

Adaptive Cloud Radio Access Networks: Compression and Optimization

Thang X. Vu, *Member, IEEE*, Hieu Duy Nguyen, *Member, IEEE*, Tony Q. S. Quek, *Senior Member, IEEE*,
and Sumei Sun, *Fellow, IEEE*

Abstract—Future mobile networks are facing with exponential data growth due to the proliferation of diverse mobile equipment and data-hungry applications. Among promising technology candidates to overcome this problem, cloud radio access network (C-RAN) has received much attention. In this paper, we investigate the design of fronthaul in C-RAN uplink by focusing on the compression and optimization in fronthaul uplinks based on the statistics of wireless fading channels. First, we derive the system block error rate (BLER) under Rayleigh fading channels. In particular, upper and lower bounds of the BLER union bound are obtained in closed-form. From these bounds, we gain insight in terms of diversity order and limits of the BLER. Next, we propose adaptive compression schemes to minimize the fronthaul transmission rate subject to a BLER constraint. Furthermore, a fronthaul rate allocation is proposed to minimize the system BLER. It is shown that the uniform rate allocation approaches the optimal scheme as the total fronthauls' bandwidth increases. Finally, numerical results are presented to demonstrate the effectiveness of our proposed optimizations.

Index Terms—Cloud radio access networks, adaptive compression, optimization, Rayleigh fading.

I. INTRODUCTION

AMONG promising technology candidates for future mobile networks, cloud radio access network (C-RAN) has received much attention [1]. In C-RAN, one centralized processor or baseband unit (BBU) communicates with users via distributed remote radio heads (RRHs). The RRHs are connected to the BBU via high capacity, low latency fronthaul links and have minimal functioning since most baseband processing tasks are centralized at the BBU. In this way, C-RAN enables adaptive load balancing via virtual base station pool [2]

Manuscript received January 14, 2016; revised August 3, 2016 and September 20, 2016; accepted October 7, 2016. Date of publication October 13, 2016; date of current version November 4, 2016. The associate editor coordinating the review of this paper and approving it for publication was Dr. Walaa Hamouda. This work was supported in part by the Agency for Science Technology and Research, SERC under Grant 1224104048 and in part by the MOE ARF Tier 2 under Grant MOE2014-T2-2-002. This paper was presented in part at the IEEE International Conference on Communications, Kuala Lumpur, Malaysia, May 2016.

T. X. Vu is with the Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg, Luxembourg L-1359, Luxembourg (e-mail: thang.vu@uni.lu).

H. D. Nguyen is with the Swiss Re, Singapore 018961 (e-mail: ngduyhieu@gmail.com).

T. Q. S. Quek is with the Singapore University of Technology and Design, Singapore 487382 (e-mail: tonyquek@sutd.edu.sg).

S. Sun is with the Institute of Informatics Research, Agency for Science Technology and Research, Singapore 138632 (e-mail: sunsm@i2r.a-star.edu.sg).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TSP.2016.2617826

and effective inter-cell interference management via multi-cell processing [3]–[6]. However, since the baseband processing is executed at the BBU, it requires enormous transmission rate on the fronthaul links to transfer in-phase/quadrature-phase (I/Q) samples, which represent the radio signals obtained through the sampling of complex baseband signals. Reducing this rate is extremely important in the implementation of C-RAN since the fronthaul links' capacity is generally limited by cost and deployment constraints.

As a result, there are recent interests in tackling the compression issues in C-RAN [7]–[13]. The compression process is implemented via a test channel and the quantization noise is modelled as an independent Gaussian random variable with variance linked to the test channel capacity. It is shown, in general, that the joint design of precoding and quantization noise matrices can significantly improve the system sum rate over separate design [8]. Such improvement is achieved due to the correlation among the RRHs when distributed source coding is applied [3]. The quality of the received signal at one RRH can be enhanced by exploiting other RRHs' received signal as the side information. In [9], a hybrid compression and message-sharing strategy is proposed for downlinks. It is shown that the hybrid solution achieves a better rate region than the pure method of compression or data-sharing. In [11], an optimum compression scheme is derived for sensor networks. From the practical system point of view, various compression techniques have been studied in both time- and frequency- domains (sub-carrier compression) [1], which exploit the structure of common public radio interface (CPRI) package to minimize redundancy within control signal [14]. Lossless compression is proposed to achieve a good compression ratio by adding two nodes at the fronthaul ends to optimize the redundancy in both time and frequency domains [15], [16]. Statistical multiplexing gain is achieved since: i) only information data of active users are transmitted via the fronthaul links, ii) minimum information needed for the control signal reconstruction is locally generated, and iii) a reduced set of the precoding matrix is transferred. A similar time-domain compression technique is proposed in [17]–[20]. However, a limitation of all the above works is the assumption that the instantaneous channel state information (CSI) is required for the design and optimization. This will impose a significant overhead and delay for large-scale C-RAN deployment.

In this paper, we study compression and the corresponding performance of C-RAN uplinks in Rayleigh fading channels. Our first goal is to derive the performance metrics for C-RAN systems based only on the statistics of the channels. In particular, block error rate (BLER) under Rayleigh fading channels

is analysed via union bound analysis, through which some insights of C-RAN are drawn: i) full diversity order of N (number of RRHs) is achieved with respect to signal-to-compression-plus-noise ratio; and ii) the BLER is limited by either compression or Gaussian noise. This result is different from our previous work [21], [22] which derives the system BLER for a specific channel realization. Given our derived BLER expressions, we aim at minimizing the transmission rate on the fronthaul links given an acceptable signal distortion so that the BBU can support a maximum number of RRHs. Our objective comes from practical demands where most applications can tolerate a non-zero BLER. Specifically, we propose two adaptive compression schemes to maximize compression efficiency while satisfying a predefined BLER. This design criterion is different from that in [3], which aims to fully utilize the fronthaul link capacity. More importantly, our proposed framework only requires the statistics of the channels, which significantly reduces the overhead in terms of the compression optimization compared to previous works that required instantaneous CSI [3], [21]. Such overhead reduction becomes more important in C-RAN systems which are designed to support a large number of users. Furthermore, a fronthaul rate allocation is proposed to minimize the system BLER. Numerical results are presented to demonstrate the effectiveness of our proposed schemes.

Notations: $[x_1, \dots, x_n]$ denotes a row vector with elements x_1, \dots, x_n . $(\cdot)^T$, $(\cdot)^H$, $\text{erfc}(\cdot)$, and $\mathbb{E}\{\cdot\}$ denote the transpose operation, conjugate transpose, complementary error function, and expectation operator, respectively. $\text{diag}(\mathbf{x})$ denotes a square matrix whose diagonal is vector \mathbf{x} and all-zero elsewhere. $\mathcal{R}(x)$ and $\mathcal{I}(x)$ represent the real and imaginary parts of x , respectively. $\lceil x \rceil$ denotes the smallest integer that is larger than or equal to x , and $\lfloor x \rfloor$ denotes the closest integer to x .

The rest of the paper is organized as follows. Section II describes the system model and the receiver's structure. Section III investigates the instantaneous pair-wise error probability (PEP). Section IV analyses the BLER performance under Rayleigh fading channels. In Section V, we propose two optimization schemes to minimize the fronthaul's transmission rate based on either a PEP lower- or upper-bound. The minimization of the BLER is presented in Section VI. Section VII presents numerical results. Finally, Section VIII concludes our paper with some remarks.

II. SYSTEM MODEL

The C-RAN under consideration consists of M users $m \in \{1, 2, \dots, M\}$, N RRHs $n \in \{1, 2, \dots, N\}$, and one BBU, as shown in Figure 1. The users communicate with the RRHs via wireless medium, while the RRHs connect to the BBU via optical fibre or wireless fronthaul links. A distinguished feature of the RRH is that its function is much simpler than that of a traditional base station since all baseband processing functions are executed at the BBU. Therefore, a RRH can be seen as a "soft" relaying node that forwards I/Q signals to the BBU. Each user and RRH is assumed to equip with a single antenna. In practical systems, a multiple-antenna RRH can be seen as a

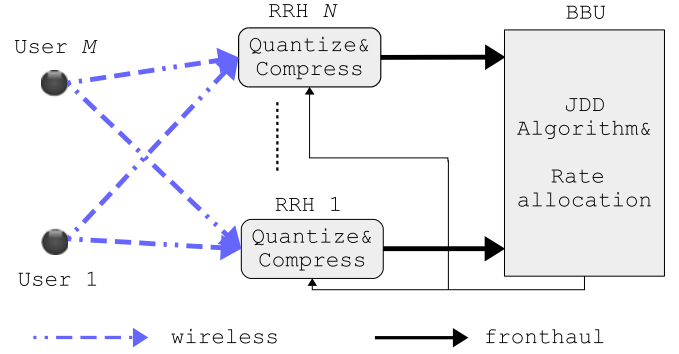


Fig. 1. Block diagram of uplinks in C-RAN. The proposed compression scheme optimizes sampling rate needed and then feedbacks them to the RRHs.

band of single-antenna RRHs which are subject to a sum rate constraint. Due to limited capacity on the fronthaul links, I/Q signals need to be compressed before being sent to the BBU [3]. The BBU decompresses the signals received from the RRHs and then performs further processing.

We assume that all nodes are time-synchronous and all wireless channels are subject to block Rayleigh fading. Denote c_m as a modulated symbol emitted by user m . The modulated symbol c_m , $1 \leq m \leq M$, belongs to the source codebook $\mathcal{S} = \{s_1, \dots, s_{|\mathcal{S}|}\}$, which has average unit power, e.g., $\mathbb{E}_{s \in \mathcal{S}} |s|^2 = 1$, where $|\mathcal{S}|$ denotes the cardinality of set \mathcal{S} . The symbols transmitted by the sources are aggregated into a codeword $\mathbf{c} = [c_1, \dots, c_M]^T$. The received signal at RRH n is given by

$$y_n = \sum_{m=1}^M h_{nm} \sqrt{P_{nm}} c_m + z_n = \mathbf{h}_n \mathbf{\Lambda}_n \mathbf{c} + z_n, \quad (1)$$

where $\mathbf{\Lambda}_n = \text{diag}([\sqrt{P_{n1}}, \dots, \sqrt{P_{nM}}])$, P_{nm} is the average received energy at RRH n from user m , including the path loss, h_{nm} is the channel fading coefficient between user m and RRH n , which is a complex Gaussian random variable with zero mean and unit variance, $\mathbf{h}_n = [h_{n1}, \dots, h_{nM}]$ is the channel vector from all users to RRH n , and z_n is independent identically distributed (i.i.d.) Gaussian noise with zero mean and variance σ^2 .

A. Uniform Compression Scheme

Upon receiving analogue signals from the users, each RRH quantizes and compresses them into digital bits and then forward these bits to the BBU. In this work, we consider uniform quantization because of its low-complexity and practical implementation [24]. This compression method can be realized by flexibly tuning the analogue-to-digital converter (ADC) resolution. Therefore, a target compression ratio can be achieved by changing the ADC's accuracy. In case of fixed ADC's resolution due to some hardware constraints, this compression method can be performed by truncating some least significant bits in the ADC's output. The compression is executed on the real and imaginary parts separately [1]. The received signal at RRH n is first normalized as $\bar{y}_n = y_n / \eta_n = \mathcal{R}(\bar{y}_n) + i\mathcal{I}(\bar{y}_n)$, where η_n is a scaling factor that restricts $\mathcal{R}(\bar{y}_n)$ and $\mathcal{I}(\bar{y}_n)$ within $[-1, 1]$ with high probability. The value of η_n can be calculated for a given

codebook \mathcal{S} and network topology. In this work, we employ the three-sigma method such that $\eta_n = 3\sqrt{(\|\mathbf{A}_n\|^2 + \sigma^2)/2}$, which is independent from the instantaneous fading channels and assumed to be known at the RRH n because its overhead is negligible compared to the data.

In the next step, the normalized signal \tilde{y}_n is quantized into \tilde{y}_n by a Q_n -bit uniform quantizer. The compressed signal can be calculated from the normalized signal as

$$\mathcal{R}(\tilde{y}_n) = \eta_n \frac{\lfloor \mathcal{R}(\tilde{y}_n) \times 2^{Q_n} \rfloor}{2^{Q_n}}, \mathcal{I}(\tilde{y}_n) = \eta_n \frac{\lfloor \mathcal{I}(\tilde{y}_n) \times 2^{Q_n} \rfloor}{2^{Q_n}}.$$

The quantization error at RRH n is given as $q_n = y_n - \tilde{y}_n$. When the absolute value of y_n is large compared to quantization step, $\mathcal{R}(q_n)$ and $\mathcal{I}(q_n)$ can be well modelled as uniform random variables with the support $[-\delta_n, \delta_n]$, where $\delta_n = \eta_n 2^{-Q_n-1}$. After compression, \tilde{y}_n is converted into a bit sequence which is later sent to the BBU via error-free fronthaul links.

