University of London, UK. His research interests fall into the broad areas of communication theory and statistical signal processing.

Dr. Elkashlan currently serves as Editor of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. He received the Best Paper Awards at the IEEE International Conference on Communications (ICC) in 2016 and 2014, the International Conference on Communications and Networking in China (CHINACOM) in 2014, and the IEEE Vehicular Technology Conference (VTC-Spring) in 2013.

Arumugam Nallanathan is Professor of Wireless Communications and Head of the Communication Systems Research (CSR) group in the School of Electronic Engineering and Computer Science at Queen Mary University of London since September 2017. He was with the Department of Informatics at Kings College London from December 2007 to August 2017, where he was Professor of Wireless Communications from April 2013 to August 2017 and a Visiting Professor from September 2017. He was an Assistant Professor in the Department of Electrical and Computer Engineering, National University of Singapore from August 2000 to December 2007. His research interests include 5G Wireless Networks, Internet of Things (IoT) and Molecular Communications. He published nearly 400 technical papers in scientific journals and international conferences. He is a co-recipient of the Best Paper Awards presented at the IEEE International Conference on Communications 2016 (ICC'2016) , IEEE Global Communications Conference 2017 (GLOBECOM'2017) and IEEE Vehicular Technology Conference 2018 (VTC'2018). He is an IEEE Distinguished Lecturer. He has been selected as a Web of Science Highly Cited Researcher in 2016.

He is an Editor for IEEE Transactions on Communications. He was an Editor for IEEE Transactions on Wireless Communications (2006-2011), IEEE Transactions on Vehicular Technology (2006-2017), IEEE Wireless Communications Letters and IEEE Signal Processing Letters. He served as the Chair for the Signal Processing and Communication Electronics Technical Committee of IEEE Communications Society and Technical Program Chair and member of Technical Program Committees in numerous IEEE conferences. He received the IEEE Communications Society SPCE outstanding service award 2012 and IEEE Communications Society RCC outstanding service award 2014.

Lajos Hanzo (http://www-mobile.ecs.soton.ac.uk) FREng, FIEEE, FIET, Fellow of EURASIP, DSc received his degree in electronics in 1976 and his doctorate in 1983. In 2009 he was awarded an honorary doctorate by the Technical University of Budapest and in 2015 by the University of Edinburgh. In 2016 he was admitted to 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 UK. Since 1986 he has been with the School of Electronics and Computer Science, University of Southampton, UK, where he holds the chair in telecommunications. He has successfully supervised 112 PhD students, co-authored 18 John Wiley/IEEE Press books on mobile radio communications totalling in excess of 10 000 pages, published 1776 research contributions at IEEE Xplore, acted both as TPC and General Chair of IEEE conferences, presented keynote lectures and has been awarded a number of distinctions. Currently he is directing an 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 (EPSRC) UK, 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 ComSoc and VTS. During 2008 - 2012 he was the Editor-in-Chief of the IEEE Press and a Chaired Professor also at Tsinghua University, Beijing. For further information on research in progress and associated publications please refer to http://www-mobile.ecs.soton.ac.uk