CSI Overhead Reduction with Stochastic Beamforming for Cloud Radio Access Networks

Yuanming Shi, Jun Zhang, and Khaled B. Letaief, *Fellow, IEEE*
Dept. of ECE, The Hong Kong University of Science and Technology
E-mail: {yshiac, eejzhang, eekhaled}@ust.hk

Abstract—**Cloud radio access network (Cloud-RAN) is a promising network architecture to meet the explosive growth of the mobile data traffic. In this architecture, as all the baseband signal processing is shifted to a single baseband unit (BBU) pool, interference management can be efficiently achieved through coordinated beamforming, which, however, often requires full channel state information (CSI). In practice, the overhead incurred to obtain full CSI will dominate the available radio resource. In this paper, we propose a unified framework for the CSI overhead reduction and downlink coordinated beamforming. Motivated by the channel heterogeneity phenomena in large-scale wireless networks, we first propose a novel CSI acquisition scheme, called *compressive CSI acquisition*, which will obtain instantaneous CSI of only a subset of all the channel links and statistical CSI for the others, thus forming the mixed CSI at the BBU pool. This subset is determined by the statistical CSI. Then we propose a new stochastic beamforming framework to minimize the total transmit power while guaranteeing quality-of-service (QoS) requirements with the mixed CSI. Simulation results show that the proposed CSI acquisition scheme with stochastic beamforming can significantly reduce the CSI overhead while providing performance close to that with full CSI.**

I. INTRODUCTION

Cloud radio access network (Cloud-RAN) [1], [2] has recently been proposed as a promising network architecture for the future evolution of cellular networks. In this architecture, the baseband processing is moved to a single baseband unit (BBU) pool, which will reduce both the network capital expenditure (CAPEX) and operating expense (OPEX). As a result, the light and cheap remote radio heads (RRH) can be deployed with a high density to improve the spatial reuse efficiency. Moreover, with the centralized signal processing at the BBU pool, the interference can be managed efficiently through coordinated beamforming [3], and the energy efficiency can also be significantly improved through RRH selection supported by group sparse beamforming [4]. However, such cooperation often requires full CSI, and its acquisition will bring excessive signaling overhead and reduce the performance gain. Therefore, effective CSI acquisition methods with low overhead are needed to exploit the full benefits of Cloud-RAN.

CSI acquisition is a key ingredient in wireless communication systems, especially after the rise of the multi-antenna transmission techniques, i.e., MIMO [5]. There has been lots of effort on addressing this issue, but still there is a lack of systematic approach for CSI overhead reduction, especially

in MIMO cooperative networks. For the MIMO broadcast channel, most existing works focus on the limited feedback design [6], which is to reduce the feedback overhead while instantaneous CSI is assumed to be available at each mobile user (MU). For coordinated cellular networks, a beamforming scheme only based on the local CSI was proposed in [7] with the help of large dimensional random matrix theory. Therefore, the backhaul overhead for CSI signaling can be reduced significantly. However, it may not necessarily reduce the CSI acquisition overhead, since all the base stations (BSs) still need to know the channel coefficients from all the users. For the recently proposed massive MIMO system [8], since the number of users is much less than the number of base station antennas, the time-division duplex (TDD) operation is promising with the fact that the CSI acquisition overhead is only proportional to the number of users in each cell. However, pilot contamination during the training phase has been well recognized as a main challenge in massive MIMO systems, which has not been well resolved [9]. For the relay-assisted wireless networks, the CSI acquisition problem is more complicated with the interference alignment method [10].

The large-scale Cloud-RAN makes CSI acquisition more challenging, and the existing CSI overhead reduction strategies are not sufficient. With a huge number of RRHs and MUs in Cloud-RAN, obtaining full CSI at the receivers through pilot training, which is a common assumption in previous works, becomes infeasible as the training period will be comparable to the channel coherence time. Moreover, feeding back all the CSI by the receivers may occupy all the uplink resources. The excessive amount of CSI in Cloud-RAN may be regarded as a type of "curse of dimensionality". Fortunately this new architecture also brings new opportunities for CSI overhead reduction. One key element that can be exploited to reduce the CSI overhead is the *channel heterogeneity* in the large scale network. Namely, different channel links will have different propagation loss and also have different effects on the performance, so the acquisition of the instantaneous information of certain links will only contribute little to the performance gain of network cooperation. Hence, we are interested in answering the following question: **Can we *directly determine* which channel coefficients are required to obtain the instantaneous values before the pilot training**?

In this paper, motivated by the channel heterogeneity in Cloud-RAN, we propose a novel CSI acquisition scheme, called *compressive CSI acquisition*, to overcome the "curse of

dimensionality" for Cloud-RAN. In this scheme, we directly determine the required instantaneous CSI at the very beginning (i.e., before the training phase). We define a novel metric, called the *compression dimension* which counts the total number of required instantaneous channel links, to measure the CSI acquisition overhead in Cloud-RAN. The proposed compressive CSI acquisition scheme can save more overhead compared to most previous works, since the channel compression dimension can be reduced significantly in large-scale Cloud-RAN. Subsequently, the BBU pool will be provided with mixed CSI, i.e., it will get a subset (denoted as Ω) of instantaneous channel coefficients and only statistical CSI for the other channel coefficients. In this paper, we assume that the statistical CSI of all the links is available at the BBU pool.

