Joint Design of Fronthauling and Hybrid Beamforming for Downlink C-RAN Systems

Jaein Kim, *Student Member, IEEE*, Seok-Hwan Park, *Member, IEEE*, Osvaldo Simeone, *Fellow, IEEE*, Inkyu Lee, *Fellow, IEEE*, and Shlomo Shamai (Shitz), *Life Fellow, IEEE*

Abstract—Hybrid beamforming is known to be a cost-effective and wide-spread solution for a system with large-scale antenna arrays. This paper studies the optimization of the analog and digital components of the hybrid beamforming solution for remote radio heads (RRHs) in a downlink cloud radio access network architecture. Digital processing is carried out at a baseband processing unit (BBU) in the "cloud," and the precoded baseband signals are quantized prior to transmission to the RRHs via finite-capacity fronthaul links. In this system, we consider two different channel state information (CSI) scenarios: 1) ideal CSI at the BBU and 2) imperfect effective CSI. The optimization of digital beamforming and fronthaul quantization strategies at the BBU as well as analog radio-frequency (RF) beamforming at the RRHs is a coupled problem since the effect of the quantization noise at the receiver depends on the precoding matrices. The resulting joint optimization problem is examined with the goal of maximizing the weighted downlink sum-rate and the network energy efficiency. Fronthaul capacity and per-RRH power constraints are enforced along with constant modulus constraint on the RF beamforming matrices. For the case of perfect CSI, a block coordinate descent scheme is proposed based on the weighted minimum-mean-square-error approach by relaxing the constant modulus constraint of the analog beamformer. Also, we present the impact of imperfect CSI on the weighted sum-rate and network energy efficiency performance, and the algorithm is extended by applying the sample average approximation. The numerical results confirm the effectiveness of the proposed scheme and show that the proposed algorithm is robust to estimation errors.

Index Terms—Cloud-RAN, massive MIMO, hybrid beamforming, fronthaul compression, imperfect CSI.

I. INTRODUCTION

MASSIVE multiple-input and multiple-output (MIMO) has been regarded as a promising technology for future wireless systems owing to its potential of improving both spectral and energy efficiency (EE) with simple signal processing [2]. This is enabled by the fact that the channel vectors for different users become orthogonal when the number of transmit antennas grows to infinity. However, with massive MIMO arrays, it is generally impractical to equip every antenna of a large array with a radio frequency (RF) chain due to hardware limitations [3]. Hybrid beamforming techniques, whereby the beamforming process consists of a low-dimensional digital beamforming followed by analog RF beamforming, has emerged as an effective means to address this problem (see, e.g., [4]–[13]). Both analog and digital components are typically designed separately and locally for a base station (BS) [3].

In a cloud radio access network (C-RAN) architecture, the baseband signal processing functionalities of multiple BSs are migrated to a baseband processing unit (BBU) in the "cloud", while RF functionalities are implemented at distributed remote radio heads (RRHs). Therefore, in the C-RAN architecture with large antenna arrays at the RRHs, digital precoding across multiple RRHs can be carried out at the BBU, while RF beamforming is performed locally at each RRH. The design problem becomes more challenging by the capacity limitations of the fronthaul links that connect the BBU to the RRHs.

In the downlink of C-RAN, the BBU performs joint encoding and precoding of the messages intended for user equipments (UEs), and then the produced baseband signals are quantized and compressed prior to being transferred to the RRH via fronthaul links. The design of precoding and fronthaul compression strategies has been studied in [14]–[16]. Specifically, the authors in [14] considered the standard point-to-point fronthaul compression strategies. In [15], [16], the authors investigated multivariate fronthaul compression.

In this work, we study the application of hybrid beamforming to the C-RAN architecture. We tackle the problem of jointly optimizing digital baseband beamforming and fronthaul compression strategies at the BBU along with RF beamforming at the RRHs with the goal of maximizing the weighted downlink sum-rate and the network EE. Fronthaul capacity and per-RRH transmit power constraints are imposed, as well as constant modulus constraint on the RF beamforming matrices which consist of analog phase shifters [3].

The limited number of RF chains determines the capability of the BBU to acquire channel state information (CSI) through conventional uplink training based on the time division duplex (TDD) operation. In particular, during the uplink training, the received baseband signal depends on a RF beamforming matrix, and hence instantaneous CSI is unavailable when designing the RF beamforming matrices. To address this limitation, the RF beamforming matrices are computed based on the second-order statistics of the downlink channel vectors, and the digital beamforming and fronthaul compression strategies are adaptive to the estimated effective channel.

A. Related Work

A hybrid beamforming design has been investigated in [4]–[13]. Specifically, in [4], a point-to-point hybrid precoding and combining algorithm was proposed that uses orthogonal matching pursuit for millimeter-wave (mmWave) systems. The authors in [6] provided a low-complexity hybrid beamforming scheme to achieve sum-rate performance close to that of the zero-forcing (ZF) method for the downlink of multiuser multiple-input single-output (MISO) systems. In this case, each RF beamforming vector for a user was determined by projecting the downlink channel onto the feasible RF space with low-dimensional ZF digital beamforming. In addition, for multi-user MIMO mmWave systems, a limited feedback hybrid beamforming scheme was presented in [7]. The work in [8] proved that in hybrid beamforming, the number of RF chains needs to be twice the number of data streams to achieve sum-rate performance equal to that of fully digital beamforming. Also, the authors in [8] considered a design of the hybrid beamforming to maximize spectral efficiency for point-to-point MIMO and multi-user MISO scenarios.

Most of works on hybrid beamforming in [4]–[10] have assumed full CSI. However, it is difficult to estimate the channel vectors across all antenna elements, since the estimation operates in the low-dimensional baseband downlink obtained after RF beamforming. To address this issue, in [11] and [12], the RF beamforming matrix was determined by using the long-term CSI, while the digital beamformer was designed based on the low-dimensional effective channel. Recently, a design of hybrid beamforming for C-RAN systems has been studied in [17] and [18]. The authors in [17] provided a two-stage algorithm that only demands low-dimensional effective CSI. In [18], a RF beamforming was computed based on a weighted sum of second-order channel statistics and the size of RF and digital beamforming matrices was determined in order to maximize the large-scale approximated sum-rate with regularized ZF digital beamforming.

Furthermore, the EE maximization problem in C-RAN has been studied in [19]–[21]. In [19], the authors considered a joint design of beamforming, virtual computing resources, RRH selection, and RRH-UE association in a limited fronthaul C-RAN, and a global optimization algorithm and a low-complexity method were presented. Also, for both single-hop and multi-hop C-RAN scenarios, the problem of EE maximization under both data-sharing and compression-based fronthaul strategies was addressed in [20]. The authors in [21]

TABLE I
DEFINITION OF VARIABLES

Variable	Definition
$\mathbf{V}_{R,i}$	RF beamforming matrix at the ith RRH
$\mathbf{v}_{D,k}$	Digital beamforming vector for the kth UE
$\mathbf{\Omega}_i$	Quantization noise covariance matrix at the ith RRH
u_k	MMSE receiver at the kth UE
$\tilde{w}_k$	Weight variable for the kth UE
$\mathbf{\Sigma}_i$	Auxiliary matrix at the ith RRH
ρ	Auxiliary variable for power consumption
α_t	Additional variable at the tth instantaneous channel sample
β_t	Auxiliary variable at the tth instantaneous channel sample

took into account a realistic power consumption model which is dependent on the data rate and dynamic power amplifier.

B. Main Contributions, Paper Organization and Notation

The main contributions of this paper are as follows:

- For downlink C-RAN with hybrid analog-digital antenna arrays, we investigate the joint design of fronthauling and hybrid beamforming with the goal of maximizing the weighted sum-rate (WSR) and the network EE.
- For the case of perfect CSI, we first decompose the problem into two sub-problems of the RF beamforming and the digital processing, which turn out to be non-convex, and then we propose an iterative algorithm based on weighted minimum-mean-square-error (WMMSE) approach by relaxing constant modulus constraint. Also, for the imperfect CSI case, we extend the solution with the perfect CSI by applying the sample average approximation (SAA) [22].
- Extensive numerical results are provided to validate the effectiveness of the proposed algorithm. In addition, in the presence of channel estimation errors, we show the robustness of the proposed scheme.

The paper is organized as follows: In Sec. II, we present the system model for the downlink of a C-RAN with hybrid digital and analog processing and finite-capacity fronthaul links. In Sec. III, for the case of perfect CSI, we describe the problems of WSR maximization and network EE maximization, and an iterative algorithm to tackle the problems is proposed. Sec. IV discusses the problem of CSI estimation in a TDD system. In addition, we introduce an uplink channel training method and provide a RF beamforming matrix design and the digital strategies. Numerical results are illustrated in Sec. V. The paper is closed with the conclusion in Sec. VI.

Throughout this paper, boldface uppercase, boldface lowercase and normal letters indicate matrices, vectors and scalars, respectively. The operators $(\cdot)^T$, $(\cdot)^H$, $\mathbb{E}(\cdot)$, $\det(\cdot)$ and $\text{tr}(\cdot)$ represent transpose, conjugate transpose, expectation, determinant and trace, respectively. A circularly symmetric complex Gaussian distribution with mean μ and covariance matrix $\mathbf{R}$ is denoted by $\mathcal{CN}(\mu, \mathbf{R})$. The set of all $M \times N$ complex matrices is defined as $\mathbb{C}^{M \times N}$. $\mathbf{I}_d$ represents an identity matrix of size d. $\otimes$ stands for the Kronecker product. The variables used in this paper are summarized in Table I.

II. SYSTEM MODEL

As illustrated in Fig. 1, we consider the downlink of a C-RAN system in which a BBU communicates with N_U
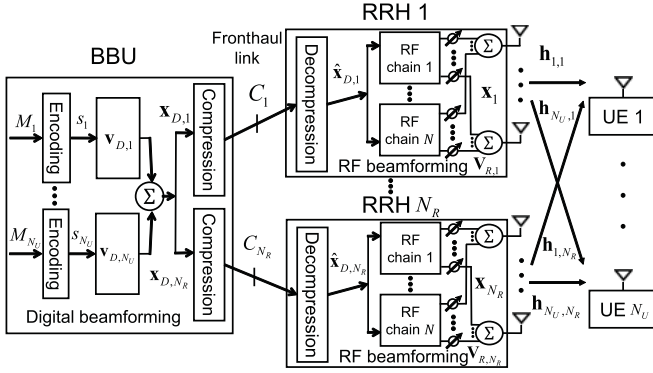
Fig. 1. Illustration of the downlink of a C-RAN with hybrid digital and analog processing.

single-antenna UEs through N_R RRHs, each equipped with M transmit antennas. We assume that the ith RRH is connected to the BBU via an error-free digital fronthaul link of capacity C_i bps/Hz [23], [24], and each RRH is equipped with $N \leq M$ RF chains due to cost limitations. This implies that fully digital beamforming across M transmit antennas of each RRH is not enabled [4]–[8] and thus hybrid analog-digital solutions are in order. For convenience, we define the sets $\mathcal{R} \triangleq \{1, \cdots, N_R\}$, $\mathcal{K} \triangleq \{1, \cdots, N_U\}$, $\mathcal{M} \triangleq \{1, \cdots, M\}$ and $\mathcal{N} \triangleq \{1, \cdots, N\}$. We assume that user scheduling is predetermined and hence all the N_U UEs are active. For rate allocation, the priority among N_U active UEs can be controlled by adjusting the weights of the weighted sum-rate in Sec. III-A. The transmit-side baseband processing is centralized at the BBU based on CSI reported by the RRHs on the fronthaul links. Assuming a TDD operation, each RRH obtains its local CSI by means of uplink channel training [25], [26].

