IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 26, 2020, accepted October 8, 2020, date of publication October 12, 2020, date of current version October 22, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3030132

Rateless Coded Multi-User Uplink Transmission With Distributed Fronthaul Compression in Cloud RAN

YU ZHANG[1,2], (Member, IEEE), ZHEHAO FAN[1], AND LIMIN MENG[1]
[1]College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China
[2]National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China

Corresponding authors: Limin Meng (mlm@zjut.edu.cn) and Yu Zhang (yzhang@zjut.edu.cn)

This work was supported in part by the Open Research Fund of National Mobile Communications Research Laboratory, Southeast University under Grant 2020D10, and in part by the National Natural Science Foundation of China under Grant 61871348, Grant 61871349, and Grant 61803336.

ABSTRACT In this paper, we design the rateless coded uplink transmission in cloud radio access network (C-RAN) with two users, two remote radio heads (RRHs), and one baseband processing unit (BBU) pool under block fading channel. Each user sends the Raptor coded signals continuously until receiving ACK from the BBU pool. A distributed fronthaul compression scheme with LDPC code is proposed for the RRHs which can reduce the compression loss without increasing the fronthaul traffic, through leveraging the correlation between the received signals at both RRHs. To guarantee the compression performance, the LDPC code profile is optimized based on extrinsic information transfer (EXIT) analysis. Furthermore, the Raptor code profiles applied at the two users are also jointly optimized to improve the average system throughput. Simulation results show that the proposed transmission scheme with the optimized LDPC compression code and Raptor code can achieve good BER and throughput performance.

INDEX TERMS Cloud radio access network, distributed fronthaul compression, LDPC code, Raptor code, degree profile optimization.

I. INTRODUCTION

In the future communication networks, massive users and traffic demands under limited available communication resource impose a stringent requirement on the spectrum efficiency and energy efficiency. C-RAN is a promising RAN technology considering the above challenges. In C-RAN, all the BBUs are backward centralized into a cloud computing resource pool and each RRH is located closer to the users. BBU pool and RRHs are connected through high-speed fronthaul links. The main advantage of C-RAN lies in that the coordinated multiple points transmission (CoMP) can be realized inherently which can greatly improve the system spectrum efficiency [1], and the Capital Expenditure (CAPEX) and Operating Expense (OPEX) are reduced compared with the conventional RAN [2].

In C-RAN, the RRHs and BBU pool communicate through the fronthaul with restricted capacity. Therefore, the signals conveyed through the fronthaul should be compressed first to meet the capacity limit. The conventional compression scheme is scalar quantization [3]. Due to the fact that the received signals at neighbor RRHs have correlations, more efficient compression can be deployed at the RRH according to the theory of distributed source coding (DSC, i.e., Slepian-wolf coding and Wyner-ziv coding) [4], [5], which is named 'distributed fronthaul compression' in the literature. In practice, DSC can be realized by channel codes such as turbo code [6]–[8], LDPC code [9]–[12], trellis code [13], and Raptor code [14], etc. There are a plethora of works, e.g., [15]–[21], which investigated the distributed fronthaul compression based on DSC for the uplink C-RAN. In [15]–[17], the authors provided theoretical optimization frameworks for distributed fronthaul compression. In [18], it was proved that, distributed fronthaul compression can reach higher sum rate even compared with the case where the RRHs can communicate with each other. In [19], [20], the authors optimized the fronthaul compression noise covariance and investigated the decompression process at BBU pool for the uplink C-RAN. In [21], the authors investigated the joint optimization of multi-antenna precoding for users and fronthaul compression noise covariance to maximize the achievable sum rate in the MIMO C-RAN. Nevertheless, all

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

the above works are mainly from the network information theoretic aspect, where the design of distributed fronthaul compression was mainly focused on the compression noise covariance optimization. To the best of authors' knowledge, the practical code design for the distributed fronthaul compression based on DSC in C-RAN has not been addressed yet.

On the other hand, rateless code is a class of channel codes (e.g, LT code and Raptor code [22]) which is quite different from the conventional fixed-rate channel code. With rateless code, there is no need for the transmitter to determine the coding rate before transmission. It generates codewords of infinite length and continuously sends the coded bits until the receiver successfully decodes and feedbacks an ACK. Rateless code can approach the channel capacity when its degree profile is optimized, solely with the statistical channel state information (CSI) [22], [23]. Therefore, system overhead for CSI acquisition at the transmitter can be greatly reduced. Rateless codes have been considered for various communications systems, e.g., distributed antenna systems [24], [25], relay systems [26], [27] and wireless broadcast systems [28]. As for C-RAN, in [29], a rateless coded transmission scheme for the multi-user downlink was proposed. In [30], [32], we designed the rateless coded transmission for single-user and multi-user uplink in C-RAN, where the degree profiles of rateless code were optimized. However, note that the above works solely considered simple scalar quantizer for fronthaul compression.

In this paper, we design the rateless coded uplink transmission with distributed fronthaul compression for C-RAN with two users and two RRHs under block fading channel. In each round of transmission, each user continuously transmits the Raptor coded messages to both RRHs. Then a novel fronthaul compression scheme is adopted at the RRHs which utilizes the correlation between the received signals at the two RRHs. Explicitly, both RRHs first quantize the received signals to meet the fronthaul capacity limit. Then one of the RRHs is selected to further compress the quantization bit sequence using LDPC code, and the generated syndrome bit sequence is sent to the BBU pool instead. The remaining fronthaul capacity is used to transmit an extra quantization bit sequence, which is generated by quantizing a part of the received signals with a higher precision. The BBU pool first decompresses the uploaded signals from the RRHs, and then performs iterative multi-user (MU) detection and decoding to recover both users' messages.

The main contributions of this paper are summarized as follows:

(1) We design the rateless coded multi-user uplink transmission in C-RAN, including the distributed fronthaul compression at RRHs, the decompression and iterative MU detection and decoding at the BBU pool.

(2) Resorting to EXIT [33], we analyze the decoding process for signal decompression at the BBU pool, based on which we optimize the degree profile of the LDPC compression code adopted at the RRH to improve the compression performance.

(3) Based on EXIT, we analyze the iterative MU detection and decoding at the BBU pool. Under the block fading channel assumption, we jointly consider the selection of RRH which executes the LDPC-based compression and the degree profile optimization for the Raptor codes applied at the two users, in order to maximize the average system throughput over all possible channel states. Specifically, the above optimization only relies on statistical CSI.

Compared with our previous works [29]–[32], as for the considered C-RAN model, [29]–[31] are different from this work while [32] is the same. Nevertheless, in this work we propose a novel transmission design wherein the key idea is to deploy the distributed fronthaul compression at the RRH, while the works in [29]–[32] solely considered the conventional scalar quantizer. The proposed transmission scheme brings several new design issues, which has never been investigated in [29]–[32]. The first issue is the selection of the appropriate RRH to execute the LDPC-based compression. The second is the optimization of LDPC code profiles for compression. The third is the design of MU detector at the BBU pool which should consider the fact that the quantization signals from the RRH have non-uniform quantization precisions. Furthermore, the above fact makes the optimization of Raptor code profiles obviously distinguishable from that in [32].

The rest of the paper is arranged as follows. In Section II, we introduce the system model. In Section III, the rateless coded uplink transmission scheme with distributed fronthaul compression is given. In Section IV, degree profile optimization for LDPC compression code and Raptor code is discussed. Simulation results are presented in Section V. Section VI concludes the paper.

II. SYSTEM MODEL

As shown in Fig. 1, we consider the uplink transmission in C-RAN where there are two users, two RRHs and one BBU pool. Both users and RRHs are equipped with single antenna.

The channel gain of the link from user $i, i = 1, 2$, to RRH $j, j = 1, 2$, is denoted as h_{ji}. We assume that each link experience independent block fading, i.e., h_{ji} is independent on each other and remains unchanged in one transmission round. During each round, each user keeps sending rateless codewords until the BBU pool successfully decodes and feedbacks an ACK. The specific transmission process is as follows.

Firstly, user i, $i = 1, 2$, applies Raptor code on its message of K bits to generate the coded bits w_i, and modulates the coded bits into symbols x_i which are sent to both RRHs. Note that the degree profiles of Raptor codes are fixed over all transmission rounds, thus the users are not required to be aware of the instant CSI. On the other hand, RRHs and BBU pool are assumed to have the full CSI. Each RRH pre-processes the received signals to obtain the baseband signals, which can be expressed as:

$$\boldsymbol{y} = \boldsymbol{H}\sqrt{P}\boldsymbol{x} + \boldsymbol{n}, \tag{1}$$

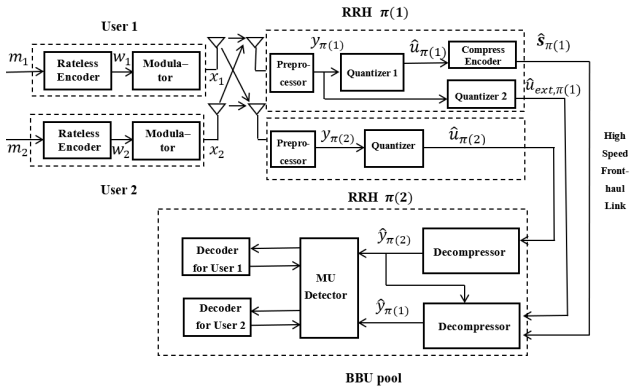


FIGURE 1. 2-user uplink in C-RAN with distributed fronthaul compression.

where $\boldsymbol{y} \triangleq [y_1, y_2]^T$, $\mathbf{H} \triangleq \left[h_{ji}\right]_{2\times 2}$, $\boldsymbol{x} \triangleq [x_1, x_2]^T$, $\boldsymbol{n} \triangleq [n_1, n_2]^T$, $y_j, j = 1, 2$, is the baseband signal received at RRH j, P is the transmission power of each user, n_j is the independent additive white Gaussian noise (AWGN) with mean of 0 and variance of σ_0^2 at RRH j.

Then RRH $j, j = 1, 2$ quantizes the baseband signals into a quantization bit sequence $\hat{u}_j$. The quantization precision depends on the fronthaul capacity. From (1), it can be observed that y_1 and y_2 are correlated, so are the two sequences $\hat{u}_1$ and $\hat{u}_2$. Therefore, we adopt the idea of Slepian-Wolf (SW) coding to further compress the quantization bit sequence. Let RRH π (1) be selected to perform the SW compression,[1] where π (1) $\in \{1, 2\}$ and the selection remains unchanged over all transmission rounds. Explicitly, RRH π (1) compression $\hat{u}_{\pi(1)}$ with LDPC code to obtain the syndrome bit sequence $\hat{\boldsymbol{s}}_{\pi\ (1)}$, whose length is shorter than the original sequence.[2] Therefore, the remaining fronthaul capacity can be used to transfer an extra quantization bit sequence $\hat{u}_{ext,\pi\ (1)}$ which is generated by quantizing a part of the received signal at RRH π (1) with a higher precision.

RRH π (1) and RRH π (2) (where π (2) $= 3 - \pi$ (1)) transmit the bit sequences $\{\hat{\boldsymbol{s}}_{\pi\ (1)}, \hat{u}_{ext,\pi\ (1)}\}$ and $\hat{u}_{\pi(2)}$ to the BBU pool through the fronthaul, respectively. At the BBU pool, the quantized bit sequence $\hat{u}_{\pi(1)}$ is firstly recovered from $\hat{u}_{\pi(2)}$ and $\hat{\boldsymbol{s}}_{\pi\ (1)}$ by belief propagation (BP) algorithm. With $\hat{u}_{\pi(1)}$, $\hat{u}_{\pi(2)}$ and $\hat{u}_{ext,\pi\ (1)}$, iterative multi-user detector (MU detector) and decoding based on BP algorithm are performed to recover both users' messages. When the decoding is successful, an ACK is fed back to both users.

III. RATELESS CODED MULTI-USER UPLINK TRANSMISSION WITH DISTRIBUTED COMPRESSION

In this section, we discuss the proposed rateless coded uplink transmission scheme with distributed fronthaul compression in detail. We first introduce the rateless encoding at the user, then the compression scheme at the RRH and finally the detailed algorithms for signal decompression and message recovery at the BBU pool.