The main design challenges involve determining the set Ω and the coordinated beamforming with mixed CSI. To determine the set Ω, a naive scheme is based only on large-scale fading, i.e., letting the set Ω consist of the indices of the channel links with the highest large-scale fading. However, such simple scheme will not work well, as large-scale fading cannot fully reflect the effect of each link. Instead, we propose to determine the set Ω based on a certainty-equivalent (CE) approach [11]. For the coordinated beamforming with mixed CSI, one may simply set the unobserved channels to zeros, which, however, is highly suboptimal and cannot guarantee the system QoS requirements. Instead, we propose a stochastic beamforming framework based on the chance-constrained programming (CCP) [12] by exploiting the statistical CSI to guarantee the QoS requirements with a tolerated system outage probability. Simulation results show that the proposed compressive CSI acquisition scheme with stochastic beamforming can significantly reduce the CSI overhead while providing performance close to that with full CSI.

II. System Model and Problem Formulation

We consider a downlink Cloud-RAN with L remote radio heads (RRHs), where the l-th RRH is equipped with N_l antennas, and K single-antenna mobile users (MUs), as shown in Fig. 1. In this architecture, all the base band units (BBUs) are moved into a single BBU pool, creating a set of shared processing resources, and enabling efficient interference management and mobility management. All the RRHs are connected to the BBU pool through the high-capacity transport links. All the user data are available at the BBU pool.

The propagation channel from the l-th RRH to the k-th MU is denoted as $\mathbf{h}_{kl} \in \mathbb{C}^{N_l}, \forall k, l$. Denote $\mathbf{h}_k = [\mathbf{h}_{k1}^T, \mathbf{h}_{k2}^T, \ldots, \mathbf{h}_{kL}^T]^T \in \mathbb{C}^N$, as the channel vector from all the RRHs to the MU k, where $N = \sum_{l=1}^L N_l$. The received signal $y_k \in \mathbb{C}$ at MU k is given by

$$y_k = \sum_{l=1}^L \mathbf{h}_{kl}^H \mathbf{v}_{lk} s_k + \sum_{i \neq k} \sum_{l=1}^L \mathbf{h}_{kl}^H \mathbf{v}_{li} s_i + n_k, \forall k, \quad (1)$$

where s_k is the encoded information symbol for MU k with $\mathbb{E}[|s_k|^2] = 1$, $\mathbf{v}_{lk} \in \mathbb{C}^{N_l}$ is the beamforming vector from the l-th RRH to the k-th MU, and $n_k \sim \mathcal{CN}(0, \sigma_k^2)$ is the additive Gaussian noise at MU k. We assume s_k's and n_k's are mutually independent. We assume all the users apply single
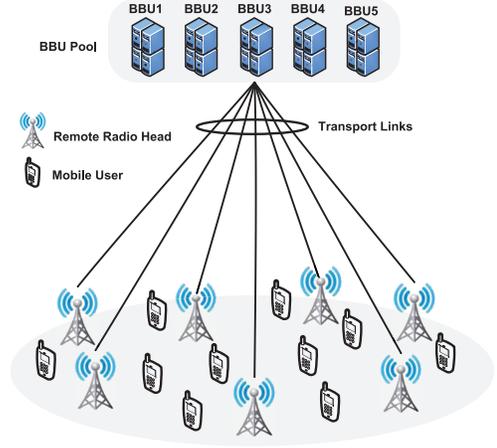

Fig. 1. The architecture of Cloud-RAN, in which, all the RRHs are connected to a BBU pool through high-capacity transport links.

user detection (i.e., treating interference as noise). Then the corresponding signal-to-interference-plus-noise ratio (SINR) for MU k is given by

$$\mathsf{SINR}_k(\mathbf{v}; \mathbf{h}_k) = \frac{|\mathbf{h}_k^H \mathbf{v}_k|^2}{\sum_{i \neq k} |\mathbf{h}_k^H \mathbf{v}_i|^2 + \sigma_k^2}, \forall k. \quad (2)$$

where $\mathbf{v}_k = [\mathbf{v}_{1k}^T, \mathbf{v}_{2k}^T, \ldots, \mathbf{v}_{Lk}^T]^T \in \mathbb{C}^N$ is the beamforming vector from all the RRHs to the MU k, and $\mathbf{v} = (\mathbf{v}_{lk})_{1 \leq l \leq L, 1 \leq k \leq K}$ is the collection of all the beamforming vectors. Each RRH has its own transmit power constraint, given as

$$\sum_{k=1}^K \|\mathbf{v}_{lk}\|^2 \leq P_l, \forall l, \quad (3)$$

where P_l is the maximum transmit power for the l-th RRH. Define the total transmit power as