B. Decoding at the BBU

The BBU employs a joint decompressing and detecting (JDD) algorithm [21]. The JDD exploits the structure of the quantizer and the codebook to perform decompression and detection for the source codeword simultaneously. The BBU is assumed to know the CSI of all wireless links. In practical systems, the CSI can be obtained via pilot-assisted training. Given the compressed signals, the BBU optimally estimates the source codeword by using the maximum a posteriori (MAP) receiver as follows:

$$\hat{\mathbf{c}} = \arg \max_{\mathbf{c}} \Pr\{\mathbf{c}\} \prod_{n=1}^N \Pr\{\tilde{y}_n|\mathbf{c}\}, \quad (2)$$

where (2) results from the fact that $\Pr\{\tilde{y}_1, \dots, \tilde{y}_N\}$ is constant for any codeword, and the noise z_n 's and compressed signals are independent given the source codeword.

In (2), $\Pr\{\tilde{y}_n|\mathbf{c}\}$ is the probability that the quantizer outputs \tilde{y}_n from the observation y_n . It is worth mentioning that for a real signal, the linear quantizer outputs y if the distance between the input and y is less than or equal to the quantization error. For complex signal y_n , the quantizer outputs \tilde{y}_n if both $|\mathcal{R}(y_n) - \mathcal{R}(\tilde{y}_n)|$ and $|\mathcal{I}(y_n) - \mathcal{I}(\tilde{y}_n)|$ are less than the quantization error. Because the quantization is performed independently for the real and the imaginary parts, we have

$$\begin{aligned} \Pr\{\tilde{y}_n|\mathbf{c}\} &= \Pr\{\mathcal{R}(y_n) \in \Omega_1 \cap \mathcal{I}(y_n) \in \Omega_2\} \\ &= \Pr\{\mathcal{R}(y_n) \in \Omega_1\} \times \Pr\{\mathcal{I}(y_n) \in \Omega_2\}, \end{aligned}$$

where $\Omega_1 = [\mathcal{R}(\tilde{y}_n) - \delta_n, \mathcal{R}(\tilde{y}_n) + \delta_n]$ and $\Omega_2 = [\mathcal{I}(\tilde{y}_n) - \delta_n, \mathcal{I}(\tilde{y}_n) + \delta_n]$.

To derive the above probability, we remind that for the given codeword and the fading channels, $\mathcal{R}(y_n)$ and $\mathcal{I}(y_n)$ are Gaussian distributed with the same variance $\sigma^2/2$ and means $\mathcal{R}(\mathbf{h}_n \mathbf{A}_n \mathbf{c})$ and $\mathcal{I}(\mathbf{h}_n \mathbf{A}_n \mathbf{c})$, respectively. Therefore, the conditional probability density function (PDF) of $\mathcal{R}(y_n)$ and $\mathcal{I}(y_n)$ are, respectively, given by $f_1(x|\mathbf{c}) = \frac{1}{\sqrt{\pi}\sigma} \exp$

$$\left(-\frac{|x - \mathcal{R}(\mathbf{h}_n \mathbf{A}_n \mathbf{c})|^2}{\sigma^2}\right), \text{ and } f_2(x|\mathbf{c}) = \frac{1}{\sqrt{\pi}\sigma} \exp\left(-\frac{|x - \mathcal{I}(\mathbf{h}_n \mathbf{A}_n \mathbf{c})|^2}{\sigma^2}\right).$$

By substituting these PDFs into $\Pr\{\tilde{y}_n|\mathbf{c}\}$ we obtain

$$\begin{aligned} \Pr\{\tilde{y}_n|\mathbf{c}\} &= \int_{\Omega_1} f_1(u|\mathbf{c}) du \int_{\Omega_2} f_2(v|\mathbf{c}) dv = \frac{1}{4} \left(\operatorname{erfc}\left(\frac{\varphi_1 - \delta_n}{\sigma}\right) \right. \\ &\quad \left. - \operatorname{erfc}\left(\frac{\varphi_1 + \delta_n}{\sigma}\right) \right) \left(\operatorname{erfc}\left(\frac{\varphi_2 - \delta_n}{\sigma}\right) - \operatorname{erfc}\left(\frac{\varphi_2 + \delta_n}{\sigma}\right) \right), \quad (3) \end{aligned}$$

where $\varphi_1 = \mathcal{R}(\tilde{y}_n) - \mathcal{R}(\mathbf{h}_n \mathbf{A}_n \mathbf{c})$, $\varphi_2 = \mathcal{I}(\tilde{y}_n) - \mathcal{I}(\mathbf{h}_n \mathbf{A}_n \mathbf{c})$. The derivation of $\Pr\{\tilde{y}_n|\mathbf{c}\}$ in (3) is exact. However, under high SNR regime and due to the fading effects, the argument of function $\operatorname{erfc}(\cdot)$ in (3) can be very large, resulting in over buffer and erroneous decoding. In order to avoid such problem, an approximation using first-order Taylor's series can be used instead: $f(x) \simeq f(x_0) + f'(x_0)(x - x_0)$, where x_0 is any feasible point. Applying to the function $\operatorname{erfc}(\cdot)$ in (3) with $x_0 = (\mathcal{R}(\tilde{y}_n) - \mathcal{R}(\mathbf{h}_n \mathbf{A}_n \mathbf{c}))/\sigma$ for the real part and $x_0 = (\mathcal{I}(\tilde{y}_n) - \mathcal{I}(\mathbf{h}_n \mathbf{A}_n \mathbf{c}))/\sigma$ for the imaginary part, the probability $\Pr\{\tilde{y}_n|\mathbf{c}\}$ can be written in a simplified form as follows:

$$\Pr\{\tilde{y}_n|\mathbf{c}\} \simeq \frac{\delta_n}{\sqrt{\pi}\sigma} \exp\left(-\frac{|\tilde{y}_n - \mathbf{h}_n \mathbf{A}_n \mathbf{c}|^2}{\sigma^2}\right). \quad (4)$$

Substituting (4) into (2), we obtain a decoding rule for codeword $\hat{\mathbf{c}}$.

III. PAIR-WISE ERROR PROBABILITY ANALYSIS

This section analyses the PEP performance of the receiver derived in the previous section using (4). The PEP will be used for further analysis in the following sections. Let $\text{PEP}_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}$ denote the probability of receiving codeword $\tilde{\mathbf{c}}$ when \mathbf{c} was transmitted, and $\tilde{\mathbf{c}}$ being the only candidate. Let $\hat{\mathbf{c}}$ be the detected codeword, we have

$$\text{PEP}_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}} = \Pr\{\hat{\mathbf{c}} = \tilde{\mathbf{c}}|\mathbf{c}\}. \quad (5)$$

To evaluate the PEP, we model the quantization effect by an uniformly distributed random variable that is independent of the input. This assumption can be well justified when the absolute value of the input is much larger than the quantization step. Under such assumption, the compressed signal from the RRH n is modelled as

$$\tilde{y}_n = \mathbf{h}_n \mathbf{A}_n \mathbf{c} + z_n + q_n, \quad (6)$$

where $q_n = q_n^R + iq_n^I$ is the quantization noise at the RRH n ($q_n^R \triangleq \mathcal{R}(q_n)$, $q_n^I \triangleq \mathcal{I}(q_n)$). Since both q_n^R and q_n^I are uniformly distributed in $[-\delta_n, \delta_n]$ (see Section II for more details), it is straightforward to verify that q_n has zero mean and variance $\sigma_{q_n}^2 = \frac{2\delta_n^2}{3}$.

Denote $\mathbb{M}(\mathbf{c}) = \prod_{n=1}^N \Pr\{\tilde{y}_n|\mathbf{c}\}$ as the detection metric for codeword \mathbf{c} , where $\Pr\{\tilde{y}_n|\mathbf{c}\}$ is computed in (4). A pair-wise error occurs if the metric of the transmitted codeword is smaller than that of another candidate:

$$\text{PEP}_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}} = \Pr\{\mathbb{M}(\mathbf{c}) < \mathbb{M}(\tilde{\mathbf{c}})\}. \quad (7)$$

Substituting (4) into $\mathbb{M}(\mathbf{c})$ we obtain $\mathbb{M}(\mathbf{c}) = K \exp(-\mathbb{D}(\mathbf{c}))$, where $K = \prod_{n=1}^N \delta_n^2 / (\pi\sigma^2)^N$ is a constant and $\mathbb{D}(\mathbf{c}) = \sum_{n=1}^N |\tilde{y}_n - \mathbf{h}_n \mathbf{A}_n \mathbf{c}|^2$.

¹We observe via intensive simulations that with the three-sigma rule, such assumption is still feasible even with a small number of quantization bits.

After some algebraic manipulations and by using (6), the PEP is expressed as follows:

$$\text{PEP}_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}} = \Pr\{Z_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}^1 + Z_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}^2 - \psi > 0\}, \quad (8)$$

where $Z_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}^1 \triangleq \sum_{n=1}^N [z_n^H \mathbf{h}_n \mathbf{\Lambda}_n (\tilde{\mathbf{c}} - \mathbf{c}) + (\tilde{\mathbf{c}} - \mathbf{c})^H \mathbf{\Lambda}_n \mathbf{h}_n^H z_n]$, $Z_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}^2 \triangleq \sum_{n=1}^N [q_n^H \mathbf{h}_n \mathbf{\Lambda}_n (\tilde{\mathbf{c}} - \mathbf{c}) + (\tilde{\mathbf{c}} - \mathbf{c})^H \mathbf{\Lambda}_n \mathbf{h}_n^H q_n]$, and $\psi \triangleq \sum_{n=1}^N |\mathbf{h}_n \mathbf{\Lambda}_n (\tilde{\mathbf{c}} - \mathbf{c})|^2$.

Because each z_n is a complex Gaussian random variable with zero mean and variance σ^2 , and z_n 's are mutually independent, $Z_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}^1$ is also a Gaussian random variable with zero mean and variance

$$\sigma_{Z_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}^1}^2 = 2\sigma^2 \sum_{n=1}^N |\mathbf{h}_n \mathbf{\Lambda}_n (\tilde{\mathbf{c}} - \mathbf{c})|^2.$$

On the other hand, the exact computation of the joint PDF of $Z_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}^2$ is challenging as q_n is uniformly distributed. For ease of analysis and supported by the central limit theorem, we model $Z_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}^2$ by a Gaussian variable \bar{Z}_2 having similar mean and variance as those of $Z_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}^2$, i.e., $\bar{Z}_2 \sim \mathcal{N}(\mu_{Z_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}^2}, \sigma_{Z_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}^2}^2)$, where $\mu_{\bar{Z}_2} = \mathbb{E}\{Z_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}^2\} = 0$ and

$$\sigma_{\bar{Z}_2}^2 = \mathbb{E}\{|Z_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}^2|^2\} = \frac{4}{3} \sum_{n=1}^N \delta_n^2 |\mathbf{h}_n \mathbf{\Lambda}_n (\tilde{\mathbf{c}} - \mathbf{c})|^2.$$

Then the sum $Z = Z_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}^1 + Z_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}^2$ is also a Gaussian random variable with zero mean and variance $\sigma_Z^2 = \sigma_{Z_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}^1}^2 + \sigma_{\bar{Z}_2}^2$. Therefore we can compute the PEP as follows:

$$\text{PEP}_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}} = \Pr\{Z > \psi\} = \frac{1}{2} \times \text{erfc} \left(\frac{\sum_{n=1}^N |\mathbf{h}_n \mathbf{\Lambda}_n (\tilde{\mathbf{c}} - \mathbf{c})|^2}{\sqrt{4\sigma^2 \sum_{n=1}^N |\mathbf{h}_n \mathbf{\Lambda}_n (\tilde{\mathbf{c}} - \mathbf{c})|^2 + \frac{8}{3} \sum_{n=1}^N \delta_n^2 |\mathbf{h}_n \mathbf{\Lambda}_n (\tilde{\mathbf{c}} - \mathbf{c})|^2}} \right). \quad (9)$$

It is observed from (9) that the PEP depends on the relative distance between \mathbf{c} and $\tilde{\mathbf{c}}$ distorted by the fading channels, thermal noise power σ^2 , and compression noise power δ_n^2 .

