IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received August 31, 2021, accepted September 14, 2021, date of publication September 29, 2021, date of current version October 27, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3116375

Precoder and Spatial Compression Filter Designs for Uplink Cloud Radio Access Networks

GUAN-YU CHEN, SHANG-HO TSAI, (Senior Member, IEEE), AND XIANG-QUAN SER
Department of Electrical Engineering, National Chiao Tung University, Hsinchu 30010, Taiwan

Corresponding author: Shang-Ho Tsai (shanghot@alumni.usc.edu)

This work was supported by the Ministry of Science and Technology, Taiwan, under Grant MOST 109-2221-E-009-106 and Grant MOST 109-2218-E-009-003.

ABSTRACT This work considers an uplink cloud radio access network (C-RAN), and proposes a MIMO spatial compression filtering and which compressed symbols to upload at RRH side as well as MIMO precoding at UE (user equipment) side. The error rate performance of this proposed scheme is analyzed and closed-form analytical results are obtained. Based on the analytical results, we propose solutions for the spatial compression filter, compressed symbols and UE precoder. Fronthaul loading can be significantly reduced by the proposed MIMO spatial compression filter and compressed symbols; meanwhile this proposed scheme can greatly outperform conventional schemes under the same settings of fronthaul loading. Moreover, when the UE has multiple antennas, the performance can further be significantly improved by using the proposed UE precoder. Simulation results are provided to show the accuracy of the analytical results as well as the advantages of the proposed solutions in terms of error rate performance and reduced fronthaul loading.

INDEX TERMS C-RAN, CoMP, 5G/B5G, spatial compression filter, joint decompression and detection, multiple antennas, MIMO, precoding, fronthaul loading reduction.

I. INTRODUCTION

Cloud Radio Access Network (C-RAN) has received intensive attention in recent years, and it is a potential candidate for 5G/B5G networks to support the increasing demand for high data rate transmission. In a C-RAN, most baseband processing functionalities are moved to baseband units (BBUs), and the conventional full-functionality base stations are replaced by low-cost remote radio heads (RRHs), which handles data transmission and reception with user equipments (UEs) [1], [2]. The UEs communicate with distributed RRHs that are connected to the BBUs via high capacity and low latency fronthaul networks. The centralized BBU has global information for the whole system and it can process most of the computational tasks [3]. Such architectures lead to many advantages including low deployment cost, better energy efficiency, and improved spectrum utilization [4], [5].

Fronthaul loading between BBUs and RRHs in a C-RAN is usually limited due to the capability of fiber bandwidth and considerations of deployment cost. Extensive research has been conducted in studying fronthaul loading issues. One research direction is to optimize certain performance metrics under a limited link capability. The authors in [6] proposed precoders to maximize energy efficiency. Precoders were proposed in [7], [8] to minimize power consumption by flexible functional partitioning between RRHs and the BBU. In [9] and [10], the authors used a point-to-point compression [11] to achieve the maximum fronthaul capacity; while in [12] and [13], the authors maximized the sum rate via optimizing the quantization noise from an information-theoretic perspective. Also, the authors in [14] consider the imperfect channel state information at the transmitter due to finite-bandwidth feedback links to design the precoders for maximizing the sum rate. Another research direction is to reduce the fronthaul loading by compressing the data. Several methods can be used to reduce the transmission rate of the fronthaul IQ data, including encoder, quantization, and spatial compression. To name a few. The authors in [15] used the Free Lossless Audio Codec (FLAC) to compress the IQ data. In [16], the authors proposed encoding methods as 1) Unused significant bit removal (USBR) [17], 2) adaptive arithmetic coding (AC) [18], 3) Elias-gamma coding [19], and 4) least

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

significant bits (LSB) to compress the IQ data. The authors in [20]–[23] used different quantization methods to compress the IQ data by adjusting the quantizer resolution. The work in [24] proposed a spatial filter to reduce symbols before transmitting from the RRHs to the BBU. In [20], the authors proposed a new joint decompression and detection (JDD) algorithm that can significantly improve the error rate performance.

From the introduction above, several methods were studied to solve the fronthaul loading issues. For example, most research outcomes were developed for MIMO channels. Hence, optimal precoding and spatial compression schemes were proposed *e.g.,* see [10], [12], [24]. However, few have been studied for optimal symbol detections. Although a clever joint decompression and detection algorithm was proposed in [20], it was only for SISO systems and thus there are no precoder designs and spatial compression. Hence the corresponding designs of UE precoder and spatial compression are still unknown. Moreover extension from SISO to MIMO channel also increases fronthaul loading because more received symbols from increased antennas at RRHs need to be processed. Hence new theoretical results and algorithms need to be developed.

Contributions: The contributions of this work are summarized below: 1.) Proposed a MIMO C-RAN architecture for uplink transmission, which consists of precoders at the UE side, spatial compression filtering at the RRH side, and the joint decompression and detection at the BBU side. More specifically, we analyze the performance of the proposed scheme and derive closed-form expressions for the error rate performance. The derived results are accurate, observed from the well-matched simulation results. Based on the analytical results, we propose closed-form solutions for the spatial compression filters, determine which symbols to compress and upload to the BBU, and design UE precoders to improve system performance.

2.) Great performance improvement and fronthaul loading reduction achieved by the proposed MIMO SCF and compressed symbols. Simulation results show that the proposed scheme significantly outperforms conventional schemes. More specifically, when the UE has only one transmit antenna while the BBU has multiple antennas, the participation of the proposed spatial compression filter can greatly outperform the conventional SISO scheme in [20] by around 7.5 dB under the same compression ratio for data from the RRHs to the BBU. In addition, the proposed MIMO spatial compression filter enables more compression ratios than the conventional SISO scheme, and thus it can well leverage between fronthaul loading and performance by controlling the compression ratio. Moreover, the proposed MIMO spatial compression filter also outperforms several conventional schemes, *e.g.,* around 3 dB gap when comparing to the widely used MIMO maximum ratio transmission (MRT).

3.) Performance further improved by the proposed UE precoders. When the UEs have multiple transmit antennas, the performance can be further improved by using the proposed precoders at the UE side. Simulation results show that around 3 dB additional performance gain can be achieved using the proposed precoders when the number of UE antennas increases from one to two, and more performance improvement can be attained with more UE antennas. Moreover, thanks to the optimality of the proposed precoders, system performance always improves by increasing the number of UE antennas. On the contrary, there is no guarantee for performance improvement in the conventional precoders with the increase of UE antennas, as will be demonstrated in Experiment 7 later.

There are several differences between the conventional and the proposed schemes, which are stated below: It is worth pointing out that joint decompression and detection has been proposed for SISO channels in [20]. However, extending the joint decompression and detection from SISO to MIMO channels is not trivial. Here, several technical bounds and approximations are needed to obtain closed-form results. Based on the results, one can further develop appropriate solutions, *e.g.,* the spatial compression filters and UE precoders in this work. Without the developed results in this work, the joint decompression and detection does not necessarily lead to a good performance in MIMO channels that will be demonstrated later in the simulation results. In addition, although the concept of the SCF was proposed in [24], there are several differences summarized as follows: First, the SCF in [24] demands that each RRH transmits a minimum number of symbols to the BBU, which equals to the minimum one of either the number of antennas at RRH side or the total user symbols. On the other hand, there is no such constraint in the proposed work once the number of total symbols from all RRHs is greater than the number of user symbols. As a result, the fronthaul loading in the proposed SCF can be more efficiently reduced than that in [24]. Second, the SCF in [24] considers SINR as the design criterion; while the performance metric of the proposed SCF is the error rate due to the use of the joint decompression and detection. Consequently, the derivations, used techniques and the obtained solutions are totally different from those in [24].

The remaining parts of this paper are organized as follows: In Sec. II, system model, problem formulation and background review are provided. Performance is analyzed and closed-form expressions are derived in Sec. III. The proposed solutions spatial compression filter, symbols to be compressed and uploaded, and UE precoders are then introduced in Sec. IV. Simulation results and performance comparisons are demonstrated in Sec. V. Conclusion remarks are provided in Sec. VI.

Notations: Boldfaced lowercases and boldfaced uppercases denote vectors and matrices, respectively. $\mathbb{E}\{x\}$ is the expectation of the random variable x. $\mathbf{A}^*$ and $\mathbf{A}^T$ denote the conjugate and transpose of $\mathbf{A}$, respectively. $\mathbf{A}^H$ is the Hermitian matrix of $\mathbf{A}$. $\|\mathbf{A}\|$ is the 2-norm of $\mathbf{A}$. σ_x^2 is the variance of x. N is the number of RRHs and M is the number of antennas per RRH, respectively. K is the number of UEs and J is the number of antennas per UE, respectively. L is the

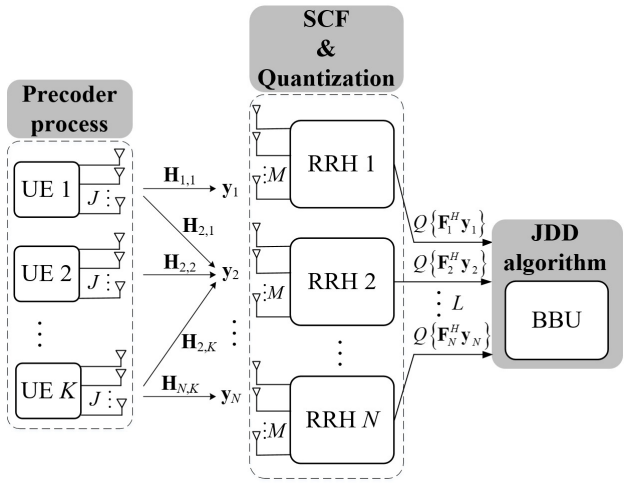

FIGURE 1. The proposed uplink C-RAN architecture.

number of symbols uploaded to BBU and Q is the number of quantization bits, respectively.

II. SYSTEM MODEL, PROBLEM FORMULATION AND BACKGROUND REVIEW

A. PROPOSED C-RAN ARCHITECTURE

Consider an uplink C-RAN system in Fig. 1. Assume that there are K UEs with J antennas per UE, N RRHs with M antennas per RRH, and one BBU. Each UE links to the RRHs via the wireless network and the RRHs links to the BBU via the fronthaul network. We assume that the BBU has full channel state information (CSI) between the users and the RRHs, and the BBU can calculate the precoders. Then the precoders are broadcasted from the BBU or via RRHs to the users through a reliable control channel. Thus, the proposed scheme is a closed-loop system. Although UE precoding increases the overhead, it can improve the system performance significantly as later will be shown in the simulation results.

Let us look the details of the system. First at the UE side, each UE transmits one data stream to the RRHs. The received symbols from all UEs at the nth RRH can be expressed by

$$\mathbf{y}_n = \sum_{k=1}^{K} \mathbf{H}_{n,k}\mathbf{g}_k\sqrt{P_k}c_k + \mathbf{z}_n = \mathbf{H}_n\mathbf{G}\mathbf{P}\mathbf{c} + \mathbf{z}_n, \tag{1}$$

where $c_k \in \mathbb{C}^{1\times 1}$ is the transmitted symbol of the kth UE with unit power constraint, *i.e.*, $\mathbb{E}\{c_k c_k^*\} = 1, \forall k = 1, \cdots, K$, and $\mathbf{c} = [c_1, \cdots, c_K]^T$ is the data vector from all UEs to the RRHs. $\mathbf{P} = diag([\sqrt{P_1}, \cdots, \sqrt{P_K}])$ is a diagonal matrix that denotes the transmit power of all UEs. $\mathbf{g}_k \in \mathbb{C}^{J\times 1}$ is the precoding vector of the kth UE, and $\mathbf{G}$ is a block diagonal matrix with dimension $\mathbf{G} \in \mathbb{C}^{JK\times K}$, consisting of the precoding vectors for all UEs, given by

$$\mathbf{G} = \begin{bmatrix} \mathbf{g}_1 & & & \mathbf{0} \\ & \mathbf{g}_2 & & \\ & & \ddots & \\ \mathbf{0} & & & \mathbf{g}_K \end{bmatrix}.$$

$\mathbf{H}_{n,k} \in \mathbb{C}^{M\times J}$ is the channel matrix between the nth RRH and the kth UE and we define the channel matrix from all UEs to the nth RRH as

$$\mathbf{H}_n = [\mathbf{H}_{n,1}, \cdots, \mathbf{H}_{n,K}]. \tag{2}$$

Each element of the channel matrix is with independent and identically distributed (i.i.d.) complex Gaussian distribution. $\mathbf{z}_n = [z_1, \cdots, z_M]^T$ is the noise vector of the nth RRH, which has i.i.d. complex Gaussian distribution with zero mean and variance σ^2.