A. Channel Model

For the downlink channel from the RRHs to the UEs, we adopt a frequency-flat fading channel model such that the received signal y_k at the kth UE is given as

$$y_k = \sum_{i \in \mathcal{R}} \mathbf{h}_{k,i}^H \mathbf{x}_i + z_k = \mathbf{h}_k^H \mathbf{x} + z_k, \qquad (1)$$

where $\mathbf{x}_i \in \mathbb{C}^{M \times 1}$ is the transmitted signal of the ith RRH subject to the transmit power constraint $\mathbb{E}\|\mathbf{x}_i\|^2 \leq P_i$, $\mathbf{h}_{k,i} \in \mathbb{C}^{M \times 1}$ equals the channel vector from the ith RRH to the kth UE, which is distributed as $\mathbf{h}_{k,i} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_{k,i})$ with $\mathbf{R}_{k,i}$ being the transmit-side correlation, $\mathbf{x} = [\mathbf{x}_1^H \cdots \mathbf{x}_{N_R}^H]^H \in \mathbb{C}^{MN_R \times 1}$ represents the signal vector transmitted by all RRHs, $\mathbf{h}_k = [\mathbf{h}_{k,1}^H \cdots \mathbf{h}_{k,N_R}^H]^H \in \mathbb{C}^{MN_R \times 1}$ indicates the channel vector from all RRHs to the kth UE, and $z_k \sim \mathcal{CN}(0, \sigma_D^2)$ denotes the additive noise at the kth UE.

B. Digital Beamforming and Fronthaul Compression

We define the message intended for the kth UE as $M_k \in \{1, \cdots, 2^{nR_k}\}$, where n stands for the coding block length and R_k is the rate of M_k. The BBU encodes the message M_k into the baseband signal $s_k \sim \mathcal{CN}(0, 1)$ for $k \in \mathcal{K}$ using a standard random Gaussian channel code. Then, in order to

manage inter-UE interference, the signals $\{s_k \text{ for } k \in \mathcal{K}\}$ are linearly precoded as

$$\mathbf{x}_D = [\mathbf{x}_{D,1}^H \cdots \mathbf{x}_{D,N_R}^H]^H = \sum_{k \in \mathcal{K}} \mathbf{v}_{D,k} s_k, \qquad (2)$$

where $\mathbf{v}_{D,k} \in \mathbb{C}^{NN_R \times 1}$ is the *digital beamforming* vector across all the RRHs for the kth UE, and $\mathbf{x}_{D,i} \in \mathbb{C}^{N \times 1}$ represents the ith subvector of $\mathbf{x}_D \in \mathbb{C}^{NN_R \times 1}$ corresponding to the signal transmitted by the ith RRH. Defining the shaping matrices $\boldsymbol{\Xi}_i = [\mathbf{0}_{N \times N(i-1)}^H \quad \mathbf{I}_N \quad \mathbf{0}_{N \times N(N_R-i)}^H]^H$, the ith subvector $\mathbf{x}_{D,i}$ can be expressed as $\mathbf{x}_{D,i} = \sum_{k \in \mathcal{K}} \boldsymbol{\Xi}_i^H \mathbf{v}_{D,k} s_k$.

Since the BBU communicates with the ith RRH via a fronthaul link of finite capacity, the signal $\mathbf{x}_{D,i}$ is quantized and compressed prior to being transferred to the RRH. Following the approaches in [27] and [28], we model the impact of the compression by writing the quantized signal $\hat{\mathbf{x}}_{D,i}$ as

$$\hat{\mathbf{x}}_{D,i} = \mathbf{x}_{D,i} + \mathbf{q}_i, \qquad (3)$$

where the quantization noise $\mathbf{q}_i \in \mathbb{C}^{N \times 1} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Omega}_i)$ is independent of the signal $\mathbf{x}_{D,i}$. As for the standard information-theoretic formulation, the covariance matrix $\boldsymbol{\Omega}_i$ describes the effect of the quantizer. From [29, Ch. 3], the quantized signal $\hat{\mathbf{x}}_{D,i}$ can be reliably recovered at the ith RRH, if the following condition is satisfied

$$g_i(\mathbf{V}_D, \boldsymbol{\Omega}_i) \triangleq I(\mathbf{x}_{D,i}; \hat{\mathbf{x}}_{D,i}) \qquad (4)$$
$$= \log_2 \det \Big(\sum_{k \in \mathcal{K}} \boldsymbol{\Xi}_i^H \mathbf{v}_{D,k} \mathbf{v}_{D,k}^H \boldsymbol{\Xi}_i + \boldsymbol{\Omega}_i \Big) - \log_2 \det(\boldsymbol{\Omega}_i) \leq C_i,$$

where we define the set of the digital beamforming vectors as $\mathbf{V}_D \triangleq \{\mathbf{v}_{D,k} \text{ for } k \in \mathcal{K}\}$.

C. RF Beamforming

The quantized signal vector $\hat{\mathbf{x}}_{D,i}$ decompressed at the ith RRH is of dimension N, and is input to one of N RF chains. In order to fully utilize $M > N$ transmit antennas, the ith RRH applies *analog RF beamforming* to the signal $\hat{\mathbf{x}}_{D,i}$ via the beamforming matrix $\mathbf{V}_{R,i} \in \mathbb{C}^{M \times N}$. The RF beamforming obtains M signals for the antenna as a combination of N output of the RF chains. As a result, the transmitted signal $\mathbf{x}_i$ from M transmit antennas is given as

$$\mathbf{x}_i = \mathbf{V}_{R,i} \hat{\mathbf{x}}_{D,i} = \sum_{k \in \mathcal{K}} \mathbf{V}_{R,i} \boldsymbol{\Xi}_i^H \mathbf{v}_{D,k} s_k + \mathbf{V}_{R,i} \mathbf{q}_i. \qquad (5)$$

As summarized in [30], the RF beamforming can be implemented using analog phase shifters and switches. Accordingly, each RF chain is connected to a specific set of transmit antennas through a phase shifter. In this paper, we consider a fully connected phase shifter architecture, whereby each RF chain is connected to all transmit antennas via a separate phase shifter. In the fully connected phase shifter architecture, the (a, b)th element of the RF beamforming matrix $\mathbf{V}_{R,i}$ is expressed as $\mathbf{V}_{R,i,a,b} = e^{j\theta_{i,a,b}}$ for $a \in \mathcal{M}$ and $b \in \mathcal{N}$, where $\theta_{i,a,b}$ indicates the phase shift between the signals $\hat{\mathbf{x}}_{D,i,b}$ and $\mathbf{x}_{i,a}$. Therefore, when designing the RF beamforming matrix $\mathbf{V}_{R,i}$, one should satisfy constant modulus constraint $|\mathbf{V}_{R,i,a,b}|^2 = 1$ for $a \in \mathcal{M}$ and $b \in \mathcal{N}$ (see, e.g., [8]).

III. DESIGN WITH PERFECT CSI

In this section, we discuss the problem of jointly designing the beamforming matrices for RF and digital beamforming, along with the fronthaul quantization noise covariance matrices. As we will see, these problems are interdependent, since the impact of the quantization noise on the receivers' performance depends on the beamforming matrices. Here, we first consider the case of perfect CSI, while the system with imperfect CSI will be addressed in Sec. IV.

To measure the achievable rate for each UE k, we rewrite the signal y_k in (1) under the transmission model (5) as

$$y_k = \sum_{l \in \mathcal{K}} \mathbf{h}_k^H \bar{\mathbf{V}}_R \mathbf{v}_{D,l} s_l + \mathbf{h}_k^H \bar{\mathbf{V}}_R \mathbf{q} + z_k, \qquad (6)$$

where $\bar{\mathbf{V}}_R \triangleq \mathrm{diag}(\mathbf{V}_{R,i}, \cdots, \mathbf{V}_{R,N_R})$ indicates the effective RF beamforming matrix across all RRHs and $\mathbf{q} \triangleq [\mathbf{q}_1^H \cdots \mathbf{q}_{N_R}^H]^H \in \mathbb{C}^{NN_R \times 1}$ stands for the vector of all the quantization noise signals, which is distributed as $\mathbf{q} \sim \mathcal{CN}(\mathbf{0}, \bar{\mathbf{\Omega}})$ with $\bar{\mathbf{\Omega}} \triangleq \mathrm{diag}(\mathbf{\Omega}_1, \cdots, \mathbf{\Omega}_{N_R})$.

Assuming that UE k decodes the message M_k by treating the interference signals as the additive noise, the achievable rate R_k for the kth UE is given as

$$\begin{aligned} R_k &= f_k(\mathbf{V}_R, \mathbf{V}_D, \mathbf{\Omega}) = I(s_k; y_k) \\ &= \log_2 \det(|\mathbf{h}_k^H \bar{\mathbf{V}}_R \mathbf{v}_{D,k}|^2 + \zeta_k(\mathbf{V}_R, \mathbf{V}_D, \mathbf{\Omega})) \\ &\quad - \log_2 \det(\zeta_k(\mathbf{V}_R, \mathbf{V}_D, \mathbf{\Omega})), \end{aligned} \qquad (7)$$

where $\mathbf{V}_R \triangleq \{\mathbf{V}_{R,i} \text{ for } i \in \mathcal{R}\}$, $\mathbf{\Omega} \triangleq \{\mathbf{\Omega}_i \text{ for } i \in \mathcal{R}\}$, and we denote the function $\zeta_k(\mathbf{V}_R, \mathbf{V}_D, \mathbf{\Omega}) \triangleq \sum_{l \in \mathcal{K} \backslash \{k\}} |\mathbf{h}_k^H \bar{\mathbf{V}}_R \mathbf{v}_{D,l}|^2 + \mathbf{h}_k^H \bar{\mathbf{V}}_R \bar{\mathbf{\Omega}} \bar{\mathbf{V}}_R^H \mathbf{h}_k + \sigma_D^2$.

Considering the power consumption in the C-RAN, the total power consumption can be modeled as [31]

$$\begin{aligned} P_T(\mathbf{V}_R, \mathbf{V}_D, \mathbf{\Omega}) &\triangleq \sum_{i \in \mathcal{R}} p_i(\mathbf{V}_{R,i}, \mathbf{V}_D, \mathbf{\Omega}) \\ &\quad + N_U P_{N_U} + NN_R P_{RF}, \end{aligned} \qquad (8)$$

where the transmission power of the ith RRH is obtained as

$$\begin{aligned} p_i(\mathbf{V}_{R,i}, \mathbf{V}_D, \mathbf{\Omega}_i) &\triangleq \mathbb{E}\|\mathbf{x}_i\|^2 &&(9) \\ &= \sum_{k \in \mathcal{K}} \mathrm{tr}(\mathbf{V}_{R,i} \mathbf{\Xi}_i^H \mathbf{v}_{D,k} \mathbf{v}_{D,k}^H \mathbf{\Xi}_i \mathbf{V}_{R,i}^H) + \mathrm{tr}(\mathbf{V}_{R,i} \mathbf{\Omega}_i \mathbf{V}_{R,i}^H), \end{aligned}$$

P_{N_U} is the circuit power consumed by a UE, and P_{RF} represents the circuit power consumption at each RRH, which is proportional to the number of RF chains.

A. Weighted Sum-Rate Maximization

In this work, we tackle the problem of maximizing the WSR $\sum_{k \in \mathcal{K}} w_k R_k$ of the UEs while satisfying the per-RRH transmit power, fronthaul capacity and constant modulus constraints, where $w_k \geq 0$ is a weight denoting the priority for the kth UE. The problem is stated as

$$\underset{\mathbf{V}_R, \mathbf{V}_D, \mathbf{\Omega}}{\text{maximize}} \sum_{k \in \mathcal{K}} w_k f_k(\mathbf{V}_R, \mathbf{V}_D, \mathbf{\Omega}) \qquad (10\text{a})$$

$$\text{s.t. } g_i(\mathbf{V}_D, \mathbf{\Omega}_i) \leq C_i, \ i \in \mathcal{R}, \qquad (10\text{b})$$

$$p_i(\mathbf{V}_{R,i}, \mathbf{V}_D, \mathbf{\Omega}_i) \leq P_i, \ i \in \mathcal{R}, \qquad (10\text{c})$$

$$|\mathbf{V}_{R,i,a,b}|^2 = 1, \ a \in \mathcal{M}, \ b \in \mathcal{N}, \ i \in \mathcal{R}. \qquad (10\text{d})$$

Problem (10) is non-convex due to the objective function (10a) and the constraints (10b), (10c) and (10d). In the next subsection, we present an iterative algorithm that computes an efficient solution of the problem. To address problem (10), we propose an iterative algorithm based on block coordinate descent (BCD), whereby the RF beamforming matrices $\mathbf{V}_R$ and the digital processing strategies $\{\mathbf{V}_D, \mathbf{\Omega}\}$ are alternately optimized. We first describe the optimization of the digital part $\mathbf{V}_D$ and $\mathbf{\Omega}$ for a fixed RF beamforming $\mathbf{V}_R$, and then introduce the optimization of the latter.