IV. PERFORMANCE ANALYSIS OVER RAYLEIGH FADING CHANNELS

This section analyses the BLER of C-RAN uplinks in Rayleigh fading channels. The BLER is defined as the probability of receiving codeword $\hat{\mathbf{c}}$ when a codeword $\mathbf{c} \neq \hat{\mathbf{c}}$ was transmitted. Note that a block error event occurs when at least one out of M symbols $\{c_m\}_{m=1}^M$ is decoded with error. Since the exact BLER is difficult to investigate (even for single-user, point-to-point pulse-amplitude modulation), we instead resort to the union bound on the BLER as follows:

$$\text{BLER} \leq \frac{1}{|\mathcal{S}|^M} \sum_{\mathbf{c}, \hat{\mathbf{c}} \in \mathcal{S}^M, \mathbf{c} \neq \hat{\mathbf{c}}} \overline{\text{PEP}}_{\mathbf{c} \rightarrow \hat{\mathbf{c}}}, \quad (10)$$

where $\overline{\text{PEP}}_{\mathbf{c} \rightarrow \hat{\mathbf{c}}} \triangleq \mathbb{E}\{\text{PEP}_{\mathbf{c} \rightarrow \hat{\mathbf{c}}}\}$ denotes the expectation over the channel fading coefficients of $\text{PEP}_{\mathbf{c} \rightarrow \hat{\mathbf{c}}}$, which is given in (9).

Before deriving $\overline{\text{PEP}}_{\mathbf{c} \rightarrow \hat{\mathbf{c}}}$, we observe that $\{\mathbf{h}_n \mathbf{\Lambda}_n (\tilde{\mathbf{c}} - \mathbf{c})\}_{n=1}^N$ are i.i.d. random variables with $\mathcal{CN}(0, \|\mathbf{\Lambda}_n (\tilde{\mathbf{c}} - \mathbf{c})\|^2)$ distribution. Define $G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}} t_n \triangleq |\mathbf{h}_n \mathbf{\Lambda}_n (\tilde{\mathbf{c}} - \mathbf{c})|^2$, where $G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}} \triangleq \|\mathbf{\Lambda}_n (\tilde{\mathbf{c}} - \mathbf{c})\|^2$. Thus, $\{t_n\}_{n=1}^N$ are i.i.d. exponential random variables each with distribution e^{-t} , $0 < t < \infty$. We express the argument inside function $\text{erfc}()$ in (9) as

$$\frac{\sum_{n=1}^N G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}} t_n}{\sqrt{\sum_{n=1}^N (4\sigma^2 + \frac{8}{3} \delta_n^2) G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}} t_n}} = \frac{\sum_{n=1}^N G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}} t_n}{\sqrt{\sum_{n=1}^N \beta_n G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}} t_n}},$$

where $\beta_n \triangleq 4\sigma^2 + \frac{8}{3} \delta_n^2$.

We will consider two cases: identical quantization noises, i.e., $\delta_1 = \delta_2 = \dots = \delta_N$, and non-identical quantization noises.

A. Identical Quantization Noises

In this subsection, we assume identical quantization noise power δ_n^2 at every RRH. Such scenario often occurs in symmetric networks with homogeneous fronthauls' capacity. For ease of notation, denote $\delta_n = \delta, \forall n$, and thus $\beta_n = 4\sigma^2 + \frac{8}{3} \delta_n^2 = \beta, \forall n$.

The average PEP in this case is computed as follows:

$$\begin{aligned} \overline{\text{PEP}}_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}} &= \frac{1}{2} \int \dots \int_{\{t_n\}_{n=1}^N: 0 < t_n < \infty} \text{erfc} \\ &\times \left(\frac{1}{\sqrt{\beta}} \sqrt{\sum_{n=1}^N G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}} t_n} \right) \times \exp \left(- \sum_{n=1}^N t_n \right) dt_1 \dots dt_N. \end{aligned} \quad (11)$$

To facilitate the computation of (11), we use a tight approximation of the error complementary function as [25]

$$\text{erfc}(x) \approx \frac{1}{6} e^{-x^2} + \frac{1}{2} e^{-4x^2/3}. \quad (12)$$

Applying (12) to (11), we obtain

$$\begin{aligned} \overline{\text{PEP}}_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}} &\simeq \frac{1}{12} \int_0^\infty \dots \int_0^\infty \exp \left(\sum_{n=1}^N \left(\frac{G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}}{\beta} + 1 \right) t_n \right) dt_1 \dots dt_N \\ &+ \frac{1}{4} \int_0^\infty \dots \int_0^\infty \exp \left(\sum_{n=1}^N \left(\frac{4G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}}{3\beta} + 1 \right) t_n \right) dt_1 \dots dt_N \\ &= \frac{1}{12} \frac{1}{\prod_{n=1}^N \left(\frac{G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}}{\beta} + 1 \right)} + \frac{1}{4} \frac{1}{\prod_{n=1}^N \left(\frac{4G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}}{3\beta} + 1 \right)}. \end{aligned} \quad (13)$$

It is observed from (13) that the average PEP depends on $G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}$, the distance between two codewords scaled by the slow fading metric, and the aggregated Gaussian and compression noises β . Substituting (13) into (10) we obtain the union bound for BLER.

B. Non-Identical Quantization Noises

In this general case, β_n at the RRHs can be different. By changing variable t_n to $v_n = \beta_n G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}} t_n$, the average PEP over

the fading channels is evaluated as follows:

$$\overline{\text{PEP}}_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}} = \frac{1}{2 \prod_{n=1}^N \beta_n G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}} \int_0^\infty \dots \int_0^\infty \text{erfc} \left(\frac{\sum_{n=1}^N \frac{v_n}{\beta_n}}{\sqrt{\sum_{n=1}^N v_n}} \right) \times \exp \left(- \sum_{n=1}^N \frac{v_n}{\beta_n G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}} \right) dv_1 \dots dv_N. \quad (14)$$

The exact computation of (14) over the set $\{v_n\}_{n=1}^N$ is challenging for arbitrary set $\{\beta_n\}_{n=1}^N$. Therefore, we investigate lower bound and upper bound of (14).

1) *Lower Bound:* Let \mathcal{P}_{ind} denote the set of all permutations of $\{1, 2, \dots, N\}$, i.e.,

$$\mathcal{P}_{ind} = \left\{ \{i_1, i_2, \dots, i_N\} : i_m \neq i_q, \forall m \neq q; \{i_1, i_2, \dots, i_N\} \equiv \{1, 2, \dots, N\} \right\}. \quad (15)$$

Furthermore, the function $f(x) \triangleq \text{erfc}(\alpha x) \exp(-x)$ is convex in $(0, +\infty)$ for $\alpha > 0$ (see Appendix A). Therefore, consider any given value set $\{v_n\}_{n=1}^N$, we have

$$\begin{aligned} & \sum_{\mathcal{P}_{ind}} \text{erfc} \left(\frac{\sum_{n=1}^N \frac{v_n}{\beta_n}}{\sqrt{\sum_{n=1}^N v_n}} \right) \exp \left(- \sum_{n=1}^N \frac{v_n}{\beta_n G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}} \right) \\ & \geq N! \text{erfc} \left(\frac{\frac{1}{N!} \sum_{\{i_n\}_{n=1}^N \in \mathcal{P}_{ind}} \sum_{n=1}^N \frac{1}{\beta_n} v_{i_n}}{\sqrt{\sum_{n=1}^N v_n}} \right) \\ & \quad \times \exp \left(- \frac{1}{N!} \sum_{\{i_n\}_{n=1}^N \in \mathcal{P}_{ind}} \sum_{n=1}^N \frac{1}{\beta_n G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}} v_{i_n} \right) \\ & = N! \text{erfc} \left(\left(\frac{1}{N} \sum_{n=1}^N \frac{1}{\beta_n} \right) \sqrt{\sum_{n=1}^N v_n} \right) \\ & \quad \times \exp \left(- \left(\frac{1}{N} \sum_{n=1}^N \frac{1}{\beta_n G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}} \right) \sum_{n=1}^N v_n \right). \end{aligned} \quad (16)$$

Combining (14) and (16), we bound the average PEP below as

$$\begin{aligned} \overline{\text{PEP}}_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}} & \geq \frac{1}{2 \prod_{n=1}^N \beta_n G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}} \int_0^\infty \dots \int_0^\infty \\ & \quad \times \text{erfc} \left(\left(\frac{1}{N} \sum_{n=1}^N \frac{1}{\beta_n} \right) \sqrt{\sum_{n=1}^N v_n} \right) \\ & \quad \times \exp \left(- \left(\frac{1}{N} \sum_{n=1}^N \frac{1}{\beta_n G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}} \right) \sum_{n=1}^N v_n \right) dv_1 \dots dv_N. \end{aligned} \quad (17)$$

Similar to the previous subsection, we deploy the approximation (12) to compute (17) as

$$\begin{aligned} \overline{\text{PEP}}_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}} & \geq \frac{1}{2 \prod_{n=1}^N \beta_n G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}} \int_0^\infty \dots \int_0^\infty \left[\frac{1}{6} \exp \left(- \left(\frac{1}{N} \sum_{n=1}^N \frac{1}{\beta_n} \right)^2 \sum_{n=1}^N v_n \right) \right. \\ & \quad \left. + \frac{1}{2} \exp \left(- \frac{4}{3} \left(\frac{1}{N} \sum_{n=1}^N \frac{1}{\beta_n} \right)^2 \sum_{n=1}^N v_n \right) \right] \\ & \quad \times \exp \left(- \left(\frac{1}{N} \sum_{n=1}^N \frac{1}{\beta_n G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}} \right) \sum_{n=1}^N v_n \right) dv_1 \dots dv_N \\ & = \frac{1}{12 \prod_{n=1}^N \beta_n G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}} \left(\left(\frac{1}{N} \sum_{n=1}^N \frac{1}{\beta_n} \right)^2 + \frac{1}{N} \sum_{n=1}^N \frac{1}{\beta_n G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}} \right)^{-N} \\ & \quad + \frac{1}{4 \prod_{n=1}^N \beta_n G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}} \left(\frac{4}{3} \left(\frac{1}{N} \sum_{n=1}^N \frac{1}{\beta_n} \right)^2 + \frac{1}{N} \sum_{n=1}^N \frac{1}{\beta_n G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}} \right)^{-N}. \end{aligned} \quad (18)$$

Substituting (18) into (10), we obtain the lower bound of the union bound in Rayleigh fading channels.

2) *Upper Bound:* In this subsection, we derive an upper bound for (14). Denoting $\beta_{\max} = \max_{n \in \{1, \dots, N\}} \beta_n$, we have

$$\sum_{n=1}^N \frac{v_n}{\beta_n} \geq \sqrt{\sum_{n=1}^N \frac{v_n}{\beta_n}} \sqrt{\sum_{n=1}^N v_n}. \quad (19)$$

Since $\text{erfc}(x)$ is a decreasing function, applying (19) to (14) we obtain

$$\begin{aligned} \overline{\text{PEP}}_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}} & \leq \frac{1}{2 \prod_{n=1}^N \beta_n G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}} \int_0^\infty \dots \int_0^\infty \text{erfc} \left(\frac{1}{\sqrt{\beta_{\max}}} \right. \\ & \quad \times \sqrt{\sum_{n=1}^N \frac{v_n}{\beta_n}} \left. \right) \times \exp \left(- \sum_{n=1}^N \frac{v_n}{\beta_n G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}} \right) dv_1 \dots dv_N \\ & \stackrel{(a)}{\approx} \frac{1}{2 \prod_{n=1}^N \beta_n G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}} \int_0^\infty \dots \int_0^\infty \left[\frac{1}{6} \exp \left(- \sum_{n=1}^N \left(\frac{1}{\beta_{\max}} + \frac{1}{G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}} \right) \right. \right. \\ & \quad \times \left. \left. \frac{v_n}{\beta_n} \right) + \frac{1}{2} \exp \left(- \sum_{n=1}^N \left(\frac{4}{3\beta_{\max}} + \frac{1}{G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}} \right) \frac{v_n}{\beta_n} \right) \right] dv_1 \dots dv_N \\ & = \frac{1}{12 \prod_{n=1}^N \left(\frac{G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}}{\beta_{\max}} + 1 \right)} + \frac{1}{4 \prod_{n=1}^N \left(\frac{4G_n^{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}}{3\beta_{\max}} + 1 \right)}. \end{aligned} \quad (20)$$

where (a) results from the approximation (12).

Substituting (20) into (10), we obtain the upper bound of the BLER in Rayleigh fading channels. We observe from (18) and (20) that C-RAN achieves full diversity of order N with respect to signal-to-compression-plus-noise ratio. When the quantization noise powers at the RRHs are identical, i.e., $\delta_1 = \dots = \delta_N$ and therefore $\beta_1 = \dots = \beta_N$, the lower bound and upper bound completely coincide and equal to (13) in Section IV-A.