For a compression purpose, we reduce the dimension of the received MIMO vector $\mathbf{y}_n$ at the nth RRH via spatial compression filtering (SCF). More specifically, referring to Eq. (1), after the SCF process, the received symbols $\mathbf{y}'_n$ can be expressed as

$$\mathbf{y}'_n = \mathbf{F}_n^H\mathbf{y}_n, \tag{3}$$

where $\mathbf{y}'_n \in \mathbb{C}^{L_n\times 1}$, and L_n is the number of compressed symbols at the nth RRH. Note that the dimension of the received symbol vector is reduced from M to L_n. $\mathbf{F}_n \in \mathbb{C}^{M\times L_n}$ is the SCF matrix of the nth RRH.

B. SCALAR QUANTIZATION

Before transmitting the spatially compressed symbols to the BBU, scalar quantization is applied to further reduce data size. We apply the "three-sigma" rule [25], which was also applied in [20], to bound the compressed symbols within $[-1, 1]$. The lth element of the compressed symbols $\bar{\mathbf{y}}_n$ can be bounded as

$$[\bar{\mathbf{y}}_n]_l = \left[\mathbf{y}'_n\right]_l / \eta_{n,l}, \tag{4}$$

where

$$\eta_{n,l} = 3\sqrt{\left\|[\mathbf{F}_n]_{:,l}^H \mathbf{H}_n\mathbf{G}\mathbf{P}\right\|^2 + \sigma^2 [\mathbf{F}_n]_{:,l}^H [\mathbf{F}_n]_{:,l}}, \tag{5}$$

is a factor obtained by the "three-sigma" rule [25] to bound the compressed symbols. Also, $[\mathbf{A}]_{:,l}$ denotes the l column of matrix $\mathbf{A}$. Then, according to [26], the quantized symbols can be expressed as

$$\left[\tilde{\mathbf{y}}_n\right]_l = \eta_{n,l}\frac{round\left([\bar{\mathbf{y}}_n]_l \times 2^{Q_n-1}\right)}{2^{Q_n-1}}, \tag{6}$$

where $round(\cdot)$ is the rounding process, and Q_n is the quantization bits at the nth RRH. The quantization error can be calculated as $[\mathbf{q}_{e,n}]_l = \left[\mathbf{y}'_n\right]_l - \left[\tilde{\mathbf{y}}_n\right]_l = \Re\{[\mathbf{q}_{e,n}]_l\} + \Im\{[\mathbf{q}_{e,n}]_l\}$. From experiments, the quantization error $\Re\{[\mathbf{q}_{e,n}]_l\}$, $\Im\{[\mathbf{q}_{e,n}]_l\}$ can be approximated by uniform random variables within $\left[-\delta_{n,l}, \delta_{n,l}\right]$, where $\delta_{n,l} = \eta_{n,l}2^{-Q_n}$.

C. REVIEW OF JOINT DECOMPRESSION AND DETECTION

In this subsection, we review a joint decompression and detection (JDD) algorithm proposed in [20], which is applied in the proposed system. It is worth pointing out that this JDD algorithm was proposed for SISO channels. Hence there were no precoding and spatial compression filtering in [20].

The fundamental ideal of the JDD algorithm is the maximum a posteriori probability (MAP) detection, and the detected codeword $\hat{\mathbf{c}}$ can be expressed as [27]

$$\begin{aligned}\hat{\mathbf{c}} &= \arg\max_{\mathbf{c}} \mathrm{P}\left\{\mathbf{c}\,\middle|\,[\tilde{\mathbf{y}}_1]_1, \cdots, [\tilde{\mathbf{y}}_1]_{L_n}, [\tilde{\mathbf{y}}_2]_1, \cdots, [\tilde{\mathbf{y}}_N]_{L_n}\right\} \\ &\overset{(\star)}{=} \arg\max_{\mathbf{c}} \mathrm{P}\left\{\mathbf{c}, [\tilde{\mathbf{y}}_1]_1, \cdots, [\tilde{\mathbf{y}}_1]_{L_n}, [\tilde{\mathbf{y}}_2]_1, \cdots, [\tilde{\mathbf{y}}_N]_{L_n}\right\} \\ &\overset{(\star\star)}{=} \arg\max_{\mathbf{c}} \mathrm{P}\{\mathbf{c}\} \prod_{n=1}^{N}\prod_{l=1}^{L_n} \mathrm{P}\left\{[\tilde{\mathbf{y}}_n]_l\,|\mathbf{c}\right\}, \end{aligned} \tag{7}$$

where $(\star)$ is because $\mathrm{P}\left\{[\tilde{\mathbf{y}}_1]_{L_n}, [\tilde{\mathbf{y}}_2]_1, \cdots, [\tilde{\mathbf{y}}_N]_{L_n}\right\}$ is a constant for any codeword, and $(\star\star)$ is that $[\tilde{\mathbf{y}}_n]_l$'s are mutually independent. The mutual independency of $[\tilde{\mathbf{y}}_n]_l$'s can be shown using the facts that the columns of the proposed SCF $\mathbf{F}_n$ at the nth RRH are orthonormal (see Proposition 1 introduced later), and similar assumptions in [20] that the elements of $\mathbf{y}_n$ are independent and Gaussian.

$\mathrm{P}\left\{[\tilde{\mathbf{y}}_n]_l\,|\mathbf{c}\right\}$ in Eq. (7) is the probability of the quantized outputs $[\tilde{\mathbf{y}}_n]_l$ via the observation of $[\mathbf{y}_n]_l$. It can be interpreted as that the quantizer output is $[\tilde{\mathbf{y}}_n]_l$ if both $|\Re\{[\mathbf{y}_n]_l\} - \Re\{[\tilde{\mathbf{y}}_n]_l\}|$ and $|\Im\{[\mathbf{y}_n]_l\} - \Im\{[\tilde{\mathbf{y}}_n]_l\}|$ are less than the quantization error $\delta_{n,l}$. Since the quantization is performed independently for the real and the imaginary parts, $\mathrm{P}\left\{[\tilde{\mathbf{y}}_n]_l\,|\mathbf{c}\right\}$ can be written as follows:

$$\mathrm{P}\left\{[\tilde{\mathbf{y}}_n]_l\,|\mathbf{c}\right\} = \mathrm{P}\{\mathbb{A}\cap\mathbb{B}\} = \mathrm{P}\{\mathbb{A}\}\,\mathrm{P}\{\mathbb{B}\}, \tag{8}$$

where $\mathbb{A} = [\mathbf{y}_n^R]_l \in \left[[\tilde{\mathbf{y}}_n^R]_l - \delta_{n,l}, [\tilde{\mathbf{y}}_n^R]_l + \delta_{n,l}\right]$, $\mathbb{B} = [\mathbf{y}_n^I]_l \in \left[[\tilde{\mathbf{y}}_n^I]_l - \delta_{n,l}, [\tilde{\mathbf{y}}_n^I]_l + \delta_{n,l}\right]$, and y^R and y^I are respectively the real and the imaginary parts of y.

Then, since $[\mathbf{y}_n^R]_l$ and $[\mathbf{y}_n^I]_l$ can be approximated by the Gaussian random variables with the same variance $\frac{\sigma^2}{2}$ and mean $\Re\{[\mathbf{F}_n]_{:,l}^H\mathbf{H}_n\mathbf{G}\mathbf{P}\mathbf{c}\}$ and $\Im\{[\mathbf{F}_n]_{:,l}^H\mathbf{H}_n\mathbf{G}\mathbf{P}\mathbf{c}\}$, respectively, the probability density functions (PDFs) of $[\mathbf{y}_n^R]_l$ and $[\mathbf{y}_n^I]_l$ are given respectively by

$$\begin{aligned} &f\left(\left[\mathbf{y}_n^R\right]_l|\mathbf{c}\right) \\ &= \frac{1}{\sqrt{\pi}\sigma}\exp\left(\frac{-\left|[\mathbf{y}_n^R]_l - \Re\left\{[\mathbf{F}_n]_{:,l}^H\mathbf{H}_n\mathbf{G}\mathbf{P}\mathbf{c}\right\}\right|^2}{\sigma^2}\right), \end{aligned} \tag{9}$$

$$\begin{aligned} &f\left(\left[\mathbf{y}_n^I\right]_l|\mathbf{c}\right) \\ &= \frac{1}{\sqrt{\pi}\sigma}\exp\left(\frac{-\left|[\mathbf{y}_n^I]_l - \Im\left\{[\mathbf{F}_n]_{:,l}^H\mathbf{H}_n\mathbf{G}\mathbf{P}\mathbf{c}\right\}\right|^2}{\sigma^2}\right). \end{aligned} \tag{10}$$

Substituting Eqs. (9)-(10) into Eq. (8), $\mathrm{P}\left\{[\tilde{\mathbf{y}}_n]_l\,|\mathbf{c}\right\}$ is defined as the block error rate (BLER) given by

$$\begin{aligned} &\int_{[\tilde{\mathbf{y}}_n^R]_l-\delta_{n,l}}^{[\tilde{\mathbf{y}}_n^R]_l+\delta_{n,l}} f\left(\left[\mathbf{y}_n^R\right]_l|\mathbf{c}\right) d\left[\mathbf{y}_n^R\right]_l \\ &\cdot\int_{[\tilde{\mathbf{y}}_n^I]_l-\delta_{n,l}}^{[\tilde{\mathbf{y}}_n^I]_l+\delta_{n,l}} f\left(\left[\mathbf{y}_n^I\right]_l|\mathbf{c}\right) d\left[\mathbf{y}_n^I\right]_l \\ &= \frac{1}{4}\times\left[\mathrm{erfc}\left(\theta_1 - \frac{\delta_{n,l}}{\sigma}\right) - \mathrm{erfc}\left(\theta_1 + \frac{\delta_{n,l}}{\sigma}\right)\right] \\ &\times\left[\mathrm{erfc}\left(\theta_2 - \frac{\delta_{n,l}}{\sigma}\right) - \mathrm{erfc}\left(\theta_2 + \frac{\delta_{n,l}}{\sigma}\right)\right], \end{aligned} \tag{11}$$

where

$$\theta_1 = \frac{\Re\left\{[\tilde{\mathbf{y}}_n]_l\right\} - \Re\{\alpha\}}{\sigma},\ \theta_2 = \frac{\Im\left\{[\tilde{\mathbf{y}}_n]_l\right\} - \Im\{\alpha\}}{\sigma},$$

and $\alpha = [\mathbf{F}_n]_{:,l}^H\mathbf{H}_n\mathbf{G}\mathbf{P}\mathbf{c}$. Substituting Eq. (11) into Eq. (7), we have the detection rule for the transmitted symbol vector $\mathbf{c}$.