A. RATELESS ENCODER AT THE USER

Each user $i, i = 1, 2$, encodes its message m_i into Raptor coded bits w_i, where m_i is firstly encoded by an LDPC code with rate R_P, and then by an LT code. LT code is characterized by its output node degree profile:

$$\Omega_i(x) = \sum_{d=1}^{d_c} \Omega_{i,d} x^d, i = 1, 2, \tag{2}$$

where d_c is the maximum output node degree, and $\Omega_{i,d}$ represents the probability that a LT coded bit (hereinafter referred to as the output bit) is with degree d. For an output bit with degree d, d bits are randomly selected from the LDPC coded bits (hereinafter referred to as the input bit) and XORed to generate the value of the output bit. Through the above encoding process, Raptor coded bits can be continuously generated. For simplicity, we use binary modulation on the coded bits, i.e., bits 0 and 1 are mapped to 1 and -1, respectively.

B. DISTRIBUTED FRONTHAUL COMPRESSION AT THE RRH

According to (1), the pre-pocessed baseband signal at RRH j is given by:

$$y_j = h_{j1}\sqrt{P}x_1 + h_{j2}\sqrt{P}x_2 + n_j, \quad j = 1, 2. \tag{3}$$

Its variance can be calculated as follows:

$$D\left(y_j\right) = E\left(y_j^2\right) - E^2(y_j) = P\sigma_{j1}^2 + P\sigma_{j2}^2 + \sigma_0^2, \tag{4}$$

where σ_{ji}^2 represents the variance of the channel gain h_{ji}.

Then each RRH first quantizes the baseband signal to meet the fronthaul capacity limit. Scalar quantization is utilized, wherein the quantization threshold is fixed during the transmission, since it is related to the hardware implementation. According to [34], it can be regarded that the value of y_j is almost distributed in the range $\left(-3D_j, 3D_j\right), D_j = \sqrt{D\left(y_j\right)}$. Assume that the fronthaul capacity between each RRH and the BBU pool is B bits/symbol. Then the number of quantization levels should satisfy $M_B = 2^B$. The quantized signal $\hat{y}_{j,B}$ is generated from the signal y_j according to the following rules:

$$\begin{aligned}\hat{y}_{j,B} &= Q_j(y_j) \\ &= \begin{cases} q_{j,1}, & y_j < -3D_j + \Delta_{j,B} \\ q_{j,k}, & -3D_j + (k-1)\Delta_{j,B} \le y_j < -3D_j + k\Delta_{j,B}, \\ & \qquad k = 2, \ldots, M_B - 1 \\ q_{j,M_B}, & y_j \ge 3D_j - \Delta_{j,B} \end{cases}\end{aligned} \tag{5}$$

where $q_{j,k} = -3D_j + \left(k - \frac{1}{2}\right)\Delta_{j,B}, k = 1, \ldots, M_B$ is the quantized value and $\Delta_{j,B} = \frac{3D_j}{M_B}$ is the quantization interval.

Then each quantized signal $\hat{y}_{j,B} = q_{j,k}$ is mapped to a quantization bit sequence l_j of length B:

$$l_j \triangleq \left\{l_j[1], \ldots, l_j[B]\right\}, \tag{6}$$

[1] Note that RRH selection will affect the system achievable rate. We will discuss the RRH selection scheme in Section IV.

[2] Note that the degree profile of the LDPC compression code may be changed in each transmission round according to the current channel state.

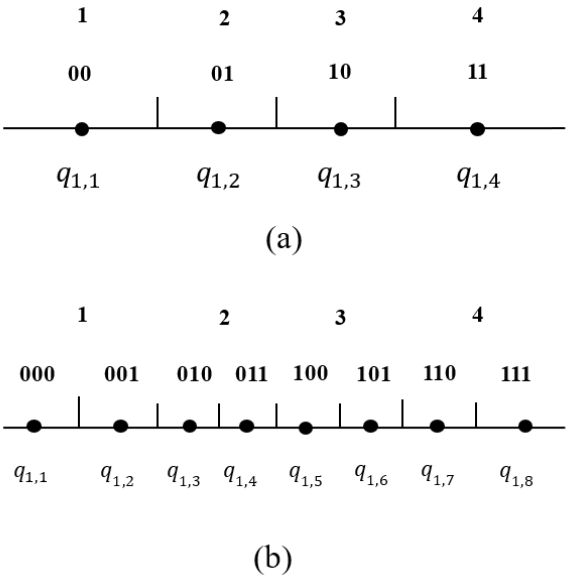


FIGURE 2. Illustration of 2-bit and 3-bit scalar quantizers.

where l_j is generated by decimal-to-binary conversion of k. According to (4), the received signals at both RRHs are correlated, so are l_1 and l_2. The transition probability from $l_2[r]$ to $l_1[r]$ is given by:

$$P(l_1[r]|l_2[r]) = \frac{P(l_1[r], l_2[r])}{P(l_2[r])} = \frac{1}{2}\sum_{w_1,w_2\in\psi} P(l_1[r]|w_1, w_2)P(l_2[r]|w_1, w_2), \tag{7}$$

where $r = 1, 2, \ldots, B$, $P(l_2[r]) = \frac{1}{2}$, ψ represents the set $\{0, 1\}$. Taking the case $\{w_1 = 0, w_2 = 0\}$ as an example, $P(l_1[r]|w_1, w_2)$ is calculated as follows:

$$P(l_1[r]|w_1 = 0, w_2 = 0) = \int_{\Re_r(l_1[r])} \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{(x-h_{11}-h_{12})^2}{2\sigma_0^2}} dx, \tag{8}$$

where the integral range $\Re_r(l_1[r])$ consists of all the quantization intervals that the r-th bit of the corresponding quantization bit sequence is $l_1[r]$. An illustration is given in Fig. 2(a), where the quantization bit number $B = 2$. In the case of $l_1[1] = 1$, $\Re_1(l_1[1])$ represents the third and fourth quantization intervals; If $l_1[2] = 1$, $\Re_2(l_1[2])$ includes the second and fourth quantization intervals. Calculation of $P(l_2[r]|w_1, w_2)$ is similar.

Each RRH j, $j = 1, 2$, packs the quantization bits corresponding to each N quantized signals into a bit sequence $\hat{u}_j$ of length NB. That is, $\hat{u}_j = [l_{j1}, \ldots, l_{jN}]$, where $l_{jk}, k = 1, \ldots, N$, denotes the quantization bit sequence corresponding to the kth signal in the block. RRH $\pi(1)$ compresses the bit sequence $\hat{u}_{\pi(1)}$ using LDPC code. Explicitly, $\hat{u}_{\pi(1)}$ is modular multiplied with the LDPC check matrix $\boldsymbol{A}_{m\times NB}$ to obtain a syndrome bit sequence $\hat{\boldsymbol{s}}_{\pi(1)} = [s_1, \ldots, s_m]$ of length m. Now the remaining fronthaul capacity for RRH $\pi(1)$ is $(B - \frac{m}{N})$ bits/symbol. To fully utilized it, $(BN - m)$ out of each N received signals $y_{\pi(1)}$ are randomly selected which are quantized with a finer quantizer with quantization bit number $B + 1$.

Remark 1: In Section IV-A, we will discuss how to determine the length of the syndrome bit sequence, i.e., m. Note that $(BN - m)$ may be greater than N under some channel states. In this case, RRH $\pi(1)$ can select a part of received signals and apply a more precise quantizer on the selected signals, e.g., with quantization bit number $B + 2$. In the following, we mainly discuss the case that $BN - m < N$ for simplicity.

For the finer quantizer with quantization bit number $B+1$, the number of quantization levels $M_{B+1} = 2^{B+1}$. Then the quantized signal $\hat{y}_{\pi(1),B+1}$ with quantization bit number $B+1$ is generated according to the following rule:

$$\begin{aligned}&\hat{y}_{\pi(1),B+1}\\ &= Q_{1,B+1}(y_{\pi(1)})\\ &\triangleq \begin{cases} q_{1,1}, & y_{\pi(1)} < -3D_1 + \Delta_{1,B+1}\\ q_{1,k}, & -3\sqrt{D_1}+(k-1)\Delta_{1,B+1} \le y_{\pi(1)} < -3\sqrt{D_1}\\ & \quad +k\Delta_{1,B+1} \quad k = 2, \ldots, M_{B+1}-1\\ q_{1,M_{B+1}}, & y_{\pi(1)} \ge 3D_1 - \Delta_{1,B+1}\end{cases}\end{aligned} \tag{9}$$

where $\Delta_{1,b+1} = \frac{3D_1}{M_{B+1}}$. The quantized signal $\hat{y}_{\pi(1),B+1}$ with quantization bit number $B+1$ is mapped to a new bit sequence l'_j of length $B + 1$ according the same rule in (6). Note that the first B bits l'_1 are exactly the original bit sequence l_1 of length B, as illustrated in Fig. 2(b). All the additional quantization bits corresponding to the selected $(BN - m)$ signals from an extra bit sequence $\hat{u}_{ext,\pi(1)}$ of length $(BN - m)$. RRH $\pi(1)$ transmits $\hat{\boldsymbol{s}}_{\pi(1)}$ and $\hat{u}_{ext,\pi(1)}$ to the BBU pool, without violating the capacity constraint of the fronthaul. On the other hand, RRH $\pi(2)$ directly transmits the bit sequence $\hat{u}_{\pi(2)}$ to the BBU.

C. DECOMPRESSION AND ITERATIVE MU DETECTION AND DECODING AT THE BBU POOL

As shown in Fig. 1, upon receiving the bit sequences from both RRHs, the BBU pool firstly recovers $\hat{u}_{\pi(1)}$ according to $\hat{\boldsymbol{s}}_{\pi(1)}$ from RRH $\pi(1)$ and $\hat{u}_{\pi(2)}$ from RRH $\pi(2)$. Then the MU detector and Raptor decoding are iteratively performed to recover both users' messages according to $\hat{u}_{\pi(1)}$, $\hat{u}_{\pi(2)}$ and $\hat{u}_{ext,\pi(1)}$.

1) DECOMPRESSION AT THE BBU POOL

The LDPC decoder recovers $\hat{u}_{\pi(1)} = [l_{\pi(1)1}, \ldots, l_{\pi(1)N}]$ which is the quantization bit sequence of length NB corresponding to N received signals at RRH $\pi(1)$. Fig. 3 shows the LDPC decoding graph for decompression, where the variable nodes correspond to $\hat{u}_{\pi(1)}$ to be recovered and the check nodes correspond to the syndrome bit sequence $\hat{\boldsymbol{s}}_{\pi(1)}$ from RRH $\pi(1)$. According to the correlation between $\hat{u}_{\pi(1)}$ and $\hat{u}_{\pi(2)}$ given by (7), the input log-likelihood-ratio (LLR) of the variable node corresponding to the r-th quantization bit of

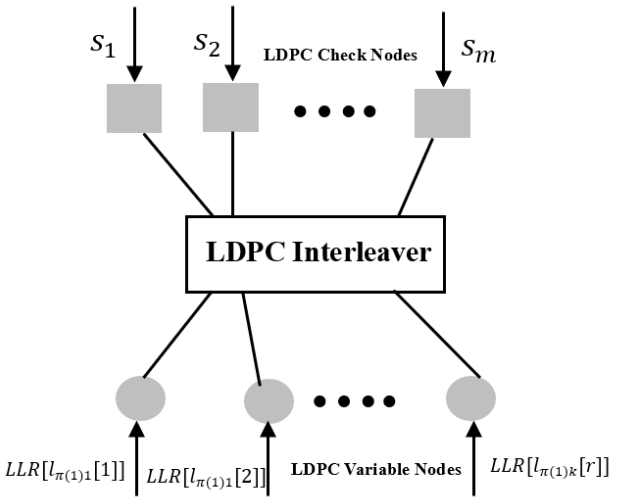


FIGURE 3. LDPC decoding graph for decompression at the BBU pool.

the k-th signal in $\hat{u}_{\pi(1)}$, i.e., $l_{\pi(1)k}[r]$, is expressed as (where $r = 1, \ldots, B, k = 1, \ldots, N$):

$$\begin{aligned} &LLR[l_{\pi(1)k}[r]] \\ &= ln\frac{P(l_{\pi(1)k}[r]=0|l_{\pi(2)k}[r])}{P(l_{\pi(1)k}[r]=1|l_{\pi(2)k}[r])} \\ &= ln\frac{\sum_{w_1,w_2\in\hat{E}\psi}P(l_{\pi(2)k}[r]|w_1,w_2)P(l_{\pi(1)k}[r]=0|w_1,w_2)}{\sum_{w_1,w_2\in\hat{E}\psi}P(l_{\pi(2)k}[r]|w_1,w_2)P(l_{\pi(1)k}[r]=1|w_1,w_2)}, \end{aligned} \tag{11}$$

which can be calculated according to (7).