V. MINIMIZATION OF FRONTHAUL TRANSMISSION RATE

In practical systems, different applications might require different QoS. For example, a voice message can usually tolerate a higher BLER than a video call. From the service provider's perspectives, it is always beneficial to minimize the network resources as long as the required QoS is satisfied. This motivates us to propose adaptive compression schemes to minimize the fronthaul transmission rate (maximize the compression efficiency) so that a fronthaul link can support a maximum number of antennas. Such schemes are desirable for systems which support large fronthaul feedback and/or require stringent BLER requirements. We emphasize that the proposed adaptive compressions are based on the average BLER, and require only the statistics of the fading channels, which reveals the advantages of the proposed optimizations over the existing C-RAN studies which require instantaneous CSI [3], [21]. Since the actual transmission rate at the n -th RRH is equal to the sampling resolution at that RRH, we refer Q_n as the transmission rate for convenience.

For a given QoS constraint ζ , we want to minimize the total fronthauls' transmission rate. The corresponding optimization is formulated as follows:

$$\begin{aligned} & \underset{\{Q_n : Q_n \geq 1\}}{\text{minimize}} && \sum_{n=1}^N Q_n \\ & \text{s.t.} && \frac{1}{|\mathcal{S}|^M} \sum_{\mathbf{c} \in \mathcal{S}^M} \sum_{\tilde{\mathbf{c}} \neq \mathbf{c}} \overline{\text{PEP}}_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}} \leq \zeta, \end{aligned} \quad (21)$$

where $\overline{\text{PEP}}_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}$ is given in the previous section. In the following, we propose two optimization frameworks for (21) based on the lower and upper bounds of $\overline{\text{PEP}}_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}$.

A. Lower Bound-Based Minimization

This minimization uses the lower bound of the union bound (18) as the QoS constraint. Denote $L = |\mathcal{S}|^M (|\mathcal{S}|^M - 1)$ and let $\{G_n^l\}_{l=1}^L$ represent $\|\mathbf{A}_n(\tilde{\mathbf{c}} - \mathbf{c})\|^2$, $\forall \tilde{\mathbf{c}} \neq \mathbf{c}$. Recalling that $\beta_n = 4\sigma^2 + \frac{2}{3}\eta_n^2 2^{-2Q_n}$ and $\{\eta_n\}_{n=1}^N$ are fixed. The problem of minimizing $\sum_{n=1}^N Q_n$ subject to the constraint $\text{BLER} \leq \zeta$ is equivalent to

$$\begin{aligned} & \underset{\{\beta_n\}_{n=1}^N}{\text{maximize}} && \sum_{n=1}^N \log(\beta_n - 4\sigma^2) \\ & \text{s.t.} && \frac{1}{|\mathcal{S}|^M} \sum_{l=1}^L \left(\frac{1}{12 \prod_{n=1}^N \beta_n G_n^l} \left[\left(\frac{1}{N} \sum_{n=1}^N \frac{1}{\beta_n} \right)^2 \right. \right. \\ & && \left. \left. + \frac{1}{N} \sum_{n=1}^N \frac{1}{\beta_n G_n^l} \right]^{-N} + \frac{1}{4 \prod_{n=1}^N \beta_n G_n^l} \left[\frac{4}{3} \left(\frac{1}{N} \sum_{n=1}^N \frac{1}{\beta_n} \right)^2 \right. \right. \\ & && \left. \left. + \frac{1}{N} \sum_{n=1}^N \frac{1}{\beta_n G_n^l} \right]^{-N} \right) \leq \zeta, \\ & && 0 < \beta_n - 4\sigma^2 \leq \frac{1}{6}\eta_n^2, \forall n. \end{aligned} \quad (22)$$

By introducing auxiliary positive variables $\{A_l\}_{l=1}^L$, $\{B_l\}_{l=1}^L$, $\{E_l\}_{l=1}^L$, and D , with $A_l, B_l, E_l, D \in \mathbb{R}^+$, we can further reformulate (22) as

$$\underset{\{\beta_n\}_{n=1}^N, D, \{A_l, B_l, E_l\}_{l=1}^L}{\text{maximize}} \quad \sum_{n=1}^N \log(\beta_n - 4\sigma^2) \quad (23)$$

$$\text{s.t.} \quad \frac{1}{|\mathcal{S}|^M} \sum_{l=1}^L (A_l + B_l) \leq \zeta, \quad (23a)$$

$$\begin{aligned} & -\log(12) - \sum_{n=1}^N \log(G_n^l) \leq \log(A_l) + \\ & \sum_{n=1}^N \log(\beta_n) + N \log(D + E_l), \forall l, \end{aligned} \quad (23b)$$

$$\begin{aligned} & -\log(4) - \sum_{n=1}^N \log(G_n^l) \leq \log(B_l) + \\ & \sum_{n=1}^N \log(\beta_n) + N \log\left(\frac{4}{3}D + E_l\right), \forall l, \end{aligned} \quad (23c)$$

$$\sqrt{D} \leq \frac{1}{N} \sum_{n=1}^N \frac{1}{\beta_n}, \quad (23d)$$

$$E_l \leq \frac{1}{N} \sum_{n=1}^N \frac{1}{\beta_n G_n^l}, \forall l, \quad (23e)$$

$$0 < \beta_n - 4\sigma^2 \leq \frac{\eta_n^2}{6}, \forall n. \quad (23f)$$

The problem (23) is difficult to solve due to the non-convexity of constraints (23d) and (23e). Therefore, we propose to solve an equivalent problem as given in Theorem 1.

Theorem 1: The problem (23) is equivalent to the following optimization:

$$\begin{aligned} & \underset{\{\beta_n, \gamma_n\}_{n=1}^N, D, \theta, \{A_l, B_l, E_l\}_{l=1}^L}{\text{maximize}} && \sum_{n=1}^N \Theta_n + 2D - \theta^2 - \frac{D^2}{\theta^2} \\ & \text{s.t.} && (23a) - (23c), (23f), \\ & && \frac{N}{2} \left(\theta + \frac{D}{\theta} \right) \leq 2 \sum_{n=1}^N \gamma_n - \sum_{n=1}^N \beta_n \gamma_n^2, \\ & && NE_l \leq 2 \sum_{n=1}^N \frac{\gamma_n}{G_n^l} - \sum_{n=1}^N \frac{\beta_n}{G_n^l} \gamma_n^2, \forall l. \end{aligned} \quad (24)$$

where $\Theta_n = \log(\beta_n - 4\sigma^2) + \left(2\gamma_n - \beta_n \gamma_n^2 - \frac{1}{\beta_n}\right) + \sum_{l=1}^L \frac{1}{G_n^l} \left(2\gamma_n - \beta_n \gamma_n^2 - \frac{1}{\beta_n}\right)$.

Proof: See Appendix B. ■

We observe that (24) is a convex optimization problem in $\{\beta_n\}_{n=1}^N$ and D for any given set of $\{\gamma_n\}_{n=1}^N$ and θ . Furthermore, (24) is a convex optimization problem in $\{\gamma_n\}_{n=1}^N$ and θ for any given set of $\{\beta_n\}_{n=1}^N$ and D . This remark mo-

TABLE I
ALGORITHM TO SOLVE (25)

1. Initialize $\{\gamma_n\}_{n=1}^N$ and θ .
2. Solve (25) given $\{\gamma_n\}_{n=1}^N$ and θ .
3. Update $\gamma_n = 1/\beta_n$, $\theta = \sqrt{D}$.
4. Repeat steps 2 and 3 until the objective value converges within a prescribed accuracy or a predefined maximum number of iterations is reached.

tivates the adoption of alternating optimization to solve (24). In particular, for given $\{\gamma_n\}_{n=1}^N$ and θ , we optimize (24) with variables $\{\beta_n\}_{n=1}^N$ and D and get $\{\beta_n^*\}_{n=1}^N$ and D^* . On the other hand, given $\{\beta_n^*\}_{n=1}^N$ and D^* , we optimize (24) with variables $\{\gamma_n\}_{n=1}^N$ and θ , and obtain $\gamma_n^* = \frac{1}{\beta_n^*}$, $\theta^* = \sqrt{D^*}$ (see Appendix B). The detail algorithm is given in Table I. Note that the function $g(x) \triangleq \log(\alpha x^2 + y)$ is concave in $x > 0$ and $y > 0$ for $\alpha > 0$. Therefore, the problem (25) is convex and thus can be solved efficiently in polynomial time.

$$\begin{aligned} & \underset{\substack{\{\beta_n\}_{n=1}^N, D, \\ \{A_l, B_l\}_{l=1}^L}}{\text{maximize}} && \sum_{n=1}^N \Theta_n + 2D - \theta^2 - \frac{D^2}{\theta^2} \\ & \text{s.t.} && \text{the constraints of (24)}. \end{aligned} \quad (25)$$

It is worthy noting that (25) is not necessarily equivalent to the original problem (21) since it is based on the lower bound of the union bound. However, it will be showed later that this optimization strategy provides optimum solution when the total fronthaul bandwidth is sufficiently large.

Remark 1: The integer constraint of the original problem (21) can be obtained from β_n simply by choosing the smallest integer of $\hat{Q}_n = \frac{1}{2}(\log_2(\frac{2}{3}\eta_n^2) - \log_2(\beta_n - 4\sigma^2))$, i.e., $\lceil \hat{Q}_n \rceil$. In general, there is no bound for the optimality loss of such approximation. However, as the constraint threshold BLER tends to zero, the loss also converges to zero. The reason is that each Q_n becomes large in such cases, which leads to a small ratio $\frac{\lceil \hat{Q}_n \rceil - \hat{Q}_n}{\hat{Q}_n}$.

B. Upper Bound-Based Minimization

In this subsection, we use the upper bound of the BLER in (20) as the constraint. The corresponding optimization problem is formulated as follows:

$$\begin{aligned} & \underset{\{\beta_n\}_{n=1}^N}{\text{maximize}} && \sum_{n=1}^N \log(\beta_n - 4\sigma^2) \\ & \text{s.t.} && \frac{1}{|\mathcal{S}|^M} \sum_{l=1}^L \left(\frac{1}{12 \prod_{n=1}^N \left(\frac{G_n^l}{\beta_{\max}} + 1 \right)} \right. \\ & && \left. + \frac{1}{4 \prod_{n=1}^N \left(\frac{4G_n^l}{3\beta_{\max}} + 1 \right)} \right) \leq \zeta, \\ & && 0 < \beta_n - 4\sigma^2 \leq \frac{\eta_n^2}{6}, \forall n, \end{aligned} \quad (26)$$

TABLE II
ALGORITHM TO SOLVE (27)

1. Initialize A_H , A_L , and the accuracy ϵ .
2. $A_M = (A_H + A_L)/2$.
3. Given $A = A_M$, if (27) is feasible, then $A_L := A_M$. Otherwise $A_H := A_M$.
4. Repeat step 2 and 3 until $|A_H - A_L| \leq \epsilon$.

where β_{\max} has been defined in Section IV-B2.

By introducing an auxiliary positive variable A , the above problem is written equivalently as follows:

$$\begin{aligned} & \underset{\{\beta_n\}_{n=1}^N, A > 0}{\text{maximize}} && \sum_{n=1}^N \log(\beta_n - 4\sigma^2) \\ & \text{s.t.} && \frac{1}{|\mathcal{S}|^M} \sum_{l=1}^L \log \left(\frac{1}{12 \prod_{n=1}^N (G_n^l A + 1)} \right. \\ & && \left. + \frac{1}{4 \prod_{n=1}^N (\frac{4}{3} G_n^l A + 1)} \right) \leq \zeta, \\ & && \beta_n \leq \frac{1}{A}, \forall n, \\ & && 0 < \beta_n - 4\sigma^2 \leq \frac{\eta_n^2}{6}, \forall n. \end{aligned} \quad (27)$$

It is observed that for a given A , the problem (27) is convex and thus efficiently solvable. We therefore resort to bisection method to solve (27). The steps are detailed in Table II.

C. Optimality of Identical Quantization Noise

In this subsection, we will show that a sampling scheme $\{Q_n\}_{n=1}^N$ which leads to identical quantization noise is the optimal solution of the problem (21), as the BLER threshold ζ goes to zero. The formal statement of this result is given in two following propositions.

Proposition 1: As the QoS threshold $\zeta \rightarrow 0$, the solution of (22) based on the PEP lower bound satisfies the identical quantization noise.

Proof: See Appendix C. ■

Proposition 2: As the QoS threshold $\zeta \rightarrow 0$, the solution of (26) based on the PEP upper bound satisfies the identical compression noise.

Proof: See Appendix D. ■

Propositions 1 and 2 are not strong enough to state that the optimal solution of (21) satisfies identical compression noise as the BLER threshold decreases since (22) and (26) are based on the PEP bounds. Nevertheless, Propositions 1 and 2 provide a justification for implementing the sampling that imposes identical compression noise, especially under delay-constrained systems where sophisticated adaptive sampling might not be applicable.