D. PROBLEM FORMULATION

Although the SCF and quantization compress the data and thus reduce the fronthaul loading, they degrade system performance. It is desired to design the SCF to minimize the performance degradation under a fixed compression ratio, which relates to the transmission rate budget of fiber equipment. Moreover, designing the precoders at the UE side can further improve the performance. Hence we would like to design the SCF and the precoders as well as determine which symbols to transmit to the BBU to minimize the BLER. The problem can be formulated as follows:

Problem to be solved: Designing precoders and spatial compression filter as well as determining which symbols to upload to minimize BLER.

$$\begin{aligned} (\hat{\mathbf{F}}, \hat{\mathbf{G}}, \hat{L}_n) = &\arg\min_{\mathbf{F},\mathbf{G},L_n} \mathrm{BLER}, \\ &\text{s.t.}\ \left\|[\mathbf{F}]_{:,l}\right\| = 1 \quad \forall l = 1, 2, \cdots, L, \\ &\left\|[\mathbf{G}]_{:,k}\right\| = 1 \quad \forall k = 1, 2, \cdots, K, \\ &\sum_{n=1}^{N} L_n = L, \end{aligned} \tag{12}$$

where recall that $\mathbf{G}$ is the precoding matrix of all UEs, $\mathbf{F} = \left[\mathbf{F}_1\ \mathbf{F}_2\ \cdots\ \mathbf{F}_N\right] \in \mathbb{C}^{M\times L}$ is the matrix consisting of the SCFs at all N RRHs, and L_n is the number of compressed symbols at the nth RRH, which is to be uploaded to the BBU. It implies that each RRH may transmit different numbers of compressed symbols to the BBU.

III. PERFORMANCE ANALYSIS

In this section, we analyze the BLER performance of the proposed system and obtain corresponding closed-form expressions. Then, solutions are proposed according to the analytical results in the next section.

Referring to the problem formulation in (12), the BLER in the problem formulation can be expressed as [20]

$$\mathrm{BLER} \triangleq \frac{1}{|S|^K} \sum_{\mathbf{c}\in|S|^K} \sum_{\mathbf{c}\neq\tilde{\mathbf{c}}} \mathrm{P}\{\mathbf{c} \to \tilde{\mathbf{c}}\}, \tag{13}$$

where $|S|^K$ is the size of symbol constellation, *e.g.,* if the modulation is 4-QAM, the size of constellation is 4^K, where K is the number of UEs. $\mathrm{P}\{\mathbf{c} \to \tilde{\mathbf{c}}\}$ is the instantaneous error when $\mathbf{c}$ is transmitted and $\tilde{\mathbf{c}}$ is detected.

To compute the instantaneous error, the quantization effect should be included. Referring to Eqs. (3)-(5), at the nth RRH, the received symbols after the SCF process is given by

$$\tilde{\mathbf{y}}_n = \mathbf{F}_n^H \left(\mathbf{H}_n\mathbf{G}\mathbf{P}\mathbf{c} + \mathbf{z}_n\right) + \mathbf{q}_{e,n}, \tag{14}$$

where $\tilde{\mathbf{y}}_n \in \mathbb{C}^{L_n\times 1}$ is the compressed symbols, which are to be transmitted to the BBU, at the nth RRH. The term $\mathbf{q}_{e,n} = \Re\{\mathbf{q}_{e,n}\} + i\Im\{\mathbf{q}_{e,n}\} \in \mathbb{C}^{L_n\times 1}$ is the quantization error of the nth RRH, and $\Re\{\mathbf{q}_{e,n}\}$, $\Im\{\mathbf{q}_{e,n}\}$ are mutually independent.

As mentioned in Subsection II-B, the quantization noise can be approximated as the uniform random variables, and $\left[\Re\{\mathbf{q}_{e,n}\}\right]_l$, $\left[\Im\{\mathbf{q}_{e,n}\}\right]_l$ are in $U\left[-\delta_{n,l}, \delta_{n,l}\right], \forall l = 1, \cdots, L$. The variance of $q_{e,n}$ can be computed as [28]

$$\left[\sigma^2_{q_{e,n}}\right]_l = \mathrm{Var}\{\left[\Re\{\mathbf{q}_{e,n}\}\right]_l\} + \mathrm{Var}\{\left[\Im\{\mathbf{q}_{e,n}\}\right]_l\} = \frac{2\delta^2_{n,l}}{3}. \tag{15}$$

Assume that the constellations of individual UEs are uniformly distributed. Hence $\mathrm{P}\{\mathbf{c}\}$ in (7) is a constant, and the MAP detection reduces to the maximum likelihood detection. The instantaneous error occurs if the likelihood function of $\mathbf{c}$ is less than that of $\tilde{\mathbf{c}}$, which can be expressed as

$$\mathrm{P}\{\mathbf{c} \to \tilde{\mathbf{c}}\} \triangleq \mathrm{P}\{\mathbb{M}(\mathbf{c}) < \mathbb{M}(\tilde{\mathbf{c}})\}, \tag{16}$$

where

$$\mathbb{M}(\mathbf{c}) = \prod_{n=1}^{N}\prod_{l=1}^{L_n} \mathrm{P}\left\{\left[\tilde{\mathbf{y}}_n\right]_l | \mathbf{c}\right\}. \tag{17}$$

Eq. (17) is the likelihood function used in the JDD algorithm, and one can directly compute Eq. (17) by Eq. (11). Since Eq. (11) is too complicated to tackle, the first order Taylor approximation can be used to simplify the problem. The first Taylor approximation is that $f(x) \simeq f(x_0) + f'(x_0)(x - x_0)$ with x_0 being a feasible point, and applying it to Eq. (11) with $[\mathbf{x}_0]_l = \Re\{\tilde{\mathbf{y}}_n - \left[\mathbf{F}_n^H\mathbf{H}_n\mathbf{G}\mathbf{P}\mathbf{c}\right]_l\}/\sigma$ for the real part and $[\mathbf{x}_0]_l = \Im\{\tilde{\mathbf{y}}_n - \left[\mathbf{F}_n^H\mathbf{H}_n\mathbf{G}\mathbf{P}\mathbf{c}\right]_l\}/\sigma$ for the imaginary part, where σ is the noise standard deviation. Eq. (11) can be approximated by

$$\mathrm{P}\{\left[\tilde{\mathbf{y}}_n\right]_l | \mathbf{c}\} \simeq 4\frac{\delta^2_{n,l}}{\pi\sigma^2} \exp\left(-\frac{\left|\left[\tilde{\mathbf{y}}_n\right]_l - \left[\mathbf{F}_n^{\mathbf{H}}\mathbf{H}_n\mathbf{G}\mathbf{P}\mathbf{c}\right]_l\right|^2}{\sigma^2}\right), \tag{18}$$

where l is the index of the data stream at the nth RRH. Then, substitute Eq. (18) into Eq. (17), we have

$$\mathbb{M}(\mathbf{c}) \simeq 4\frac{\delta^2_{n,l}}{\pi\sigma^2} \exp\left(-\frac{\sum_{n=1}^{N}\sum_{l=1}^{L_n}\left|\left[\tilde{\mathbf{y}}_n\right]_l - \left[\mathbf{F}_n^{\mathbf{H}}\mathbf{H}_n\mathbf{G}\mathbf{P}\mathbf{c}\right]_l\right|^2}{\sigma^2}\right). \tag{19}$$

We assume that the symbols after the SCF process are mutually independent. Since the quantization is also performed independently for individual symbols, the term $\left[\tilde{\mathbf{y}}_n\right]_l$'s are mutually independent.

Lemma 1: Let $\mathbf{e} \triangleq \left(\tilde{\mathbf{c}} - \mathbf{c}\right)$. *The instantaneous error can be approximated as*:

$$P\{\mathbf{c} \to \tilde{\mathbf{c}}\} \approx \frac{1}{2}erfc\left(\underbrace{\frac{f_1\left([\mathbf{F}_n]_{:,l}, \mathbf{H}_n\right)}{\sqrt{4\sigma^2 f_2\left([\mathbf{F}_n]_{:,l}, \mathbf{H}_n\right) + \frac{8}{3}f_3\left([\mathbf{F}_n]_{:,l}, \mathbf{H}_n\right)}}}_{x}\right), \tag{20}$$

where

$$f_1\left([\mathbf{F}_n]_{:,l}, \mathbf{H}_n\right) = \sum_{n=1}^{N}\sum_{l=1}^{L_n}\left|[\mathbf{F}_n]_{:,l}^H \mathbf{H}_n\mathbf{G}\mathbf{P}\mathbf{e}\right|^2, \tag{21}$$

$$f_2\left([\mathbf{F}_n]_{:,l}, \mathbf{H}_n\right) = \sum_{n=1}^{N}\sum_{l=1}^{L_n}\left\|[\mathbf{F}_n]_{:,l}[\mathbf{F}_n]_{:,l}^H \mathbf{H}_n\mathbf{G}\mathbf{P}\mathbf{e}\right\|^2, \tag{22}$$

$$f_3\left([\mathbf{F}_n]_{:,l}, \mathbf{H}_n, \delta^2_{n,l}\right) = \sum_{n=1}^{N}\sum_{l=1}^{L_n}\delta^2_{n,l}\left|[\mathbf{F}_n]_{:,l}^H \mathbf{H}_n\mathbf{G}\mathbf{P}\mathbf{e}\right|^2. \tag{23}$$

Proof: See Appendix A. ■

Eq. (20) indicates that when the Gaussian noise σ^2 and the quantization noise $\delta^2_{n,l}$ are larger, the BLER performance is poor. Substituting Eq. (20) into Eq. (13), one can obtain the theoretical BLER. Later in Sec. V, we will show by simulation results that the theoretical BLER indeed well approximates the simulated BLER.

IV. PROPOSED SOLUTIONS

In this section, we propose solutions according to the analytical results in Sec. III. The problem can be solved in two steps. First, we can design the SCF and determine which compressed symbols to upload first. Then, the precoders can be optimized with the determined SCF and compressed symbols. We iterate the two steps until it converges. It is worthwhile to point out closed-form solutions are derived for both steps. Hence the corresponding computational complexity is low since there is no need to use optimization toolboxes.

A. PROPOSED SOLUTIONS FOR SCF AND COMPRESSED SYMBOLS

In this subsection, we design the SCF and determine the compressed symbols. From Eq. (13), since $\mathbf{c}$ is the symbols to be detected, in practice, one cannot obtain the instantaneous error $\mathrm{P}\{\mathbf{c} \to \tilde{\mathbf{c}}\}$. For practical implementation, one may treat the BLER as the averaged error of $\mathrm{P}\{\mathbf{c} \to \tilde{\mathbf{c}}\}$ in terms of all different detection errors, and rewrite Eq. (13) as

$$\frac{1}{|S|^K} \sum_{\mathbf{c}\in|S|^K} \sum_{\mathbf{c}\neq\tilde{\mathbf{c}}} \mathrm{P}\{\mathbf{c} \to \tilde{\mathbf{c}}\} = \mathbb{E}_{\mathbf{e}}\{\mathrm{P}\left\{\mathbf{c} \to \tilde{\mathbf{c}}\right\}\}, \tag{24}$$

where recall that $\mathbf{e} \triangleq \left(\tilde{\mathbf{c}} - \mathbf{c}\right)$. Since x in Eq. (20) is always positive, Eq. (20) is a convex function. By Jensen's inequality, we have

$$\mathbb{E}_{\mathbf{e}}\{\mathrm{P}\left\{\mathbf{c} \to \tilde{\mathbf{c}}\right\}\} = \frac{1}{2}\mathbb{E}_{\mathbf{e}}\left\{\mathrm{erfc}\left(x\right)\right\} \geq \frac{1}{2}\mathrm{erfc}\left(\mathbb{E}_{\mathbf{e}}\left\{x\right\}\right). \tag{25}$$

We can use the right side term of Eq. (25) to approximate $\mathbb{E}_{\mathbf{e}}\{\mathrm{P}\left\{\mathbf{c} \to \tilde{\mathbf{c}}\right\}\}$. Later simulation results in Section V will show that this approximation is accurate. Since the complementary error function is monotonic decreasing, the problem can be reformulated as

$$\begin{aligned}
&(\hat{\mathbf{F}}_n, \hat{L}_n) \\
&= \underset{\mathbf{F}_n, L_n}{\arg\max} \\
&\quad \mathbb{E}_{\mathbf{e}}\left\{\frac{f_1\left([\mathbf{F}_n]_{:,l}, \mathbf{H}_n\right)}{\sqrt{4\sigma^2 f_2\left([\mathbf{F}_n]_{:,l}, \mathbf{H}_n\right) + \frac{8}{3}f_3\left([\mathbf{F}_n]_{:,l}, \mathbf{H}_n, \delta_{n,l}^2\right)}}\right\}, \\
&\quad \text{s.t.} \left\|[\mathbf{F}_n]_{:,l}\right\| = 1 \quad \forall l = 1, 2, \cdots, L, \\
&\quad \sum_{n=1}^{N} L_n = L.
\end{aligned} \tag{26}$$

The following two lemmas help in obtaining the solution for this problem.