$$P(\mathbf{v}) = \sum_{l=1}^L \sum_{k=1}^K \|\mathbf{v}_{lk}\|^2. \quad (4)$$

The beamforming vectors $\mathbf{v}_{lk}, \forall l, k$ are designed to minimize the total transmit power while satisfying the QoS requirements for all the MUs, i.e.,

$$\mathscr{P} : \underset{\mathbf{v}}{\text{minimize}} \quad P(\mathbf{v})$$
$$\text{subject to} \quad \mathsf{SINR}_k(\mathbf{v}; \mathbf{h}_k) \geq \gamma_k, \forall k,$$
$$\sum_{k=1}^K \|\mathbf{v}_{lk}\|^2 \leq P_l, \forall l, \quad (5)$$

where γ_k is the target QoS requirement for MU k. This problem can be reformulated as a second order conic programming (SOCP) problem, which is convex and can be solved efficiently (e.g. via interior-point methods). Please refer to [4] for more details. However, solving problem $\mathscr{P}$ requires full CSI, which is impractical for Cloud-RAN. Thus, in the next subsection, we propose a novel CSI acquisition scheme, called *compressive CSI acquisition*, which can be regarded as a first attempt to address the CSI acquisition issue for Cloud-RAN.

A. Compressive CSI Acquisition

We first define the channel model for the Cloud-RAN. Denote $\mathbf{H} = [H_{kn}] \in \mathbb{C}^{K \times N}$ as the full channel matrix:

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}_{11}^{\mathsf{H}} & \cdots & \mathbf{h}_{1L}^{\mathsf{H}} \\ \vdots & \ddots & \vdots \\ \mathbf{h}_{K1}^{\mathsf{H}} & \cdots & \mathbf{h}_{KL}^{\mathsf{H}} \end{bmatrix}. \tag{6}$$

In this paper, we assume the coefficients H_{kn}'s are mutually independent with complex Gaussian distribution. Specifically, we assume $H_{kn} = D_{kn}G_{kn}$, where D_{kn} is the large-scale fading coefficient and $G_{kn} \sim \mathcal{CN}(0,1)$ models the small-scale fading from the n-th antenna to the k-th MU, thus $H_{kn} \sim \mathcal{CN}(0, D_{kn}^2), \forall k, n$. Therefore, the full channel matrix $\mathbf{H}$ can be represented as the Hadamard product of the large-scale fading channel matrix $\mathbf{D} = [D_{kn}]$ and the small-scale fading channel matrix $\mathbf{G} = [G_{kn}]$, i.e., $\mathbf{H} = \mathbf{D} \circ \mathbf{G}$. Due to the *channel heterogeneity*, that is, different links will have different propagation loss and also have different effects on the performance, it is not necessary to obtain the instantaneous information of all the links, as some links will only contribute little to the performance gain of network cooperation. Therefore, in the proposed compressive CSI acquisition scheme, we will only obtain a subset of the instantaneous channel coefficients $H_{kn}, (k,n) \in \Omega$ with the size $|\Omega| = m$. Specifically, at the BBU pool, there are two different types of CSI:

1) *Incomplete Instantaneous CSI*: The BBU pool only knows a subset of instantaneous channel coefficients;
2) *Statistical CSI*: The BBU pool knows the statistics of $H_{kn}, \forall k, n$, e.g., the distribution of H_{kn} and large-scale fading coefficients D_{kn}.

As a result, mixed CSI is available at the BBU pool, which will then be used for the design of coordinated beamforming. Actually, the proposed compressive CSI acquisition scheme follows a similar philosophy with "compressive sensing" [13], [14], in which, instead of first obtaining all the data and then removing the redundancy, we can directly extract useful information with much fewer samples. In our context, similarly, instead of letting all the users first obtain all the instantaneous CSI through downlink training and then probably selectively feed back some of them, we *directly determine* which channel coefficients are critical for the performance and required to obtain the instantaneous values before pilot training.

Although we do not explicitly consider detailed CSI acquisition steps, the proposed framework can be applied in a wide setting: for the FDD system, both the downlink training and feedback overhead can be reduced; for the TDD system, the uplink training overhead can be reduced. Specifically, we define the *compression dimension* as follows:

Definition 1 (Compression Dimension): For the full channel matrix $\mathbf{H} = [H_{kn}] \in \mathbb{C}^{K \times N}$, we define the compression dimension $\mathcal{D}$ as the total number of entries that can be observed, i.e.,

$$\mathcal{D} = |\Omega|, \tag{7}$$

where Ω is the set such that H_{kn} is observed if and only if $(k,n) \in \Omega$ and $|\cdot|$ represents the cardinality operator.

The compression dimension can be regarded as a first-order measurement of the CSI acquisition overhead for Cloud-RAN. For instance, in order to guarantee a constant CSI distortion D, the total CSI feedback bits B should scale as $\mathcal{O}(|\Omega| \log(\frac{1}{D}))$ [6]. We further define $\alpha \triangleq \frac{|\Omega|}{KN}$ as the measurement of the proposed compressive CSI acquisition scheme: the smaller the value of α is, the more significance the compressive CSI acquisition scheme will make on the CSI overhead reduction. In this paper, we will assume a fixed α, while how to determine α will be treated in a subsequent study.