BP algorithm is executed on the decoding graph in Fig. 3. In iteration 0, the LLR message $q_{vc}^{(0)}$ from the variable node v to the check node c is the input LLR (11):

$$q_{vc}^{(0)} = LLR[l_{\pi(1)k}[r]], \tag{12}$$

where node v corresponds to the r-th quantization bit corresponding to the k-th signal.

In the t-th iteration, the message from the variable node v to the check node c is given by:

$$q_{vc}^{(t)} = q_{vc}^{(0)} + \sum\nolimits_{c'\in\hat{E}C_v\setminus\{c\}} q_{c'v}^{(t-1)}, \tag{13}$$

where C_v denotes a set of check nodes connected to variable node v, $q_{c'v}^{(t-1)}$ represents the message passed from the check node c' to the variable node v in the previous round. The message from the check node c to the variable node v is updated as follows:

$$tanh\left(\frac{q_{cv}^{(t)}}{2}\right) = (1-2s_c)\prod\nolimits_{v'\neq v} tanh\left(\frac{q_{v'c}^{(t)}}{2}\right), \tag{15}$$

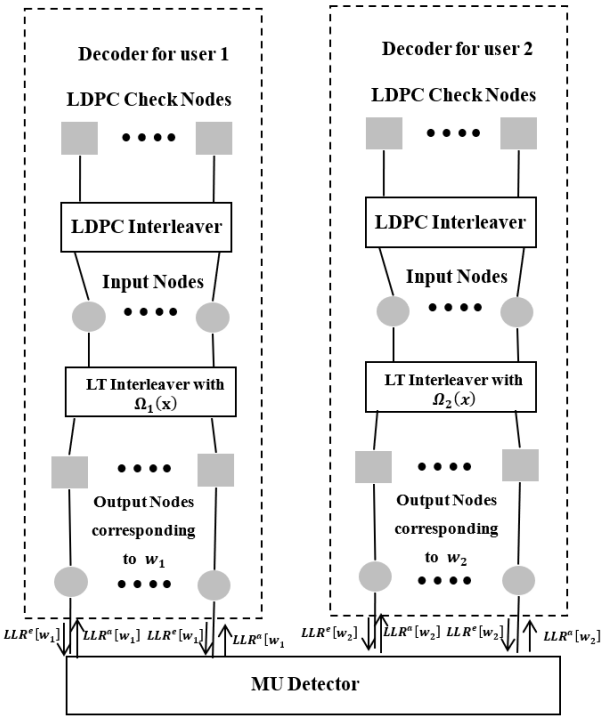


FIGURE 4. Decoding graph for MU detection and Raptor decoding at the BBU pool.

where s_c represents the syndrome bit in $\hat{s}$ corresponding to the check node c, v' denotes the variable node connected to the check node c except the variable node v, $q_{v'c}^{(t)}$ represents the message passed from the variable node v' to the check node c.

The above iterations end when the check equation $\hat{u}_1' A_{m\times NB}^T = \hat{s}_1$ is satisfied or the maximum iteration number is reached, where $\hat{u}_1'$ is the recovered bit sequence corresponding to the variable nodes. In the case of successfully decoding, $\hat{u}_1' = \hat{u}_{\pi(1)}$. Note that in the case of unsuccessful decoding, the LDPC decoder will still output the results to the following MU detector and decoder. Absolutely, the wrong quantization bits will degrade the performance of MU detector and decoder.

2) JOINT MU DETECTION AND DECODING AT THE BBU POOL

With the sequence $\hat{u}_{\pi(1)}$ and $\hat{u}_{ext,\pi(1)}$, the BBU can reconstruct the quantized received signals at RRH $\pi(1)$ which can be divided into two classes according to the quantization bit number, i.e., $\hat{y}_{\pi(1),B}$ and $\hat{y}_{\pi(1),B+1}$ (recalling (5) and (9)). Then the iterative MU detection and decoding is performed to recover users' messages m_i, $i = 1, 2$, from the quantized signals at both RRHs. The decoding graph is given in Fig. 4,

$$\begin{aligned} LLR^a[w_i] &= ln\frac{p(w_i=0|\hat{y}_{\pi(1),B},\hat{y}_{\pi(2),B},LLR^e[w_{i'}])}{p(w_i=1|\hat{y}_{\pi(1),B},\hat{y}_{\pi(2),B},LLR^e[w_{i'}])} \\ &= ln\frac{p(\hat{y}_{\pi(1),B},\hat{y}_{\pi(2),B}|w_i=0,w_i=1)+\frac{p(w_{i'}=0)}{p(w_{i'}=1)}p(\hat{y}_{\pi(1),B},\hat{y}_{\pi(2),B}|w_i=0,w_{i'}=0)}{p(\hat{y}_{\pi(1),B},\hat{y}_{\pi(2),B}|w_i=1,w_{i'}=1)+\frac{p(w_{i'}=0)}{p(w_{i'}=1)}p(\hat{y}_{\pi(1),B},\hat{y}_{\pi(2),B}|w_i=1,w_{i'}=0)} \end{aligned} \tag{14}$$

where the two subgraph in the upper part and middle part correspond to the LDPC precoder and the LT code of the Raptor code and the lower part is the MU detector. The input nodes correspond to input bits (LDPC coded bits) and the output nodes correspond to the (Raptor coded) output bits w_i for each user. For simplicity, each output node and the LT check node are regarded as a whole and referred to as 'output node' in the following. The MU detector calculates the LLR of the output bit w_i of user $i, i = 1, 2$, based on the quantized received signals $\hat{y}_{\pi(1),B}$ or $\hat{y}_{\pi(1),B+1}$ at RRH $\pi(1)$, $\hat{y}_{\pi(2),B}$ at RRH $\pi(2)$, and the output $LLR^e[w_i]$ from the Raptor decoders, where $i' = 3 - i$. For example, when the quantization bit number of the received signal at RRH $\pi(1)$ corresponding to the output bit w_i is B, the output LLR from the MU detector is given as follows:

where $\frac{p(w_{i'}=0)}{p(w_{i'}=1)}$ equals 1 in the first decoding iteration and equals $e^{LLR^e[w_{i'}]}$ in the following iterations. Taking the case $\{i = 1, w_1 = 1, w_2 = 1\}$ as an example, according to (5), $P(\hat{y}_{\pi(1),B}, \hat{y}_{\pi(2),B}|w_1 = 1, w_2 = 1$ is calculated as follows:

$$\begin{aligned} &p(\hat{y}_{\pi\ (1),B}, \hat{y}_{\pi\ (2),B}|w_1 = 1, w_2 = 1) \\ &= \int_{\Re(\hat{y}_{\pi\ (1),B})} \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{(x+h_{11}-h_{12})^2}{2\sigma_0^2}} dx \\ &\quad \times \int_{\Re(\hat{y}_{\pi\ (2),B})} \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{(x+h_{21}-h_{22})^2}{2\sigma_0^2}} dx, \end{aligned} \tag{16}$$

where $\Re(\hat{y}_{\pi(1),B})$ and $\Re(\hat{y}_{\pi(2),B})$ respectively represent the quantization intervals corresponding to $\hat{y}_{\pi(1),\mathrm{B}}$ and $\hat{y}_{\pi(2),\mathrm{B}}$ according to (5) and Fig. 2. The other three conditional probability in (14), as shown at the bottom of the previous page, can be derived accordingly.

The iterative detection and decoding process can be divided into two stage. In the first stage, the decoding iterations are performed on the whole decoding graph in Fig. 4. The first stage ends when the mean of the absolute LLR of the input nodes of both users exceeds threshold L_{th}, which is necessary for the successful decoding of the LDPC precoder [35]. In the second stage, decoding iterations are executed individually on the LDPC decoding graph of each user to eliminate remaining errors. The detailed decoding procedure in the t-th iteration of stage 1 is given as follows (taking the decoding for user 1 as an example):

Step 1: In the decoding graph for user 1, the message conveyed from the input node i to the LDPC check node c is updated by:

$$m_{ic}^{(t)} = \sum\nolimits_o m_{oi}^{(t-1)}, \tag{17}$$

where $m_{oi}^{(t-1)}$ represents the message transmitted from the output node to the input node in the $(t-1)$ round.

Step 2: The message from the LDPC check node c to the input node i is updated as follows:

$$\tanh\left(\frac{m_{ci}^{(t)}}{2}\right) = \prod\nolimits_{i'\neq i} \tanh\left(\frac{m_{i'c}^{(t)}}{2}\right), \tag{18}$$

where $m_{i'c}^{(t)}$ means that the message is from all input nodes (except i) to the LDPC check node c.

Step 3: The message from the input node i to the LT output node o is:

$$m_{io}^{(t)} = \sum\nolimits_{o'\neq o} m_{o'i}^{(t-1)} + \sum\nolimits_c m_{ci}^{(t)}, \tag{19}$$

where $m_{o'i}^{(t-1)}$ represents the message transmitted from the output node o' to the input node i in the $(t-1)$ round. $m_{ci}^{(l)}$ represents the message from LDPC check node c to the input node i.

Step 4: $LLR^e[w_1]$ transferred to the MU detector is given by:

$$LLR^e[w_1] = \sum\nolimits_o m_{oi}^{(t-1)} + \sum\nolimits_c m_{ci}^{(t)}, \tag{20}$$

Then the MU detector transmitted the above LLR message to the decoding graph of user 2. The decoding procedure is similar to step 1-4.

Step 5: LT output node o returns the message to the input node i:

$$\tanh\left(\frac{m_{oi}^{(t)}}{2}\right) = \tanh\left(\frac{z_0}{2}\right) \prod\nolimits_{i'\neq i} \tanh\left(\frac{m_{i'o}^{(t)}}{2}\right), \tag{21}$$

where $z_0 = LLR^a[w_1]$ represents the message from the MU detector which is calculated by (14).

Step 6: After each iteration, the LLR of the input bit corresponding to the input node i is:

$$m_i^{(t)} = \sum\nolimits_o m_{oi}^{(t)}. \tag{22}$$

The second stage of decoding is performed individually on the LDPC decoding graph for each user. The process is similar to step 1 and step 2.

IV. EXIT ANALYSIS AND DEGREE PROFILE OPTIMIZATION

In this section, we first present the optimization scheme for the variable node degree profile of the LDPC compression code applied at the RRH. Then we discuss the selection of the RRH to perform LDCP-based compression and the optimization of the output node degree profiles of Raptor code for each user to maximize the average throughput of the system. Recalling that since RRHs and BBU pool are assumed to be aware of the instant CSI, the selected RRH can adopt the optimized LDPC compression code for the current channel state. On the other hand, each user does not acquire the instant CSI and solely applies the pre-determined Raptor code degree profile over all transmission rounds. So the Raptor code profiles should be optimized to accommodate all possible channel states.

For all the above code optimization, we resort to the EXIT analysis. We investigate the extrinsic information (EI) passing on the decoding graph during the decoding iteration, where EI is defined as mutual information of the message bit and its corresponding LLR, and derive the EI update rules. Based on this, the optimization problems are formulated.

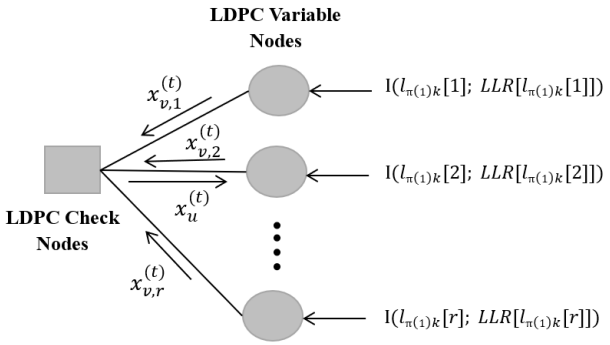


FIGURE 5. EI transfer on the LDPC decoding graph for decompression.