TABLE III
ALGORITHM TO SOLVE (30)

1.	Initialize A_H , A_L , and the accuracy ϵ .
2.	$A_M = (A_H + A_L)/2$.
3.	Given $A = A_M$, if (30) is feasible, then $A_L := A_M$. Otherwise $A_H := A_M$.
4.	Repeat step 2 and 3 until $ A_H - A_L \leq \epsilon$.

VI. MINIMIZATION OF BLER

In this section, we minimize the BLER for a given total fronthaul bandwidth Q_{sum} bits. Specifically, we will optimally allocate sampling resolution Q_n to RRH n to achieve the smallest BLER. The optimization problem is formulated as follows:

$$\begin{aligned} & \underset{\{Q_n : Q_n \geq 1\}_{n=1}^N}{\text{minimize}} && \text{BLER} \\ & \text{s.t.} && \sum_{n=1}^N Q_n \leq Q_{sum}, \end{aligned} \quad (28)$$

where BLER is given in (10).

In order to guarantee an effective optimization in (28), the BLER is calculated based on the upper bound of PEP given in (20). By changing variable to $\beta_n = 4\sigma^2 + \frac{2}{3}\eta_n^2 2^{-2Q_n}$ and recalling that $\beta_{\max} \triangleq \max_n \{\beta_n\}$, the resulting optimization problem is as

$$\begin{aligned} & \underset{\{\beta_n\}_{n=1}^N}{\text{minimize}} && \sum_{l=1}^L \left(\frac{1}{12} \prod_{n=1}^N \frac{1}{\frac{G_n^l}{\beta_{\max}} + 1} + \frac{1}{4} \prod_{n=1}^N \frac{1}{\frac{4G_n^l}{3\beta_{\max}} + 1} \right) \\ & \text{s.t.} && \sum_{n=1}^N \log_2(\beta_n - 4\sigma^2) \geq \sum_{n=1}^N \log_2\left(\frac{2}{3}\eta_n^2\right) - 2Q_{sum}, \\ & && 0 < \beta_n \leq \frac{1}{6}\eta_n^2, \forall n. \end{aligned} \quad (29)$$

By introducing an auxiliary positive variable $A \in \mathbb{R}^+$, the problem (29) is equivalent to

$$\begin{aligned} & \underset{\substack{\{\beta_n : \beta_n > 0\}_{n=1}^N \\ A > 0}}{\text{minimize}} && \sum_{l=1}^L \left(\frac{1}{12} \prod_{n=1}^N \frac{1}{G_n^l A + 1} + \frac{1}{4} \prod_{n=1}^N \frac{1}{\frac{4}{3}G_n^l A + 1} \right) \\ & \text{s.t.} && \beta_n \leq \frac{1}{A}, \forall n \end{aligned} \quad (30)$$

$$\text{s.t.} \quad \beta_n \leq \frac{1}{A}, \forall n \quad (30a)$$

$$\sum_{n=1}^N \log_2(\beta_n - 4\sigma^2) \geq \sum_{n=1}^N \log_2\left(\frac{2}{3}\eta_n^2\right) - 2Q_{sum}, \quad (30b)$$

$$\beta_n \leq 4\sigma^2 + \frac{1}{6}\eta_n^2, \forall n. \quad (30c)$$

We observe that (30) is a convex optimization problem on $\{\beta_n\}_{n=1}^N$ for a given A . Therefore, (30) can efficiently be solved and the steps to solve (30) are given in Table III.

In general, it is difficult to obtain the exact formula for the optimal solution of (28). Under certain circumstances, however, a closed-form solution can be derived in the following theorem.

Theorem 2: If there exist $\{q_n : q_n \geq 1\}_{n=1}^N$ such as $\eta_1 2^{-q_1} = \eta_2 2^{-q_2} = \dots = \eta_N 2^{-q_N}$ and $\sum_{n=1}^N q_n = Q_{sum}$, then the solution of problem (30) satisfies identical quantization noises and is given as

$$\beta_n^* = 2^{\frac{1}{N} \left(\sum_{n=1}^N \log_2\left(\frac{2}{3}\eta_n^2\right) - 2Q_{sum} \right)} + 4\sigma^2, \forall n.$$

Consequently, the optimal fronthaul rate allocation $\{Q_n^*\}_{n=1}^N$ is given as

$$Q_n^* = \frac{1}{2} \log_2\left(\frac{2}{3}\eta_n^2\right) + \frac{1}{N} Q_{sum} - \frac{1}{2N} \sum_{n=1}^N \log_2\left(\frac{2}{3}\eta_n^2\right).$$

Proof: See Appendix E. ■

Corollary 1: For symmetric C-RAN systems, i.e., $\eta_1 = \eta_2 = \dots = \eta_N$, which employ the quantization and the receiver as in Section II, the uniform sampling $\{Q_n = Q_{sum}/N\}_{n=1}^N$ achieves the minimum BLER.

Proof: The proof is obtained straightforwardly from Theorem 2 by using $\eta_1 = \eta_2 = \dots = \eta_N$. ■

VII. NUMERICAL RESULTS

The simulation is evaluated for a C-RAN system under block Rayleigh fading channel, i.e., h_{mn} 's are i.i.d. random variables, each distributed as $\mathcal{CN}(0, 1)$. Unless otherwise stated, we assume $M = 3$, $N = 3$ and QPSK modulation with the codebook $\mathcal{S} = \{-1 - 1i, -1 + 1i, 1 - 1i, 1 + 1i\}/\sqrt{2}$. Furthermore, $P_{nm} = P, \forall m, n$ and the average SNR is defined as P/σ^2 . The BBU is assumed to know CSI of all wireless channels.

A. Accuracy of the Bounds

Figure 2 shows effects of modulation order (a) and the number of RRHs (b) on the performance of C-RAN. Symmetric network topology is assumed, i.e., $\eta_1 = \eta_2 = \dots = \eta_N$. The sampling rate is equally allocated, i.e., $Q_1 = Q_2 = Q_3 = 6$ bits. Since the quantization noise at every RRH is identical, the upper and lower bounds of the PEP given in Section IV-A coincide. Figure 2(a) compares the bounds with simulation results for $N = 3$ RRHs and three different modulations, i.e., BPSK, QPSK, and 16-QAM. For all cases, the derived bounds closely match simulation results at the medium and high SNR regimes. It is shown that the bound is closer to the simulation for BPSK than 16-QAM. This can be explained from the fact that the derived bound's accuracy depends on both the modulation order and the operating SNR. Higher modulation order results looser union bound of the BLER because the compound constellation is more diverse. Moreover, the approximation in (12) is more tight in the medium and high SNR regime [25]. Figure 2(b) presents the theoretical and simulation results for different number of RRHs under QPSK modulation. Similarly, the derived bounds closely match the simulations.

Figure 3(a) shows the effect of the fronthaul rate on the BLER. Figure 3(b) presents the BLER as a function of fronthaul

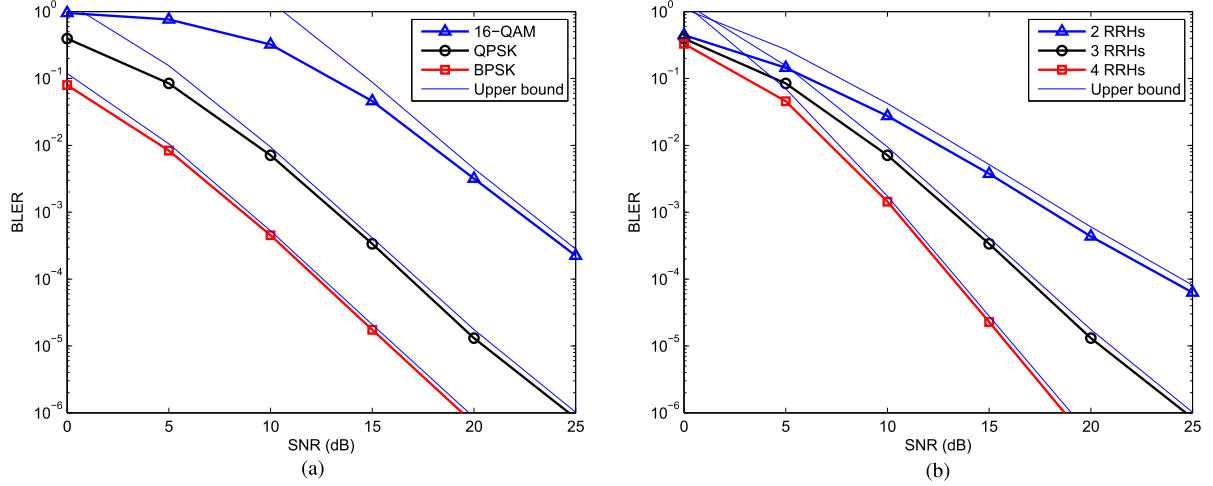


Fig. 2. Performance of the C-RAN in Rayleigh fading channels for (a) various modulation orders and (b) different number of RRHs. Solid and marked lines correspond to analytical upper bound and numerical results, respectively. The system parameters are: $M = N = 3$, $Q_1 = Q_2 = Q_3 = 6$ bits.

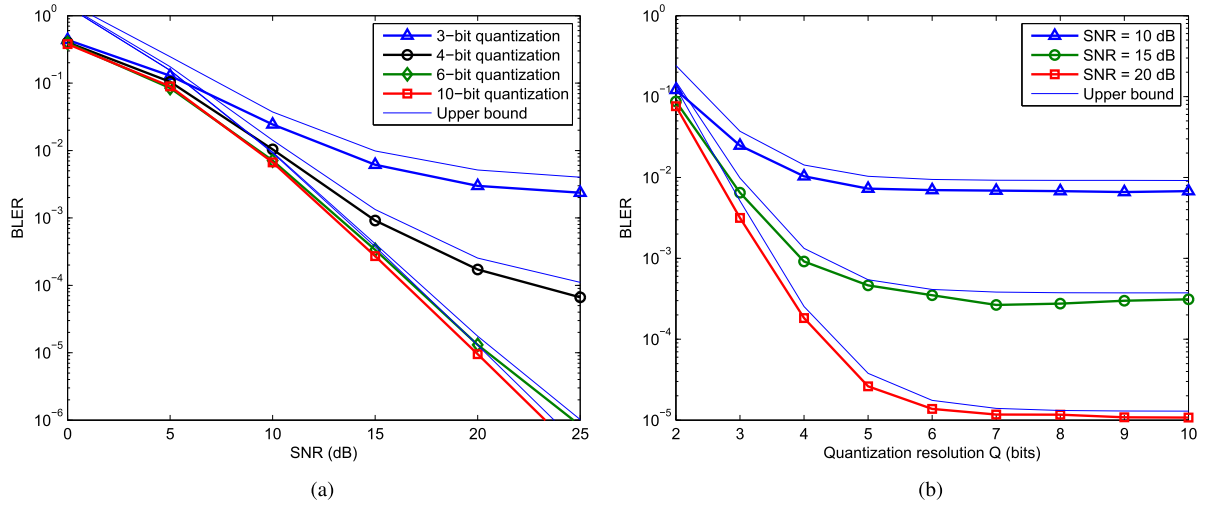


Fig. 3. Performance of the C-RAN in Rayleigh fading channels for (a) different quantizer's resolutions and (b) different operating SNRs. Solid and marked lines correspond to analytical upper bound and numerical results, respectively. The system parameters are: QPSK modulation, $M = N = 3$, $Q_1 = Q_2 = Q_3 = Q$.

transmission rate for difference operating SNRs. For all cases, the theoretical upper bound is close to the simulation.

Figure 4 shows simulation results and the corresponding bounds under non-identical quantization noise scenario. As SNR increases, the analysed BLER based on upper bound of the PEP and based on lower bound of the PEP diverges. Despite of this mismatch, we will show later that the optimal rate allocation based on either lower bound or upper bound approaches that of the identical quantization noise.