Lemma 2: The term $\delta_{n,l}^2$ in the objective function of (26) has a Gamma distribution given by

$$\delta_{n,l}^2 \sim \Gamma\left(K, 9 \times P_k \times 2^{-2Q_n}\right). \tag{27}$$

Moreover, in the high SNR regime, as the number K of UEs approaches ∞, the term $\delta_{n,l}^2$ can be approximated by a Gaussian distribution given by

$$\delta_{n,l}^2 \sim \mathcal{N}\left(K \times 9 \times P_k \times 2^{-2Q_n}, K \times 81 \times P_k^2 \times 2^{-4Q_n}\right). \tag{28}$$

Furthermore, if the number Q_n of quantization bits approaches ∞, $\delta_{n,l}^2$ tends to be a constant.

Proof: See Appendix B. ■

Let us see an example for Lemma 2. Let the value of SNR be 10 dB. Fig. 2 shows the Gaussian approximated and the exact probability density functions in Lemma 2. Observe that the approximation is generally accurate and some curves are nearly overlapping. Moreover, when the value of K increases from 20 to 100, the approximation becomes more accurate. Furthermore, when the number Q_n of quantization bits increases, the term $\delta_{n,l}^2$ indeed has a small variance and tends to be a constant.

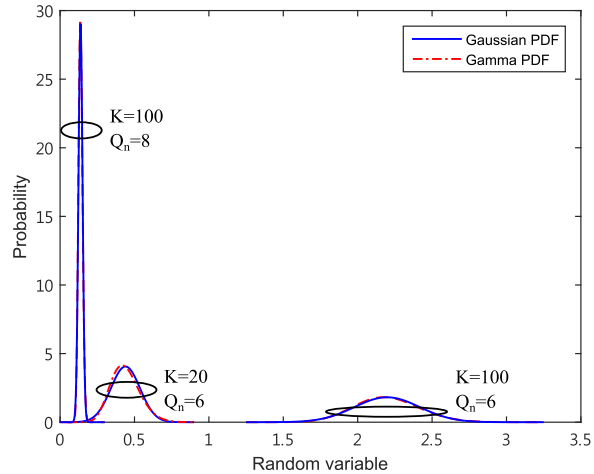

FIGURE 2. The Gamma distribution $\delta_{n,l}^2$ and corresponding Gaussian distribution with different parameter settings in 10 dB.

Let us introduce the second lemma to help us to simplify the objective function in (26).

Lemma 3: The objective function in (26) can be approximated by

$$\mathbb{E}_{\mathbf{e}}\left\{\frac{f_1\left([\mathbf{F}_n]_{:,l}, \mathbf{H}_n\right)}{\sqrt{4\sigma^2 f_2\left([\mathbf{F}_n]_{:,l}, \mathbf{H}_n\right) + \frac{8}{3}f_3\left([\mathbf{F}_n]_{:,l}, \mathbf{H}_n, \delta_{n,l}^2\right)}}\right\} \approx \sum_{n=1}^{N}\sum_{l=1}^{L_n}\left\|[\mathbf{F}_n]_{:,l}^H \mathbf{H}_n \mathbf{P}\right\|^2. \tag{29}$$

Proof: See Appendix C. ■

Using the result of Lemma 3, the problem can be reformulated as

$$\begin{aligned}
(\hat{\mathbf{F}}_n, \hat{\mathbf{G}}, \hat{L}_n) &= \underset{\mathbf{F}_n, \mathbf{G}, L_n}{\arg\max} \sum_{n=1}^{N}\sum_{l=1}^{L_n}\left\|[\mathbf{F}_n]_{:,l}^H \mathbf{H}_n \mathbf{G}\mathbf{P}\right\|^2, \\
&\text{s.t.} \left\|\left[\mathbf{F}_n\right]_{:,l}\right\| = 1 \quad \forall l = 1, 2, \cdots, L_n, \\
&\quad \left\|\left[\mathbf{G}\right]_{:,k}\right\| = 1 \quad \forall k = 1, 2, \cdots, K, \\
&\quad \sum_{n=1}^{N} L_n = L.
\end{aligned} \tag{30}$$

The proposed solution for this problem is introduced in the following proposition.

Proposition 1: Assume that the precoder $\mathbf{G}$ is determined. Let the L largest singular values of the N matrices $\mathbf{H}_1\mathbf{G}\mathbf{P}, \mathbf{H}_2\mathbf{G}\mathbf{P}, \cdots, \mathbf{H}_N\mathbf{G}\mathbf{P}$, $\sum_{n=1}^{N} L_n = L$, be denoted by $[\mathbf{V}_1]_{(1)} \cdots [\mathbf{V}_1]_{(L_1)}$, $[\mathbf{V}_2]_{(1)} \cdots [\mathbf{V}_2]_{(L_2)}$, $\cdots, [\mathbf{V}_N]_{(1)} \cdots [\mathbf{V}_N]_{(L_N)}$. Then the optimal SCF for the problem in (30) is given by

$$\mathbf{F}_n = \left[\,[\mathbf{V}_n]_{(1)} \cdots [\mathbf{V}_n]_{(L_n)}\,\right], \quad n = 1, 2, \cdots, N. \tag{31}$$

The nth RRH compresses L_n symbols and sends them to the BBU, where L_n is the number of singular values of $\mathbf{H}_n\mathbf{P}$ belong to the L largest singular values of the N matrices $\mathbf{H}_1\mathbf{GP}, \mathbf{H}_2\mathbf{GP}, \cdots, \mathbf{H}_N\mathbf{GP}$.

Proof: See Appendix D. ■

Remark 1: From Proposition 1, the numbers of symbols L_n to be uploaded from individual RRHs to the BBU can be different. Since the total number of uploaded symbols is fixed to be L, an RRH, say RRH n, uploads more symbols if its corresponding equivalent channel $\mathbf{H}_n\mathbf{GP}$ has more singular values that are among the L largest singular values of all N matrices $\mathbf{H}_1\mathbf{GP}, \mathbf{H}_2\mathbf{GP}, \cdots, \mathbf{H}_N\mathbf{GP}$.

B. PROPOSED PRECODER DESIGN

Now we consider the precoder design assuming that the SCF and compressed symbols are determined according to the results in the previous subsection. According to the result in Lemma 3 and Eq. (30), when the SCF and compressed symbols are determined, the problem can be formulated as

$$\begin{aligned}\hat{\mathbf{G}} &= \arg\max_{\mathbf{G}} \sum_{n=1}^{N}\sum_{l=1}^{L_n} \left\|[\mathbf{F}_n]_{:,l}^H \mathbf{H}_n\mathbf{GP}\right\|^2,\\ &\text{s.t. } \left\|[\mathbf{G}]_{:,k}\right\| = 1 \quad \forall k = 1, 2, \cdots, K,\end{aligned} \tag{32}$$

where recall that $\mathbf{G}$ is the diagonal matrix that represents the precoders of all UEs defined in (II-A). Let us introduce the following lemmas that help in solving this problem.

Lemma 4: The objective function in (32) can be expressed as

$$\sum_{n=1}^{N}\sum_{l=1}^{L_n} \left\|[\mathbf{F}_n]_{:,l}^H \mathbf{H}_n\mathbf{GP}\right\|^2 = P_k tr\left(\mathbf{G}_e^H\mathbf{r}\mathbf{r}^H\mathbf{G}_e\right), \tag{33}$$

where

$$\mathbf{G}_e = \begin{bmatrix}\mathbf{G}^H & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \ddots & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{G}^H\end{bmatrix}, \tag{34}$$

$$\mathbf{r}^H \triangleq \left[\,[\mathbf{F}_1]_{:,1}^H\mathbf{H}_1 \;\cdots\; [\mathbf{F}_N]_{:,L_N}^H\mathbf{H}_N\,\right], \tag{35}$$

and $\mathbf{r}$ is with dimension $\mathbb{C}^{(J\times K\times L)\times 1}$.

Proof: See Appendix E. ■

Let us further manipulate Eq. (33). Since $\mathbf{G}_e$ in Eq. (34) is the precoders of all UEs, it can be expressed as

$$\mathbf{G}_e = \left[\mathbf{u}_1\ \mathbf{u}_2\ \cdots\ \mathbf{u}_{K\times L}\right] \in \mathbb{C}^{(J\times K\times L)\times(K\times L)}, \tag{36}$$

where

$$\mathbf{u}_i = \left[\,[\mathbf{0}]_{1\times(J\times(i-1))}^T\ \mathbf{g}_k^T\ [\mathbf{0}]_{1\times(J\times(K(L-l-k)))}^T\,\right]^T, \tag{37}$$

for $i = K\times(l-1)+k$.

For example, let $K = 3$, $J = 2$, and $L = 3$. $\mathbf{u}_1 \sim \mathbf{u}_9$ are all with dimension $\mathbb{C}^{18\times 1}$ and should be

$$\begin{aligned}\mathbf{u}_1 &= \left[\,\mathbf{g}_1^T\ [\mathbf{0}^T]_{1\times 16}\,\right]^T,\\ \mathbf{u}_2 &= \left[\,[\mathbf{0}^T]_{1\times 2}\ \mathbf{g}_2^T\ [\mathbf{0}^T]_{1\times 14}\,\right]^T,\\ \mathbf{u}_3 &= \left[\,[\mathbf{0}^T]_{1\times 4}\ \mathbf{g}_3^T\ [\mathbf{0}^T]_{1\times 12}\,\right]^T,\\ \mathbf{u}_4 &= \left[\,[\mathbf{0}^T]_{1\times 6}\ \mathbf{g}_1^T\ [\mathbf{0}^T]_{1\times 10}\,\right]^T,\\ &\vdots\\ \mathbf{u}_9 &= \left[\,[\mathbf{0}^T]_{1\times 16}\ \mathbf{g}_3^T\,\right]^T.\end{aligned}$$

Given $\mathbf{W} \in \mathbb{C}^{a\times b}$, $\mathbf{T} \in \mathbb{C}^{a\times a}$, then

$$tr\left(\mathbf{W}^H\mathbf{TW}\right) = \sum_{i=1}^{b}[\mathbf{W}]_{:,i}^H\,\mathbf{T}\,[\mathbf{W}]_{:,i}. \tag{38}$$

We explain how to obtain Eq. (38) below. $tr(\cdot)$ is a process to sum all diagonal elements of a matrix. Dividing the matrices $\mathbf{W}$, $\mathbf{T}$ in Eq. (38) into column-wise terms, and substituting them into $tr\left(\mathbf{W}^H\mathbf{TW}\right)$ leads to the result of Eq. (38).