A. DEGREE PROFILE OPTIMIZATION FOR LDPC COMPRESSION CODE

Firstly we introduce some notations for the LDPC decoding graph given in Fig. 3. The LDPC variable node degree distribution is defined as $\xi(x) = \sum_{d=1}^{d_v} \xi_d x^d$, where ξ_d represents the ratio of variable nodes with degree d to all variable nodes, d_v is the maximum degree of the variable node. LDPC variable node edge distribution is defined as $\tilde{\xi}(x) = \sum_{d=1}^{d_v} \tilde{\xi}_d x^{d-1}$, where $\tilde{\xi}_d$ is the probability of the edges connected to the variable node with degree d to all edges. LDPC check node edge distribution is defined as $\tilde{\rho}(x) = \sum_{d=1}^{d_c} \tilde{\rho}_d x^{d-1}$, where $\tilde{\rho}_d$ is the probability of the edges connected to the check node with degree d to all edges. The LDPC check node degree distribution is defined as $\rho(x) = \sum_{d=1}^{d_c} \rho_d x^d$, where ρ_d represents the ratio of check nodes with degree d to all check nodes. Here we consider LDPC code with identical check node degree $\tilde{\rho}$, i.e., $\tilde{\rho}(x) = \rho(x) = x^{\tilde{\rho}}$.

Recalling (5) and (6), the LDPC variable nodes corresponding to the quantization bit sequence $\hat{u}_{\pi(1)} = [l_{\pi(1)1}, \ldots, l_{\pi(1)N}]$ can be divided into B groups each of which consists of N nodes. Explicitly, each node in the r-th group ($r = 1, \ldots, B$) corresponds to the r-th bit of $l_{\pi(1)k}$, $k = 1, \ldots, N$.

Fig. 5 illustrate the EI update between the check nodes and the variable nodes in the LDPC decoding graph given in Fig. 3. For the t-th round of decoding iteration, the EI is updated as follows:

Step 1: EI transferred from the variable node in group r, $r = 1, \ldots, B$, to the check node is given by:

$$x_{v,r}^{(t)} = \sum_{d=1}^{d_v} \tilde{\xi}_d J\left((i-1)J^{-1}\left(x_u^{(t-1)}\right) + J^{-1}\left(I(l_{\pi(1)k}[r]; LLR[l_{\pi(1)k}[r]])\right)\right), \quad (23)$$

where $x_u^{(t-1)}$ represents EI transferred from the check node to the variable node in round $(t-1)$, $J(\tau)$ is the EI corresponding to the LLR message satisfying symmetric Gaussian distribution with mean τ [36]:

$$J(\tau) = 1 - \frac{1}{\sqrt{4\pi\tau}} \int_{-\infty}^{+\infty} \log_2(1+e^{-v}) e\left(-\frac{(v-\tau)^2}{4\tau}\right) dv. \quad (24)$$

The input EI $I(l_{\pi(1)k}[r]; LLR[l_{\pi(1)k}[r]])$ can be derived as follows:

$$\begin{aligned} &I\left(l_{\pi(1)k}[r]; LLR\left[l_{\pi(1)k}[r]\right]\right) \\ &= H\left(l_{\pi(1)k}[r]\right) - H\left(l_{\pi(1)k}[r] \,|\, LLR\left[l_{\pi(1)k}[r]\right]\right) \\ &= 1 - H_r, \end{aligned} \quad (25)$$

where $H_r \triangleq H(l_{\pi(1)k}[r]|LLR[l_{\pi(1)k}[r]])$.

Note that the B groups have the same number of variable nodes, the average EI passed from all variable nodes to check nodes is:

$$x_v^{(t)} = \frac{1}{B} \sum_{r=1}^{B} x_{v,r}^{(t)}. \quad (26)$$

Step 2: The EI from the variable nodes to the check nodes is:

$$x_u^{(t)} = 1 - \sum_{d=1}^{d_c} \tilde{\rho}_d J\left((j-1)J^{-1}\left(1-x_v^{(t)}\right)\right). \quad (27)$$

Substituting (23) and (25) into (27), we can get the update rule of $x_u^{(t)}$, which is denoted as a function $\Phi(\cdot)$:

$$x_u^{(t)} = \Phi\left(x_u^{(t-1)}, I(l_{\pi(1)k}[r]; LLR[l_{\pi(1)k}[r]]), \hat{\rho}, \{\tilde{\xi}_d\}\right). \quad (28)$$

Obviously, (28) is linear with the variable node edge distribution $\{\tilde{\xi}_d\}$.

Recalling the distributed fronthaul compression scheme proposed in Section III-B, a shorter syndrome bit sequence $\hat{s}_{\pi(1)} = [s_1, \ldots, s_m]$ results in larger remaining fronthaul capacity. Therefore, we optimize $\{\tilde{\xi}_d\}$ to minimize the required syndrome bits number m (or, equivalently, the number of check nodes) for successfully decompression at the BBU pool. Since the number of check nodes is given by $NB(\hat{\rho}\sum_{d=1}^{d_v} \frac{\tilde{\xi}_d}{d})^{-1}$, the optimization problem is formulated as follows:

$$\begin{aligned} &\max_{\hat{\rho}\{\tilde{\xi}_d\}} \ \hat{\rho} \sum_{=1}^{d_v} \frac{\tilde{\xi}_d}{d} \\ &\begin{cases} C1: \sum_{d=1}^{d_v} \tilde{\xi}_d = 1 \\ C2: \xi_d \geq 0, \quad d = 1, \ldots, D \\ C3: \Phi\left(x_u, I\left(l_{\pi(1)k}[r]; LLR\left[l_{\pi(1)k}[r]\right]\right), \quad \bar{\rho}, \{\xi_d\}\right) > x_u \end{cases} \end{aligned} \quad (29)$$

where $C1$ indicates that the sum of the edge distribution coefficients $\{\tilde{\xi}_d\}$ is 1, $C2$ means that coefficients $\{\tilde{\xi}_d\}$ are no smaller than 0, $C3$ means that EI should keep ascending during the decoding iteration. Note that x_u in $C3$ is continuous. To make it tractable, x_u is discretized into N equal-spaced values within the range [0, 1].

To solve the above problem, firstly we can use LP to obtain the optimal variable node edge distribution $\tilde{\xi}(x)$ under an arbitrary fixed $\bar{\rho}$. Then the optimal $\bar{\rho}$ can be obtained by exhaustive searching. Finally, the variable node degree profile $\xi(x)$ of the LDPC code can be derived according to the equality $\tilde{\xi}(x) = \frac{\xi'(x)}{\xi'(1)}$.

Remark 2: Note that under fixed channel state, no matter which RRH is selected to perform the compression, the resulted optimized LDPC code profile is the same. Recalling the input EI given in (25), we have the following equality:

$I\left(l_{\pi(1)k}[r]; LLR\left[l_{\pi(1)k}[r]\right]\right) = I\left(l_{\pi(1)k}[r]; l_{\pi(2)k}[r]\right) = I\left(l_{\pi(2)k}[r]; LLR\left[l_{\pi(2)k}[r]\right]\right)$. That indicates no matter RRH 1 or RRH 2 is selected to perform the compression, the input EI is the same for the proposed LDPC code optimization problem, which results in the same optimization result. Nevertheless, the value of $\pi(1)$ affects the Raptor code optimization as shown in the next subsection.

Remark 3: The proposed optimization is based on infinite LDPC code length. In practice, the code length NB is quite limited since it results in the signal processing latency at the RRH $\pi(1)$. To guarantee a satisfactory decompression performance at the BBU pool, we choose a slightly larger syndrome bits number than the optimized result in (28) is given as follows:

$$m = NB(\hat{\rho}\sum\nolimits_{=1}^{d_v}\frac{\tilde{\xi}_d}{d})^{-1}(1+\vartheta), \tag{30}$$

where ϑ is a small positive number. Using PEG algorithm [37], the LDPC check matrix $\boldsymbol{A}_{m\times NB}$ for compressing $\hat{u}_1$ into $\hat{\boldsymbol{s}}_{\pi(1)}$ is generated according to the node degree profile $\xi(x)$ and the syndrome bit number m. Finally, in the case that the m is larger than NB, RRH $\pi(1)$ will not apply the LDPC compression.

B. DEGREE PROFILE OPTIMIZATION FOR RAPTOR CODE

In this subsection, we jointly optimize the output node degree profiles of Raptor code for both users.

Firstly we introduce some notations for the decoding graph given in Fig. 4. The LDPC variable node degree distribution is defined as $\zeta(x) = \sum_{d=1}^{d_v}\zeta_d x^d$, where ζ_d represents the ratio of variable nodes with degree d to all variable nodes, d_v is the maximum degree. LDPC check node edge distribution is defined as $\tilde{\varrho}(x) = \sum_{d=1}^{d_c}\tilde{\varrho}_d x^{d-1}$, where $\tilde{\varrho}_d$ is the probability of an edge connected to the check node with degree d. The input node degree distribution of user i, $i = 1, 2$, is defined as $\alpha_i(x) = \sum_{d=1}^{d_v'}\alpha_{i,d}x^d$, where d_v' is the maximum degree of the input node and $\alpha_{i,d}$ is the probability of the input node of user i with degree d. The input node edge distribution of user i is defined as $\tilde{\alpha}_i(x) = \sum_{d=1}^{d_v'}\tilde{\alpha}_{i,d}x^{d-1}$, where $\tilde{\alpha}_{i,d}$ is the probability of an edges connected to the input node with degree d. According to [38], $\alpha_i(x)$ and $\tilde{\alpha}_i(x)$ can be approximate by Poisson distribution with mean $\bar{\alpha}_i$. The output node edge distribution of user i is defined as $\omega_i(x) = \sum_{d=1}^{d_c'}\omega_{i,d}x^{d-1}$, where $\omega_{i,d}$ is the probability of an edge connected to the output node with degree d. The relationship between the output node degree distribution $\Omega_i(x)$ and $\omega_i(x)$ can be expressed as:

$$\omega_i(x) = \frac{\Omega_i'(x)}{\Omega_i'(1)}. \tag{31}$$

Recall that the iterative MU detection and decoding has two stages. In the first stage, the decoding iteration is executed on the whole graph in Fig. 4. Then in the second stage, decoding is individually performed on the LDPC subgraph. Fig. 6 shows the update process of EI transferred between each node in the first decoding stage. According to the proposed distributed fronthaul compression scheme, for each N

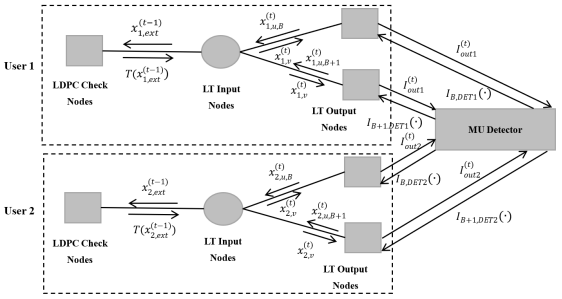


FIGURE 6. EI transfer on the decoding graph for MU detection and Raptor decoding at the BBU pool.

received signals $y_{\pi(1)}$ at RRH $\pi(1)$, $(BN - m)$ signals are quantized by $B+1$ bits, while the other by B bits. Therefore, the output nodes in the Raptor decoding graph for each use should be divided into two parts corresponding to the different quantization bit numbers.

Now we analyze the EI update in Fig. 6. In the t-th round, the EI update in the decoding subgraph for user 1 is given as follows, which can be derived accordingly for user 2.

Step 1: The LT input nodes transmit the LLR messages to the LDPC check nodes, and the corresponding EI is:

$$x_{1,ext}^{(t-1)} = \sum\nolimits_{d=1}^{d_v'}\alpha_{1,d}J(dJ^{-1}\left(x_{1,u}^{(t-1)}\right)), \tag{32}$$

where $x_{1,u}^{(t-1)}$ is the average EI transmitted from the output nodes to the input nodes in the round $(t-1)$.