1) Optimization of Digital Beamforming and Fronthaul Compression: For a given RF beamforming $\mathbf{V}_R = \mathbf{V}_R'$, problem (10) with respect to the digital beamforming $\mathbf{V}_D$ and the fronthaul compression strategies $\mathbf{\Omega}$ can be written as

$$\underset{\mathbf{V}_D, \mathbf{\Omega}}{\text{maximize}} \sum_{k \in \mathcal{K}} w_k f_k(\mathbf{V}_R', \mathbf{V}_D, \mathbf{\Omega}) \qquad (11\text{a})$$

$$\text{s.t. } g_i(\mathbf{V}_D, \mathbf{\Omega}_i) \leq C_i, \ i \in \mathcal{R}, \qquad (11\text{b})$$

$$p_i(\mathbf{V}_{R,i}', \mathbf{V}_D, \mathbf{\Omega}_i) \leq P_i, \ i \in \mathcal{R}, \qquad (11\text{c})$$

where we eliminate the constant modulus constraint (10d) which is independent of the digital variables $\mathbf{V}_D$ and $\mathbf{\Omega}$. Problem (11) is still non-convex due to the non-convex objective function (11a) and constraint (11b).

To solve this problem, we extend the WMMSE-based algorithm in [32]. To this end, we introduce two convex lower bounds on (11a) and (11b) by applying a similar approach in [32]. Denoting $x^{(\kappa)}$ as the quantity x obtained at the κth iteration of the BCD, a lower bound on the function $f_k(\mathbf{V}_R', \mathbf{V}_D, \mathbf{\Omega})$ in (11a) is written as

$$f_k(\mathbf{V}_R', \mathbf{V}_D, \mathbf{\Omega}) \geq \frac{1}{\ln 2} \gamma_k(\mathbf{V}_R', \mathbf{V}_D, \mathbf{\Omega}, u_k^{(\kappa)}, \tilde{w}_k^{(\kappa)}), \quad (12)$$

where we define the function

$$\begin{aligned} &\gamma_k(\mathbf{V}_R', \mathbf{V}_D, \mathbf{\Omega}, u_k^{(\kappa)}, \tilde{w}_k^{(\kappa)}) \\ &= \ln \tilde{w}_k^{(\kappa)} - \tilde{w}_k^{(\kappa)} e_k(\mathbf{V}_R', \mathbf{V}_D, \mathbf{\Omega}, u_k^{(\kappa)}) + 1, \quad (13) \end{aligned}$$

with arbitrary parameters $\tilde{w}_k^{(\kappa)} \geq 0$ and $u_k^{(\kappa)}$, and the mean squared error (MSE) function is denoted as

$$\begin{aligned} e_k(\mathbf{V}_R', \mathbf{V}_D, \mathbf{\Omega}, u_k^{(\kappa)}) &= |1 - (u_k^{(\kappa)})^* \mathbf{h}_k^H \bar{\mathbf{V}}_R' \mathbf{v}_{D,k}|^2 \\ &\quad + |u_k^{(\kappa)}|^2 \zeta_k(\mathbf{V}_R', \mathbf{V}_D, \mathbf{\Omega}). \quad (14) \end{aligned}$$

Note that the lower bound in (12) is satisfied with equality when the variables $u_k^{(\kappa)}$ and $\tilde{w}_k^{(\kappa)}$ are equal to

$$\begin{aligned} u_k^{(\kappa)} &= \tilde{u}_k(\mathbf{V}_R', \mathbf{V}_D, \mathbf{\Omega}) \\ &\triangleq \frac{\mathbf{h}_k^H \bar{\mathbf{V}}_R' \mathbf{v}_{D,k}}{|\mathbf{h}_k^H \bar{\mathbf{V}}_R' \mathbf{v}_{D,k}|^2 + \zeta_k(\mathbf{V}_R', \mathbf{V}_D, \mathbf{\Omega})}, \quad (15) \end{aligned}$$

$$\tilde{w}_k^{(\kappa)} = \frac{1}{e_k(\mathbf{V}_R', \mathbf{V}_D, \mathbf{\Omega}, u_k^{(\kappa)})}. \quad (16)$$

Furthermore, an upper bound of the function $g_i(\mathbf{V}_D, \mathbf{\Omega}_i)$ in the constraint (11b) is given as

$$g_i(\mathbf{V}_D, \mathbf{\Omega}) \leq \tilde{g}_i(\mathbf{V}_D, \mathbf{\Omega}, \mathbf{\Sigma}_i^{(\kappa)}), \quad (17)$$

for any arbitrary positive definite matrix $\boldsymbol{\Sigma}_i^{(\kappa)}$, where $\tilde{g}_i(\mathbf{V}_D, \boldsymbol{\Omega}, \boldsymbol{\Sigma}_i^{(\kappa)})$ is represented as

$$\tilde{g}_i(\mathbf{V}_D, \boldsymbol{\Omega}_i, \boldsymbol{\Sigma}_i^{(\kappa)}) = \log_2 \det(\boldsymbol{\Sigma}_i^{(\kappa)}) - \log_2 \det(\boldsymbol{\Omega}_i)$$
$$+ \frac{\text{tr}\big((\boldsymbol{\Sigma}_i^{(\kappa)})^{-1}(\sum_{k \in \mathcal{K}} \boldsymbol{\Xi}_i^H \mathbf{v}_{D,k} \mathbf{v}_{D,k}^H \boldsymbol{\Xi}_i + \boldsymbol{\Omega}_i)\big)}{\ln 2} - \frac{N}{\ln 2}. \quad (18)$$

Here, the matrix $\boldsymbol{\Sigma}_i^{(\kappa)}$ that achieves equality in (17) is written as

$$\boldsymbol{\Sigma}_i^{(\kappa)} = \tilde{\boldsymbol{\Sigma}}_i(\mathbf{V}_D, \boldsymbol{\Omega}) \triangleq \sum_{k \in \mathcal{K}} \boldsymbol{\Xi}_i^H \mathbf{v}_{D,k} \mathbf{v}_{D,k}^H \boldsymbol{\Xi}_i + \boldsymbol{\Omega}_i. \quad (19)$$

Based on the inequalities (12) and (17), we formulate the problem

$$\underset{\mathbf{V}_D, \boldsymbol{\Omega}, \mathbf{u}^{(\kappa)}, \tilde{\mathbf{w}}^{(\kappa)}, \boldsymbol{\Sigma}^{(\kappa)}}{\text{maximize}} \sum_{k \in \mathcal{K}} \frac{w_k}{\ln 2} \gamma_k(\mathbf{V}_R', \mathbf{V}_D, \boldsymbol{\Omega}, u_k^{(\kappa)}, \tilde{w}_k^{(\kappa)}) \quad (20a)$$

$$\text{s.t.} \ \tilde{g}_i(\mathbf{V}_D, \boldsymbol{\Omega}_i, \boldsymbol{\Sigma}_i^{(\kappa)}) \leq C_i, \ i \in \mathcal{R}, \quad (20b)$$

$$p_i(\mathbf{V}_{R,i}', \mathbf{V}_D, \boldsymbol{\Omega}_i) \leq P_i, \ i \in \mathcal{R}, \quad (20c)$$

where $\mathbf{u}^{(\kappa)} \triangleq \{u_k^{(\kappa)} \text{ for } k \in \mathcal{K}\}$, $\tilde{\mathbf{w}}^{(\kappa)} \triangleq \{\tilde{w}_k^{(\kappa)} \text{ for } k \in \mathcal{K}\}$, and $\boldsymbol{\Sigma}^{(\kappa)} \triangleq \{\boldsymbol{\Sigma}_i^{(\kappa)} \text{ for } i \in \mathcal{R}\}$. Although problem (20) is still non-convex, it is convex with respect to $\{\mathbf{V}_D, \boldsymbol{\Omega}\}$ when the variables $\{\mathbf{u}^{(\kappa)}, \tilde{\mathbf{w}}^{(\kappa)}, \boldsymbol{\Sigma}^{(\kappa)}\}$ are fixed and vice versa. As proved in [32], since each variable update yields a non-decreasing objective value in (20a), solving problem (20) alternately over these two sets of variables would yield a solution that is guaranteed to converge to a stationary point. This is detailed in Algorithm 1 below.

Algorithm 1 Algorithm for Updating $\mathbf{V}_D$ and $\boldsymbol{\Omega}$

Set $\kappa = 1$ and initialize $\mathbf{V}_D^{(\kappa)}$ and $\boldsymbol{\Omega}^{(\kappa)}$ satisfying the constraints (11b)-(11c).
Repeat
    Update $u_k^{(\kappa)} = \tilde{u}_k(\mathbf{V}_R', \mathbf{V}_D^{(\kappa)}, \boldsymbol{\Omega}^{(\kappa)})$ for $k \in \mathcal{K}$.
    Update $\tilde{w}_k^{(\kappa)} = 1/e_k(\mathbf{V}_R', \mathbf{V}_D^{(\kappa)}, \boldsymbol{\Omega}^{(\kappa)}, u_k^{(\kappa)})$ for $k \in \mathcal{K}$.
    Update $\boldsymbol{\Sigma}_i^{(\kappa)} = \tilde{\boldsymbol{\Sigma}}_i(\mathbf{V}_D^{(\kappa)}, \boldsymbol{\Omega}^{(\kappa)})$ for $i \in \mathcal{R}$.
    Update $\{\mathbf{V}_D^{(\kappa+1)}, \boldsymbol{\Omega}^{(\kappa+1)}\}$ as a solution of problem (20) for the given $\{\mathbf{u}^{(\kappa)}, \tilde{\mathbf{w}}^{(\kappa)}, \boldsymbol{\Sigma}^{(\kappa)}\}$.
    Set $\kappa \leftarrow \kappa + 1$.
Until convergence.

---

2) Optimization of RF Beamforming: We now discuss the optimization of the RF beamformers $\mathbf{V}_R$ in problem (10) for fixed digital variables $\mathbf{V}_D = \mathbf{V}_D'$ and $\boldsymbol{\Omega} = \boldsymbol{\Omega}'$. The problem can be stated as

$$\underset{\mathbf{V}_R}{\text{maximize}} \sum_{k \in \mathcal{K}} w_k f_k(\mathbf{V}_R, \mathbf{V}_D', \boldsymbol{\Omega}') \quad (21a)$$

$$\text{s.t.} \ p_i(\mathbf{V}_{R,i}, \mathbf{V}_D', \boldsymbol{\Omega}_i') \leq P_i, \ i \in \mathcal{R}, \quad (21b)$$

$$|\mathbf{V}_{R,i,a,b}|^2 = 1, \ a \in \mathcal{M}, \ b \in \mathcal{N}, \ i \in \mathcal{R}. \quad (21c)$$

The presence of the constant modulus constraint (21c) makes it difficult to solve problem (21). To address this issue, as in [10, Sec. III-A], we relax the condition (21c) to the convex constraint $|\mathbf{V}_{R,i,a,b}|^2 \leq 1$. Then, the obtained

---

Algorithm 2 Algorithm for Updating $\mathbf{V}_R$

Set $\kappa = 1$ and initialize $\mathbf{V}_R^{(\kappa)}$ satisfying the constraints (21b)-(21c).
Repeat
    Update $u_k^{(\kappa)} = \tilde{u}_k(\mathbf{V}_R^{(\kappa)}, \mathbf{V}_D', \boldsymbol{\Omega}')$ for $k \in \mathcal{K}$.
    Update $\tilde{w}_k^{(\kappa)} = 1/e_k(\mathbf{V}_R^{(\kappa)}, \mathbf{V}_D', \boldsymbol{\Omega}')$ for $k \in \mathcal{K}$.
    Update $\mathbf{V}_R^{(\kappa)}$ as a solution of problem (22) for the given $\{\mathbf{u}^{(\kappa)}, \tilde{\mathbf{w}}^{(\kappa)}\}$.
    Set $\kappa \leftarrow \kappa + 1$.
Until convergence.