Using the result of Eq. (38) and substituting Eqs. (36)-(37) into Eq. (33), we have

$$P_k tr\left(\mathbf{G}_e^H\mathbf{r}\mathbf{r}^H\mathbf{G}_e\right) = P_k\sum_{i=1}^{KL}\mathbf{u}_i^H\mathbf{r}\mathbf{r}^H\mathbf{u}_i. \tag{39}$$

Since P_k is a constant, the problem in (32) is equivalent to

$$\begin{aligned}\hat{\mathbf{u}}_i &= \arg\max_{\mathbf{u}_i} \sum_{i=1}^{KL}\mathbf{u}_i^H\mathbf{r}\mathbf{r}^H\mathbf{u}_i,\\ &\text{s.t. } \|\mathbf{u}_i\| = 1 \quad \forall i = 1, 2, \cdots, KL,\\ &\text{The constriant in Eq. (37).}\end{aligned} \tag{40}$$

The objective function in (40) is a quadratic form of the precoders with a special constraint (37). There are two facts for the objective function of (40), introduced in the following two lemmas.

The objective function in (40) can be divided into different cost functions of different precoders of UEs given by

$$\sum_{i=1}^{KL}\mathbf{u}_i^H\mathbf{r}\mathbf{r}^H\mathbf{u}_i = \sum_{k=1}^{K}f(\mathbf{g}_k), \tag{41}$$

where the cost function for the precoder of UE k can be expressed as

$$f(\mathbf{g}_k) = \sum_{l=1}^{L}\mathbf{u}_{K\times(l-1)+k}^H\mathbf{r}\mathbf{r}^H\mathbf{u}_{K\times(l-1)+k}. \tag{42}$$

Eqs. (41)-(42) can be derived below. The objective function in (40) can be expanded as

$$\sum_{i=1}^{KL}\mathbf{u}_i^H\mathbf{r}\mathbf{r}^H\mathbf{u}_i = \mathbf{u}_1^H\mathbf{r}\mathbf{r}^H\mathbf{u}_1 + \mathbf{u}_2^H\mathbf{r}\mathbf{r}^H\mathbf{u}_2 \cdots + \mathbf{u}_{KL}^H\mathbf{r}\mathbf{r}^H\mathbf{u}_{KL}, \tag{43}$$

according to the definition of $\mathbf{u}_i$ in Eq. (37), we can directly rearrange Eq. (43) as Eq. (41).

The problem in (32) can be reformulated as

$$\hat{\mathbf{g}_k} = \arg\max_{\mathbf{g}_k} \sum_{k=1}^{K}f(\mathbf{g}_k),$$

$$\text{s.t. } \|\mathbf{g}_k\| = 1 \quad \forall k = 1, 2, \cdots, K, \tag{44}$$

where

$$f(\mathbf{g}_k) = \mathbf{g}_k^H \left(\sum_{l=1}^{L} \left[\mathbf{r}\mathbf{r}^H\right]_{K\times(l-1)+k} \right) \mathbf{g}_k, \tag{45}$$

and $\left[\mathbf{r}\mathbf{r}^H\right]_{K\times(l-1)+k} \in \mathbb{C}^{J\times J}$ is the $(K \times (l-1)+k)$th block diagonal matrix of $\mathbf{r}\mathbf{r}^H$. Eqs. (44)-(45) can be explained below. From Lemma 4 and Eqs. (38)-(42), the objective function in (32) can be rewritten as Eq. (41). Then Eq. (41) can be directly rearranged as Eq. (45), and Eqs. (41)-(42) shows that one can obtain the precoder $\mathbf{g}_k$ independently. Hence, we can obtain the result of Eqs. (44)-(45).

Let us give an example for $\left[\mathbf{r}\mathbf{r}^H\right]_{K\times(l-1)+k}$ in Eqs. (44)-(45). Let $K = 3$, $J = 2$, and $L = 3$. In this case, $f(\mathbf{g}_1) = \mathbf{g}_1^H \left(\left[\mathbf{r}\mathbf{r}^H\right]_1 + \left[\mathbf{r}\mathbf{r}^H\right]_4 + \left[\mathbf{r}\mathbf{r}^H\right]_7\right) \mathbf{g}_1$, and $\left[\mathbf{r}\mathbf{r}^H\right]_{K\times(l-1)+k}$ is defined as

$$\mathbf{r}\mathbf{r}^H = \begin{bmatrix} \left[\mathbf{r}\mathbf{r}^H\right]_1 & & & \\ & \left[\mathbf{r}\mathbf{r}^H\right]_2 & & \\ & & \ddots & \\ & & & \left[\mathbf{r}\mathbf{r}^H\right]_9 \end{bmatrix},$$

where $\left[\mathbf{r}\mathbf{r}^H\right]_i \in \mathbb{C}^{2\times 2}$ is block diagonal matrix of matrix $\mathbf{r}\mathbf{r}^H$.

An important result from these two lemmas is that one can design the precoders $\mathbf{g}_1, \cdots, \mathbf{g}_K$ independently. Since the term $\mathbf{r}\mathbf{r}^H$ in (40) can be regarded as a covariance matrix, which is positive semidefinite, the precoder designs for individual UEs are convex problems and can be solved independently by

$$(\hat{\mathbf{g}}_k, \hat{L}_n) = \arg\min_{\mathbf{g}_k, L_n} -\sum_{k=1}^{K} f(\mathbf{g}_k), \quad \text{s.t. } \|\mathbf{g}_k\| = 1 \quad \forall k = 1, 2, \cdots, K. \tag{46}$$

Proposition 2: The optimal precoders in (32) are obtained by letting the precoder $\mathbf{g}_k$ be the eigenvector corresponding to the largest eigenvalue of $\sum_{l=1}^{L} \left[\mathbf{r}\mathbf{r}^H\right]_{K\times(l-1)+k}$, where $\mathbf{r}$ is defined in (35).

Proof: See Appendix F. ■

According to Propositions 1 and 2, we conclude the proposed solutions for the SCF, compressed symbols to BBU and the precoders of all UEs in Algorithm 1.

Remark 2: It is worth pointing out that the precoders should be designed jointly for all UEs. However, from the derivation in Propositions 1 and 2, it turns out that each UE's precoder can be designed independently in the first iteration once the CSI of all UEs are available. Then, the SCFs in individual RRHs are designed jointly for all UEs after the UE precoders are determined.

Remark 3: We have proved that the UE precoding problem is convex in (46) when the SCF is fixed. On the hand, the SCF problem can be approximated to be an eigen problem when the UE precoder is fixed. The two steps of UE precoding and SCF designs iterate until the objective function converges. Later in the simulation results (Experiment 6), we will show that it usually converges quickly within 2-3 iterations.

Algorithm 1 Proposed SCFs and Precoders

Input: The number of UEs K; Average transmit power P_k; Channels of all RRHs $\mathbf{H}_1, \mathbf{H}_2, \cdots, \mathbf{H}_N$; The number of compressed symbols L; a predetermined threshold δ.

Output: The proposed precoders and the SCFs

1: **Initialization:** $\left[\mathbf{F}_{n=1,\cdots,N}\right]_{l=1,\cdots,L_n}^{(i)} = [\mathbf{1}]_{M\times 1}$, $i = 1$
2: **for** $i = 1 : \infty$ **do**
3: Substitute $\left[\mathbf{F}_{n=1,\cdots,N}\right]_{l=1,\cdots,L_n}^{(i+1)}$ into Eq. (33).
4: Use the result of Proposition 2 to solve $\mathbf{g}_1, \cdots, \mathbf{g}_K$, which forms $\mathbf{G}^{(i+1)}$.
5: Substitute $\mathbf{g}_1, \cdots, \mathbf{g}_K$ into the objective function in (30).
6: Obtain $\left[\mathbf{F}_{n=1,\cdots,N}\right]_{l=1,\cdots,L_n}^{(i+1)}$ using Proposition 1.
7: $\left[\mathbf{F}_{n=1,\cdots,N}\right]_{l=1,\cdots,L_n}^{(i+1)} = \left[\mathbf{F}_{n=1,\cdots,N}\right]_{l=1,\cdots,L_n}^{(i)}$
8: **if** $\sum_{k=1}^{K} f^{(i+1)}(\mathbf{g}_k) < \sum_{k=1}^{K} f^{(i)}(\mathbf{g}_k) + \delta$, **break;**
9: **else**, **go to step 2**
10: **end for**
11: $\mathbf{G} = \mathbf{G}^{(i+1)}$, $\mathbf{F} = \mathbf{F}^{(i+1)}$

V. SIMULATION RESULTS

In this section, we provide simulation results to show the accuracy of the analytical results and the performance of the proposed SCF and precoders. In these experiments, we consider three UEs, three RRHs, and one BBU. Equal transmit power and 4-QAM symbols are used for each UE. The channel is with i.i.d. Rayleigh fading unless it is specifically mentioned, and we run more than 10^4 realizations.

Experiment 1 Theoretical vs. Simulated Results: In this experiment, we compare the simulated and the analyzed BLER performances derived in Lemma 1. Several different parameters settings are used as follows: The number J of transmit antennas for each UE is 1. The number Q of quantization bits for compressed symbols is 6. The total number L of compressed symbols is 3. The numbers M of receive antennas for each RRH are 4, 6 and 8. The corresponding results are shown in Fig. 3. The horizontal axis is the transmitted SNR and the vertical axis is the BLER. From the figure, the analytical results are accurate. Hence the proposed optimal SCF and precoder solutions based on this analytical result are also accurate.

Experiment 2 Performance Comparison Between Proposed and Conventional Schemes: In this experiment, we show the performance improvement of the proposed SCF and precoder. Conventional SISO scheme in [20], MRT (maximum ratio transmission) SCF and random SCF (without design) are used for comparison. The reason why we chose MRT SCF and random SCF for comparison is addressed as follows. In many massive MIMO systems, where the interference decreases as the number of antennas increases, the MRT is

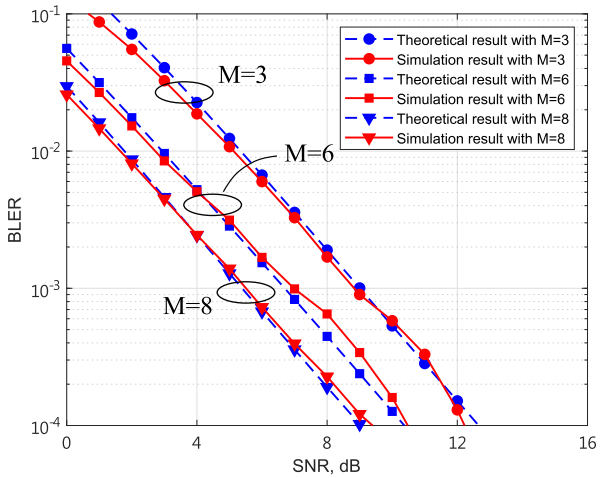

FIGURE 3. Theoretical and simulated results for the system with $J = 1$, $L = 3$, $Q = 6$, and various values of $M = 4, 6$, and 8.