The proposed compressive CSI method involves two main design problems:

1) How to determine the set Ω based on the statistical CSI?
2) How to design the beamformers based on the mixed CSI?

The design problem 1) will be addressed in Section III while problem 2) will be addressed in Section IV.

III. COMPRESSIVE CSI ACQUISITION VIA CERTAINTY-EQUIVALENT FORMULATION

In this section, we will propose a novel CSI acquisition scheme, i.e., compressive CSI acquisition. As this design step is before actually obtaining any instantaneous CSI, it must be based on statistical CSI. One way to approach this goal is the certainty-equivalent (CE) formulation by replacing the random channel gains with their expectation [11]. Specifically, define $\mathbf{Q}_k = \mathbf{v}_k \mathbf{v}_k^{\mathsf{H}}$ with $\mathrm{rank}(\mathbf{Q}_k) = 1$ and $\mathbf{\Lambda}_k = \mathrm{diag}\{\lambda_{k1}, \ldots, \lambda_{kN}\} \in \mathbb{C}^{N \times N}$, where $\lambda_{kn} = D_{kn}^2, \forall k, n$ denotes the channel gain between the k-th MU and the n-th transmit antenna, then the CE SINR at MU k is given by

$$\overline{\mathsf{SINR}}_k = \frac{\mathbf{\Lambda}_k \bullet \mathbf{Q}_k}{\sum_{i \neq k} \mathbf{\Lambda}_k \bullet \mathbf{Q}_i + \sigma_k^2}, \forall k, \tag{8}$$

where $\mathbf{A} \bullet \mathbf{B} \triangleq \mathrm{tr}(\mathbf{AB})$.

We then define the CE QoS constraints as

$$\min_k \frac{\overline{\mathsf{SINR}}_k}{\gamma_k} \geq \mu, \tag{9}$$

where μ is a regularized parameter to measure the penalty due to the certainty equivalent approximation. Therefore, the CE QoS constraints are given by

$$\mathcal{C}_1(\mathbf{\Lambda}): \mathbf{\Lambda}_k \bullet \mathbf{Q}_k - (\mu\gamma_k) \sum_{i \neq k} \mathbf{\Lambda}_k \bullet \mathbf{Q}_i \geq (\mu\gamma_k)\sigma_k^2, \forall k,$$

where $\mathbf{\Lambda} \triangleq (\mathbf{\Lambda}_k)_{1 \leq k \leq K}$. The above QoS constraints can be rewritten as

$$\sum_{n=1}^{N} \pi_{kn} \geq (\mu\gamma_k)\sigma_k^2, \forall k, \tag{10}$$

where π_{kn} is defined as the *network channel gain* between the n-th transmit antenna and the k-th MU to measure the impact of the corresponding channel gain λ_{kn}:

$$\pi_{kn}(\mathbf{Q}) = \lambda_{kn} \underbrace{\left(\mathbf{I}_N^n \bullet \mathbf{Q}_k - (\mu\gamma_k) \sum_{i \neq k} \mathbf{I}_N^n \bullet \mathbf{Q}_i \right)}_{\text{channel gain encoding}}, \forall k, n, \tag{11}$$

with $\mathbf{Q} = (\mathbf{Q}_k)_{k=1}^K$ and $\mathbf{I}_N^n$ as an $N \times N$ matrix with one at the n-th diagonal entry and zeros otherwise. Intuitively, given the precoding matrices $\mathbf{Q}_k$'s, the smaller the coefficients $\pi_{kn}(\mathbf{Q})$ are, the less the corresponding channels (k, n) will contribute to the QoS constraints. Therefore, we propose the compressive CSI scheme based on this intuition. Specifically, define the "encoded" channel gain matrix as $\mathbf{\Pi}(\mathbf{Q}) = [\pi_{kn}(\mathbf{Q})]$, then the set Ω of size $\mathcal{D}$ is determined by including the indices of the $\mathcal{D}$ largest coefficients $\pi_{kn}(\mathbf{Q})$'s.

In order to determine the precoding matrices $\mathbf{Q}_k$'s, we minimize the total transmit power with the CE QoS constraints:

$$\mathscr{P}_{\text{CE}} : \underset{\mathbf{Q}_1,\ldots,\mathbf{Q}_K}{\text{minimize}} \quad \sum_{k=1}^K \mathbf{I} \bullet \mathbf{Q}_k$$
$$\text{subject to} \quad \mathcal{C}_1(\mathbf{\Lambda}), \mathbf{Q}_k \succeq \mathbf{0}, \text{rank}(\mathbf{Q})_k = 1, \forall k,$$
$$\sum_{k=1}^K \mathbf{B}_{lk} \bullet \mathbf{Q}_k \leq P_l, \forall l, \tag{12}$$