---

problem can be solved by again applying the WMMSE method in [32]. The procedure for solving problem (21) is summarized in Algorithm 2, where the convex problem of the original problem (21) is stated as

$$\underset{\mathbf{V}_R, \mathbf{u}^{(\kappa)}, \tilde{\mathbf{w}}^{(\kappa)}}{\text{maximize}} \sum_{k \in \mathcal{K}} \frac{w_k}{\ln 2} \gamma_k(\mathbf{V}_R, \mathbf{V}_D', \boldsymbol{\Omega}', u_k^{(\kappa)}, \tilde{w}_k^{(\kappa)}) \quad (22a)$$

$$\text{s.t.} \ p_i(\mathbf{V}_{R,i}, \mathbf{V}_D', \boldsymbol{\Omega}_i') \leq P_i, \ i \in \mathcal{R}, \quad (22b)$$

$$|\mathbf{V}_{R,i,a,b}|^2 \leq 1, \ a \in \mathcal{M}, \ b \in \mathcal{N}, \ i \in \mathcal{R}. \quad (22c)$$

Since the RF beamforming matrices computed from Algorithm 2, denoted as $\tilde{\mathbf{V}}_R$, may not satisfy the constraint (21c), we propose to obtain a feasible RF beamformer $\mathbf{V}_R$ by projecting $\tilde{\mathbf{V}}_R$ onto the feasible space [10, Sec. III-A]. Specifically, we find the RF beamformer $\mathbf{V}_{R,i}$ such that the distance $\|\mathbf{V}_{R,i} - \tilde{\mathbf{V}}_{R,i}\|_F^2$ is minimized. As a result, the beamformer $\mathbf{V}_{R,i}$ is calculated as $\exp(j\angle\tilde{\mathbf{V}}_{R,i,a,b})$ for $a \in \mathcal{M}, b \in \mathcal{N}$ and $i \in \mathcal{R}$ [10, eq. (14)]. In summary, for a joint design of the digital beamforming $\mathbf{V}_D$, the fronthaul compression $\boldsymbol{\Omega}$ and the RF beamforming strategies $\mathbf{V}_R$, we run Algorithm 1 and 2 alternately. We note that while both Algorithm 1 and Algorithm 2 are individually convergent in the absence of modulus constraint for the RF beamforming, due to the projection step in the update of RF beamforming, the overall alternating optimization algorithm is not guaranteed to converge. This is also the case for the related algorithms in [33]. Therefore, we will observe the convergence behavior of the proposed algorithm in Sec. V.

B. Network Energy Efficiency Maximization

We now consider jointly designing RF and digital beamforming along with fronthaul compression with the aim of maximizing the overall network EE. The network EE is defined as the ratio of the WSR to the corresponding power consumption. Accordingly, the problem is formulated as

$$\underset{\mathbf{V}_R, \mathbf{V}_D, \boldsymbol{\Omega}}{\text{maximize}} \ \frac{\sum_{k \in \mathcal{K}} w_k f_k(\mathbf{V}_R, \mathbf{V}_D, \boldsymbol{\Omega})}{P_T(\mathbf{V}_R, \mathbf{V}_D, \boldsymbol{\Omega})} \quad (23a)$$

$$\text{s.t. (10b), (10c), (10d).} \quad (23b)$$

Problem (23) is also non-convex due to the objective function (23a) and the constraints (23b). In the following subsection, similar to Sec III-A, we adopt alternating optimization to tackle problem (23).

1) Optimization of Digital Beamforming and Fronthaul Compression: For a given RF beamforming $\mathbf{V}_R = \mathbf{V}'_R$, the digital beamforming $\mathbf{V}_D$ and the fronthaul compression strategies $\mathbf{\Omega}$ are optimized by solving the following problem

$$\underset{\mathbf{V}_D, \mathbf{\Omega}}{\text{maximize}} \quad \frac{\sum_{k \in \mathcal{K}} w_k f_k(\mathbf{V}'_R, \mathbf{V}_D, \mathbf{\Omega})}{P_T(\mathbf{V}'_R, \mathbf{V}_D, \mathbf{\Omega})} \tag{24a}$$

$$\text{s.t. (11b), (11c).} \tag{24b}$$

Since problem (24) is non-convex, we also apply a similar approach proposed in Sec. III-A.1. To make problem (24) more tractable, we first introduce a new objective function as a natural logarithm of the objective function (24a)

$$\ln \Big(\sum_{k \in \mathcal{K}} w_k f_k(\mathbf{V}'_R, \mathbf{V}_D, \mathbf{\Omega}) \Big) - \ln \big(P_T(\mathbf{V}'_R, \mathbf{V}_D, \mathbf{\Omega}) \big). \tag{25}$$

Then, we consider a convex lower bound of the function (25) as

$$\ln(\sum_{k \in \mathcal{K}} w_k f_k(\mathbf{V}'_R, \mathbf{V}_D, \mathbf{\Omega})) - \ln(P_T(\mathbf{V}'_R, \mathbf{V}_D, \mathbf{\Omega}))$$
$$\geq \epsilon(\mathbf{V}'_R, \mathbf{V}_D, \mathbf{\Omega}, \mathbf{u}^{(\kappa)}, \tilde{\mathbf{w}}^{(\kappa)}, \rho^{(\kappa)}), \tag{26}$$

where we define the function

$$\epsilon(\mathbf{V}'_R, \mathbf{V}_D, \mathbf{\Omega}, \mathbf{u}^{(\kappa)}, \tilde{\mathbf{w}}^{(\kappa)}, \rho^{(\kappa)})$$
$$= \ln \Big(\sum_{k \in \mathcal{K}} \frac{w_k}{\ln 2} \gamma_k(\mathbf{V}'_R, \mathbf{V}_D, \mathbf{\Omega}, u_k^{(\kappa)}, \tilde{w}_k^{(\kappa)}) \Big) - \ln(\rho^{(\kappa)})$$
$$- \frac{P_T(\mathbf{V}'_R, \mathbf{V}_D, \mathbf{\Omega})}{\rho^{(\kappa)}} + 1, \tag{27}$$

with arbitrary parameters $u_k^{(\kappa)}$, $\tilde{w}_k^{(\kappa)} \geq 0$ and $\rho^{(\kappa)} \geq 0$.

One can show that for fixed $\{\mathbf{V}'_R, \mathbf{V}_D, \mathbf{\Omega}\}$, the lower bound in (26) holds with equality when the variables $\mathbf{u}^{(\kappa)}$, $\tilde{\mathbf{w}}^{(\kappa)}$ and $\rho^{(\kappa)}$ are given as

$$u_k^{(\kappa)} = \tilde{u}_k(\mathbf{V}'_R, \mathbf{V}_D, \mathbf{\Omega}), \ k \in \mathcal{K}, \tag{28}$$

$$\tilde{w}_k^{(\kappa)} = \frac{1}{e_k(\mathbf{V}'_R, \mathbf{V}_D, \mathbf{\Omega}, u_k^{(\kappa)})}, \ k \in \mathcal{K}, \tag{29}$$

$$\rho^{(\kappa)} = P_T(\mathbf{V}'_R, \mathbf{V}_D, \mathbf{\Omega}). \tag{30}$$

Based on the bounds (17) and (26), the problem is formulated as

$$\underset{\mathbf{V}_D, \mathbf{\Omega}, \mathbf{u}^{(\kappa)}, \tilde{\mathbf{w}}^{(\kappa)}, \mathbf{\Sigma}^{(\kappa)}, \rho^{(\kappa)}}{\text{maximize}} \quad \epsilon(\mathbf{V}'_R, \mathbf{V}_D, \mathbf{\Omega}, \mathbf{u}^{(\kappa)}, \tilde{\mathbf{w}}^{(\kappa)}, \rho^{(\kappa)}) \tag{31a}$$

$$\text{s.t. (20b), (20c).} \tag{31b}$$

Similar to Algorithm 1, to obtain a solution $\{\mathbf{V}_D, \mathbf{\Omega}\}$, we alternately update the sets of variables $\{\mathbf{V}_D, \mathbf{\Omega}\}$ and $\{\mathbf{u}^{(\kappa)}, \tilde{\mathbf{u}}^{(\kappa)}, \mathbf{\Sigma}^{(\kappa)}, \rho^{(\kappa)}\}$ until convergence.

2) Optimization of RF Beamforming: In this subsection, for fixed digital variables $\mathbf{V}_D = \mathbf{V}'_D$ and $\mathbf{\Omega} = \mathbf{\Omega}'$, we focus on optimizing the RF beamforming by solving the following non-convex problem

$$\underset{\mathbf{V}_R}{\text{maximize}} \quad \frac{\sum_{k \in \mathcal{K}} w_k f_k(\mathbf{V}_R, \mathbf{V}'_D, \mathbf{\Omega}')}{P_T(\mathbf{V}_R, \mathbf{V}'_D, \mathbf{\Omega}')} \tag{32a}$$

$$\text{s.t. (21b), (21c).} \tag{32b}$$

To solve problem (32), by using the bound (26) and relaxing the modulus constraint (21c), we express the relaxed problem as

$$\underset{\mathbf{V}_R, \mathbf{u}^{(\kappa)}, \tilde{\mathbf{w}}^{(\kappa)}, \rho^{(\kappa)}}{\text{maximize}} \quad \epsilon(\mathbf{V}_R, \mathbf{V}'_D, \mathbf{\Omega}', \mathbf{u}^{(\kappa)}, \tilde{\mathbf{w}}^{(\kappa)}, \rho^{(\kappa)}) \tag{33a}$$

$$\text{s.t. (22b), (22c).} \tag{33b}$$

Similar in Sec. III-A.2, the sets of variables $\mathbf{V}_R$ and $\{\mathbf{u}^{(\kappa)}, \tilde{\mathbf{w}}^{(\kappa)}, \rho^{(\kappa)}\}$ are alternately updated until convergence, and then the obtained RF beamforming matrices are projected onto the feasible space to satisfy the modulus constraint (21c). To sum up, the digital beamforming $\mathbf{V}_D$, the fronthaul compression $\mathbf{\Omega}$ and the RF beamforming $\mathbf{V}_R$ are jointly obtained by optimizing alternately $\{\mathbf{V}_D, \mathbf{\Omega}\}$ and $\mathbf{V}_R$. The effectiveness of the proposed algorithm will be confirmed by numerical results in Sec. V.

IV. DESIGN WITH IMPERFECT CSI

In the previous section, we have assumed that the instantaneous channel vectors $\mathbf{h} \triangleq \{\mathbf{h}_k \text{ for } k \in \mathcal{K}\}$ are perfectly known at the BBU. In this section, we study a more practical case in which low-dimensional effective CSI $\{\bar{\mathbf{V}}_R^H \mathbf{h}_k \text{ for } k \in \mathcal{K}\}$ is acquired by the RRHs via uplink channel training in a TDD operation. The key challenge is that the analog beamforming matrices affect the signal received on the uplink during the training phase. Therefore, the design of the analog beamforming cannot rely on the knowledge of full CSI $\mathbf{h}$. Instead, it is assumed that only the covariance matrices $\{\mathbf{R}_{k,i} \text{ for } k \in \mathcal{K}, \ i \in \mathcal{R}\}$ of the channel vectors are available at the BBU when designing analog precoding. In practice, this long-term CSI can be estimated by means of time average if the fading channels are stationary for a sufficiently long time [11], [34].

A. Uplink Channel Training

In the TDD operation, the downlink CSI is obtained based on the uplink training signals by leveraging reciprocity between downlink and uplink channels. The channel matrix $\mathbf{H}_i = [\mathbf{h}_{1,i}, \cdots, \mathbf{h}_{N_U,i}] \in \mathbb{C}^{M \times N_U}$ between all UEs and the ith RRH is estimated at the RRH and forwarded to the BBU. Importantly, since channel estimation is performed based on the low-dimensional output of RF beamforming, the design of the RF beamforming $\mathbf{V}_R$ affects the channel estimation as well as the WSR performance. For the rest of this subsection, we describe the relationships between $\mathbf{V}_R$ and the channel estimation error.