Step 2: The EI from the LDPC check nodes to the input nodes of user i is:

$$T\left(x_{1,ext}^{(t-1)}\right) = \sum\nolimits_{d=1}^{d_v}\zeta_d J(dJ^{-1}\left(1-\sum\nolimits_{j=1}^{d_c'}\tilde{\varrho}_j J((j-1)J^{-1}\left(1-x_{1,ext}^{(t-1)}\right)\right)\right)). \tag{33}$$

Step 3: The EI from the input nodes to the output nodes is:

$$x_{1,v}^{(t)} = \sum_{d=1}^{d_v'}\tilde{\alpha}_{1,d}J((d-1)J^{-1}\left(x_{1,u}^{(t-1)}\right) + J^{-1}(T(x_{1,ext}^{(t-1)}))). \tag{34}$$

Step 4: The EI from LT output nodes for user 1 to MU detector is:

$$I_{out1}^{(t)} = 1 - \sum\nolimits_{d=1}^{d_c'}\Omega_{1,d}J(dJ^{-1}\left(1-x_{1,v}^{(t-1)}\right)). \tag{35}$$

Step 5: The EI from the output nodes corresponding to quantization bit number B to the input nodes is:

$$x_{1,u,B}^{(t)} = 1 - \sum\nolimits_{d=1}^{d_c'}\omega_{1,d}J\left((d-1)J^{-1}\left(1-x_{1,v}^{(t)}\right) + J^{-1}\left(1-I_{B,DET1}(I_{out2}^{(t-1)};\mathbf{H})\right)\right). \tag{36}$$

The EI from the output nodes corresponding to quantization bit number $B+1$ to the input nodes is:

$$x_{1,u,B+1}^{(t)} = 1 - \sum\nolimits_{d=1}^{d_c'}\omega_{1,d}J\left((d-1)J^{-1}\left(1-x_{1,v}^{(t)}\right) + J^{-1}\left(1-I_{B+1,DET1}(I_{out2}^{(t-1)};\mathbf{H})\right)\right). \tag{37}$$

In (35) and (36), $I_{B,DET1}(I_{out2}^{(t-1)};\mathbf{H})$ and $I_{B+1,DET1}(I_{out2}^{(t-1)};\mathbf{H})$ represent the output EI from the MU detector to the output nodes of user 1 where the corresponding quantization bit number for $y_{\pi(1)}$ is B and B + 1, respectively. The derivation of $I_{B,DET1}(I_{out2}^{(t-1)};\mathbf{H})$ and $I_{B+1,DET1}(I_{out2}^{(t-1)};\mathbf{H})$ will be discussed later.

Step 6: The average EI from all output nodes to the input nodes in the subgraph for user 1 is given by:

$$x_{1,u}^{(t)} = (B - m/N)x_{1,u,B+1}^{(t)} + (1 - B + m/N)x_{1,u,B}^{(t)}. \quad (38)$$

In Step 5, $I_{B,DET1}(I_{out2};\mathbf{H})$ and $I_{B+1,DET1}(I_{out2};\mathbf{H})$ are functions of the input EI I_{out2} from the output nodes of user 2 to the MU detector in t-th iteration and the channel matrix $\mathbf{H}$. Both functions can be evaluated by Monte Carlo method [39]. In the following we take $I_{B+1,DET1}(I_{out2};\mathbf{H})$ as an example. Firstly, set the transmitted signals x_1 of user 1 to be 1, and randomly generate the transmitted symbols x_2 of user 2, i.e., either 1 or -1, with uniform probability. According to (1), under fixed $\mathbf{H}$, we generate Gaussian noise samples and get y_1 and y_2. Then according to (9), we obtain $\hat{y}_{\pi(1),B+1}$ and $\hat{y}_{\pi(2),B}$. For an arbitrary I_{out2}, we randomly generate the corresponding $LLR^e[w_2]$ from symmetric Gaussian distribution with mean τ and $-\tau$ when the corresponding x_2 equals 1 and -1, respectively, where τ is calculated by (24). $LLR^a[w_1]$ can be obtained according to (14). In conclusion, by generating a large number of x_2 and Gaussian noise sample, we can approximately obtain the conditional p.d.f. (i.e., probability distribution function) of $LLR^a[w_1]$ when $x_1 = 1$, i.e., $p(LLR^a[w_1]\,|\,x_1 = 1)$. Then $p(LLR^a[w_1]\,|\,x_1 = 1, x_2 = 1)$ and $p(LLR^a[w_1]\,|\,x_1 = 1, x_2 = -1)$ can be obtained similarly. With this, $I_{B+1,DET1}(I_{out2};\mathbf{H})$ is given by:

$$I_{B+1,DET1}(I_{out2};\mathbf{H}) = \sum_{x_2=1,-1} p\big(LLR^a[w_1]\,|\,x_1=1,x_2\big)\log\frac{p(LLR^a[w_1]\,|\,x_1=1,x_2)}{p(LLR^a[w_1]\,|\,x_1=1)} \quad (39)$$

Furthermore, from the numerical results, we find that under fixed $\mathbf{H}$, $I_{B+1,DET1}(I_{out2},\mathbf{H})$ can be well approximated by a linear function:

$$I_{B+1,DET1} = aI_{out2} + b, \quad (40)$$

where a and b can be obtained by evaluating two special cases:

$$I_{B+1,DET1}(0;\mathbf{H}) = I\left(x_1;\hat{y}_{\pi(1),B+1},\hat{y}_{\pi(2),B}\right), \quad (41)$$

$$I_{B+1,DET1}(1;\mathbf{H}) = I(x_2;\hat{y}_{\pi(1),B+1},\hat{y}_{\pi(2),B}|x_2). \quad (42)$$

Fig. 7 gives the numerical $I_{B+1,DET1}(I_{out2},\mathbf{H})$ and its linear approximation, named 'MC' and 'LA', respectively, when $B = 2$, $SNR = \frac{P}{\sigma_0} = -1.58$ dB and $\pi(1) = 1$. Three channel matrices $\mathbf{H}$ are considered, i.e., case 1: $\mathbf{H} = \begin{bmatrix} 1.8672 & 0.4413 \\ 1.3802 & 0.6236 \end{bmatrix}$, case 2: $\mathbf{H} = \begin{bmatrix} 0.8937 & 1.3841 \\ 0.6928 & 1.9839 \end{bmatrix}$, and case 3: $\mathbf{H} = \begin{bmatrix} 2.9021 & 0.4483 \\ 0.3402 & 1.2053 \end{bmatrix}$. From Fig. 7, we can find that the linear approximation is considerably accurate.

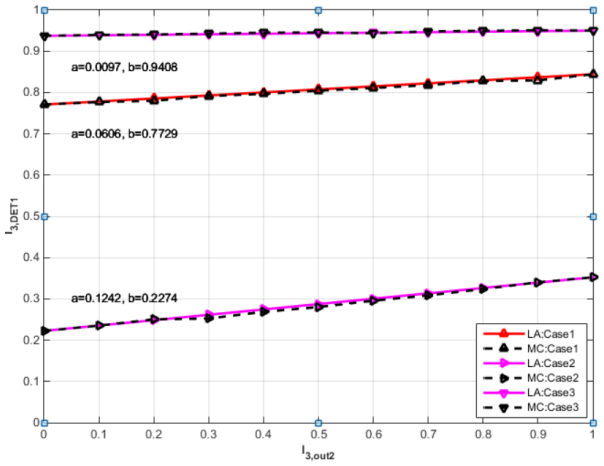


FIGURE 7. The relationship between the output and input EI for the MU detector.

For successfully decoding, the conditions $x_{i,u}^{(t+1)} > x_{i,u}^{(t)}$ and $x_{i,u}^{(l_{max})} > x_{i,u}^{th}$, $i = 1, 2$, should be satisfied, where l_{max} is the maximal decoding iteration times, $x_{i,u}^{th}$ is the minimal EI required by successful decoding of the LDPC codes in the second stage and can be calculated according to L_{th} given in Section III-C with the method given in [35].

Based on the above EXIT analysis, we optimize the output node degree profile of the Raptor code used at each user as well as the RRH selection, i.e., $\pi(1)$. Recall that under the block fading channel, the channel state matrix $\mathbf{H}$ remains unchanged during each round of transmission and changes independently round by round. In one round of transmission, both users continue to send Raptor codewords simultaneously until the BBU pool successfully recover users' messages. The length of the Raptor codewords for each user in each round of transmission is related to the current channel state matrix $\mathbf{H}$ and the RRH selection $\pi(1)$, which can be expressed as:

$$L_i(\mathbf{H},P,\pi(1)) = \frac{K}{R_p}\bar{\alpha}_i(\mathbf{H},P,\pi(1))\sum_{d=1}^{d_c'}\frac{\omega_{i,d}}{d}, \quad (43)$$

where $\bar{\alpha}_i(\mathbf{H},P,\pi(1))$ is the average LT input node degree for user $i, i = 1, 2$, under channel matrix $\mathbf{H}$ and P is the user transmit power. In each round transmission, both users start and stop transmission simultaneously, therefore let $L_1(\mathbf{H},P,\pi(1)) = L_2(\mathbf{H},P,\pi(1)) \triangleq L(\mathbf{H},P,\pi(1))$.

Usually the channel state space is continuous. To make the problem tractable, we discretize the space into Q states: $\mathbf{H}_q$, $q = 1,\ldots,Q$. The probability of each state is defined as $\Pr(\mathbf{H}_q)$. Then the average Raptor codeword length over all Q state can be expressed as:

$$\bar{L}_i\left(\mathbf{H}_q,P,\pi(1)\right) = \sum_{q=1}^{Q}\Pr(\mathbf{H}_q)\frac{K}{R_p}\bar{\alpha}_i(\mathbf{H}_q,P,\pi(1))\sum_{d=1}^{d_c'}\frac{\omega_{i,d}}{d}. \quad (44)$$

For all channel states $\mathbf{H}_q$, $\bar{\alpha}_i(\mathbf{H}_q,P,\pi(1))$ can be approximated as $\bar{\alpha}_i\left(\mathbf{H}_q,P,\pi(1)\right) = \bar{\alpha}_{i,0}C_i^{-1}(\mathbf{H}_q,P,\pi(1))$ [40], where $\bar{\alpha}_{i,0}$ is a constant and independent on the channel state, $C_i(\mathbf{H}_q,P,\pi(1))$ represents the theoretical achievable rate of

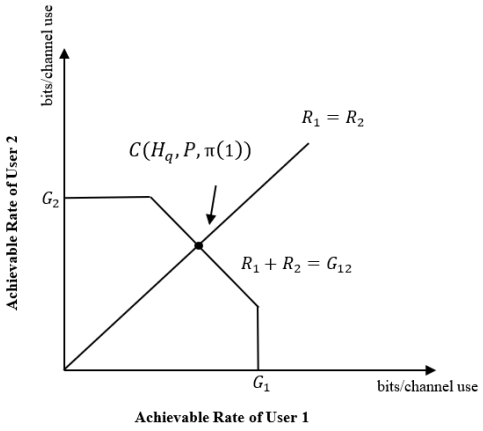


FIGURE 8. Illustration of the theoretical achievable rate region for the two-user C-RAN uplink.

user i under channel matrix $\mathbf{H}_q$. In the following, we derive $C_i(\mathbf{H}_q, P, \pi\ (1))$.