where $\mathbf{B}_{lk}$ is a diagonal matrix with the l-th diagonal block matrix as $\mathbf{I}_{N_l}$. We can always assume the problem $\mathscr{P}_{\text{CE}}$ is feasible by choosing an appropriate parameter μ. One key difficulty for problem $\mathscr{P}_{\text{CE}}$ is the rank-one constraints for $\mathbf{Q}_k$'s. By the idea of semi-definite relaxation (SDR), we first drop the rank-one constraints of $\mathbf{Q}_k$'s to relax the above problem to a semi-definite programming (SDP) problem. If the rank-one solution is failed to be obtained, the Gaussian randomization method [15] will be employed to obtain a rank-one approximation solution to $\mathscr{P}_{\text{CE}}$. We define the rank-one solution of $\mathscr{P}_{\text{CE}}$ as $\mathbf{Q}_k^\star, \forall k$, thus the network channel gain for the channel link (k, n) is given by $\pi_{kn}(\mathbf{Q}^\star)$.

Although the compressive CSI acquisition based on the CE formulation cannot provide any performance guarantee, i.e., it is difficult to tell the performance loss compared to the case with full CSI, it provides a systematic approach to determine the set Ω. Moreover, through simulation, it will be shown that this design is quite effective and significantly outperforms a naive scheme only based on large-scale fading.

IV. STOCHASTIC BEAMFORMING VIA CHANCE-CONSTRAINED PROGRAMMING

In the last section, we proposed a CE formulation based compressive CSI acquisition scheme to identify the set Ω. The BBU pool will then obtain instantaneous CSI of the channel links indexed by the set Ω, and in this paper we assume such information is perfect. With the mixed CSI, the system can not guarantee the QoS requirements for all the channel realizations. The robust optimization methodology may be adopted to handle the uncertainty in the channel state information by assuming the unobserved channel coefficients lie in a certain bounded set, called the *uncertainty set*. However, the solutions obtained by using the robust optimization methodology can be very conservative in general and the uncertainty set is also difficult to define especially for the general network setting.

Actually, in practical wireless systems, most applications do not require to satisfy the QoS requirements all the time, i.e., most systems can tolerate some outage in QoS. This motivates

us to impose the following probabilistic QoS constraint for the system:

$$\Pr\{\text{SINR}_k(\mathbf{v}; \boldsymbol{\omega}, \boldsymbol{\xi}) \geq \gamma_k, \forall k\} \geq 1 - \epsilon, \tag{13}$$

where $\epsilon \in [0, 1)$ is to indicate that the system should guarantee the QoS requirements with probability at least $1 - \epsilon$, $\boldsymbol{\omega} = (H_{kn})_{(k,n)\in\Omega} \in \mathbb{C}^{\mathcal{D}}$ is a deterministic vector containing the available instantaneous CSI of links indexed by the set Ω of size $\mathcal{D}$ and $\boldsymbol{\xi} = (H_{kn})_{(k,n)\notin\Omega} \in \mathbb{C}^{KN-\mathcal{D}}$ is a random vector containing the statistical CSI of other links. Therefore, we propose the following stochastic beamforming framework to minimize the total transmit power while satisfying the QoS requirements with probability at least $1 - \epsilon$:

$$\mathscr{P}_{\text{CCP}} : \underset{\mathbf{v}\in\mathcal{V}}{\text{minimize}} \quad P(\mathbf{v})$$
$$\text{subject to} \quad \Pr\{\text{SINR}_k(\mathbf{v}; \boldsymbol{\omega}, \boldsymbol{\xi}) \geq \gamma_k, \forall k\} \geq 1 - \epsilon,$$

which is called the chance-constrained programming (CCP) [16] in stochastic optimization literature. Define the set $\mathcal{V}$ as the collection of $\mathbf{v}_{lk}$'s that satisfies the per-RRH power constraints:

$$\mathcal{V} := \left\{ \mathbf{v}_{lk} : \sum_{k=1}^K \|\mathbf{v}_{lk}\|^2 \leq P_l, \forall l, k \right\}, \tag{14}$$

which is a convex set.

Although there are enough practical insights behind this formulation, solving $\mathscr{P}_{\text{CCP}}$ has two major challenges in general. First, the probabilistic constraint has no closed-form in general and is difficult to evaluate. Second, the convexity of the feasible set formed by the probabilistic constraint is difficult to verify even though the set formed by the QoS constraint is convex in $\mathbf{v}$. To circumvent these difficulties, in this paper, we propose to use a safe tractable approximation method, called the *scenario approach* [12], to find a feasible solution with high probability to the problem $\mathscr{P}_{\text{CCP}}$.