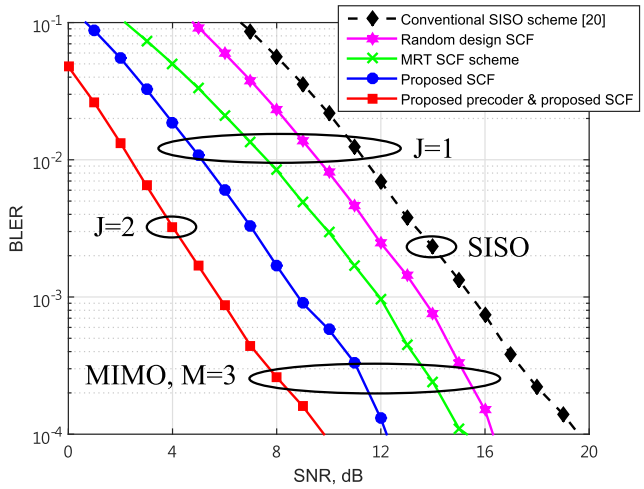

FIGURE 4. Performance comparison between the conventional schemes and proposed SCF and precoding schemes for $M = 3$ and $L = 3$.

applied because it maximizes the SNR in interference-free environments. On the other hand, the random SCF can be regarded as a system without sophisticated designs. By showing the random SCF, one can more clearly see the performance improvement contributed by the proposed scheme. The parameters are $L = 3$, which means that for all schemes, the RRHs transmit totally three symbols to the BBU no matter SISO or MIMO systems. $M = 3$ for the proposed system and the random SCF. $J = 1$ means that there is no precoding at the UE side, while $J = 2$ means that the proposed precoder is used at the UE side with two transmit antennas. The performance comparisons are shown in Fig. 4. Observe that the proposed SCF without precoding outperforms the conventional SISO scheme by around 7.5 dB (see the circled and diamond curves), outperforms the MRT SCF scheme (see the cross curve) by around 3 dB, and outperforms the random SCF scheme (see the star curve) by around 4 dB. Moreover, when the proposed precoder is applied (see the rectangular curve), the performance can be further improved by around 3 dB. These show that the performance can be improved significantly the proposed SCF and precoder.

Experiment 3 The Effect of Quantization Bits: Quantization bits for the compressed symbols directly relate to the fronthaul loading between the RRHs and the BBU. Let $J = 1$, $M = 3$ and $L = 3$ (following Experiment 2). Fig. 5. shows the performance of the proposed scheme with various quantization bits $Q = 1, 2, 3, 5, 6$, and 12. From the figure, severe performance degradation occurred with low quantization bits such as $Q = 1, 2$, and 3. It also shows that $Q = 6$ is sufficient to sustain satisfactory system performance. Observe that although increasing the number of quantization bits improves the performance, however, it saturates, *e.g.,* for $Q = 6$, in this example. This means that one does not need to use many bits to quantize the compressed symbols after the SCF. In this example, the total number of bits for all RRHs to the BBU is $L \times Q = 18$ bits to achieve satisfactory performance.

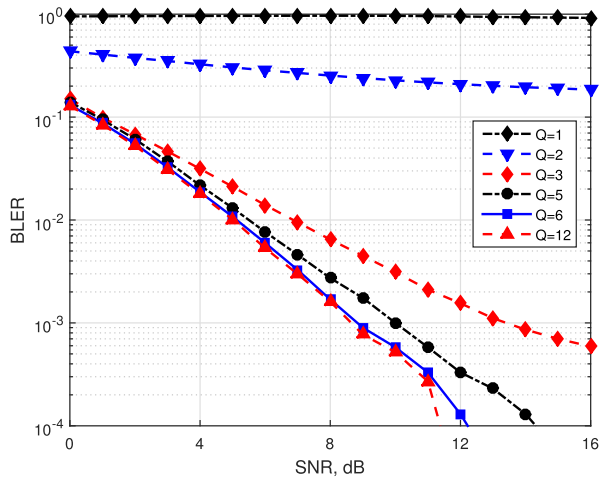

FIGURE 5. The effect of quantization bits for $Q = 1, 2, 3, 5, 6$, and 12 with $M = 3$.

Experiment 4 SCF With Various Compression Ratios: In this experiment, we would like to evaluate the performance of different compression ratios *i.e*, $L/(N \times M)$, used by the SCF. Let $J = 1$ and $M = 3$ (following Experiment 2), and now let the numbers L of compressed symbols after the SCF be 3, 4, 6, 8 and 9. Note that 9 means no compression, because each of the 3 RRHs has received $M = 3$ symbols, and total symbols are 9 without compression. The results are shown in Fig. 6. We see that using the SCF and control the compression ratio can well leverage between performance and fronthaul loading. On one hand, if the fronthaul loading is limited, using a small number of compressed symbols, *e.g.,* $L = 3$, can outperform conventional schemes as shown in Experiment 2. On the other hand, if the fronthaul loading issue is not that critical, using a moderate number of compressed symbols, *e.g.,* $L = 6$, can achieve close performance as that without compression, observed from Fig. 6.

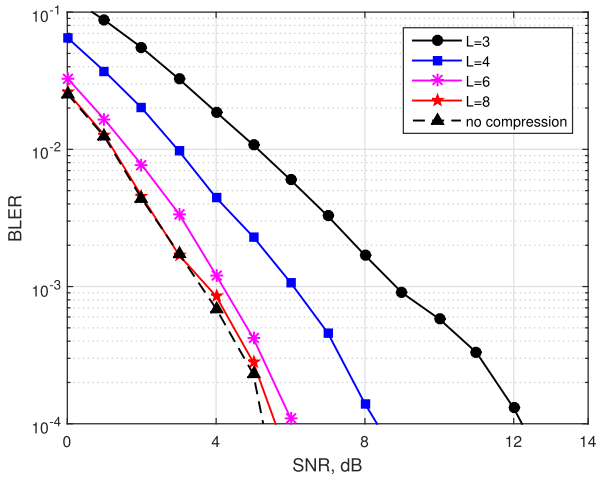

FIGURE 6. Performance of SCF with various compression ratios for $L = 3, 4, 6, 8$ and without compression; $J = 1$ and $M = 3$.

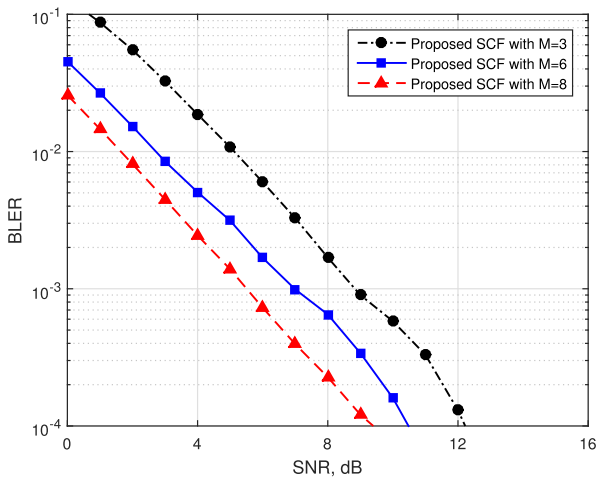

FIGURE 7. Performance of SCF with various numbers of antennas at RRH for $M = 3, 6$, and 8; $J = 1$ and $L = 3$.

Experiment 5 SCF With Various Numbers of Antennas at RRH: Let $J = 1$ and $L = 3$ (following Experiment 2). Now let the numbers of antennas at each RRH be $M = 3, 6$, and 8. Fig. 7 shows the performance. Observed from the figure that increasing the number of antennas at RRHs improves the performance due to array gain. It is worth pointing out in this experiment, the number L of compressed symbols is fixed to be 3. Hence the fronthaul loading from the RRHs to the BBU is fixed, even if the number of antennas at RRHs increases, thanks to the proposed MIMO SCF scheme.

Experiment 6 Converging Speed of Proposed Scheme: When $J > 1$ and the proposed precoder is to be used together with the proposed SCF, iteration is needed according to Algorithm 1. In this example, we evaluate how many iterations are needed to converge. Let $J = 2$, $M = 3$ and $L = 3$ (following Experiment 2). Fig. 8 shows 20 of the 10^4 channel realizations in terms of converging speed for implementing Algorithm 1.

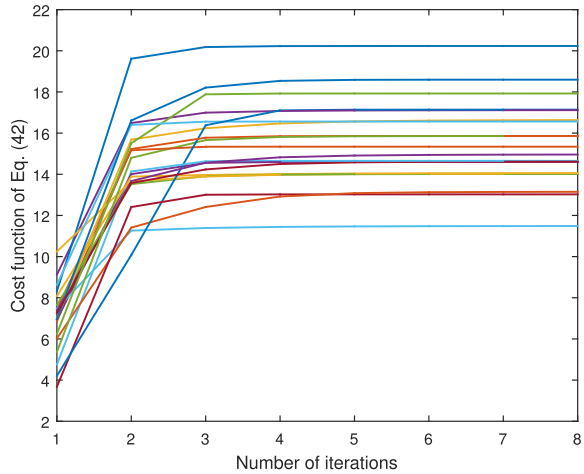

FIGURE 8. Converging speed with 20 channel realizations; $J = 2$, $L = 3$, and $M = 3$.

Observe that most of them converge in 2-4 iterations. This shows the fast converging speed of the proposed algorithm.

Experiment 7 Performance Comparison of Proposed and Conventional Precoders in Cellular Networks: In these experiments, we would like to show the performance improvement contributed from the proposed precoder in Proposition 2. To be fair, the proposed SCF is applied to all the precoders used in this experiment. The conventional precoder used here, which were intuitively regarded as optimal solution from Eq. (1) but it was not, is the right singular vector corresponding to the largest singular value of matrix $\mathbf{H}_k = \left[\mathbf{H}_{1,k}^T\ \mathbf{H}_{2,k}^T \cdots \mathbf{H}_{N,k}^T\right]^T$, where $\mathbf{H}_k$ is the channel matrix from the kth UE to all RRHs. We consider i.i.d. large scale fading for a practical cellular environment [30]. More specifically, the formula is expressed as:

$$\overline{PL}(\mathrm{d})\,[\mathrm{dB}] = -10\log\left[\frac{G_t G_r \lambda^2}{(4\pi)^2 d^2}\right] + 10\mathrm{n}\log\left(\frac{d}{d_0}\right),$$

where G_t and G_r are the gain of transmitting and receiving antenna, respectively. d_0 and d respectively represent the reference and the transmitter-receiver separation distances. We set G_t, G_r, and d_0 equal to one, and consider carrier frequency is 3.5 GHz and free space environment, *i.e.*, $n = 2$. d is set as Fig. 9, and users (UEs) are randomly distributed among remote radio heads (RRHs) within 60 m. In this case, the channel variances between the UEs and the RRHs are not equal.

Let the numbers J of transmit antennas at UE side be 2 and 3, and the numbers L of compressed symbols be 3. Figs. 10 and 11 show the results for the numbers M of receive antennas at each RRH being 3 and 8, respectively. In Figs. 10-11, we demonstrate the performance comparison of proposed precoder and the conventional precoder respectively for $M = 3, 6$, and 8 in $J = 2, 3$, and $L = 3$. Observe that the proposed precoder outperforms the conventional precoder. It is worth pointing out that the proposed precoder with $J = 2$ can even outperform the conventional precoder with $J = 3$

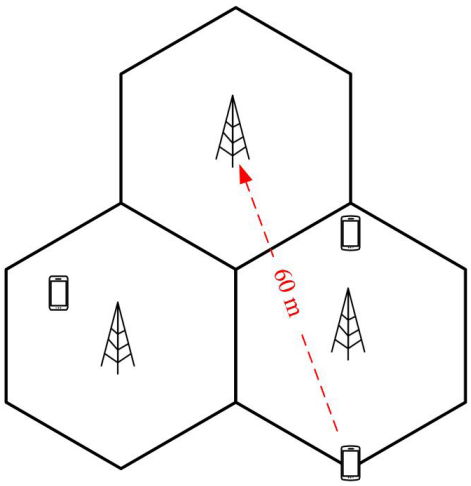

FIGURE 9. A cellular network structure.