Note that $(BN - m)$ out of N received signals at RRH $\pi\ (1)$ is quantized by $B + 1$ bits. In this case, the theoretical achievable rate region for user 1 and user 2 is given by:

$$R_1 \le I\left(x_1; \hat{y}_{\pi(1),B+1}, \hat{y}_{\pi(2),B} \,|\, x_2\right) \triangleq G_{1,B+1}, \quad (45)$$

$$R_2 \le I\left(x_2; \hat{y}_{\pi(1),B+1}, \hat{y}_{\pi(2),B} \,|\, x_1\right) \triangleq G_{2,B+1}, \quad (46)$$

$$R_1 + R_2 \le I\left(x_1, x_2; \hat{y}_{\pi(1),B+1}, \hat{y}_{\pi(2),B}\right) \triangleq G_{12,B+1}, \quad (47)$$

The other received signals at RRH $\pi\ (1)$ are quantized by B bits, and the corresponding theoretical achievable rate region is $R_1 \le G_{1,B}$, $R_2 \le G_{2,B}$ and $R_1 + R_2 \le G_{12,B}$. Therefore, the overall theoretical achievable rate region under RRH selection $\pi\ (1)$ for the current round of transmission is: $R_1 \le G_1$, $R_2 \le G_2$, and $R_1 + R_2 \le G_{12}$, where G_1, G_2 and G_{12} can be calculated as follows:

$$G_1 = (B - m/N)\, G_{1,B+1} + (1 - B + m/N) G_{1,B}, \quad (48)$$

$$G_2 = (B - m/N)\, G_{2,B+1} + (1 - B + m/N) G_{2,B}, \quad (49)$$

$$G_{12} = (B - m/N)\, G_{12,B+1} + (1 - B + m/N) G_{12,B}. \quad (50)$$

Fig. 8 illustrate the above region. Since the messages of both users are of the same length, and the Raptor code length of both users are the same as well, the achievable rate of both user should be equal. Therefore, the maximal achievable rate for both users is the intersection point of the line $R_1 = R_2$ and the boundary of the achievable region, and we have:

$$C_1\left(\mathbf{H}_q, P, \pi\ (1)\right) = C_2\left(\mathbf{H}_q, P, \pi\ (1)\right) = C\left(\mathbf{H}_q, P, \pi\ (1)\right) \triangleq \min\left(G_1, G_2, \frac{1}{2} G_{12}\right). \quad (51)$$

Remark 4: As discussed in Remark 3, in practice, there may be some channel states under which the distributed fronthaul compression is inapplicable. For these channel states, the theoretical achievable rates of both users are given by:

$$C_{1,con}\left(\mathbf{H}_q, P\right) = C_{2,con}\left(\mathbf{H}_q, P\right) = C_{con}\left(\mathbf{H}_q, P\right). \triangleq \min\left(G_{1,B}, G_{2,B}, \frac{1}{2} G_{12,B}\right). \quad (52)$$

We define $T = \frac{2K}{\bar{L}(\mathbf{H}_q, P, \pi(1))}$ as the average sum throughput, which is inverse to the average Raptor code length of both users. Our goal is to optimize both users' Raptor code degree profiles to maximize T, or equivalently, minimize $\bar{L}\left(\mathbf{H}_q, P, \pi\ (1)\right)$ under the constraint that BBU pool successfully decodes both users' messages under all possible channel states. This problem can not be solved by linear programming method as in [31]. Instead, we resort to searching degree profiles to minimize the required user transmit power while ensuring successful decoding under fixed average code length $\bar{L}\left(\mathbf{H}_q, P, \pi\ (1)\right)$. Note that this strategy has been used for AWGN multiple access relay channel [32] [39]. Here we set the fixed average code length to be the theoretical code length required for successful recovery, which is given by:

$$L = \frac{K}{R_p} \sum\nolimits_{q=1}^{Q} \Pr(\mathbf{H}_q) C^{-1}(\mathbf{H}_q, P, \pi\ (1)). \quad (53)$$

Overall, the optimization problem can be formulated as:

$$\min_{P_{th}, \bar{\alpha}_{i,0}, \{\omega_{i,d}\}, \pi(1)} P_{th}$$

$$\begin{cases} C1: \sum\limits_{d=1}^{d_c'} \omega_{i,d} = 1, & i = 1, 2, \\ C2: \omega_{i,1} > \varepsilon, & i = 1, 2, \\ C3: x_{i,u}^{(l_{\max})} > x_{i,u}^{th}, & \text{for all channel matrices } H_q, \\ & i = 1, 2, \\ C4: \dfrac{K}{R_p} \sum\limits_{q=1}^{Q} \Pr(H_q) \bar{\alpha}_{i,0} C^{-1}(H_q, P_{th}, \pi\ (1)) \sum\limits_{d=1}^{d_c'} \dfrac{\omega_{i,d}}{d} = L, & \\ & i = 1, 2, \end{cases} \quad (54)$$

where $C1$ indicates that the sum of the edge distribution coefficients $\{\omega_{i,d}\}$ for each user is 1, $C2$ is the BP starting condition. $C3$ ensures successful decoding under all channel states. $C4$ is derived from (43) and (53), i.e., the average code length is fixed to L. To solve the above problem, firstly we can use differential evolution (DE) method [41] to obtain the corresponding optimization edge profile $\{\omega_{i,d}\}$ under an arbitrary fixed $\bar{\alpha}_{i,0}$ and $\pi\ (1)$. Then the optimal $\bar{\alpha}_{i,0}$ and $\pi\ (1)$ can be obtained by exhaustive searching. Finally, the optimal output node degree distributions $\Omega_{i,opt}\ (x)$ are obtained through (30).

C. COMPLEXITY ANALYSIS

The computational complexity of the proposed scheme should be discussed from two aspects. The first is the complexity for optimizing the degree profiles of the LDPC compression codes and Raptor codes, and the RRH selection. The second is the complexity to implement the proposed transmission scheme with the optimized code profiles.

For the first aspect, firstly consider the complexity for the optimization of the LDPC compression code, i.e., the problem (28). The complexity is $O(l_{LP} q Q)$, wherein $O\,(l_{LP})$ is the complexity of LP [42] (usually a polynomial of the variable number l_{LP}), q denotes the number of the discrete values of $\bar{\rho}$ for exhaustive searching, and Q is the number of

channel states (since (28) should be solved for each channel state). Secondly, consider the degree profile optimization for Raptor code, i.e., problem (53). The complexity is $O(2l_{DE}b)$, wherein $O(l_{DE})$ is the complexity of DE [43] (which is the product of the number of the variables, the generation number and the number of agents), b denotes the number of the discrete values of $\overline{\alpha}_{i,0}$ for exhaustive searching, and 2 is the number of possible values for π (1).

Now we discuss the complexity for the implementation of the proposed scheme with the optimized code profiles. Firstly, consider the fronthaul compression operation at the selected RRH. The complexity is $O(NBme)$, where $O(NBm)$ is the complexity for the modular multiplication of $\hat{u}_{\pi(1)}$ and the LDPC check matrix $\boldsymbol{A}_{m\times NB}$ and e denotes the number of the blocks of length N into which one codeword can be divided. Secondly, consider the complexity for the LDPC decompression algorithm and the iterative MU detection and decoding algorithm. Both are based on the BP algorithm. Its complexity has been widely discussed in the existing literature, e.g., [44], which is the product of the maximal iteration number and a polynomial of the node number and edge number in the decoding graph.

Finally, compared with [32], the additional complexity of this work comes from: (1) The optimization for LDPC compression code, which is executed only once. (2) The fronthaul compression executed at one of the RRHs, which is just modular multiplication. (4) The decompression at the BBU pool, which is actually a BP-based LDPC decoding algorithm.

D. EXTENSION TO THE SCENARIOS WITH MORE USERS AND RRHs

We discuss the extension of this work to the scenarios with more users and RRHs. Actually, the extension is non-trivial. In the following we discuss some related issues.

The first issue is the decompression order at the BBU pool. For the scenario with more than two RRHs, sequential decompression [15] should be applied, with which the signals from each RRH are decompressed by a certain order and the decompression for the current RRH utilizes the decompression results before, similar to (11). The decompression order affects the overall system performance, as well as the code profile optimization for the LDPC code and Raptor code. The decompression order should be optimized, and in [15], solely a sub-optimal heuristic order selection scheme was proposed.

The second issue is on the MU detector at the BBU pool. In this work, when deriving the output LLR for the MU detector, there are two possible combinations of quantization bit numbers for the signals from the two RRHs, i.e., $\{B, B\}$ and $\{B, B+1\}$, (recalling (14)). However, for the scenario with more RRHs, since each RRH randomly selects the signals for higher quantization precision, there will be considerably more combinations of quantization precisions. This fact also affects the EXIT analysis of the iterative MU detector and decoder, wherein the Raptor output nodes should be divided into more groups than that in Fig. 6.

The third issue is the EXIT analysis on the MU detector. According to [39], for the three-user multiple access channel, the output EI of MU detector can be approximated by some closed-form function, as (40) for the two-user case. Nevertheless, for the scenarios with more users and RRHs where the signals from the RRHs have non-uniform compression precisions, it is also an open question whether the polynomial approximation in [39] is accurate.

As for the computational complexity, it is obvious that the extension to the scenarios of more users and more RRHs will lead to higher complexity, due to the sequential decompression at the BBU pool and the fact that the code profile optimization for Raptor code and LDPC compression code becomes more complex and the possible decompression order is factorial of the number of RRHs.

V. SIMULATION RESULTS

In the simulation, we consider the scenario with two users and two RRHs. The users are not aware of the instant CSI while RRHs and BBU pool are assumed to have the full CSI. The message length for both users is set to $K = 9500$ bits. For the Raptor code, a regular LDPC code with rate $R_p = 0.95$ is applied as the precoder. The codebook of the regular LDPC code is generated by PEG algorithm [45]. The variance of Gaussian noise at each RRH is $\sigma_0^2 = 1.2^2$. The fronthaul capacity for both RRHs is set to 2 bits/symbol, which is a simplified system setting. In the simulation, we assume that there are three channel states with uniform probability. Each entry of the corresponding channel matrices are independently and randomly generated according to Rayleigh distribution with parameter 1: $\mathbf{H}_1 = \begin{bmatrix} 1.8672 & 0.4413 \\ 1.3802 & 0.6236 \end{bmatrix}$, $\mathbf{H}_2 = \begin{bmatrix} 0.8937 & 1.3841 \\ 0.6928 & 1.9839 \end{bmatrix}$, $\mathbf{H}_3 = \begin{bmatrix} 2.9021 & 0.4483 \\ 0.3402 & 1.2053 \end{bmatrix}$.

A. BER PERFORMANCE OF LDPC COMPRESSION CODE

Firstly, we simulate the performance of the LDPC compression code under the three channel states $\mathbf{H}_1$, $\mathbf{H}_2$ and $\mathbf{H}_3$, respectively. According to the optimization result of the problem (53), RRH 2 is selected to perform the proposed compression scheme. The transmit power of each user is $P = 1.616$. In Fig. 9, the overhead is defined as:

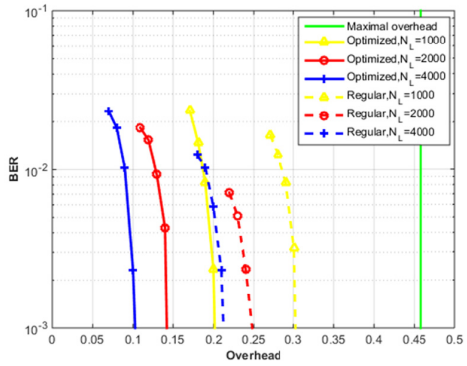
$$overhead = \frac{m}{\sum_{r=1}^{B} H_r N} - 1, \tag{55}$$

where H_r is given by (25) and $\sum_{r=1}^{B} H_r N$ is the theoretically required syndrome bit length for successful decompression. The optimized LDPC code variable node degree profiles for three channel states are listed here:

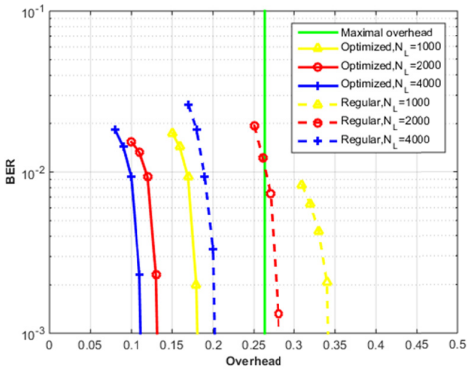
$$\xi_{\mathbf{H}_1}(x) = 0.4904x^2 + 0.3430x^3 + 0.1044x^6 + 0.0161x^7 + 0.0462x^{20}. \tag{56}$$

$$\xi_{\mathbf{H}_2}(x) = 0.5544x^2 + 0.3051x^3 + 0.1067x^6 + 0.0338x^{20}. \tag{57}$$

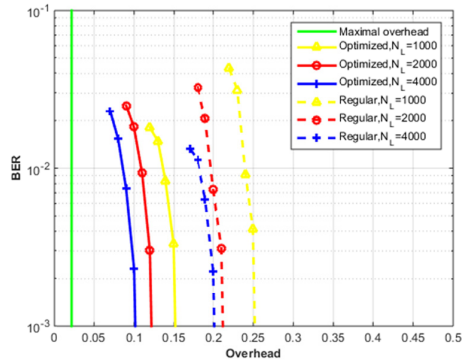
$$\xi_{\mathbf{H}_3}(x) = 0.8061x^2 + 0.1136x^3 + 0.0360x^4 + 0.0355x^5 + 0.0089x^{20}. \tag{58}$$



(a)



(b)



(c)

FIGURE 9. Compression performance of the optimized LDPC and regular LDPC under different channel states: (a) $\mathbf{H}_1$, (b) $\mathbf{H}_2$, (c) $\mathbf{H}_3$.