A. The Scenario Approach for Stochastic Beamforming

Denote $\xi_j, j = 1, 2, \ldots, J$ as the independent samples of $\boldsymbol{\xi}$ that can be generated from Monte Carlo simulation. Given the J samples $\mathcal{H}_J = (\xi_j)_{1\leq j\leq J}$, the probabilistic constrain can be approximated by the following KJ QoS constraints:

$$\text{SINR}_k(\mathbf{v}; \boldsymbol{\omega}, \xi_j) \geq \gamma_k, \forall k, j, \tag{15}$$

which indicate that the QoS requirements for all the J channel realizations $(\boldsymbol{\omega}, \xi_j)$ must be satisfied. Then we need to solve the following scenario approximation problem with the given multiple channel samples $\mathcal{H}_J$:

$$\mathscr{P}_{\text{CCP}}(\mathcal{H}_J) : \underset{\mathbf{v}\in\mathcal{V}}{\text{minimize}} \quad P(\mathbf{v})$$
$$\text{subject to} \quad \text{SINR}_k(\mathbf{v}; \boldsymbol{\omega}, \xi_j) \geq \gamma_k, \forall k, j, \tag{16}$$

which can also be reformulated as an SOCP problem. Define the optimal solution as $\hat{\mathbf{v}}(\mathcal{H}_J)$, which is a unique solution to $\mathscr{P}_{\text{CCP}}(\mathcal{H}_J)$. Intuitively, if the number of samples is reasonably large, the solution $\hat{\mathbf{v}}(\mathcal{H}_J)$ will satisfy the probabilistic constraint in $\mathscr{P}_{\text{CCP}}$ with a high probability. Specifically, given

the beamformer $\mathbf{v}$, define a function $f : \mathcal{V} \to [0,1]$ as the probability of the QoS satisfiability:

$$f(\mathbf{v}) = \text{Pr}\{\text{SINR}_k(\mathbf{v};\boldsymbol{\omega},\boldsymbol{\xi}) \geq \gamma_k, \forall k\}, \qquad (17)$$

where the probability is calculated over the random vector $\boldsymbol{\xi}$ conditioned on vectors $\mathbf{v}$ and $\boldsymbol{\omega}$. Since the solution $\hat{\mathbf{v}}(\mathcal{H}_J)$ depends on random realizations ξ_j's, $f(\hat{\mathbf{v}}(\mathcal{H}_J))$ is a random variable. Now, we are interested in evaluating the quantity $\text{Pr}\{f(\hat{\mathbf{v}}(\mathcal{H}_J)) \geq 1 - \epsilon\}$ over all the possible realization $\mathcal{H}_J$.

In the following theorem, we give a bound for the probability quantity $\text{Pr}\{f(\hat{\mathbf{v}}(\mathcal{H}_J)) \geq 1 - \epsilon\}$ based on the result in the following theorem.

Theorem 1: [12] Let $\hat{\mathbf{v}}(\mathcal{H}_J)$ be the unique solution to $\mathscr{P}_{\text{CCP}}(\mathcal{H}_J)$, and $\epsilon \in [0,1)$ be the maximum outage probability that the system can tolerate, then we have

$$\text{Pr}\{f(\hat{\mathbf{v}}(\mathcal{H}_J)) \geq 1 - \epsilon\} \geq 1 - \sum_{i=1}^{NK-1} \binom{J}{i} \epsilon^i (1-\epsilon)^{J-i},$$

where the probability is calculated over all possible realizations of the random samples $\mathcal{H}_J$.

Let $\beta \in (0,1)$ satisfy the following equation:

$$\sum_{i=1}^{NK-1} \binom{J}{i} \epsilon^i (1-\epsilon)^{J-i} = \beta, \qquad (18)$$

where the left-hand side of the equation is the cumulative distribution of a binomial random variable and denote the solution as J^*. This theorem indicates that the solution $\hat{\mathbf{v}}(\mathcal{H}_{J*})$ will be feasible for $\mathscr{P}_{\text{CCP}}$ with probability at least $1 - \beta$.

In order to derive an analytic expression for the minimum J^*, we use the Chernoff's inequality [17] to yield an upper bound for $\sum_{i=1}^{NK-1} \binom{J}{i} \epsilon^i (1-\epsilon)^{J-i}$ as $\exp\{-(J\epsilon - JK + 1)^2/2J\epsilon\}$ when $NK \leq J\epsilon$. Thus, by solving the following equation:

$$\beta = \exp\left\{-\frac{(J\epsilon - NK + 1)^2}{2J\epsilon}\right\}, \qquad (19)$$

we obtain the approximately minimum value for J as

$$J^\star = \frac{1}{\epsilon}\left[(NK-1) + \ln\frac{1}{\beta} + \sqrt{\ln^2\frac{1}{\beta} + 2(NK-1)\ln\frac{1}{\beta}}\right].$$

Finally, we can conclude that with $J \geq J^\star$ realizations to form the scenario approximation problem $\mathscr{P}_{\text{CCP}}(\mathcal{H}_J)$, the solution $\hat{\mathbf{v}}(\mathcal{H}_J)$ will be feasible for $\mathscr{P}_{\text{CCP}}$ with probability at least $1 - \beta$, which is called the *confidence level*.