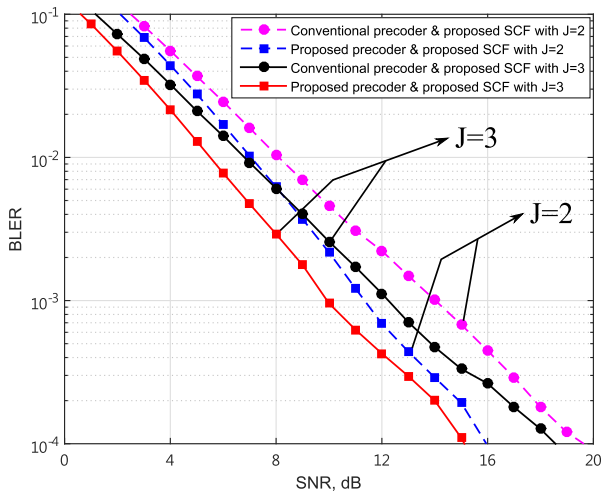

FIGURE 10. Performance comparison of the proposed precoders and the conventional precoders for $J = 2$ and $J = 3$ with $L = 3$ and $M = 3$.

in interested SNR and BLER regions (see solid-circle and dash-rectangular curves). In addition, the performance gap increases as SNR increase. Moreover, because the proposed precoder is the optimal solution, increasing the number of transmit antennas always improves the performance. On the other hand, there is no guarantee for the conventional precoder. For example, from Fig. 11, although J is increased from 2 to 3, using more transmit antennas does not necessarily lead to a better performance in the conventional scheme.

VI. CONCLUSION

We have proposed an integrated uplink C-RAN architecture, which consists of MIMO precoders, spatial compression filters, and joint decompression and detection. Optimal solutions have been derived for the proposed scheme according to the analytical results. The integrated techniques have significantly improved the system performance and enjoyed several nice properties. More specifically, a good trade-off between

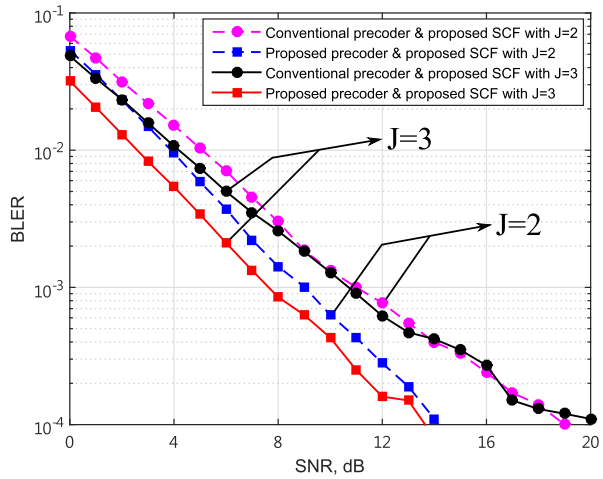

FIGURE 11. Performance comparison of the proposed precoders and the conventional precoders for $J = 2$ and $J = 3$ with $L = 3$ and $M = 8$.

fronthaul loading and system performance can be achieved by well controlling the compression ratio of the proposed spatial compression filters. In addition, the optimal precoders have been shown to be able to design separately at the UE sides despite that the join decompression and detection have been applied. As a result, the computational complexity is low, while the corresponding performance improvement is significant. Simulation results have shown to corroborate the analytical results, and demonstrate the good properties of the proposed scheme.

APPENDIX A
PROOF OF LEMMA 1

Proof: Substituting Eq. (18) into Eq. (16) yields

$$\mathrm{P}\{\mathbf{c} \to \tilde{\mathbf{c}}\} \triangleq \mathrm{P}\{\mathbb{D}(\mathbf{c}) - \mathbb{D}(\tilde{\mathbf{c}}) > 0\}, \tag{47}$$

where $\mathbb{D}(\mathbf{c}) = \sum_{n=1}^{N}\sum_{l=1}^{L_n} \left|[\tilde{\mathbf{y}}_n]_l - [\mathbf{F}_n^H \mathbf{H}_n \mathbf{G}\mathbf{P}\mathbf{c}]_l\right|^2$. For tidy expression, we define $\mathbb{I}(\mathbf{c}, \tilde{\mathbf{c}}) = \mathbb{D}(\mathbf{c}) - \mathbb{D}(\tilde{\mathbf{c}})$, *i.e.,*

$$\mathbb{I}(\mathbf{c}, \tilde{\mathbf{c}}) = \sum_{n=1}^{N}\sum_{l=1}^{L_n} \left[f\left([\tilde{\mathbf{y}}_n]_l, \mathbf{c}\right) - f\left([\tilde{\mathbf{y}}_n]_l, \tilde{\mathbf{c}}\right)\right], \tag{48}$$

where $f\left([\tilde{\mathbf{y}}_n]_l, \mathbf{c}\right) = \left|[\tilde{\mathbf{y}}_n]_l - [\mathbf{F}_n^H \mathbf{H}_n \mathbf{G}\mathbf{P}\mathbf{c}]_l\right|^2$ and $f\left([\tilde{\mathbf{y}}_n]_l, \tilde{\mathbf{c}}\right) = \left|[\tilde{\mathbf{y}}_n]_l - [\mathbf{F}_n^H \mathbf{H}_n \mathbf{G}\mathbf{P}\tilde{\mathbf{c}}]_l\right|^2$. Substituting Eq. (14) into Eq. (48), we have

$$\mathbb{I}(\mathbf{c}, \tilde{\mathbf{c}}) = \gamma_1 + \gamma_2 - \varepsilon, \tag{49}$$

where $\mathbf{e} \triangleq (\tilde{\mathbf{c}} - \mathbf{c})$,

$$\gamma_1 = \sum_{n=1}^{N}\sum_{l=1}^{L_n} 2\Re\{\mathbf{z}_n^H [\mathbf{F}_n]_{:,l} [\mathbf{F}_n]_{:,l}^H \mathbf{H}_n \mathbf{G}\mathbf{P}\mathbf{e}\},$$

$$\gamma_2 = \sum_{n=1}^{N}\sum_{l=1}^{L_n} 2\Re\{\mathbf{q}_n^H [\mathbf{F}_n]_{:,l}^H \mathbf{H}_n \mathbf{G}\mathbf{P}\mathbf{e}\},$$

$$\varepsilon = \sum_{n=1}^{N}\sum_{l=1}^{L_n}\left|[\mathbf{F}_n]_{:,l}^H \mathbf{H}_n\mathbf{G}\mathbf{P}\mathbf{e}\right|^2.$$

Since $\mathbf{z}_n$ is a complex Gaussian random variables with zero mean and covariance matrix $\sigma^2\mathbf{I}$, the term γ_1 is also a Gaussian random variable with zero mean and variance given by

$$\sigma_{\gamma_1}^2 = 2\sigma^2\sum_{n=1}^{N}\sum_{l=1}^{L_n}\left\|[\mathbf{F}_n]_{:,l}[\mathbf{F}_n]_{:,l}^H\mathbf{H}_n\mathbf{P}\mathbf{e}\right\|^2. \tag{50}$$

We approximate γ_2 as a Gaussian random variable according to the Central Limit Theory, which has zero mean and the variance can be computed using Eq. (15) as

$$\sigma_{\gamma_2}^2 = \frac{4}{3}\sum_{n=1}^{N}\sum_{l=1}^{L_n}\delta_{n,l}^2\left|[\mathbf{F}_n]_{:,l}^H\mathbf{H}_n\mathbf{P}\mathbf{e}\right|^2, \tag{51}$$

Then $\gamma_1+\gamma_2$ can also be approximated by a Gaussian random variable with zero mean and the variance given by

$$\mathrm{Var}\left(\gamma_1+\gamma_2\right) = \sigma_{\gamma_1}^2+\sigma_{\gamma_2}^2. \tag{52}$$

From Eq. (49), the term $\gamma_1+\gamma_2$ is a Gaussian random variable and ε is a deterministic scalar as long as $(\mathbf{c},\tilde{\mathbf{c}})$ is determined. Hence, one can resort to the complementary error function to derive $\mathrm{P}\{\mathbf{c}\to\tilde{\mathbf{c}}\}$ [29]. ■

APPENDIX B
PROOF OF LEMMA 2

Proof: With high SNR assumption, Eq. (5) can be approximated as

$$\eta_{n,l} \approx 3\sqrt{\left\|[\mathbf{F}_n]_{:,l}^H\mathbf{H}_n\mathbf{G}\mathbf{P}\right\|^2}. \tag{53}$$

Recall that $\delta_{n,l} = \eta_{n,l}2^{-Q_n}$ in Section II-B, we have

$$\begin{aligned}\delta_{n,l}^2 &= 9\times 2^{-2Q_n}\left\|[\mathbf{F}_n]_{:,l}^H\mathbf{H}_n\mathbf{G}\mathbf{P}\right\|^2\\ &= 9\times 2^{-2Q_n}\sum_{k=1}^{K}\underbrace{\left|\left[[\mathbf{F}_n]_{:,l}^H\mathbf{H}_n\mathbf{G}\mathbf{P}\right]_k\right|^2}_{S_k}.\end{aligned} \tag{54}$$

Since the element of $\mathbf{H}_n$ is with i.i.d. complex Gaussian distribution, S_k is an exponential random variable [29], and $\delta_{n,l}^2$ is a summation of the independent exponential random variables, which has the gamma distribution. By the Central Limit Theorem, when the number K of UEs approaches to ∞, the Gamma distribution $\Gamma(K, 9\times P_k\times 2^{-2Q_n})$ can be approximated to the Gaussian distribution with mean and variance as shown in this lemma. ■

APPENDIX C
PROOF OF LEMMA 3

Proof: From (26), using the constraint $\|[\mathbf{F}_n]_{:,l}\| = 1$, the first term in the denominator can be rewritten as

$$\begin{aligned}&\left\|[\mathbf{F}_n]_{:,l}[\mathbf{F}_n]_{:,l}^H\mathbf{H}_n\mathbf{P}\mathbf{e}\right\|^2\\ &= \mathbf{e}^H\mathbf{P}^H\mathbf{H}_n^H[\mathbf{F}_n]_{:,l}[\mathbf{F}_n]_{:,l}^H[\mathbf{F}_n]_{:,l}[\mathbf{F}_n]_{:,l}^H\mathbf{H}_n\mathbf{P}\mathbf{e}\\ &= \left|[\mathbf{F}_n]_{:,l}^H\mathbf{H}_n\mathbf{P}\mathbf{e}\right|^2.\end{aligned} \tag{55}$$

Using the result of Lemma 2, the term $\delta_{n,l}^2$ can be approximated as a constant. Together with Eq. (55), we can approximate the objective function in (26) as

$$\begin{aligned}&\mathbb{E}_{\mathbf{e}}\left\{\frac{f_1\left([\mathbf{F}_n]_{:,l},\mathbf{H}_n\right)}{\sqrt{4\sigma^2 f_2\left([\mathbf{F}_n]_{:,l},\mathbf{H}_n\right)+\frac{8}{3}f_3\left([\mathbf{F}_n]_{:,l},\mathbf{H}_n,\delta_{n,l}^2\right)}}\right\}\\ &\approx \mathbb{E}_{\mathbf{e}}\left\{\frac{f_1\left([\mathbf{F}_n]_{:,l},\mathbf{H}_n\right)}{\sqrt{\left(4\sigma^2+\frac{8}{3}\delta_{n,l}^2\right)f_1\left([\mathbf{F}_n]_{:,l},\mathbf{H}_n\right)}}\right\}\\ &= \mathbb{E}_{\mathbf{e}}\left\{\sqrt{\frac{f_1\left([\mathbf{F}_n]_{:,l},\mathbf{H}_n\right)}{4\sigma^2+\frac{8}{3}\delta_{n,l}^2}}\right\}.\end{aligned} \tag{56}$$