B. Performance of the Proposed Optimizations

Figure 5 presents the performance of the proposed adaptive compressions versus SNR. For a given BLER target, we want to maximize the compression efficiency, or equivalently to minimize the actual fronthaul transmission rate. Two adaptive compression schemes based on the lower bound and the upper bound of the union bound in Section V are presented. For reference, the scheme without QoS constraint which fully utilizes

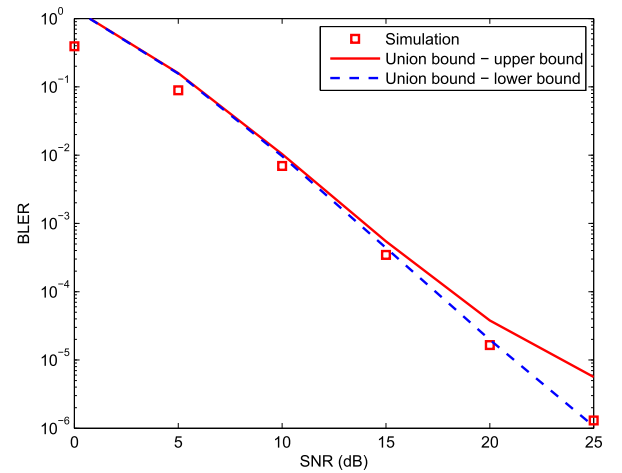


Fig. 4. Performance of the C-RAN in Rayleigh fading channels with non-identical quantization noises. System parameters: QPSK modulation, $M = N = 3$, $\{Q_1, Q_2, Q_3\} = \{5, 6, 7\}$.

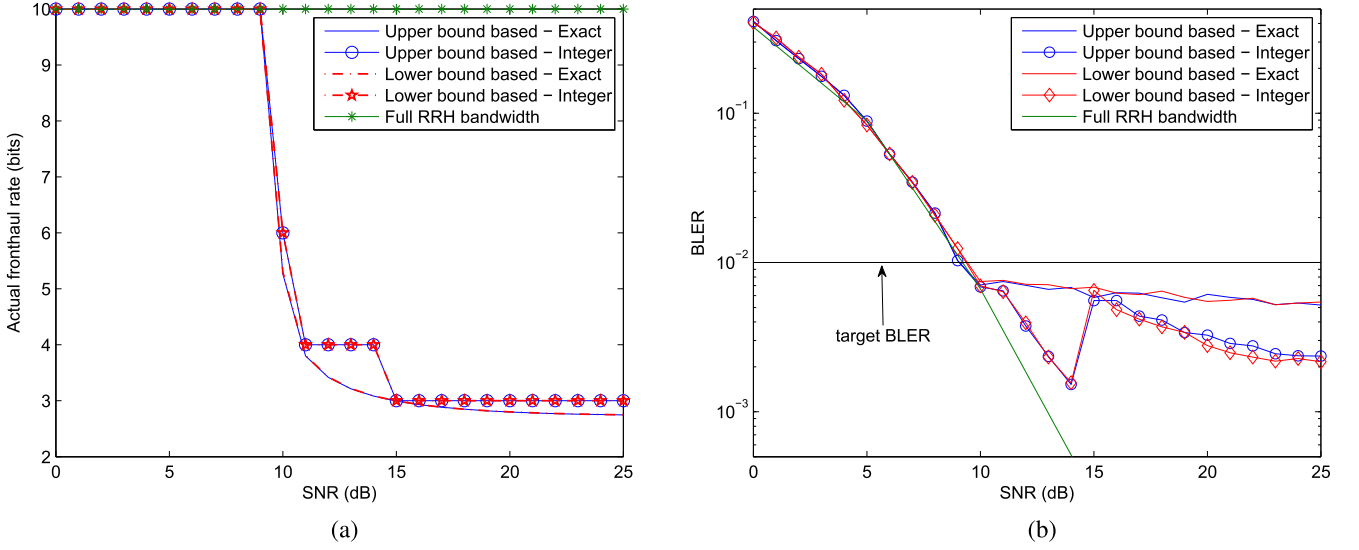


Fig. 5. Performance of the optimization with QoS constraint proposed in Section V for two schemes based on lower and upper bounds of the APEP. The target BLER equals 10^{-2} . The total fronthaul bandwidth $Q_{sum} = 30$ bits, $M = N = 3$.

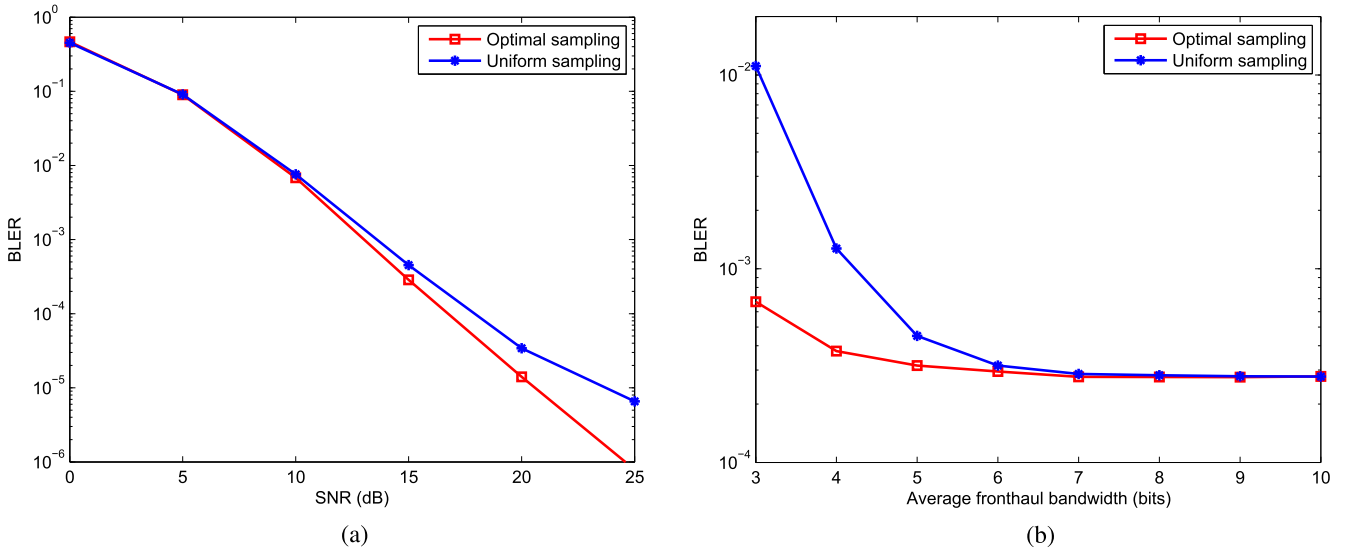


Fig. 6. Performance of the optimal rate allocation with QPSK modulation, $M = N = 3$: (a) performance versus SNR, $Q_{sum} = 15$ bit; (b) performance versus average per-RRH bandwidth, SNR = 15 dB.

the fronthaul bandwidth is also plotted. In addition, to provide full details on the proposed optimizations, curves corresponding to exact optimum solutions of (25) and (27) are drawn. These curves are mark as “- exact” in the figure. Curves corresponding to integer $\{Q_n\}$ are marked as “- integer”. The threshold ζ is equal to $1e-2$. Note that these optimizations are carried out for each operating SNR only once and does not depend on the instantaneous CSI. The compression efficiency is presented in Figure 5(a) and the BLER performance is shown in Figure 5(b). In the low SNR regime, the BLER does not satisfy the target QoS because the channel is so poor. Even using all 10 bits for quantization does not satisfy the target BLER. Therefore, the optimizations are infeasible and all schemes consume full fronthaul bandwidth, as shown in Figure 5(a). When SNR increases,

the optimizations are activated (from 10 dB in the figure). Both optimization schemes meet the target QoS while significantly improve the compression efficiency. Because the full-bandwidth scheme always uses 10 bits for quantization, its fronthaul rate is 10 bits per sample for all SNRs. On the other hand, a compression efficiency of 330% is observed by both schemes, which only require 3 bits per sample to achieve a BLER less than $1e-2$. One important observation is that the optimizations based on lower bound or upper bound achieve identical compression efficiencies. Such observation is in line with Section V-C, which shows the optimality of uniform compression noise rate allocation under asymptotic regimes. Note that the integer constraint on $\{Q_n\}_{n=1}^N$ results in lower BLER than the exact value of Q_n because $Q_n^{integer} = \lceil Q_n^{exact} \rceil$ is the smallest integer that is

larger than or equal to Q_n^{exact} . The sharp step in BLER curves of integer Q_n results from the $\lceil \cdot \rceil$ operation.

The above results illustrate the minimization of the actual fronthaul rate with the BLER constraint. Based on the analysed BLER, a reciprocal problem is how to allocate the fronthaul bandwidth $\{Q_n\}_{n=1}^N$ to minimize the BLER for a given $Q_{sum} = \sum_{n=1}^N Q_n$. The optimization problem is described in Sec. VI. The scaling factors are set as $\{\eta_1, \eta_2, \eta_3\} = \{3, 6, 9\}$. Uniform sampling scheme, e.g., $Q_1 = Q_2 = Q_3 = Q_{sum}/3$ is also shown for reference. Figure 6(a) shows the BLER performance. The total fronthauls' bandwidth is $Q_{sum} = 15$ bits. As expected, the optimum rate allocation achieves better performance (smaller BLER) than the reference scheme. Under the small SNR regime, both uniform and optimal rate allocations get similar performance because at this SNR the thermal noise is dominant. As SNR increases, performance gain provided by the optimal rate allocation is larger.

Figure 6(b) shows the BLER performance of the optimal rate allocation versus average fronthaul's bandwidth per RRH. The operating SNR is equal to 15 dB. It is observed that the optimal scheme is more effective for small fronthaul's bandwidth. As the fronthaul's bandwidth increases, the gap between the two schemes narrows and the two curves eventually coincide. This result is in line with Theorem 2 since with large Q_{sum} the condition in Theorem 2 is always satisfied. As a result, the optimal rate allocation $\{Q_1^*, Q_2^*, Q_3^*\}$ approaches uniform sampling as Q_{sum} increases.

VIII. CONCLUSION

In this work, we investigated the design of fronthaul in C-RAN uplink by focusing on the compression and optimization in fronthaul links based on the statistics of wireless fading channels. First, we derived the system BLER under Rayleigh fading channels. In particular, we derived upper and lower bounds of the BLER in closed-form. Based on the analysed bounds, we proposed two adaptive compression schemes to minimize the fronthaul transmission rate subject to a BLER constraint. Furthermore, a fronthaul rate allocation is proposed to minimize the system BLER. We showed that the uniform rate allocation approaches the optimal scheme as the total fronthauls bandwidth increases. Based on our current work, a promising research direction is to study more realistic compression model applying to time-domain correlated signal, e.g., OFDM symbols. Beside BLER, error vector magnitude is also an important performance measurement to investigate in the future.

APPENDIX A

CONVEXITY OF SEVERAL FUNCTIONS

A. Convexity of $f(x) \triangleq \text{erfc}(\alpha x) \exp(-x)$

The first and second derivatives of $f(x)$ are given as

$$f'(x) = -\frac{2\alpha e^{-\alpha^2 x^2}}{\sqrt{\pi}} e^{-x} - \text{erfc}(\alpha x) e^{-x},$$

$$f''(x) = \frac{2\alpha}{\sqrt{\pi}} e^{-\alpha^2 x^2 - x} (2\alpha x + 1) + \frac{2\alpha e^{-\alpha^2 x^2}}{\sqrt{\pi}} e^{-x} + \text{erfc}(\alpha x) e^{-x}.$$

Since $f''(x) > 0, \forall x > 0$, we conclude that $f(x)$ is a convex function in $x > 0$.

B. Concavity of $g(x) \triangleq \log(\alpha x^2 + x)$

The first and second derivatives of $g(x)$ are given as

$$\begin{aligned} g'(x) &= \frac{1}{\alpha x^2 + x} (2\alpha x + 1), \\ g''(x) &= -\frac{1}{(\alpha x^2 + x)^2} (2\alpha x + 1)^2 + \frac{2\alpha}{\alpha x^2 + x} \\ &= \frac{-2\alpha^2 x^2 - 2\alpha x - 1}{(\alpha x^2 + x)^2}. \end{aligned}$$

Since $g''(x) < 0, \forall x > 0$, we conclude that $g(x)$ is a concave function in $x > 0$.

APPENDIX B

PROOF OF THEOREM 1

We will prove that the optimal value of (23) is achievable by (24); the optimal value of (24) is also achievable by (23); and furthermore the optimal solution $\{\beta_n^*\}_{n=1}^N$ is the same for the two optimizations. To this end, we denote the optimal values of (23) and (24) as S_1^* and S_2^* , respectively, and will prove $S_1^* = S_2^*$. A necessary fact for the proof is given as follows.

Fact 1: Consider an arbitrary positive number $x > 0$. We have

$$\frac{1}{x} \geq 2y - xy^2, \forall y \in \mathbb{R}^+,$$

where the equality is achieved if and only if $y^* = 1/x$.

A. $S_1^* \leq S_2^*$

First, assume that $\{\{\beta_{1,n}^*\}_{n=1}^N, A_l^*, B_l^*, D^*, E_l^*\}$ is an optimal solution of (23). Consider a set $\{\{\beta_{1,n}^*, \gamma_n\}_{n=1}^N, A_l^*, B_l^*, D^*, E_l^*, \theta\}$ where $\gamma_n = \frac{1}{\beta_{1,n}^*}$ and $\theta = \sqrt{D^*}$. It is straightforward to see that the value S_1^* is also achieved by (24). Therefore, the optimal value of (24) is $S_2^* \geq S_1^*$.