Combined with the proposed CSI acquisition scheme presented in the last section, we now present our proposed compressive CSI acquisition framework for Cloud-RAN as Algorithm 1. Note that, in this algorithm, the update of Step 1 and 2 will not be frequent, since decisions made in these steps are only based on the statistical CSI, which varies much slowly than small-scale fading.

V. SIMULATION RESULTS

In this section, we simulate the performance of the proposed CSI acquisition framework. We consider the following channel model:

$$\mathbf{h}_{kl} = \underbrace{10^{-L(d_{kl})/20}\sqrt{\varphi_{kl}s_{kl}}}_{D_{kl}} \mathbf{g}_{kl}, \forall k, l, \qquad (20)$$

Algorithm 1: Compressive CSI Acquisition Framework for Cloud-RAN

Step 1: Based on the statistical CSI $\mathbf{D}$, solve the CE problem $\mathscr{P}_{\text{CE}}$ and obtain the network channel gains π_{kn}'(11)

Step 2: Determine the set Ω consisting the indices of the $\mathcal{D}$ largest π_{kn}'s, where $\mathcal{D}$ is an integer with $|\Omega| = \mathcal{D}$;

Step 3: Based on the set Ω, the BBU pool will allocate the pilot symbols for the CSI acquisition;

Step 4: After the CSI acquisition procedure, based on the mixed CSI, solve the stochastic beamforming problem $\mathscr{P}_{\text{CCP}}$ and obtain the beamformers $\mathbf{v}_{lk}$'s for the downlink transmission;

End

TABLE I
SIMULATION PARAMETERS

Parameter	Value
Pathloss at distance d_{kl} (km)	$148.1 + 37.6 \log_2(d_{kl})$
Standard deviation of log-norm shadowing σ_s	8 dB
Small-scale fading distribution $\mathbf{g}_{kl}$	$\mathcal{CN}(\mathbf{0}, \mathbf{I})$
Noise power σ_k^2 (10 MHz bandwidth)	-102 dBm
Maximum transmit power of RRH P_l	10 W
Transmit antenna power gain	9 dBi

where $L(d_{kl})$ is the path-loss at distance d_{kl}, as given in Table I, s_{kl} is the shadowing coefficient, φ_{kl} is the antenna gain and $\mathbf{g}_{kl}$ is the small-scale fading coefficient. We use the standard cellular network parameters as showed in Table I. The maximum outage probability that the system can tolerate is set as $\epsilon = 0.1$ and β is set to be 0.01. Since the required number of samples of the scenario approach is in the order of $\mathcal{O}(\frac{1}{\epsilon}NK)$, in the simulations, we only consider a small size network with $L = 5$ single-antennas RRHs and $K = 5$ single-antenna MUs. The compression dimension is $|\Omega| = \mathcal{D} = 20$.

We first consider a fixed network with the large-scale fading coefficients matrix $\mathbf{D} = [D_{kl}]$ as follows:

$$\mathbf{D} = \begin{bmatrix} 32.77 & 16.09 & 11.89 & 12.63 & 15.30 \\ 8.33 & 15.76 & 14.12 & 271.34 & 72.48 \\ 1.75 & 35.46 & 11.87 & 274.75 & 143.54 \\ 26.37 & 120.35 & 2.85 & 36.18 & 45.93 \\ 5.81 & 23.69 & 16.19 & 156.73 & 33.89 \end{bmatrix}.$$

For comparison, we consider a "naive" scheme for compressive CSI acquisition, in which, the set Ω consists the indices of the $\mathcal{D}$ largest large-scale fading coefficients in $\mathbf{D}$. Therefore, for this particular setting, only the statistical CSI of the channel links indexed by the set $\{(2,1),(3,1),(3,3),(4,3),(5,1)\}$ is available. On the other hand, using our proposed compressive CSI acquisition scheme, the instantaneous CSI of the channel links indexed by the set $\{(2,3),(2,5),(3,4),(5,2),(5,4)\}$ *need not* be obtained. Fig. 2 demonstrates the total transmit power with different target SINRs using different CSI acquisition schemes. Each point of the simulation results is averaged over 100 channel realizations. This figure shows that a significant performance gain can be obtained by using our proposed compressive CSI acquisition scheme compared to the
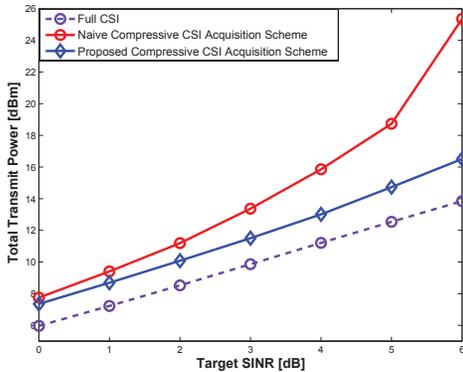
approach for the beamforming design with mixed CSI.