With the 2-bit quantizer, the length of the quantization bit sequence before LDPC compression for a block of N received signals is $N_L = 2N$. In Fig. 9, we simulate the error performance of the optimized LDPC compression code for different N_L under three channel states $\mathbf{H}_1$, $\mathbf{H}_2$ and $\mathbf{H}_3$. The regular LDPC code is also simulated for comparisons. As shown in the figure, the optimized LDPC compression code considerably outperforms the benchmark under all channel states. When $N_L = 4000$, the optimized LDPC code achieves the BER of 10^{-3} at a small overhead of 0.11. It can be observed that a better error performance is achieved for a larger code length N_L. However, note that in practice, a large N_L results in a large processing delay at the RRH since the RRH should wait for N received signals to do compression each time. This will also affect the overall system delay. Finally, it should be noted that, under each channel state, when the overhead is larger than 0.458, 0.263 and 0.022, respectively, the distributed fronthaul compression is not applicable since in these cases $m \geq NB$, i.e., the length of the compressed bit sequence exceeds the original one.

B. BER PERFORMANCE OF THE OVERALL SYSTEM

In this section, we simulate the error performance of the proposed scheme under the three channel states $\mathbf{H}_1$, $\mathbf{H}_2$ and $\mathbf{H}_3$, respectively. The transmit power of each user is $P = 1.616$. We set $N_L = 2000$, and choose m which makes the error performance is around 10^{-3} for the optimized LDPC compression code and the regular LDPC code, respectively. According to Remark 2 and Fig. 9, it can be found that under the channel state $\mathbf{H}_2$ and $\mathbf{H}_3$, the compression using regular LDPC can not work if we require BER to be 10^{-3}. Similarly, under channel state $\mathbf{H}_3$, the compression using the optimized LDPC can not be applied as well. Therefore, in the simulation, the proposed distributed fronthaul compression with the optimized LDPC code is applied under $\mathbf{H}_1$ and $\mathbf{H}_2$, while that with the regular LDPC code is only applied under $\mathbf{H}_1$.

We optimize the Raptor code degree profiles for both users considering the above situations with the method given in Section IV-C. The resulted Raptor code degree profiles are as follows:

$$\begin{aligned}\Omega_{1,opt}(x) &= 0.0117x^1 + 0.4229x^2 + 0.2648x^3 \\ &\quad + 0.0019x^5 + 0.2115x^6 + 0.0872x^{20}, \quad (59)\\ \Omega_{2,opt}(x) &= 0.0113x^1 + 0.4414x^2 + 0.2531x^3 \\ &\quad + 0.2162x^6 + 0.0780x^{20}. \quad (60)\end{aligned}$$

For comparisons, we consider the rateless coded transmission instead of fixed-rate channel coded transmission since the latter requires CSI at the transmitter and the transmission mechanism is quite different. For C-RAN with distributed fronthaul compression, the coded transmission design has not been investigated yet, except for the conference version of this paper. Therefore, we mainly consider the following reference schemes:

Case 1-A: RRH 1 applies the proposed distributed fronthaul compression with the optimized LDPC code. The optimal Raptor code degree profile for binary erasure channel (BEC) [39] is used at each user:

$$\begin{aligned}\Omega_{BEC}(x) &= 0.0008x + 0.494x^2 + 0.166x^3 \\ &\quad + 0.073x^4 + 0.083x^5 + 0.056x^8 \\ &\quad + 0.037x^9 + 0.056x^{19} + 0.025x^{65} \\ &\quad + 0.003x^{66}. \quad (61)\end{aligned}$$

Case 1-B: RRH 1 applies the proposed distributed fronthaul compression with the regular LDPC code. BEC profile in (60) is used by each users.

Case 2: Each RRH applies the conventional scalar quantization and no further compression is performed under all

channel states. BEC profile is used by each users. This case is also one of the benchmark schemes in [32].

Case 3: Each RRH applies the conventional scalar quantization and no further compression is performed. The Raptor code degree profiles are optimized over all channel states. This case is the proposed scheme in [32]. The optimal Raptor code degree profile is:

$$\Omega_{1,scalar}(x) = 0.0120x^1 + 0.4073x^2 + 0.2663x^3 + 0.2220x^6 + 0.0924x^{20}. \quad (62)$$

$$\Omega_{2,scalar}(x) = 0.0125x^1 + 0.3804x^2 + 0.2781x^3 + 0.0022x^5 + 0.2252x^6 + 0.1016x^{20}. \quad (63)$$

Fig. 10 shows the achieved sum BER of both users by the proposed scheme with the optimized LDPC compression code applied at RRH 2 and the optimized Raptor code for each user, under three channel states $\mathbf{H}_1$, $\mathbf{H}_2$ and $\mathbf{H}_3$. In the figure, 'Theoretical limit-1' and 'Theoretical limit-2' are the theoretically required length of Raptor code for successful recovery of both users' messages with and without the distributed fronthaul compression. 'Theoretical limit-1' is derived as $\frac{K}{C(\mathbf{H}_q,P,\pi(1))}$, where $C\left(\mathbf{H}_q,P,\pi(1)\right)$ is calculated from (50). 'Theoretical limit-2' is derived as $\frac{K}{C_{con}(\mathbf{H}_q,P)}$, where $C_{con}\left(\mathbf{H}_q,P\right)$ is calculated from (51).

In Fig. 10 (a) and (b), it can be observed that the proposed scheme with optimized LDPC compression code and Raptor code outperforms all the baselines. Especially, it can be found that the proposed scheme shows a considerable gain over case 2 and case 3 which deploy the conventional scalar quantizer. This is due to the fact that, the distributed fronthaul compression introduces lower quantization noise at the RRH compared with the conventional scalar quantizer. As discussed before, $\mathbf{H}_3$ is not suitable for distributed fronthaul compression even with the optimized LDPC code. Therefore, we only plot the error performance for the proposed scheme and Case 2 and 'Theoretical limit-2' in Fig. 10 (c). It can be found that the optimized Raptor code with no fronthaul compression still shows a considerable gain over case 2. The reason is that, problem (54) for Raptor code optimization has already taken into consideration the situation that the compression can not be applied under some channel states. Interestingly, it can be found that although the Raptor code profiles are optimized in the sense of the average throughput, a good error performance is achieved under all channel states. Finally, it is observed that Case 1-A outperforms Case 1-B both of which apply BEC profile. The reason lies in that the optimized LDPC compression code results in a higher compression rate, i.e., a smaller m, and a larger remaining fronthaul capacity, which enables more received signals to be quantized with a higher precision.

C. SYSTEM AVERAGE THROUGHPUT

We simulate the average system throughput under different SNR. In the simulation, we fix σ_0^2 and change the user transmission power P, and SNR is defined as P/σ_0^2.

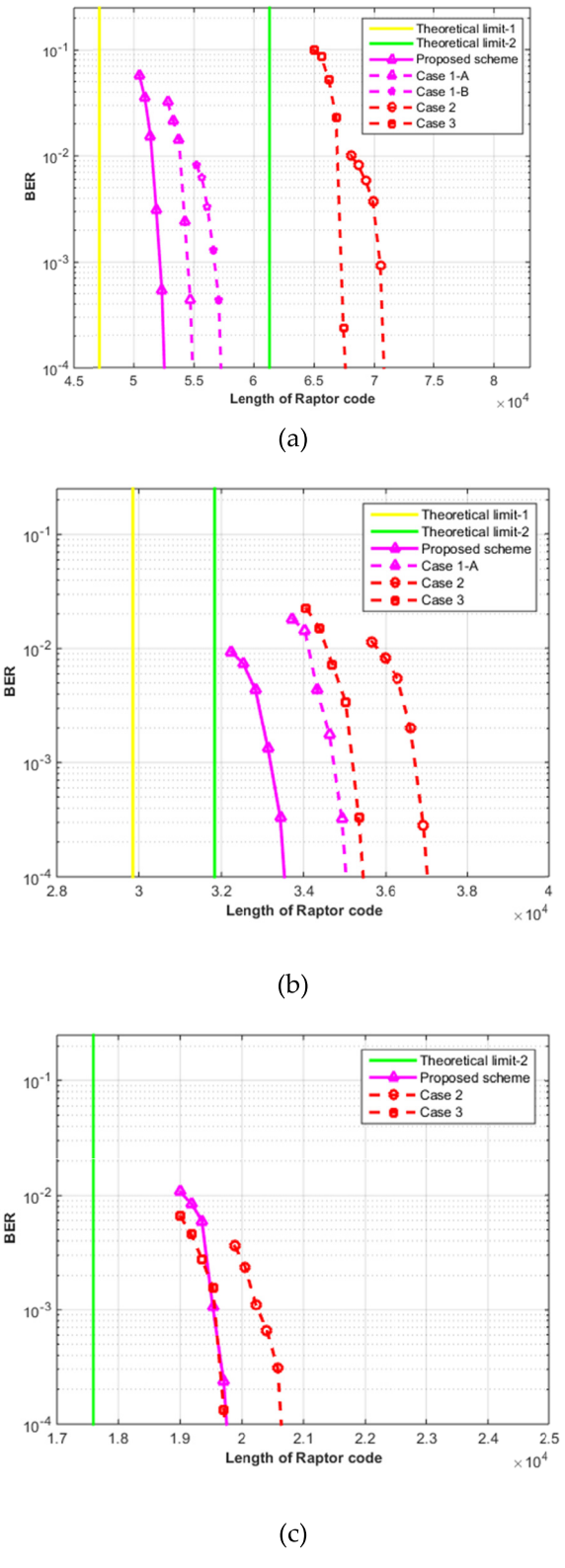


FIGURE 10. BER performance of the 2-user uplink C-RAN under different channel states: (a) $\mathbf{H}_1$, (b) $\mathbf{H}_2$, (c) $\mathbf{H}_3$.

In Fig. 11, the average system sum throughput over the three channel states achieved by the proposed scheme with the optimized LDPC compression code and Raptor code as

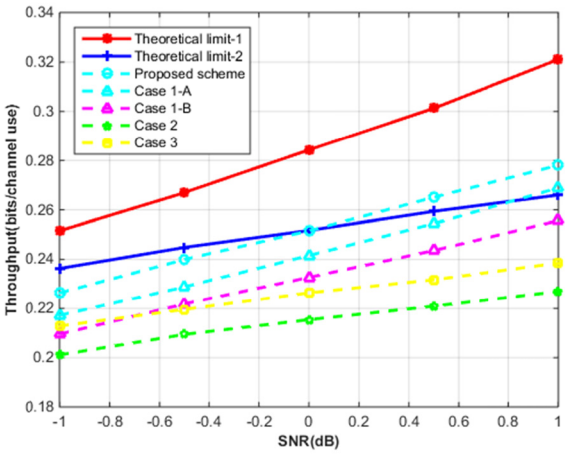


FIGURE 11. Average system sum throughput of the 2-user C-RAN uplink.

well as the other baseline schemes are simulated. The theoretically achievable sum rates with the proposed distributed fronthaul compression and conventional scalar quantizer are also plotted, respectively, which are derived as:

$$C_{lim,\pi(1)} = \frac{1}{\sum_{q=1}^{3} C^{-1}(\mathbf{H}_q, P, \pi(1))\Pr(H_q)}, \quad (64)$$

$$C_{lim,con} = \frac{1}{\sum_{q=1}^{3} C_{con}^{-1}(H_q, P)\Pr(H_q)}, \quad (65)$$

where $C\left(\mathbf{H}_q, P, \pi(1)\right)$ and $C_{con}\left(\mathbf{H}_q, P\right)$ are derived from (50) and (51), respectively.