To elaborate, on the uplink, UE k transmits the orthogonal training sequence $\psi_k \in \mathbb{C}^{L \times 1}$ of L symbols with transmit power p_k, where the condition $L \geq N_U$ is required in order to ensure the orthogonality of the training sequences. We have $\psi_k^H \psi_l = L p_k \delta_{kl}$ for $k, l \in \mathcal{K}$, where δ_{ij} denotes the Kronecker delta function. The signal matrix $\mathbf{Y}_i \in \mathbb{C}^{N \times L}$ received at the ith RRH during uplink training is given as

$$\mathbf{Y}_i = \mathbf{V}_{R,i}^H \mathbf{H}_i \mathbf{\Psi}^T + \mathbf{V}_{R,i}^H \mathbf{N}_i, \tag{34}$$

where $\mathbf{\Psi} = [\psi_1 \cdots \psi_{N_U}] \in \mathbb{C}^{L \times N_U}$ represents the orthogonal training sequence matrix with $\mathbf{\Psi}^H \mathbf{\Psi} = \text{diag}(L p_1, \cdots, L p_{N_U})$

is the matrix of training signal powers, and $\mathbf{N}_i = [\mathbf{n}_{i,1} \cdots \mathbf{n}_{i,L}] \in \mathbb{C}^{M \times L}$ indicates the additive Gaussian noise matrix at the ith RRH with $\mathbf{n}_{i,l} \in \mathbb{C}^{M \times 1} \sim \mathcal{CN}(\mathbf{0}, \sigma_U^2 \mathbf{I}_M)$ for $l \in \{1, \cdots, L\}$.

To estimate the channel $\mathbf{H}_i$ from the received signal $\mathbf{Y}_i$, we define the received signal vector $\mathbf{y}_i \in \mathbb{C}^{NL \times 1}$ of the ith RRH as

$$
\begin{aligned}
\mathbf{y}_i &= \mathrm{vec}(\mathbf{Y}_i) \\
&= (\boldsymbol{\Psi} \otimes \mathbf{V}_{R,i}^H)\mathrm{vec}(\mathbf{H}_i) + (\mathbf{I}_L \otimes \mathbf{V}_{R,i}^H)\mathrm{vec}(\mathbf{N}_i), \quad (35)
\end{aligned}
$$

where $\mathrm{vec}(\mathbf{X})$ denotes the vector obtained by stacking all columns of the matrix $\mathbf{X}$ on top of each other. Note that the signal (35) depends on the RF beamforming matrix $\mathbf{V}_{R,i}$.

Minimizing the MSE yields the estimated channel vector as

$$
\hat{\mathbf{h}}_i = [\hat{\mathbf{h}}_{1,i}^H \cdots \hat{\mathbf{h}}_{N_U,i}^H]^H = \mathbf{W}_i \mathbf{y}_i, \quad (36)
$$

where $\hat{\mathbf{h}}_{k,i} \in \mathbb{C}^{M \times 1}$ stands for the kth subvector of $\hat{\mathbf{h}}_i$ corresponding to the kth UE and $\mathbf{W}_i \triangleq \mathbf{R}_i(\boldsymbol{\Psi}^H \otimes \mathbf{V}_{R,i})((\boldsymbol{\Psi} \otimes \mathbf{V}_{R,i}^H)\mathbf{R}_i(\boldsymbol{\Psi}^H \otimes \mathbf{V}_{R,i}) + (\mathbf{I}_L \otimes \sigma_U^2 \mathbf{V}_{R,i}^H \mathbf{V}_{R,i}))^{-1}$ with $\mathbf{R}_i = \mathrm{diag}(\mathbf{R}_{1,i}, \cdots, \mathbf{R}_{N_U,i})$. RRH i sends the estimated channel vector $\hat{\mathbf{h}}_i$ to the BBU via the fronthaul link. We assume that the coherence block is sufficiently large, so that the CSI overhead is amortized over many fronthaul channel uses. As a result, the estimated channel vectors $\{\hat{\mathbf{h}}_i \text{ for } i \in \mathcal{R}\}$ are available at the BBU without additional distortion due to fronthaul transmission and does not violate the fronthaul capacity constraint [35].

B. Weighted Sum-Rate Maximization

We consider the problem of maximizing the average WSR of the UEs, while satisfying the per-RRH transmit power, fronthaul capacity and constant modulus constraints. The problem is written as

$$
\underset{\mathbf{V}_R}{\text{maximize}} \ \mathbb{E}_{\mathbf{h}}\Big(\underset{\mathbf{V}_D(\mathbf{h}),\boldsymbol{\Omega}(\mathbf{h})}{\text{maximize}} \sum_{k \in \mathcal{K}} w_k f_k(\mathbf{V}_R, \mathbf{V}_D(\mathbf{h}), \boldsymbol{\Omega}(\mathbf{h}))\Big)
$$

$$(37a)$$

$$
\text{s.t. } g_i(\mathbf{V}_D(\mathbf{h}), \boldsymbol{\Omega}_i(\mathbf{h})) \le C_i, \ i \in \mathcal{R}, \ \forall \mathbf{h}, \quad (37b)
$$

$$
p_i(\mathbf{V}_{R,i}, \mathbf{V}_D(\mathbf{h}), \boldsymbol{\Omega}_i(\mathbf{h})) \le P_i, \ i \in \mathcal{R}, \ \forall \mathbf{h}, \quad (37c)
$$

$$
|\mathbf{V}_{R,i,a,b}|^2 = 1, \ a \in \mathcal{M}, \ b \in \mathcal{N}, \ i \in \mathcal{R}. \quad (37d)
$$

In problem (37), we account for the fact that while the RF beamforming can only depend on long-term CSI, the digital beamforming and fronthaul compression can be a function of the instantaneous CSI.

1) Design of RF Beamforming: As discussed before, the RF beamforming matrix affects both the downlink rate and the quality of the estimated CSI via uplink training. In this subsection, we focus on the design of matrices $\mathbf{V}_R$ by assuming only the long-term CSI on the covariance matrices $\{\mathbf{R}_{k,i} \text{ for } k \in \mathcal{K}, \ i \in \mathcal{R}\}$. Adopting the SAA method [22], we generate T instantaneous channel samples $\tilde{\mathbf{h}} \triangleq \{\tilde{\mathbf{h}}_t \text{ for } t \in \mathcal{T} \triangleq \{1, \cdots, T\}\}$ based on the second-order statistic $\{\mathbf{R}_{k,i} \text{ for } k \in \mathcal{R}, i \in \mathcal{R}\}$ of the downlink channel vectors. By approximating the objective function (37a) with an empirical average as $\sum_{k \in \mathcal{K}} w_k \mathbb{E}_{\mathbf{h}}(f_k(\mathbf{V}_R, \mathbf{V}_D(\mathbf{h}), \boldsymbol{\Omega}(\mathbf{h}))) \approx$

$\sum_{k \in \mathcal{K}} \sum_{t \in \mathcal{T}} \frac{w_k}{T} f_k(\mathbf{V}_R, \mathbf{V}_D(\tilde{\mathbf{h}}_t), \boldsymbol{\Omega}(\tilde{\mathbf{h}}_t))$, we formulate the problem as

$$
\underset{\mathbf{V}_R,\mathbf{V}_D(\tilde{\mathbf{h}}),\boldsymbol{\Omega}(\tilde{\mathbf{h}})}{\text{maximize}} \sum_{k \in \mathcal{K}} \sum_{t \in \mathcal{T}} \frac{w_k}{T} f_k(\mathbf{V}_R, \mathbf{V}_D(\tilde{\mathbf{h}}_t), \boldsymbol{\Omega}(\tilde{\mathbf{h}}_t)) \quad (38a)
$$

$$
\text{s.t. } g_i(\mathbf{V}_D(\tilde{\mathbf{h}}_t), \boldsymbol{\Omega}_i(\tilde{\mathbf{h}}_t)) \le C_i, \ i \in \mathcal{R}, \ t \in \mathcal{T}, \quad (38b)
$$

$$
p_i(\mathbf{V}_{R,i}, \mathbf{V}_D(\tilde{\mathbf{h}}_t), \boldsymbol{\Omega}_i(\tilde{\mathbf{h}}_t)) \le P_i, \ i \in \mathcal{R}, \ t \in \mathcal{T}, \quad (38c)
$$

$$
|\mathbf{V}_{R,i,a,b}|^2 = 1, \ a \in \mathcal{M}, \ b \in \mathcal{N}, \ i \in \mathcal{R}. \quad (38d)
$$

Similar to problem (10), we update the variables $\{\mathbf{V}_D(\tilde{\mathbf{h}}), \boldsymbol{\Omega}(\tilde{\mathbf{h}})\}$ from Algorithm 1 and the analog RF beamforming matrices $\mathbf{V}_R$ from Algorithm 2 alternately until convergence.

2) Design of Digital Beamforming and Fronthaul Compression: Based on the channel estimate obtained for the RF beamforming matrices $\mathbf{V}_R$ in Sec. IV-B.1, the BBU optimizes the digital strategies $\mathbf{V}_D$ and $\boldsymbol{\Omega}$. We write the downlink received signal y_k in (6) of the kth UE as

$$
y_k = \sum_{l \in \mathcal{K}}(\hat{\mathbf{h}}_k^H \bar{\mathbf{V}}_R + \mathbf{e}_k^H)\mathbf{v}_{D,l}s_l + (\hat{\mathbf{h}}_k^H \bar{\mathbf{V}}_R + \mathbf{e}_k^H)\mathbf{q} + z_k, \quad (39)
$$

where $\hat{\mathbf{h}}_k = [\hat{\mathbf{h}}_{k,1} \ \cdots \ \hat{\mathbf{h}}_{k,N_R}] \in \mathbb{C}^{MN_R \times 1}$ represents the estimated channel vector from all RRHs to the kth UE, and $\mathbf{e}_k \triangleq \mathbf{h}_k - \hat{\mathbf{h}}_k$ indicates the estimation error vector from all RRHs to the kth UE, which is distributed as $\mathbf{e}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{E}_k)$ with $\mathbf{E}_k = \mathrm{diag}(\mathbf{E}_{k,1}, \cdots, \mathbf{E}_{k,N_R})$. Here, the error covariance matrix $\mathbf{E}_{k,i}$ is given as

$$
\mathbf{E}_{k,i} = \Big(\mathbf{I}_M + \frac{Lp_k}{\sigma_U^2}\mathbf{R}_{k,i}\mathbf{V}_{R,i}(\mathbf{V}_{R,i}^H \mathbf{V}_{R,i})^{-1}\mathbf{V}_{R,i}^H\Big)^{-1}\mathbf{R}_{k,i}. \quad (40)
$$

Assuming that the estimation error and interference are treated as the additive noise [36], the achievable rate for the kth UE is computed as

$$
\begin{aligned}
\bar{f}_k(\mathbf{V}_D, \boldsymbol{\Omega} = \log_2 \det\big(|\hat{\mathbf{h}}_k^H \bar{\mathbf{V}}_R \mathbf{v}_{D,k}|^2 + \bar{\zeta}_k(\mathbf{V}_D, \boldsymbol{\Omega})\big) \\
- \log_2 \det(\bar{\zeta}_k(\mathbf{V}_D, \boldsymbol{\Omega})), \quad (41)
\end{aligned}
$$

where

$$
\begin{aligned}
\bar{\zeta}(\mathbf{V}_D, \boldsymbol{\Omega}) &\triangleq \sum_{l \in \mathcal{K}\backslash\{k\}}|\hat{\mathbf{h}}_k^H \bar{\mathbf{V}}_R \mathbf{v}_{D,l}|^2 + \sum_{l \in \mathcal{K}}\mathbf{v}_{D,l}^H \bar{\mathbf{V}}_R^H \mathbf{E}_k \bar{\mathbf{V}}_R \mathbf{v}_{D,l} \\
&\quad + \hat{\mathbf{h}}_k^H \bar{\mathbf{V}}_R \bar{\boldsymbol{\Omega}} \bar{\mathbf{V}}_R^H \hat{\mathbf{h}}_k + \mathrm{tr}(\bar{\mathbf{V}}_R^H \mathbf{E}_k \bar{\mathbf{V}}_R \bar{\boldsymbol{\Omega}}) + \sigma_D^2. \quad (42)
\end{aligned}
$$

Then, the WSR maximization problem is formulated as

$$
\underset{\mathbf{V}_D,\boldsymbol{\Omega}}{\text{maximize}} \sum_{k \in \mathcal{K}} w_k \bar{f}_k(\mathbf{V}_D, \boldsymbol{\Omega}) \quad (43a)
$$

$$
\text{s.t. } g_i(\mathbf{V}_D, \boldsymbol{\Omega}_i) \le C_i, \ i \in \mathcal{R}, \quad (43b)
$$

$$
p_i(\mathbf{V}_{R,i}, \mathbf{V}_D, \boldsymbol{\Omega}_i) \le P_i, \ i \in \mathcal{R}. \quad (43c)
$$

As in problem (11), we adopt Algorithm 1 with minor modifications.