VI. CONCLUSIONS AND DISCUSSIONS

In this paper, we proposed a unified framework consisting of a novel compressive CSI acquisition scheme and coordinated stochastic downlink beamforming with mixed CSI for Cloud-RAN. The obtained results from this novel framework are very promising, and further investigation is needed. The proposed compressive CSI acquisition scheme demonstrated the effectiveness of exploiting the channel heterogeneity for CSI overhead reduction, and its performance with practical CSI training/feedback requires further investigation. We proposed to use the scenario approach to find a feasible solution with QoS guarantee to the stochastic beamforming problem for its simplicity. However, the required number of samples will increase rapidly for the large size networks. Therefore, more efficient algorithms are needed for practical implementations.

REFERENCES

[1] China Mobile, "C-RAN: the road towards green RAN," *White Paper, ver. 2.5*, Oct. 2011.
[2] J. Wu, "Green wireless communications: from concept to reality [industry perspectives]," *IEEE Wireless Commun.*, vol. 19, pp. 4–5, Aug. 2012.
[3] D. Gesbert, S. Hanly, H. Huang, S. Shamai Shitz, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: A new look at interference," *IEEE J. Sel. Areas Commun.*, vol. 28, pp. 1380–1408, Sep. 2010.
[4] Y. Shi, J. Zhang, and K. Letaief, "Group sparse beamforming for green Cloud-RAN," *IEEE Trans. Wireless Commun.*, to appear.
[5] L. Zheng and D. Tse, "Communication on the Grassmann manifold: a geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 2, pp. 359–383, 2002.
[6] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, pp. 5045–5060, Nov. 2006.
[7] R. Zakhour and S. V. Hanly, "Min-max power allocation in cellular networks with coordinated beamforming," *IEEE J. Sel. Areas Commun.*, vol. 31, pp. 287–302, Feb. 2013.
[8] T. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 3590–3600, Nov. 2010.
[9] J. Jose, A. Ashikhmin, T. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Commun.*, vol. 10, pp. 2640–2651, Aug. 2011.
[10] X. Chen, S. Song, and K. Letaief, "Interference alignment in dual-hop MIMO interference channel," *IEEE Trans. Wireless Commun.*, vol. PP, no. 99, pp. 1–10, 2014.
[11] S. Kandukuri and S. Boyd, "Optimal power control in interference-limited fading wireless channels with outage-probability specifications," *IEEE Trans. Wireless Commun.*, vol. 1, no. 1, pp. 46–55, 2002.
[12] M. C. Campi and S. Garatti, "The exact feasibility of randomized solutions of uncertain convex programs," *SIAM J. Optim.*, vol. 19, no. 3, pp. 1211–1230, 2008.
[13] D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, pp. 1289–1306, Apr. 2006.
[14] E. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?," *IEEE Trans. Inf. Theory*, vol. 52, pp. 5406–5425, Dec. 2006.
[15] Z.-Q. Luo, W.-K. Ma, A.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, pp. 20–34, May 2010.
[16] A. Charnes and W. W. Cooper, "Chance-constrained programming," *Management science*, vol. 6, no. 1, pp. 73–79, 1959.
[17] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," *The Annals of Mathematical Statistics*, vol. 23, no. 4, pp. 493–507, 1952.
[18] Y. Shi, J. Zhang, and K. B. Letaief, "Optimal stochastic coordinated beamforming with compressive CSI acquisition for Cloud-RAN," *arXiv preprint arXiv:1312.0363*, 2013.

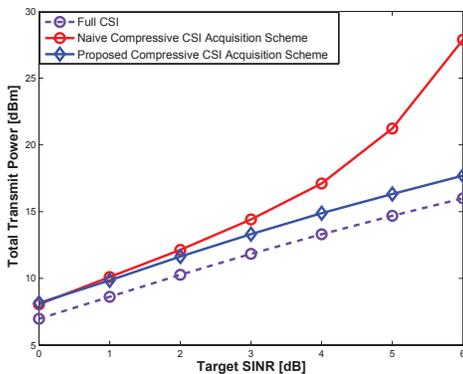Fig. 2. Total transmit power versus target SINR under a fixed network deployment.

Fig. 3. Total transmit power versus target SINR under random remote radio head positions and random mobile user positions.

naive scheme. Moreover, the proposed stochastic beamforming scheme is also efficient in terms of total transmit power minimization while guaranteeing the QoS requirements with a high probability.

We then consider a random setting, in which the positions of the RRHs and MUs are randomly distributed in the area $[-200, 200] \times [-200, 200]$ meters. Fig. 3 presents the total transmit power with different target SINRs using different compressive CSI acquisition schemes. Each point of the simulation results is averaged over 10 network topology realizations, and 10 small-scale fading realizations for each network topology realization. This figure further confirms the effectiveness of our proposed compressive CSI acquisition scheme. Note that the performance gap between the full CSI case and the stochastic beamforming case is from two parts: the first one is the lack of full CSI and the second one is that only a feasible solution (which will yield an *upper* bound for the total transmit power) of the stochastic beamforming can be obtained, i.e., the gap is partially due to the optimization algorithm itself. Due to the limited space of this paper, pleas refer to [18] for more simulation results to further justify the proposed compressive CSI acquisition method and the scenario