B. $S_2^* \leq S_1^*$

Now assume that $\{\{\beta_{2,n}^*, \gamma_n^*\}_{n=1}^N, A_l^*, B_l^*, D^*, E_l^*, \theta^*\}$ is an optimal solution of (24). Consider a new set $\{\{\beta_{2,n}^*, \frac{1}{\beta_{2,n}^*}\}_{n=1}^N, A_l^*, B_l^*, D^*, E_l^*, \sqrt{D^*}\}$. We note that $\{\{\beta_{2,n}^*, \frac{1}{\beta_{2,n}^*}\}_{n=1}^N, A_l^*, B_l^*, D^*, E_l^*, \sqrt{D^*}\}$ satisfies all constraints in (24) since

$$2 \sum_{n=1}^N \frac{\gamma_n^*}{G_n^l} - \sum_{n=1}^N \frac{\beta_{2,n}^* (\gamma_n^*)^2}{G_n^l} \leq \sum_{n=1}^N \frac{1}{\beta_{2,n}^* G_n^l}, \quad (31)$$

and $2\sqrt{D^*} \leq \theta^* + \frac{D^*}{\theta^*}$.

Furthermore, due to (31) and the fact that $2D^* \leq (\theta^*)^2 + (\frac{D^*}{\theta^*})^2$, γ_n^* and θ^* have to be equal to $\frac{1}{\beta_{2,n}^*}$ and $\sqrt{D^*}$, respectively, since otherwise S_2^* would not be the optimal value of (24). Therefore, from (24) we see that

$$S_2^* = \sum_{n=1}^N \log(\beta_{2,n}^* - 4\sigma^2), \quad (32)$$

where $\beta_{2,n}^*$ satisfies

$$\begin{aligned} & -\log(12) - \sum_{n=1}^N \log(G_n^l) \\ & \leq \log(A_l^*) + \sum_{n=1}^N \log(\beta_{2,n}^*) + N \log(D^* + E_l^*), \end{aligned} \quad (33)$$

$$\begin{aligned} & -\log(4) - \sum_{n=1}^N \log(G_n^l) \\ & \leq \log(B_l^*) + \sum_{n=1}^N \log(\beta_{2,n}^*) + N \log\left(\frac{4}{3}D^* + E_l^*\right), \end{aligned} \quad (34)$$

$$N\sqrt{D^*} \leq \sum_{n=1}^N \frac{1}{\beta_{2,n}^*}, \quad NE_l^* \leq \sum_{n=1}^N \frac{1}{\beta_{2,n}^* G_n^l}, \quad (35)$$

$$\beta_{2,n}^* - 4\sigma^2 \leq \frac{\eta_n^2}{6}. \quad (36)$$

As the last step, we observe that due to (33)-(36), the set $\{\{\beta_{2,n}^*\}_{n=1}^N, A_l^*, B_l^*, D^*, E_l^*\}$ satisfies all constraints of (23). Therefore, we see that S_2^* is achieved by (23) with the set $\{\beta_{2,n}^*\}_{n=1}^N$. That leads to $S_2^* \leq S_1^*$.

From Section B-A and B-B, we conclude that $S_1^* = S_2^*$, which also implies that they have the same optimal solution $\{\beta_n^*\}_{n=1}^N$. This concludes the proof for Theorem 1.

APPENDIX C PROOF OF PROPOSITION 1

A. Preliminary Results

Before proceeding, we first note the following facts.

Fact 2: As the QoS threshold $\zeta \rightarrow 0$, we have $\beta_n \rightarrow 4\sigma^2$, $\forall n$.

Proof: Consider the first constraint in (22). As the BLER target ζ decreases, at least one of β_n has to decrease. Note that each β_n can not keep decreasing while the other are fixed since $\beta_n > 4\sigma^2, \forall n$. Therefore, all β_n 's will decrease to $4\sigma^2$ as $\zeta \rightarrow 0$. ■

Fact 3: The function $h(x) = \log(\frac{1}{x} - \alpha)$ with $x > 0$ is concave when $2\alpha x > 1$.

Proof: The first and second derivatives of $h(x)$ are given as

$$h'(x) = \frac{\alpha}{\alpha x - 1} - \frac{1}{x}, \quad h''(x) = \frac{1 - 2\alpha x}{x^2(\alpha x - 1)^2}.$$

Since $h''(x) < 0$ when $2\alpha x > 1$ we conclude that $h(x)$ is a concave function when $2\alpha x > 1$. ■

B. Proof of Proposition 1

Recalling that $\beta_n = 4\sigma^2 + \frac{2}{3}\eta_n^2 2^{-2Q_n}$. The optimization problem (21) therefore can be written as

$$\begin{aligned} & \text{maximize} \quad \sum_{n=1}^N \log(\beta_n - 4\sigma^2) \\ & \text{s.t.} \quad \frac{1}{|S|^M} \sum_{\forall \mathbf{c} \neq \tilde{\mathbf{c}}} \overline{\text{PEP}}_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}} \leq \zeta, \\ & \quad \quad 0 < \beta_n - 4\sigma^2 \leq \frac{\eta_n^2}{6}, \forall n, \end{aligned} \quad (37)$$

where $\overline{\text{PEP}}_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}$ is given in (14). Denote $\{\beta_n^*\}_{n=1}^N$ as the optimal solution of (37), we will prove that β_n^* 's are identical as ζ is small.

We consider the following set $\{\tilde{\beta}_n\}_{n=1}^N$ where

$$\tilde{\beta}_1 = \tilde{\beta}_2 = \dots = \tilde{\beta}_N = N \left(\sum_{n=1}^N \frac{1}{\beta_n^*} \right)^{-1}. \quad (38)$$

From (14) and (17), we have

$$\begin{aligned} & \overline{\text{PEP}}_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}(\{\beta_n^*\}_{n=1}^N) \geq \frac{1}{2 \prod_{n=1}^N \beta_n^* G_n^l} \\ & \times \int_0^\infty \dots \int_0^\infty \text{erfc}\left(\|\mathbf{A}_n(\tilde{\mathbf{c}} - \mathbf{c})\| \left(\frac{1}{N} \sum_{n=1}^N \frac{1}{\beta_n^*}\right) \sqrt{\sum_{n=1}^N v_n}\right) \\ & \times \exp\left(-\left(\frac{1}{N} \sum_{n=1}^N \frac{1}{\beta_n^* G_n^l}\right) \sum_{n=1}^N v_n\right) dv_1 \dots dv_N. \end{aligned} \quad (39)$$

Now note that $\frac{1}{N} \sum_{n=1}^N \frac{1}{\beta_n} = \frac{1}{N} \sum_{n=1}^N \frac{1}{\tilde{\beta}_n}$ while

$$\frac{1}{\prod_{n=1}^N \tilde{\beta}_n} = \left(\frac{N}{\sum_{n=1}^N \frac{1}{\beta_n^*}} \right)^N \leq \frac{1}{\prod_{n=1}^N \beta_n^*}, \quad (40)$$

as the geometric mean does not exceed the arithmetic mean. Combining (39) and (40), it yields:

$$\overline{\text{PEP}}_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}(\{\beta_n^*\}_{n=1}^N) \geq \overline{\text{PEP}}_{\mathbf{c} \rightarrow \tilde{\mathbf{c}}}(\{\tilde{\beta}_n\}_{n=1}^N). \quad (41)$$

Consider the function $\log(\frac{1}{\beta_n^*} - 4\sigma^2)$. As the QoS threshold ζ approaches 0, from Fact 2 we have $\beta_n^* \rightarrow 4\sigma^2$ and therefore $2 \times 4\sigma^2 \times \frac{1}{\beta_n^*} > 1$. From Fact 3, we conclude that $\log(\frac{1}{\beta_n^*} - 4\sigma^2)$ is concave. In the light of Jensen's inequality, we have

$$\begin{aligned} & \sum_{n=1}^N \log(\beta_n^* - 4\sigma^2) = \sum_{n=1}^N \log\left(\frac{1}{\beta_n^*} - 4\sigma^2\right) \\ & \leq N \log\left(\frac{N}{\sum_{n=1}^N \frac{1}{\beta_n^*}} - 4\sigma^2\right) = \sum_{n=1}^N \log(\tilde{\beta}_n - 4\sigma^2). \end{aligned} \quad (42)$$

Finally, we note that as the QoS threshold ζ approaches 0, from Fact 2 we have $\beta_n^* \rightarrow 4\sigma^2$. It is thus straightforward to prove that

$$0 < \tilde{\beta}_n - 4\sigma^2 \leq \frac{\eta_n^2}{6}. \quad (43)$$

From (41)-(43), we observe that the set $\{\tilde{\beta}_n\}_{n=1}^N$ satisfies all the constraints of (37) while achieves a larger or equal objective function than the optimal solution $\{\beta_n^*\}_{n=1}^N$. The equality is achieved if and only if (41) and (42) have the equalities, i.e., $\beta_1^* = \dots = \beta_N^*$. This concludes the proof of Proposition 1.

APPENDIX D

PROOF OF PROPOSITION 2

Let $\{\beta_1^*, \beta_2^*, \dots, \beta_N^*, A^*\}$ denote the optimal solution of problem (30). We will show that $\{\beta_1^* = \beta_2^* = \dots = \beta_N^*\}$ as η tends to zero. Assuming that $\{\beta_n^*\}_{n=1}^N$ do not satisfy identical quantization noise, without loss of generality, let $\beta_1^* \geq \beta_2^* \geq \dots \geq \beta_N^*$.

As the QoS η tends to zero, A^* must be very large to satisfy the first constraint of problem (27). As a result, $\beta_n^* - 4\sigma^2$ is strictly smaller than $\eta_n^2/6, \forall n$. Thus, there always exists $\bar{\beta}_N^* > \beta_N^*$ such as $\bar{\beta}_N^* - 4\sigma^2 \leq \eta_N^2/6$ and $\bar{\beta}_N^* \leq \beta_1^*$. Consider a set $\{\beta_1^*, \dots, \beta_{N-1}^*, \bar{\beta}_N^*, A^*\}$ which satisfies all constraints of (27) and thus is a feasible solution. Since \log is a monotonically increasing function, this solution yields an objective value $\log(\bar{\beta}_N^* - 4\sigma^2) + \sum_{n=1}^{N-1} \log(\beta_n^* - 4\sigma^2) > \sum_{n=1}^N \log(\beta_n^* - 4\sigma^2)$, which is in contrast to the optimality assumption of $\{\beta_1^*, \beta_2^*, \dots, \beta_N^*, A^*\}$. Thus, $\{\beta_n^*\}_{n=1}^N$ must be identical.

APPENDIX E

PROOF OF THEOREM 2

Let $\{\beta_1^*, \beta_2^*, \dots, \beta_N^*, A^*\}$ denote the optimal solution of problem (30). We will show that, under the condition in Theorem 2, then

$$\beta_1^* = \beta_2^* = \dots = \beta_N^*. \quad (44)$$

Denote $F(A)$ as the objective function of problem (30). By definition, $F(A^*)$ is the smallest among all feasible sets of (30).

Now if (44) does not hold, then there is at least one β_n^* is strictly larger than the others. Without loss of generality, assuming that $\beta_1^* > \beta_2^* \geq \dots \geq \beta_N^*$. First, we will show that there will be at least one equality in (30c) does not hold for $2 \leq n \leq N$. Indeed, let us assume that $\beta_n^* = 4\sigma^2 + \frac{1}{6}\eta_n^2, 2 \leq n \leq N$. Define $Q_n^* = \frac{1}{2}(\log_2(\frac{2}{3}\eta_n^2) - \log_2(\beta_n^* - 4\sigma^2))$. Thus, we have $Q_n^* = 1$ with $2 \leq n \leq N$ and $Q_1^* = Q_{sum} - N + 1$. Consider the set $\{q_n\}_{n=1}^N$ that satisfies the condition of Theorem 2. Since $q_n \geq 1, \forall n$, we have $q_1 = Q_{sum} - \sum_{n=2}^N q_n \leq Q_{sum} - N + 1 = Q_1^*$. Because β_n is a monotonic decreasing function of Q_n , it yields $\beta_1(q_1) = 4\sigma^2 + \frac{2}{3}\eta_1^2 2^{-2q_1} \geq \beta_1^*$. For any $n \geq 2$, we have $\beta_n(q_n) \leq \beta_n^*$ as $q_{n:n \geq 2} \geq 1 = Q_{n:n \geq 2}^*$. By the condition in Theorem 2, $\beta_1(q_1) = \dots = \beta_N(q_N)$. Consequently, for $n \geq 2$, we have $\beta_n^* \geq \beta_n(q_n) = \beta_1(q_1) \geq \beta_1^*$, which is in contrast to the assumption of the largest β_1^* .