It can be observed that the proposed scheme with the optimized code profiles achieves highest throughput over all the baselines and only has 12% loss from the theoretical limit, which validates the code optimization scheme discussed in Section IV. Again, the distributed fronthaul compression shows considerable performance gain over the conventional scaler quantization, from both the theoretical and simulation results. However, it is also shown that if the LDPC compression code or the Raptor channel code are not properly optimized, the gain will be reduced.

VI. CONCLUSION

In this work, we designed the Raptor coded transmission for two-user uplink in C-RAN with two RRHs under block fading channel, including the distributed fronthaul compression with LDPC code at RRHs and joint decompression and decoding at the BBU pool. We also proposed the code profile optimization schemes for the LDPC compression code and Raptor channel code. Simulation results verify that the proposed scheme with the optimized code profiles has a good error performance under all possible channel states and achieves considerable throughput gain compared with the conventional rateless coded transmission scheme.

Note that the current work can be extended to the scenarios with more users and RRHs. However, as discussed in Section IV-D, some non-trivial issues should be investigated in the design. We leave the extension as the future work.

ACKNOWLEDGMENT

This article was presented in pat at the 2020 Information Communication Technology Conference, Nanjing, China (ICTC 2020).

REFERENCES

[1] K. Wang, K. Yang, and C. S. Magurawalage, "Joint energy minimization and resource allocation in C-RAN with mobile cloud," *IEEE Trans. Cloud Comput.*, vol. 6, no. 3, pp. 760–770, Jul. 2018.

[2] 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.

[3] I. D. Technology, "Front-haul compression for emerging C-RAN and small cell networks," I. D. Technol., Los Angeles, CA, USA, Tech. Rep., Apr. 2013.

[4] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. IT-19, no. 4, pp. 471–480, Jul. 1973.

[5] J. Wolfowitz, "The rate-distortion function for source coding with side information at the decoder," *Probality Threory Rel. Field*, vol. 50, no. 3, pp. 245–255, 1979.

[6] A. D. Liveris, Z. Xiong, and C. N. Georghiades, "Distributed compression of binary sources using conventional parallel and serial concatenated convolutional codes," in *Proc. Data Compress. Conf. (DCC)*, Snowbird, UT, USA, 2003, pp. 193–202, doi: 10.1109/DCC.2003.1194010.

[7] A. Roumy, K. Lajnef, and C. Guillemot, "Rate-adaptive turbo-syndrome scheme for Slepian-Wolf Coding," in *Proc. Conf. Rec. 41st Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, 2007, pp. 545–549, doi: 10.1109/ACSSC.2007.4487272.

[8] V. Toto-Zarasoa, A. Roumy, C. Guillemot, and C. Herzet, "Robust and fast non asymmetric distributed source coding using turbo codes on the syndrome trellis," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Apr. 2009, pp. 2477–2480, doi: 10.1109/ICASSP.2009.4960124.

[9] S. Tu, "Asymmetric slepian-wolf coding design with spatially coupled LDPC convolutional codes," in *Proc. Int. Conf. Cyber-Enabled Distrib. Comput. Knowl. Discovery*, Xi'an, China, Sep. 2015, pp. 152–155, doi: 10.1109/CyberC.2015.59.

[10] S. Eleruja, U.-F. Abdu-Aguye, M. Ambroze, M. Tomlinson, and M. Zak, "Design of binary LDPC codes for slepian-wolf coding of correlated information sources," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Nov. 2017, pp. 1120–1124.

[11] Z. Mheich and E. Dupraz, "Short length non-binary rate-adaptive LDPC codes for slepian-wolf source coding," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Barcelona, Spain, Apr. 2018, pp. 1–5, doi: 10.1109/WCNC.2018.8377291.

[12] F. Ye, Z. Mheich, E. Dupraz, and K. Amis, "Optimized short-length rate-adaptive LDPC codes for slepian-wolf source coding," in *Proc. 25th Int. Conf. Telecommun. (ICT)*, St. Malo, France, Jun. 2018, pp. 351–355, doi: 10.1109/ICT.2018.8464851.

[13] Z. Xiong, A. D. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Process. Mag.*, vol. 21, no. 5, pp. 80–94, Sep. 2004.

[14] M. Fresia and L. Vandendorpe, "Distributed source coding using raptor codes," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Washington, DC, USA, Nov. 2007, pp. 1587–1591, doi: 10.1109/GLOCOM.2007.305.

[15] 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.

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

[17] M. Peng, Y. Sun, X. Li, Z. Mao, and C. Wang, "Recent advances in cloud radio access networks: System architectures, key techniques, and open issues," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 2282–2308, 3rd Quart., 2016, doi: 10.1109/COMST.2016.2548658.

[18] 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)*, Kaohsiung, Taiwan, Nov. 2017, pp. 171–175.

[19] Y. Zhou and W. Yu, "Optimized backhaul compression for uplink cloud radio access network," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1295–1307, Jun. 2014, doi: 10.1109/JSAC.2014.2328133.

[20] Y. Zhou, Y. Xu, W. Yu, and J. Chen, "On the optimal fronthaul compression and decoding strategies for uplink cloud radio access networks," *IEEE Trans. Inf. Theory*, vol. 62, no. 12, pp. 7402–7418, Dec. 2016.
[21] 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.
[22] A. Shokrollahi, "Raptor codes," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2551–2567, Jun. 2006.
[23] J. Castura and Y. Mao, "Rateless coding over fading channels," *IEEE Commun. Lett.*, vol. 10, no. 1, pp. 46–48, Jan. 2006.
[24] X. Chen, Z. Zhang, S. Chen, and C. Wang, "Adaptive mode selection for multiuser MIMO downlink employing rateless codes with QoS provisioning," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 790–799, Feb. 2012.
[25] X. Chen and C. Yuen, "Efficient resource allocation in a rateless-coded MU-MIMO cognitive radio network with QoS provisioning and limited feedback," *IEEE Trans. Veh. Technol.*, vol. 62, no. 1, pp. 395–399, Jan. 2013.
[26] Y. Zhang and Z. Zhang, "Joint network-channel coding with rateless code over multiple access relay system," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 320–332, Jan. 2013.
[27] Y. Zhang, Z. Zhang, R. Yin, G. Yu, and W. Wang, "Joint network-channel coding with rateless code in two-way relay systems," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3158–3169, Jul. 2013.
[28] M. Luby, M. Watson, T. Gasiba, T. Stockhammer, and W. Xu, "Raptor codes for reliable download delivery in wireless broadcast systems," in *Proc. 3rd IEEE Consum. Commun. Netw. Conf. (CCNC)*, Las Vegas, NV, USA, Jan. 2006, pp. 192–197.
[29] Y. Zhang, Y. Zhang, H. Peng, L. Xie, and L. Meng, "Rateless coded multi-user downlink transmission in cloud radio access network," *Mobile Netw. Appl.*, pp. 1–10, Dec. 2018, doi: 10.1007/s11036-018-1175-z.
[30] L. Xie, Y. Zhang, Y. Zhang, J. Hua, and L. Meng, "Uplink Transmission Scheme Based on Rateless Coding in Cloud-RAN," in *Proc. Int. Conf. Mach. Learn. Intell. Commun.*, Hangzhou, China, 2018, pp. 429–437.
[31] Y. Zhang, J. Xu, H. Peng, W. Lu, and J. Hua, "Rateless code profiles design for uplink C-RAN under block fading channel," in *Proc. 10th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Hangzhou, China, Oct. 2018, pp. 1–6.
[32] Y. Zhang, J. Xu, H. Peng, W. Lu, and Z. Zhang, "Rateless coded uplink transmission design for multi-user C-RAN," *Sensors*, vol. 19, no. 13, p. 2978, Jul. 2019.
[33] L. Schmalen, S. ten Brink, and A. Leven, "Analysis and optimization of iteration schedules for LDPC coded modulation and detection," in *Proc. 7th Int. Symp. Turbo Codes Iterative Inf. Process. (ISTC)*, Aug. 2012, pp. 210–214.
[34] F. Pukelsheim, "The three sigma rule," *Amer. Statistician*, vol. 48, no. 2, pp. 88–91, May 1994.
[35] Y. Zhang, "On the capacity bounds and rateless network coding design of cooperative relay networks," Ph.D. dissertation, Dept. Inf. Sci. Electron. Eng., Zhejiang Univ., Hangzhou, China, 2013.
[36] V. Auguste, P. Charly, and D. David, "Jointly decoded raptor codes: Analysis and design for the BIAWGN channel," *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, no. 1, Art. no. 657970.
[37] X.-Y. Hu, E. Eleftheriou, and D.-M. Arnold, "Progressive edge-growth tanner graphs," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, San Antonio, TX, USA, vol. 2, Nov. 2001, pp. 995–1001.
[38] O. Etesami and A. Shokrollahi, "Raptor codes on binary memoryless symmetric channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2033–2051, May 2006.
[39] C. Gong, G. Yue, and X. Wang, "Analysis and optimization of a rateless coded joint relay system," *IEEE Trans. Wireless Commun.*, vol. 9, no. 3, pp. 1175–1185, Mar. 2010.
[40] A. Venkiah, P. Piantanida, C. Poullia, P. Duhamel, and D. Declercq, "Rateless coding for quasi-static fading channels using channel estimation accuracy," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2008, pp. 2257–2261.
[41] R. Storn and K. Price, "Differential evolution: A simple and efficient heuristic for global optimization over continuous spaces," *J. Global Optim.*, vol. 11, no. 4, pp. 341–359, 1997.
[42] A. Amirzadeh, M. H. Taieb, and J.-Y. Chouinard, "On the design of good LDPC codes with joint genetic algorithm and linear programming optimization," in *Proc. 15th Can. Workshop Inf. Theory (CWIT)*, Quebec City, QC, Canada, Jun. 2017, pp. 1–5, doi: 10.1109/CWIT.2017.7994822.
[43] S. Das and P. N. Suganthan, "Differential evolution: A survey of the State-of-the-Art," *IEEE Trans. Evol. Comput.*, vol. 15, no. 1, pp. 4–31, Feb. 2011, doi: 10.1109/TEVC.2010.2059031.
[44] C. Yao, Z. Zhang, and K. Tu, "Design of complexity-optimized raptor codes for BI-AWGN channel," in *Proc. IEEE 24th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, London, U.K., Sep. 2013, pp. 824–829, doi: 10.1109/PIMRC.2013.6666250.
[45] T. Chu, X.-Q. Jiang, J. Hou, W. Hui-Ming, and L. Kong, "Construction of multiple-rate LDPC codes using modified PEG," in *Proc. 9th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Nanjing, China, Oct. 2017, pp. 1–5, doi: 10.1109/WCSP.2017.8171095.

YU ZHANG (Member, IEEE) received the B.S. degree in communication engineering and the Ph.D. degree in communication and information systems from Zhejiang University, Hangzhou, China, in 2008 and 2013, respectively. From 2013 to 2015, he was a Postdoctoral Researcher with the Department of Information Science and Electronic Engineering, Zhejiang University. From 2013 to 2014, he was a Visiting Scholar with the Department of Electronic Engineering, City University of Hong Kong, Hong Kong. Since 2015, he has been with the College of Information Engineering, Zhejiang University of Technology, Hangzhou, where he is currently an Assistant Professor. His current research interests include rateless coding, cloud RAN, intelligent reflecting surface, massive MIMO, and cooperative communication.

ZHEHAO FAN received the bachelor's degree in communication engineering from the Zhejiang University of Science and Technology, Hangzhou, China, in 2018. He is currently pursuing the master's degree with the Zhejiang University of Technology. His main research interests include rateless coding and wireless communication.

LIMIN MENG received the B.E., M.E., and Ph.D. degrees from Zhejiang University, Hangzhou, China, in 1984, 1998, and 2003, respectively. She is currently a Professor with the College of Information Engineering, Zhejiang University of Technology, Hangzhou, China, and the Director of the Zhejiang Provincial Key Laboratory of Communication Networks and Applications. Her current research interests include wireless communication and networks, multimedia communications, and network communications and managements.