C. Network Energy Efficiency Maximization

We also address the problem of jointly optimizing RF and digital beamforming and fronthaul compression design with

the goal of maximizing the average network EE. The problem is written as

$$\underset{\mathbf{V}_R}{\text{maximize}} \quad \mathbb{E}_{\mathbf{h}}\left(\underset{\mathbf{V}_D(\mathbf{h}),\Omega(\mathbf{h})}{\text{maximize}} \frac{\sum_{k\in\mathcal{K}} w_k f_k(\mathbf{V}_R,\mathbf{V}_D(\mathbf{h}),\Omega(\mathbf{h}))}{P_T(\mathbf{V}_R,\mathbf{V}_D(\mathbf{h}),\Omega(\mathbf{h}))}\right) \tag{44a}$$

$$\text{s.t. (37b), (37c), (37d).} \tag{44b}$$

1) Design of RF beamforming: Adopting the SAA method, the approximation problem of the original problem is written as

$$\underset{\mathbf{V}_R,\mathbf{V}_D(\tilde{\mathbf{h}}),\Omega(\tilde{\mathbf{h}})}{\text{maximize}} \sum_{t\in\mathcal{T}} \frac{\sum_{k\in\mathcal{K}} w_k f_k(\mathbf{V}_R,\mathbf{V}_D(\tilde{\mathbf{h}}_t),\Omega(\tilde{\mathbf{h}}_t))}{T P_T(\mathbf{V}_R,\mathbf{V}_D(\tilde{\mathbf{h}}_t),\Omega(\tilde{\mathbf{h}}_t))} \tag{45a}$$

$$\text{s.t. (38b), (38c), (38d).} \tag{45b}$$

Since the objective function (45a) is a sum of ratios unlike the average WSR maximization problem, the proposed algorithm in Sec. III-B cannot be directly applied to the average network EE maximization problem. To tackle this issue, we introduce additional optimization variables $\boldsymbol{\alpha} \triangleq \{\alpha_t \text{ for } t \in \mathcal{T}\}$ which satisfy the non-convex constraint

$$\ln(\alpha_t) \leq \ln(\sum_{k\in\mathcal{K}} w_k f_k(\mathbf{V}_R,\mathbf{V}_D(\tilde{\mathbf{h}}_t),\Omega(\tilde{\mathbf{h}}_t)))$$
$$- \ln(P_T(\mathbf{V}_R,\mathbf{V}_D(\tilde{\mathbf{h}}_t),\Omega(\tilde{\mathbf{h}}_t))), \ t\in\mathcal{T}. \tag{46}$$

Then, the problem can be equivalently recast as

$$\underset{\mathbf{V}_R,\mathbf{V}_D(\tilde{\mathbf{h}}),\Omega(\tilde{\mathbf{h}}),\boldsymbol{\alpha}}{\text{maximize}} \sum_{t\in\mathcal{T}} \frac{\alpha_t}{T} \tag{47a}$$

$$\text{s.t. (38b), (38c), (38d), (46).} \tag{47b}$$

However, since $\ln(\alpha_t)$ in (46) is a concave function, it is still difficult to solve problem (47). To make the constraint (46) more tractable, we also consider a convex upper bound on $\ln(\alpha_t)$ as

$$\ln(\alpha_t) \leq \ln(\beta_t^{(\kappa)}) + \frac{\alpha_t}{\beta_t^{(\kappa)}} - 1. \tag{48}$$

By using the convex bounds (17), (26) and (48) and relaxing the modulus constraint (38d), we formulate the problem

$$\underset{\mathcal{A}}{\text{maximize}} \sum_{t\in\mathcal{T}} \frac{\alpha_t}{T} \tag{49a}$$

$$\text{s.t. } \tilde{g}_i(\mathbf{V}_D(\tilde{\mathbf{h}}_t),\Omega_i(\tilde{\mathbf{h}}_t),\Sigma_i^{(\kappa)}(\tilde{\mathbf{h}}_t)) \leq C_i, \ i\in\mathcal{R}, \ t\in\mathcal{T}, \tag{49b}$$

$$p_i(\mathbf{V}_{R,i},\mathbf{V}_D(\tilde{\mathbf{h}}_t),\Omega_i(\tilde{\mathbf{h}}_t)) \leq P_i, i\in\mathcal{R}, \ t\in\mathcal{T}, \tag{49c}$$

$$\epsilon(\mathbf{V}_R,\mathbf{V}_D(\tilde{\mathbf{h}}_t),\Omega(\tilde{\mathbf{h}}_t),\mathbf{u}^{(\kappa)}(\tilde{\mathbf{h}}_t),\tilde{\mathbf{w}}^{(\kappa)}(\tilde{\mathbf{h}}_t),\rho^{(\kappa)}(\tilde{\mathbf{h}}_t)) \tag{49d}$$
$$\geq \ln\beta_t^{(\kappa)} + \frac{\alpha_t}{\beta_t^{(\kappa)}} - 1, \ t\in\mathcal{T},$$

$$|\mathbf{V}_{R,i,a,b}| \leq 1, \ i\in\mathcal{R}, \ a\in\mathcal{M}, \ b\in\mathcal{N}, \tag{49e}$$

where $\mathcal{A} \triangleq \{\mathbf{V}_R,\mathbf{V}_D(\tilde{\mathbf{h}}),\Omega(\tilde{\mathbf{h}}),\mathbf{u}^{(\kappa)}(\tilde{\mathbf{h}}),\tilde{\mathbf{w}}^{(\kappa)}(\tilde{\mathbf{h}}),\Sigma^{(\kappa)}(\tilde{\mathbf{h}}),$ $\rho^{(\kappa)}(\tilde{\mathbf{h}}),\boldsymbol{\alpha},\boldsymbol{\beta}\}$ with $\boldsymbol{\beta} \triangleq \{\beta_t \text{ for } t\in\mathcal{T}\}$. As in Sec. III-B, to obtain the RF beamforming matrices $\mathbf{V}_R$, we update the variables $\{\mathbf{V}_D(\tilde{\mathbf{h}}),\Omega(\tilde{\mathbf{h}})\}$ and $\mathbf{V}_R$ alternately until convergence.

2) Design of Digital Beamforming and Fronthaul Compression: After channel estimation based on the RF beamforming matrices $\mathbf{V}_R$ in Sec. IV-C.1, the digital beamforming $\mathbf{V}_D$ and the fronthaul compression strategies Ω are computed by solving the following problem

$$\underset{\mathbf{V}_D,\Omega}{\text{maximize}} \quad \frac{\sum_{k\in\mathcal{K}} w_k \bar{f}_k(\mathbf{V}_D,\Omega)}{P_T(\mathbf{V}_R,\mathbf{V}_D,\Omega)} \tag{50a}$$

$$\text{s.t. (43b), (43c).} \tag{50b}$$

Problem (50) can be solved by the algorithm in Sec. III-B.1 with minor modifications. The performance of the proposed algorithm will be evaluated by numerical results in Sec. V.

V. NUMERICAL RESULTS

In this section, we present numerical results to validate the effectiveness of the proposed joint design of the RF and digital processing strategies. Throughout the simulation results, we consider the case of $N_U = 4$ UEs, $N_R = 2$ RRHs, and $M = 10$ RRH antennas, and evaluate the sum-rate of the UEs with $w_k = 1$ for all $k \in \mathcal{K}$. The length of pilot sequences is set to $L = N_U$, the uplink transmit power of all UEs is given as $p_k = 1$ for all $k \in \mathcal{K}$, the circuit power per RF chain and the static power consumed at the UE are respectively fixed as $P_{RF} = 1$ and $P_{N_U} = 1$, and each RRH has the same fronthaul capacity C and the same downlink transmit power P for all $i \in \mathcal{R}$, i.e., $C_i = C$ and $P_i = P$. In addition, we set the downlink noise variance to be $\sigma_D^2 = 1$ so that the downlink signal-to-noise ratio (SNR) is defined as $\text{SNR} = P$.

Following [34] and [37], we adopt the half wavelength-spaced uniform linear antenna array model of the RRH antennas such that the channel covariance matrix $\mathbf{R}_{k,i}$ is given as

$$\mathbf{R}_{k,i,a,b} = \frac{1}{2\Delta_{k,i}} \int_{\theta_{k,i}-\Delta_{k,i}}^{\theta_{k,i}+\Delta_{k,i}} e^{-j\pi(a-b)\sin\phi} d\phi, \tag{51}$$

where the angle of arrival $\theta_{k,i}$ and the angular spread $\Delta_{k,i}$ have the distributions $\theta_{k,i} \sim \mathcal{U}(-\frac{\pi}{3},\frac{\pi}{3})$ and $\Delta_{k,i} \sim \mathcal{U}(\frac{\pi}{18},\frac{2\pi}{9})$, respectively. Here, the notation $\mathcal{U}(a,b)$ represents a uniform distribution between a and b.

A. Perfect CSI Case

For the perfect CSI case, we compare the performance of the proposed scheme in Sec. III with the following baseline schemes.

- *Fully digital:* Fully digital beamforming is carried out across all RRH antennas, where beamforming is designed using Algorithm 1 with $M = N$.
- *Reduced-rank digital:* Fully digital beamforming is performed under rank constraint equal to N. This is done by running Algorithm 1 while omitting the projection step in the update of RF beamforming.
- *Random RF and optimized digital:* The phases of the RF beamforming matrices are randomly selected from an independent and identically distributed (i.i.d.) distribution $\mathcal{U}(0,2\pi)$.

Fig. 2 illustrates the convergence behavior of the proposed algorithm for one channel realization with $C = 5$ bps/Hz and
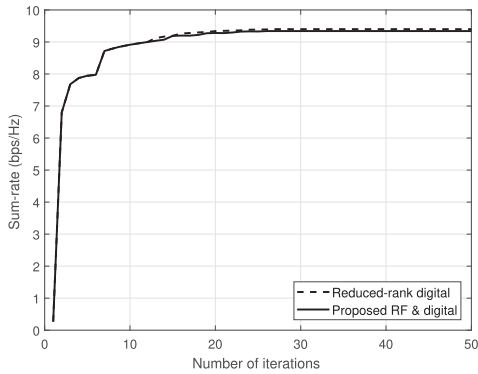
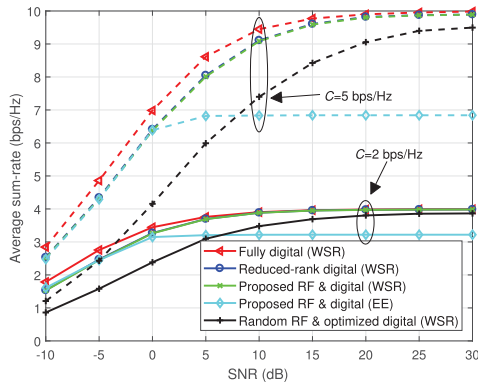Fig. 2. Convergence behavior for $N_U = 4$, $N_R = 2$, $N = 2$, $M = 10$, $C = 5$ bps/Hz and SNR $= 10$ dB.

Fig. 3. Average sum-rate performance with respect to SNR with $N_U = 4$, $N_R = 2$, $N = 2$ and $M = 10$.

Fig. 4. Average sum-rate performance with respect to N with $N_U = 4$, $N_R = 2$, $M = 10$ and $C = 5$ bps/Hz.

Fig. 5. Average EE performance with respect to SNR with $N_U = 4$, $N_R = 2$, $N = 2$ and $M = 10$.

SNR $= 10$ dB. The dashed line is obtained by the reduced-rank digital scheme, while the solid line is attained from the proposed algorithm. The figure shows that in spite of the projection step, the proposed algorithm converges within a few tens of iterations. In addition, the average per-iteration running time of the proposed algorithm and the fully digital scheme are 1.73 sec and 8.65 sec, respectively. As can be seen, the number of RF chains is a important factor of the algorithm's running time.