As a result, there is at least one strict constraint in (30c) for $n \geq 2$. Without loss of generality, assume that $\beta_2^* < 4\sigma^2 + \frac{1}{6}\eta_2^2$. Then there always exist $\bar{\beta}_1^*$ and $\bar{\beta}_2^*$ such as $\beta_1^* > \bar{\beta}_1^*, \beta_2^* < \bar{\beta}_2^* \leq 4\sigma^2 + \frac{1}{6}\eta_2^2$, and $\log_2(\beta_1^* - 4\sigma^2) + \log_2(\beta_2^* - 4\sigma^2) = \log_2(\bar{\beta}_1^* - 4\sigma^2) + \log_2(\bar{\beta}_2^* - 4\sigma^2)$. It is straightfor-

ward to verify that $\{\bar{\beta}_1^*, \bar{\beta}_2^*, \beta_3^*, \dots, \beta_N^*, \bar{A}^* \triangleq 1/\bar{\beta}_1^*\}$ satisfy all constraints of (30), and thus is a feasible solution of problem (30) with the objective function $F(\bar{A}^*)$. Because $F(x)$ is a monotonic decreasing function, then $F(\bar{A}^*) < F(A^*)$ as $\bar{A}^* > A^*$, which is in contrast to the optimal assumption of $\{\beta_1^*, \beta_2^*, \dots, \beta_N^*, A^*\}$. Therefore, $\beta_1^* = \beta_2^* = \dots = \beta_N^*$. From (30b) we obtain

$$\beta_n^* = 2^{\frac{1}{N} \left(\sum_{n=1}^N \log_2(\frac{2}{3}\eta_n^2) - 2Q_{sum} \right)} + 4\sigma^2, \quad \forall n.$$

The optimal rate allocation in Theorem 2 thus is obtained directly from the definition $\beta_n = 4\sigma^2 + \frac{1}{6}\eta_n^2 2^{-2Q_n}$.

REFERENCES

- [1] ChinaMobile, "C-RAN: The road towards green RAN," White Paper, 2011.
- [2] Z. Zhu *et al.*, "Virtual base station pool: Towards a wireless network cloud for radio access networks," in *Proc. ACM Int. Conf. Comput. Frontiers*, New York, NY, USA, Mar. 2011, pp. 34-1-34-10.
- [3] A. Sanderovich, O. Somekh, H. V. Poor, and S. Shamai, "Uplink macro diversity of limited backhaul cellular network," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3457-3478, Aug. 2009.
- [4] P. Marsch and G. Fettweis, "Uplink CoMP under a constrained backhaul and imperfect channel knowledge," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1730-1742, Jun. 2011.
- [5] T. Q. S. Quek, M. Peng, O. Simeone, and W. Yu, *Cloud Radio Access Networks: Principles, Technologies, and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2016.
- [6] J. Tang, W. P. Tay, and T. Q. S. Quek, "Cross-layer resource allocation with elastic service scaling in cloud radio access network," *IEEE Trans. Wireless Commun.*, vol. 14, no. 9, pp. 5068-5081, Sep. 2015.
- [7] Y. Zhou, W. Yu, and D. Tountakaris, "Uplink multi-cell processing: Approximate sum capacity under a sum backhaul constraint," in *Proc. IEEE Inf. Theory Workshop*, Sevilla, Spain, Sep. 2013, pp. 1-5.
- [8] S.-H. Park, O. Simeone, O. Sahin, and S. Shamai, "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, Aug. 2013.
- [9] P. Patil and W. Yu, "Hybrid compression and message-sharing strategy for the downlink cloud radio-access network," in *Proc. IEEE Inf. Theory Appl. Workshop*, San Diego, CA, USA, Feb. 2014, pp. 1-6.
- [10] J. Kang, O. Simeone, J. Kang, and S. Shamai, "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.
- [11] Y. Wang, H. Wang, and L. Scharf, "Optimum compression of a noisy measurement for transmission over a noisy channel," *IEEE Trans. Signal Process.*, vol. 62, no. 5, pp. 1279-1289, Mar. 2014.
- [12] A. del Coso and S. Simoes, "Distributed compression for MIMO coordinated networks with a backhaul constraint," *IEEE Trans. Wireless Commun.*, vol. 8, no. 9, pp. 4698-4709, Sep. 2009.
- [13] Q. Hu, M. Peng, Z. Mao, X. Xie, and H. V. Poor, "Training design for channel estimation in uplink cloud radio access networks," *IEEE Trans. Signal Process.*, vol. 64, no. 13, pp. 3324-3337, Jul. 2016.
- [14] "Common public radio interface (CPRI): Interface specification," CPRI Specification V6.0, 2013. [Online]. Available: <http://www.cpri.info>
- [15] J. Lorca and L. Cudala, "Lossless compression technique for the fronthaul of LTE/LTE-advanced Cloud-RAN architectures," in *Proc. IEEE 14th Int. Symp. Workshops World Wireless, Mobile Multimedia Netw.*, Madrid, Spain, Jun. 2013, pp. 1-9.
- [16] S. Nanba and A. Agata, "A new IQ data compression scheme for fronthaul link in centralized RAN," in *Proc. IEEE 24th Int. Symp. Pers., Indoor Mobile Radio Commun.*, London, U.K., Sep. 2013, pp. 210-214.
- [17] 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.*, Austin, TX, USA, Dec. 2013, pp. 1198-1201.
- [18] H. Si, B. L. Ng, M. S. Rahman, and J. Zhang, "A vector quantization based compression algorithm for CPRI Link," in *Proc. IEEE Global Commun. Conf.*, San Diego, CA, USA, Dec. 2015, pp. 1-6.

- [19] Y. Ren, Y. Wang, G. Xu, and Q. Huang, "A compression method for LTE-A signals transported in radio access networks," in *Proc. Int. Conf. Telecommun.*, May 2014, pp. 293–297.
- [20] J. Choi, B. L. Evans, and A. Gatherer, "Space-time fronthaul compression of complex baseband uplink LTE signals," in *Proc. IEEE Int. Conf. Commun.*, Kuala Lumpur, Malaysia, May 2016, pp. 1–6.
- [21] 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.
- [22] T. X. Vu, H. D. Nguyen, and T. Q. S. Quek, "Joint decoding and adaptive compression with QoS constraint for uplinks in cloud radio access networks," in *Proc. IEEE Global Commun. Conf.*, San Diego, CA, USA, Dec. 2015, pp. 1–6.
- [23] T. X. Vu, H. D. Nguyen, T. Q. S. Quek, and S. Sun, "Fronthaul compression and optimization for cloud radio access networks," in *Proc. IEEE Int. Conf. Commun.*, Kuala Lumpur, Malaysia, 2016, pp. 1–6.
- [24] "LTE: Feasibility study for further advancements for E-UTRA (LTE-advanced)," ETSI, Sophia Antipolis Cedex, France, Tech. Rep. ETSI TR 136 912, 2010.
- [25] M. Chiani, D. Dardari, and M. K. Simon, "New exponential bounds and approximations for the computation of error probability in fading channels," *IEEE Trans. Wireless Commun.* vol. 2, no. 4, pp. 840–845, Jul. 2003.
- [26] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.



Thang X. Vu (S'11–M'15) was born in Hai Duong, Vietnam. He received the B.S. and the M.Sc. degrees in electronics and telecommunications engineering from the VNU University of Engineering and Technology, Hanoi, Vietnam, in June 2007 and September 2009, respectively, and the Ph.D. in electrical engineering from the University Paris-Sud, Orsay, France, in January 2014.

From 2007 to 2009, he was with the Department of Electronics and Telecommunications, VNU University of Engineering and Technology, Vietnam, as

a Research Assistant. In 2010, he received the Allocation de Recherche fellowship to study Ph.D. in France. From September 2010 to May 2014, he was with the Laboratory of Signals and Systems, a joint laboratory of CNRS, CentraleSupélec and University Paris-Sud XI, France. From July 2014 to January 2016, he was a Postdoctoral Researcher with the Information Systems Technology and Design pillar, Singapore University of Technology and Design, Singapore. He is currently a Research Associate at the Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg, Luxembourg. His research interests include the field of wireless communications, with particular interests of cache-assisted 5G, cloud radio access networks, resources allocation and optimization, cooperative diversity, channel and network decoding, and iterative decoding.



Hieu Duy Nguyen (S'10–M'14) received the B.S. (First-Class Hons.) degree from Vietnam National University, Hanoi, Vietnam, in 2009, and the Ph.D. degree from National University of Singapore, Singapore, in 2013, all in electrical engineering.

From October 2013 to June 2016, he was with the Advanced Communication Technology Department, Institute for Infocomm Research, Agency for Science Technology and Research, Singapore, where he was a Research Scientist. Since July 2016, he has been with Swiss Re, Singapore, working as a Statistical Model-

ing and Risk Analyst. His research interests included wireless communications and information theory, focusing on distributed antenna systems, cloud radio access networks, large-system analysis, stochastic geometry, and interference channels.



Tony Q. S. Quek (S'98–M'08–SM'12) received the B.E. and M.E. degrees in electrical and electronics engineering from Tokyo Institute of Technology, Tokyo, Japan, and the Ph.D. degree in electrical engineering and computer science from Massachusetts Institute of Technology, Cambridge, MA, USA. He is currently a Tenured Associate Professor with the Singapore University of Technology and Design (SUTD), Singapore, and also serves as the Deputy Director of the SUTD-ZJU Innovation, Design and Entrepreneurship Alliance. His main research interests include the

application of mathematical, optimization, and statistical theories to communication, networking, signal processing, and resource allocation problems, specific current research topics include heterogeneous networks, green communications, wireless security, Internet-of-things, big data processing, and cognitive radio.

He has been actively involved in organizing and chairing sessions, and has served as a member of the Technical Program Committee as well as symposium chairs in a number of international conferences. He is serving as the Workshop Chair for IEEE Globecom in 2017 and the Special Session Chair for IEEE SPAWC in 2017. He is currently an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS and an Executive Editorial Committee Member for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He was an Editor for the IEEE WIRELESS COMMUNICATIONS LETTERS, a Guest Editor for the IEEE SIGNAL PROCESSING MAGAZINE (Special Issue on Signal Processing for the 5G Revolution) in 2014, and the IEEE WIRELESS COMMUNICATIONS MAGAZINE (Special Issue on Heterogeneous Cloud Radio Access Networks) in 2015. He is a co-author of the book "Small Cell Networks: Deployment, PHY Techniques, and Resource Allocation" published by Cambridge University Press in 2013 and the book "Cloud Radio Access Networks: Principles, Technologies, and Applications" by Cambridge University Press.

He received the 2008 Philip Yeo Prize for Outstanding Achievement in Research, the IEEE Globecom 2010 Best Paper Award, the 2012 IEEE William R. Bennett Prize, the IEEE SPAWC 2013 Best Student Paper Award, the IEEE WCSP 2014 Best Paper Award, the IEEE PES General Meeting 2015 Best Paper, and the 2015 SUTD Outstanding Education Awards – Excellence in Research.



Sumei Sun (F'16) is currently the Head of the Communications and Networks Cluster, Institute for Infocomm Research, Agency for Science, Technology, and Research, Singapore. Her research interests include energy- and spectrum-efficient communication technologies for connecting human, machines, and things. She is the Inventor and Co-Inventor of 30 granted patents and more than 30 pending patent applications, many of which have been licensed to industry. She has authored and co-authored more than 200 technical papers in prestigious IEEE journals and

conferences. She has also been actively contributing to organizing conferences in different roles. Some of her recent conference services include the Executive Vice Chair of Globecom 2017, the Symposium Co-Chair of ICC 2015 and 2016, the Track Co-Chair of IEEE VTC 2012 Spring, VTC 2014 Spring, and VTC 2016 Fall, the Publicity Co-Chair of PIMRC 2015, etc. She is an Editor for IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY (TVT) since 2011, an Editor for IEEE WIRELESS COMMUNICATIONS LETTERS from 2011 to 2016, and an Editor of IEEE COMMUNICATIONS SURVEYS AND TUTORIALS since 2015. She received the Top Editor Award in 2016, Top15 Outstanding Editors recognition in 2014, and Top Associate Editor recognition in 2013 and 2012, all from TVT. She is a Distinguished Lecturer of IEEE VEHICULAR TECHNOLOGY SOCIETY 2014–2018, a received the 16th PIMRC Best Paper Award, and Distinguished Visiting Fellow of the Royal Academy of Engineering, UK, in 2014.