In (56), $\left(4\sigma^2+\frac{8}{3}\delta_{n,l}^2\right)$ is a constant. The objective function in (26) is proportional to term $\mathbb{E}_{\mathbf{e}}\left\{\sqrt{f_1\left([\mathbf{F}_n]_{:,l},\mathbf{H}_n\right)}\right\}$. Since the term $\sqrt{f_1\left([\mathbf{F}_n]_{:,l},\mathbf{H}_n\right)}$ is a convex function, by Jensen's inequality, we have

$$\mathbb{E}_{\mathbf{e}}\left\{\sqrt{f_1\left([\mathbf{F}_n]_{:,l},\mathbf{H}_n\right)}\right\} \geq \sqrt{\mathbb{E}_{\mathbf{e}}\left\{f_1\left([\mathbf{F}_n]_{:,l},\mathbf{H}_n\right)\right\}}. \tag{57}$$

The objective function of maximizing the left side term of Eq. (57) becomes maximizing the right side term of Eq. (57). From Eq. (21), the right side term of Eq. (57) reduces to

$$\begin{aligned}&\mathbb{E}_{\mathbf{e}}\left\{\sum_{n=1}^{N}\sum_{l=1}^{L_n}|[\mathbf{F}_n]_{:,l}^H\mathbf{H}_n\mathbf{P}\mathbf{e}|^2\right\}\\ &= \sum_{n=1}^{N}\sum_{l=1}^{L_n}\mathbb{E}_{\mathbf{e}}\left\{|[\mathbf{F}_n]_{:,l}^H\mathbf{H}_n\mathbf{P}\mathbf{e}|^2\right\}\\ &= \sum_{n=1}^{N}\sum_{l=1}^{L_n}[\mathbf{F}_n]_{:,l}^H\mathbf{H}_n\mathbf{P}\mathbb{E}_{\mathbf{e}}\left\{\mathbf{e}\mathbf{e}^H\right\}\mathbf{P}^H\mathbf{H}_n^H[\mathbf{F}_n]_{:,l}\\ &\overset{(\star)}{=} \sum_{n=1}^{N}\sum_{l=1}^{L_n}[\mathbf{F}_n]_{:,l}^H\mathbf{H}_n\mathbf{P}\mathbf{P}^H\mathbf{H}_n^H[\mathbf{F}_n]_{:,l}\\ &= \sum_{n=1}^{N}\sum_{l=1}^{L_n}\left\|[\mathbf{F}_n]_{:,l}^H\mathbf{H}_n\mathbf{P}\right\|^2,\end{aligned} \tag{58}$$

where $(\star)$ is due to the fact that the elements of $\mathbf{e}$ are mutually independent, and the covariance matrix $\mathbb{E}_{\mathbf{e}}\left\{\mathbf{e}\mathbf{e}^H\right\}$ is an identity matrix, which leads to the result of Lemma 3. ■

APPENDIX D
PROOF OF PROPOSITION 1

Proof: Using the result of Lemma 3, the problem is a quadratic optimization problem for each RRH. Optimizing the summation of the quadratic optimization problem is equivalent to maximizing individual quadratic problems and find the L left singular vectors corresponding to the L largest singular values of the N matrices, $\mathbf{H}_1\mathbf{P}, \cdots, \mathbf{H}_N\mathbf{P}$. The optimal value is the summation of the L largest singular values. ■

APPENDIX E
PROOF OF LEMMA 4

Proof: The objective function in (32) can be expanded as

$$\begin{aligned}
&\sum_{n=1}^{N}\sum_{l=1}^{L_n}\left\|[\mathbf{F}_n]_{:,l}^H\,\mathbf{H}_n\mathbf{G}\mathbf{P}\right\|^2\\
&=\sum_{n=1}^{N}\sum_{l=1}^{L_n}[\mathbf{F}_n]_{:,l}^H\,\mathbf{H}_n\mathbf{G}\mathbf{P}\mathbf{P}^H\mathbf{G}^H\mathbf{H}_n^H\,[\mathbf{F}_n]_{:,l}\\
&\overset{(\star)}{=}P_k\sum_{n=1}^{N}\sum_{l=1}^{L_n}[\mathbf{F}_n]_{:,l}^H\,\mathbf{H}_n\mathbf{G}\mathbf{G}^H\mathbf{H}_n^H\,[\mathbf{F}_n]_{:,l}\,,
\end{aligned}\tag{59}$$

where $(\star)$ is because the transmit power of individual UEs are assumed to be equal, *i.e.*, $\mathbf{P}\mathbf{P}^H = P_k\mathbf{I}$, where P_k is a constant. Defining $\mathbf{r}$ as in Eq. (35), substituting it into Eq. (59), and using the property that $tr\,(\mathbf{AB}) = tr\,(\mathbf{BA})$, the result of this lemma can be obtained. ■

APPENDIX F
PROOF OF PROPOSITION 2

Proof: Eqs. (41)-(45) show that we can obtain the precoder independently. The result of Eq. (44) is a quadratic optimization problem. By (46), the precoder $\mathbf{g}_k$ is the eigenvector corresponding to the largest eigenvalue of $\sum_{l=1}^{L}\left[\mathbf{r}\mathbf{r}^H\right]_{K\times(l-1)+k}$. ■

REFERENCES

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

[2] H. Ren, N. Liu, C. Pan, M. Elkashlan, A. Nallanathan, X. You, and L. Hanzo, "Low-latency C-RAN: An next-generation wireless approach," *IEEE Veh. Technol. Mag.*, vol. 13, no. 2, pp. 48–56, Jun. 2018.

[3] J. Tang, R. Wen, T. Q. S. Quek, and M. Peng, "Fully exploiting cloud computing to achieve a green and flexible C-RAN," *IEEE Commun. Mag.*, vol. 55, no. 11, pp. 40–46, Nov. 2017.

[4] J. Wu, Z. Zhang, Y. Hong, and Y. Wen, "Cloud radio access network (C-RAN): A primer," *IEEE Netw.*, vol. 29, no. 1, pp. 35–41, Jan. 2015.

[5] Y. Cai, F. R. Yu, and S. Bu, "Cloud computing meets mobile wireless communications in next generation cellular networks," *IEEE Netw.*, vol. 28, no. 6, pp. 54–59, Nov. 2014.

[6] 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. 4955–4969, Aug. 2018.

[7] B.-S. Huang, Y.-H. Chiang, and W. Liao, "Remote radio head (RRH) deployment in flexible C-RAN under limited fronthaul capacity," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.

[8] Y. Zhou, J. Li, Y. Shi, and V. W. S. Wong, "Flexible functional split design for downlink C-RAN with capacity-constrained fronthaul," *IEEE Trans. Veh. Technol.*, vol. 68, no. 6, pp. 6050–6063, Jun. 2019.

[9] S.-H. Park, O. Simeone, and S. Shamai, "Uplink sum-rate analysis of C-RAN with interconnected radio units," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Nov. 2017, pp. 171–175.

[10] S.-H. Park, O. Simeone, O. Sahin, and S. Shamai (Shitz), "Inter-cluster design of precoding and fronthaul compression for cloud radio access networks," *IEEE Wireless Commun. Lett.*, vol. 3, no. 4, pp. 369–372, Apr. 2014.

[11] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[12] 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.

[13] J. Kang, O. Simeone, J. Kang, and S. S. 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.

[14] J. Zhang, C.-K. Wen, S. Jin, X. Gao, and K.-K. Wong, "Large system analysis of cooperative multi-cell downlink transmission via regularized channel inversion with imperfect CSIT," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 4801–4813, Oct. 2013.

[15] S. Nanba and A. Agata, "A new IQ data compression scheme for front-haul link in centralized RAN," in *Proc. IEEE 24th Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC Workshops)*, Sep. 2013, pp. 210–214.

[16] A. Vosoughi, M. Wu, and J. R. Cavallaro, "Baseband signal compression in wireless base stations," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2012, pp. 4505–4511.

[17] S. Ogg and B. Al-Hashimi, "Improved data compression for serial interconnected network on chip through unused significant bit removal," in *Proc. 19th Int. Conf. VLSI Design Held Jointly With 5th Int. Conf. Embedded Syst. Design (VLSID)*, Jan. 2006, p. 5.

[18] G. Langdon and J. Rissanen, "Compression of black-white images with arithmetic coding," *IEEE Trans. Commun.*, vol. COM-29, no. 6, pp. 858–867, Jun. 1981.

[19] P. Elias, "Universal codeword sets and representations of the integers," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 2, pp. 194–203, Mar. 1975.

[20] T. X. Vu, H. D. Nguyen, and T. Q. S. Quek, "Adaptive compression and joint detection for fronthaul uplinks in cloud radio access networks," *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 4565–4575, Nov. 2015.

[21] D. Samardzija, J. Pastalan, M. MacDonald, S. Walker, and R. Valenzuela, "Compressed transport of baseband signals in radio access networks," *IEEE Trans. Wireless Commun.* vol. 11, no. 9, pp. 3216–3225, Sep. 2012.

[22] K. F. Nieman and B. L. Evans, "Time-domain compression of complex-baseband LTE signals for cloud radio access networks," in *Proc. IEEE Global Conf. Signal Inf. Process.*, Dec. 2013, pp. 1198–1201.

[23] Y. Ren, Y. Wang, G. Xu, and Q. Huang, "A compression method for LTE—A signals transported in radio access networks," in *Proc. 21st Int. Conf. Telecommun. (ICT)*, May 2014, pp. 293–297.

[24] L. Liu and R. Zhang, "Optimized uplink transmission in multi-antenna C-RAN with spatial compression and forward," *IEEE Trans. Signal Process.*, vol. 63, no. 19, pp. 5083–5095, Oct. 2015.

[25] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Trans. Inf. Theory*, vol. IT-44, no. 6, pp. 2325–2383, Oct. 1998.

[26] S. K. Mitra, *Digital Signal Processing*, 4th ed. New York, NY, USA: McGraw-Hill, Sep. 2010.

[27] H. L. V. Trees, *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Filtering Theory*, 2nd ed. Hoboken, NJ, USA: Wiley, Apr. 2013.

[28] S. Haykin, *Communication Systems*, 4th ed. Hoboken, NJ, USA: Wiley, Jun. 2000.

[29] H. Stark and J. W. Woods, *Probability, Statistics, and Random Processes for Engineers*, 4th ed. London, U.K.: Pearson, Aug. 2011.
[30] D. C. Hogg, "Fun with the Friis free-space transmission formula," *IEEE Antennas Propag. Mag.*, vol. AP-35, no. 4, pp. 33–35, Aug. 1993.

GUAN-YU CHEN was born in Yilan, Taiwan. He received the B.S. degree in electrical engineering from Yuan Ze University (YZU), in 2017, and the M.S. degree in electrical and control engineering from the National Chiao Tung University (NCTU), Taiwan, in 2019. He is currently with Qualcomm Inc. His research interests include wireless communications and statistical signal processing.

SHANG-HO (LAWRENCE) TSAI (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the University of Southern California, CA, USA, in 2005. From 1999 to 2002, he was with Silicon Integrated Systems Corporation, where he was involved in the VLSI design for DMT-ADSL systems. From 2005 to 2007, he was with MediaTek Inc., where he was involved in the VLSI design for MIMO-OFDM systems and standard specifications for IEEE 802.11n. He was a Visiting Fellow with the Department of Electrical Engineering, Princeton University, in 2013. In 2007, he joined the Department of Electrical Engineering, National Chiao Tung University, where he is currently a Professor. His research interests include the areas of signal processing for communications, statistical signal processing, and signal processing for VLSI designs. He was a recipient of the Government Scholarship for Overseas Study from the Ministry of Education, Taiwan, for the period 2002–2005. He has been on the Editorial Board of the IEEE Signal Processing Repository, since 2018.

XIANG-QUAN SER received the B.S. and M.S. degrees in electrical and control engineering from the National Chiao Tung University (NCTU), Taiwan, in 2018 and 2021, respectively. His research interest includes wireless communications and optimization.