Fig. 3 shows the average sum-rate with respect to the downlink SNR for a C-RAN with $C \in \{2, 5\}$ bps/Hz. The proposed joint design of the RF and digital processing strategies always outperforms the random RF beamforming scheme, particularly at lower SNR, where the downlink channel sets the performance bottleneck of the system. In a similar way, the optimization of RF beamforming has a more significant impact when the fronthaul capacity is larger. As the SNR increases, the fronthaul capacity limitations become important, and thus the proposed joint design approaches the sum-rate of the fully digital scheme in spite of the limited number of RF chains. Furthermore, by comparing the proposed WSR maximization and the reduced-rank digital scheme, we can check that the performance loss caused by the projection step in the update of RF beamforming is small. In addition, it is seen that the sum-rate performance of the proposed EE maximization is consistent with the WSR approach at low SNR, but is saturated to a lower value at high SNR. This is because in this regime the additional power needed to further
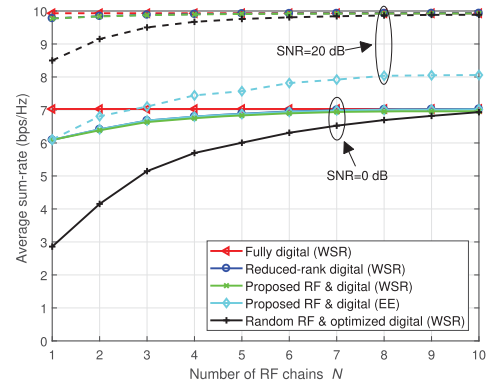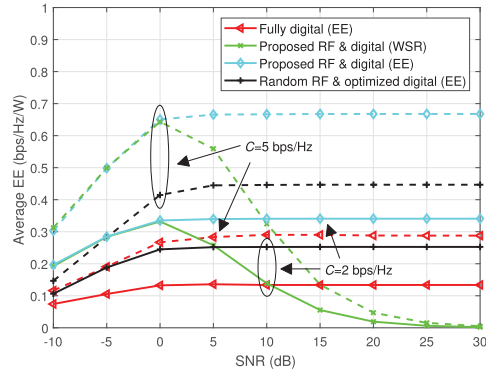
increase the sum-rate is not necessary from the viewpoint of the EE.

In Fig. 4, we plot the average sum-rate with respect to the number N of RF chains for the downlink of a C-RAN with $C = 5$ bps/Hz and SNR $\in \{0, 20\}$ dB. The sum-rate of the proposed joint design increases more rapidly with N as compared to that of the random RF beamforming scheme. Also, when N is sufficiently large, both the proposed scheme and random RF beamforming achieve sum-rate performance very close to that of the fully digital beamforming scheme. Similar to Fig. 3, the impact of RF beamforming is more pronounced when the SNR is small for fixed fronthaul capacity. It also confirms that the sum-rate performance gap between the proposed WSR maximization and EE maximization becomes larger as the SNR grows.

Next, to investigate the EE performance of the proposed algorithm, Fig. 5 depicts the average EE with respect to the downlink SNR for $C \in \{2, 5\}$ bps/Hz. The figure illustrates the fact that as the SNR increases, WSR becomes extremely inefficient in terms of energy minimization. This is in contrast to the schemes designed for EE maximization. Furthermore, the fully digital architecture shows poor EE performance due to the energy consumed by M RF chains at each RRH.

This point is further explored in Fig. 6, which shows the average EE with respect to the number N of RF chains for $C = 5$ bps/Hz and SNR $= 20$ dB. The main observation here is that increasing the number of RF chains may exhibit a negative impact on the EE, particularly when the RF beamforming is optimized.
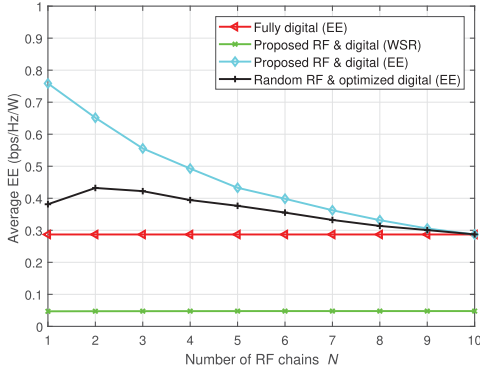
Fig. 6.   Average EE performance with respect to N with $N_U = 4$, $N_R = 2$, $M = 10$, $C = 5$ bps/Hz and SNR = 20 dB.
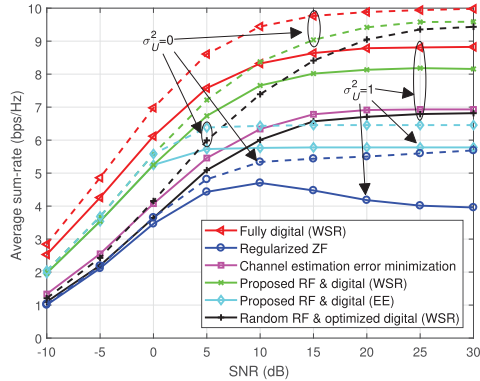

Fig. 7.   Average sum-rate performance with respect to SNR with $N_U = 4$, $N_R = 2$, $N = 2$, $M = 10$ and $C = 5$ bps/Hz.

B. Imperfect CSI Case

For the imperfect CSI case, we present the performance of two baseline schemes for comparison. The first is a regularized ZF method, whereby RF and digital beamforming is obtained from the algorithm in [11] and fronthaul compression is determined using Algorithm 1 for given digital beamforming. The second is a channel estimation error minimization scheme, where the RF beamforming matrices are computed by minimizing the MSE based on the algorithm in [30], and the digital processing is calculated from Algorithm 1.

Fig. 7 plots the average sum-rate with respect to the SNR for a C-RAN with $C = 5$ bps/Hz and $\sigma_U^2 \in \{0, 1\}$. We consider the cases of a noiseless uplink channel $\sigma_U^2 = 0$ (dashed line) and a noisy uplink channel $\sigma_U^2 = 1$ (solid line). In spite of the fact that the RF beamforming matrices are designed based on long-term CSI, we can check that the proposed WSR algorithm shows effective sum-rate performance. In the presence of the channel estimation error $\sigma_U^2 = 1$, it is clear that the proposed digital beamforming is much more robust to the estimation errors in comparison to the regularized ZF scheme. Also, although the impact of RF beamforming on the channel estimation error is considered, the channel estimation error minimization scheme is seen to yield suboptimal sum-rate performance.

In Fig. 8, we illustrate the impact of the fronthaul capacity C for a C-RAN with SNR = 5 dB and $\sigma_U^2 \in \{0, 1\}$. The proposed WSR scheme exhibits the average sum-rate which increases more rapidly with C as compared to the random RF beamforming and regularized ZF schemes. It is
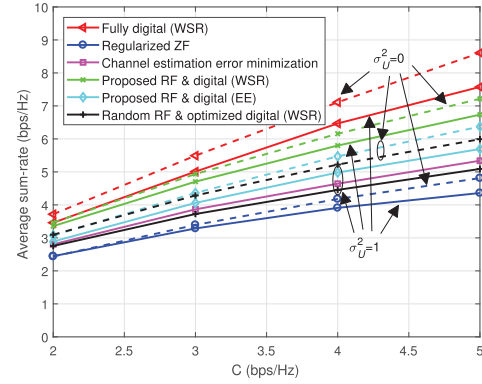

Fig. 8.   Average sum-rate performance with respect to C with $N_U = 4$, $N_R = 2$, $N = 2$, $M = 10$ and SNR = 5 dB.
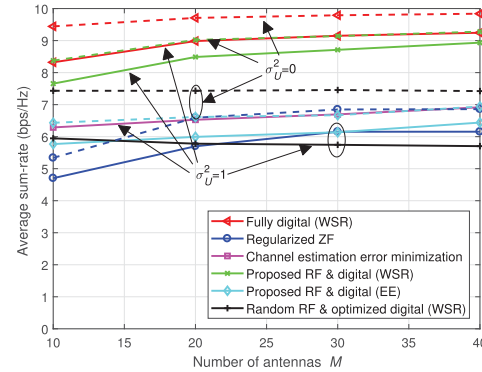

Fig. 9.   Average sum-rate performance with respect to M with $N_U = 4$, $N_R = 2$, $N = 2$, $C = 5$ bps/Hz and SNR = 10 dB.
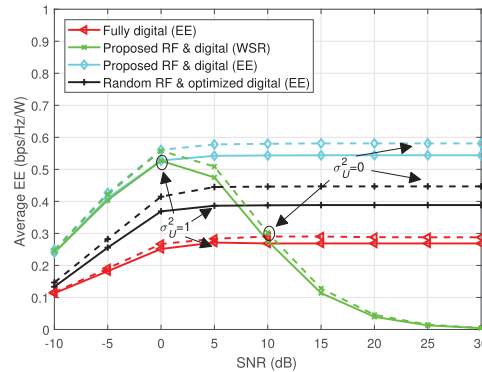

Fig. 10.   Average EE performance as a function of SNR with $N_U = 4$, $N_R = 2$, $N = 2$, $M = 10$ and $C = 5$ bps/Hz.

also observed that the proposed WSR scheme achieves the sum-rate performance close to the fully digital scheme at a small C. In addition, the performance loss caused by the channel estimation error is relatively minor for the proposed WSR scheme as compared to the fully digital and random RF beamforming strategies, although the impact of RF beamforming on the channel estimation error is ignored.

Fig. 9 depicts the average sum-rate in terms of the number M of antennas for a C-RAN with SNR = 10 dB, $C = 5$ bps/Hz and $\sigma_U^2 \in \{0, 1\}$. We can see that the average sum-rate performance gap between the fully digital scheme and the proposed WSR scheme becomes smaller as M grows. One interesting observation is that while the performance of the schemes with the optimized RF beamforming increases with M, the sum-rate of the random RF beamforming scheme

decreases. This implies that the optimization of RF beamforming is more significant when M is larger.

Fig. 10 shows the average EE as a function of the SNR with $C = 5$ bps/Hz and $\sigma_U^2 \in \{0, 1\}$. Similar to Fig. 5, we observe that when the SNR is sufficiently large, the proposed EE maximization scheme outperforms WSR. Furthermore, the proposed joint design exhibits performance robust to channel estimation errors compared to randomized beamforming.

VI. Conclusions

We have studied the joint design for the downlink of a C-RAN with hybrid analog-digital antenna arrays. Specifically, we have jointly optimized the digital beamforming, the fronthaul compression and the RF beamforming strategies with the goal of maximizing the WSR and the network EE, while satisfying the per-RRH power, fronthaul capacity and constant modulus constraints. We have proposed an iterative algorithm that achieves an efficient solution under perfect CSI. Furthermore, we have discussed the case of imperfect CSI based on the uplink channel training. Numerical results have confirmed the effectiveness of the proposed algorithm. Also, we have illustrated the impact of imperfect CSI on the downlink sum-rate and network EE performance and shown that the proposed scheme is robust to the estimation errors. As some interesting directions for future researches, we mention the development of a globally optimal algorithm and a design with low-resolution analog RF beamforming. In addition, it will be interesting to consider joint multivariate compression also for this structure where the RRH focuses on analogue processing. Furthermore, an uplink-downlink duality for this linear type pre-processing at the RRH remains as future work, where the idea is to extend the single user case [38] to multiple users.

References

[1] J. Kim, S.-H. Park, O. Simeone, I. Lee, and S. Shamai (Shitz), "Joint design of digital and analog processing for downlink C-RAN with large-scale antenna arrays," in *Proc. IEEE 18th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jul. 2017, pp. 1–5.

[2] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.

[3] A. F. Molisch *et al.*, "Hybrid beamforming for massive MIMO: A survey," *IEEE Commun. Mag.*, vol. 55, no. 9, pp. 134–141, Sep. 2017.

[4] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, Jr., "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.

[5] L. Liang, Y. Dai, W. Xu, and X. Dong, "How to approach zero-forcing under RF chain limitations in large mmWave multiuser systems?" in *Proc. IEEE/CIC Int. Conf. Commun. China*, Oct. 2014, pp. 518–522.

[6] L. Liang, W. Xu, and X. Dong, "Low-complexity hybrid precoding in massive multiuser MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 653–656, Dec. 2014.

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

[8] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 501–513, Apr. 2016.

[9] X. Yu, J.-C. Shen, J. Zhang, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 485–500, Apr. 2016.

[10] C.-S. Lee and W.-H. Chung, "Hybrid RF-baseband precoding for cooperative multiuser massive MIMO systems with limited RF chains," *IEEE Trans. Commun.*, vol. 65, no. 4, pp. 1575–1589, Apr. 2017.

[11] S. Park, J. Park, A. Yazdan, and R. W. Heath, Jr., "Exploiting spatial channel covariance for hybrid precoding in massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 65, no. 14, pp. 3818–3832, Jul. 2017.

[12] A. Liu, V. K. N. Lau, and M.-J. Zhao, "Stochastic successive convex optimization for two-timescale hybrid precoding in massive MIMO," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 3, pp. 432–444, Jun. 2018.

[13] T. Oh, C. Song, J. Kim, S. Jang, and I. Lee, "Hybrid analog-digital filter designs for mmWave multi-pair two-way relaying systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7841–7845, Aug. 2018.

[14] O. Simeone, O. Somekh, H. V. Poor, and S. Shamai (Shitz), "Downlink multicell processing with limited-backhaul capacity," *EURASIP J. Adv. Signal Process.*, vol. 2009, no. 3, pp. 1–10, Jun. 2009.

[15] S.-H. Park, O. Simeone, O. Sahin, and S. Shamai (Shitz), "Joint precoding and multivariate backhaul compression for the downlink of cloud radio access networks," *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5646–5658, Nov. 2013.

[16] S.-H. Park, O. Simeone, O. Sahin, and S. Shamai (Shitz), "Fronthaul compression for cloud radio access networks: Signal processing advances inspired by network information theory," *IEEE Signal Process. Mag.*, vol. 31, no. 6, pp. 69–79, Nov. 2014.

[17] Q. Hou, S. He, Y. Huang, H. Wang, and L. Yang, "Joint user scheduling and hybrid precoding design for MIMO C-RAN," in *Proc. 9th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Oct. 2017, pp. 1–6.

[18] J. Park, D. M. Kim, E. D. Carvalho, and C. N. Manchón. (Sep. 2017). "Hybrid precoding for massive MIMO systems in cloud RAN architecture with capacity-limited fronthauls." [Online]. Available: https://arxiv.org/abs/1709.07963

[19] P. Luong, F. Gagnon, C. Despins, and L.-N. Tran, "Joint virtual computing and radio resource allocation in limited fronthaul green C-RANs," *IEEE Trans. Wireless Commun.*, vol. 17, no. 4, pp. 2602–2617, Apr. 2018.

[20] T. T. Vu, D. T. Ngo, M. N. Dao, S. Durrani, D. H. N. Nguyen, and R. H. Middleton, "Energy efficiency maximization for downlink cloud radio access networks with data sharing and data compression," *IEEE Trans. Wireless Commun.*, vol. 17, no. 8, pp. 4955–4970, Aug. 2018.

[21] K.-G. Nguyen, Q.-D. Vu, M. Juntti, and L.-N. Tran, "Energy efficiency maximization for C-RANs: Discrete monotonic optimization, penalty, and l_0-approximation methods," *IEEE Trans. Signal Process.*, vol. 66, no. 17, pp. 4435–4449, Sep. 2018.

[22] A. Liu and V. K. N. Lau, "Impact of CSI knowledge on the codebook-based hybrid beamforming in massive MIMO," *IEEE Trans. Signal Process.*, vol. 64, no. 24, pp. 6545–6556, Dec. 2016.

[23] S.-H. Park, K.-J. Lee, C. Song, and I. Lee, "Joint design of fronthaul and access links for C-RAN with wireless fronthauling," *IEEE Signal Process. Lett.*, vol. 23, no. 11, pp. 1657–1661, Nov. 2016.

[24] Y. Jeon, S.-H. Park, C. Song, J. Moon, S. Maeng, and I. Lee, "Joint designs of fronthaul compression and precoding for full-duplex cloud radio access networks," *IEEE Wireless Commun. Lett.*, vol. 5, no. 6, pp. 632–635, Dec. 2016.

[25] G. M. Guvensen and E. Ayanoglu. (Jul. 2016). "A generalized framework on beamformer design and CSI acquisition for single-carrier massive MIMO systems in millimeter wave channels." [Online]. Available: https://arxiv.org/abs/1607.01436

[26] L. Zhao, D. W. K. Ng, and J. Yuan, "Multi-user precoding and channel estimation for hybrid millimeter wave systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1576–1590, Jul. 2017.

[27] O. Simeone, A. Maeder, M. Peng, O. Sahin, and W. Yu, "Cloud radio access network: Virtualizing wireless access for dense heterogeneous systems," *J. Commun. Netw.*, vol. 18, no. 2, pp. 135–149, Apr. 2016.

[28] S.-R. Lee, S.-H. Park, and I. Lee, "NOMA systems with content-centric multicast transmission for C-RAN," *IEEE Wireless Commun. Lett.*, vol. 7, no. 5, pp. 828–831, Oct. 2018.

[29] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2012.

[30] S. S. Ioushua and Y. C. Eldar. (Dec. 2017). "Hybrid analog-digital beamforming for massive MIMO systems." [Online]. Available: https://arxiv.org/abs/1712.03485

[31] S.-R. Lee, J. Jung, H. Park, and I. Lee, "A new energy-efficient beamforming strategy for MISO interfering broadcast channels based on large systems analysis," *IEEE Trans. Wireless Commun.*, vol. 15, no. 4, pp. 2872–2882, Apr. 2016.

[32] 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.

[33] H. Ghauch, T. Kim, M. Bengtsson, and M. Skoglund, "Subspace estimation and decomposition for large millimeter-wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 528–542, Apr. 2016.

[34] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing—The large-scale array regime," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6441–6463, Oct. 2013.

[35] J. Kang, O. Simeone, J. Kang, and S. Shamai (Shitz), "Joint signal and channel state information compression for the backhaul of uplink network MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1555–1567, Mar. 2014.

[36] B. Hassibi and M. Sharif, "Fundamental limits in MIMO broadcast channels," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1333–1344, Sep. 2007.

[37] Y. Jeon, C. Song, S.-R. Lee, S. Maeng, J. Jung, and I. Lee, "New beamforming designs for joint spatial division and multiplexing in large-scale MISO multi-user systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 3029–3041, May 2017.

[38] L. Liu, P. Patil, and W. Yu, "Channel diagonalization for cloud radio access," *IEEE Wireless Commun. Lett.*, vol. 7, no. 4, pp. 622–625, Aug. 2018.

Jaein Kim (S'16) received the B.S. and M.S. degrees in electrical engineering from Korea University, Seoul, South Korea, in 2014 and 2016, respectively, where he is currently pursuing the Ph.D. degree with the School of Electrical Engineering. His research interests include information theory, wireless communications, and optimization theory.

Seok-Hwan Park (M'11) received the B.Sc. and Ph.D. degrees in electrical engineering from Korea University, Seoul, South Korea, in 2005 and 2011, respectively.

He was with the New Jersey Institute of Technology and Samsung Electronics. Since 2015, he has been with Chonbuk National University, Jeonju, South Korea, as an Assistant Professor. His research interests include MIMO communications, optimization theory, and machine learning. He was a co-recipient of the 2018 IEEE Signal Processing Best Paper Award and the Best Paper Award of APCC 2006.

Osvaldo Simeone (F'16) received the M.Sc. (Hons.) and Ph.D. degrees in information engineering from the Politecnico di Milano, Milan, Italy, in 2001 and 2005, respectively. He is currently a Professor of information engineering with the Centre for Telecommunications Research, Department of Informatics, King's College London, London, U.K. From 2006 to 2017, he was a Faculty Member with the Electrical and Computer Engineering Department, New Jersey Institute of Technology, where he was affiliated with the Center for Wireless Information Processing. He has co-authored two monographs, an edited book with Cambridge University Press, and more than 100 research journal papers. His research interests include wireless communications, information theory, optimization, and machine learning. He is a Fellow of the IET. He was a co-recipient of the 2018 IEEE Signal Processing Best Paper Award, the 2017 JCN Best Paper Award, the 2015 IEEE Communication Society Best Tutorial Paper Award, and the Best Paper Awards of IEEE SPAWC 2007 and the IEEE WRECOM 2007. He was awarded a Consolidator Grant by the European Research Council (ERC) in 2016. His research has been supported by the U.S. NSF, by the ERC, by the Vienna Science and Technology Fund, and by a number of industrial collaborations. He currently serves on the Editorial Board of the *IEEE Signal Processing Magazine*. He is a Distinguished Lecturer of the IEEE Information Theory Society.

Inkyu Lee (S'92–M'95–SM'01–F'16) received the B.S. degree (Hons.) in control and instrumentation engineering from Seoul National University, Seoul, South Korea, in 1990, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, USA, in 1992 and 1995, respectively. From 1995 to 2001, he was a Member of Technical Staff with Bell Laboratories, Lucent Technologies, where he was involved in high-speed wireless system designs. From 2001 to 2002, he was with Agere Systems, Murray Hill, NJ, USA, as a Distinguished Member of Technical Staff. Since 2002, he has been with Korea University, Seoul, where he is currently a Professor with the School of Electrical Engineering. In 2009, he was a Visiting Professor with the University of Southern California, Los Angeles, CA, USA. He has authored over 170 journal papers in the IEEE. He holds 30 U.S. patents, granted or pending. His research interests include digital communications, signal processing, and coding techniques applied in next-generation wireless systems. He was a recipient of the IT Young Engineer Award from the IEEE, the IEEK Joint Award in 2006, and the Best Paper Award from APCC in 2006, the IEEE VTC in 2009, and ISPACS in 2013. He was also a recipient of the Best Research Award from the Korea Information and Communications Society in 2011, the Best Young Engineer Award from the National Academy of Engineering in Korea in 2013, and the Korea Engineering Award from the National Research Foundation of Korea in 2017. He has served as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS from 2001 to 2011, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2007 to 2011, the IEEE WIRELESS COMMUNICATIONS LETTERS from 2011 to 2013, and the IEEE ACCESS from 2016 to 2017. He was a Chief Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS Special Issue on 4G Wireless Systems in 2006. He is a member of National Academy of Engineering in Korea and an IEEE Distinguished Lecturer.

Shlomo Shamai (Shitz) (S'80–M'82–SM'89–F'94–LF'18) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the Technion–Israel Institute of Technology, in 1975, 1981, and 1986, respectively.

From 1975 to 1985, he was with Communications Research Laboratories as a Senior Research Engineer. Since 1986, he has been with the Department of Electrical Engineering, Technion–Israel Institute of Technology, where he is currently a Technion Distinguished Professor and holds the William Fondiller Chair of Telecommunications. His research interests span a wide spectrum of topics in information theory and statistical communications.

Dr. Shamai (Shitz) is a Union Radio Scientifique Internationale (URSI) Fellow, a Member of the Israeli Academy of Sciences and Humanities, and a Foreign Member of the U.S. National Academy of Engineering. He was a recipient of the 2011 Claude E. Shannon Award, the 2014 Rothschild Prize in Mathematics/Computer Sciences and Engineering, and the 2017 IEEE Richard W. Hamming Medal. He was a co-recipient of the 2018 Third Bell Labs Prize for Shaping the Future of Information and Communications Technology. He has been awarded the 1999 van der Pol Gold Medal of the URSI and was a co-recipient of the 2000 IEEE Donald G. Fink Prize Paper Award, the 2003 and 2004 Joint IT/COM Societies Paper Award, the 2007 IEEE Information Theory Society Paper Award, the 2009 and 2015 European Commission FP7, Network of Excellence in Wireless COMmunications (NEWCOM++ and NEWCOM#) Best Paper Awards, the 2010 Thomson Reuters Award for International Excellence in Scientific Research, the 2014 EURASIP Best Paper Award (for the *EURASIP Journal on Wireless Communications and Networking*), the 2015 IEEE Communications Society Best Tutorial Paper Award, and the 2018 IEEE Signal Processing Best Paper Award. He was listed as a Highly Cited Researcher (Computer Science) in 2013–2018. He was also a recipient of the 1985 Alon Grant for distinguished young scientists and the 2000 Technion Henry Taub Prize for Excellence in Research. He has served as an Associate Editor for the Shannon Theory of the IEEE TRANSACTIONS ON INFORMATION THEORY, and has also served twice on the Board of Governors for the IEEE Information Theory Society. He has also served on the Executive Editorial Board of the IEEE TRANSACTIONS ON INFORMATION THEORY and on the IEEE Information Theory Society Nominations and Appointments Committee, and serves on the IEEE Information Theory Society and the Shannon Award Committee.