On the Performance of Channel-Statistics-Based Codebook for Massive MIMO Channel Feedback

Wenqian Shen, Linglong Dai, Yu Zhang, Jianjun Li,
and Zhaocheng Wang

Abstract—The channel feedback overhead for massive multiple-input multiple-output systems with a large number of base station (BS) antennas is very high since the number of feedback bits of traditional codebooks scales linearly with the number of BS antennas. To reduce the feedback overhead, an effective codebook based on channel statistics has been designed, where the required number of feedback bits only scales linearly with the rank of the channel correlation matrix. However, this attractive conclusion was only proved under a particular channel assumption in the literature. To provide a rigorous theoretical proof under a general channel assumption, in this paper, we quantitatively analyze the performance of the channel-statistics-based codebook. Specifically, we first introduce the rate gap between the ideal case of perfect channel state information at the transmitter and the practical case of limited channel feedback, where we find that the rate gap depends on the quantization error of the codebook. Then, we derive an upper bound of the quantization error, based on which we prove that the required number of feedback bits to ensure a constant rate gap only scales linearly with the rank of the channel correlation matrix. Finally, numerical results are provided to verify this conclusion.

Index Terms—Channel feedback, codebook, massive multiple-input multiple-output (MIMO), performance analysis.

I. INTRODUCTION

Massive multiple-input multiple-output (MIMO) using hundreds of antennas at the base station (BS) is one of the key technologies for future 5G wireless communications [1]. To achieve the expected high spectrum efficiency and energy efficiency of massive MIMO, accurate channel state information at the transmitter (CSIT) is crucial [2], [3]. Utilizing the channel reciprocity, CSIT can be obtained from uplink channel estimation in time division duplexing (TDD) systems [4]. While most of the existing works on massive MIMO consider TDD mode due to this reason, frequency division duplexing (FDD) has many benefits over TDD and thus still dominates current cellular networks [5]. Therefore, it is important to study the CSIT acquisition problem for FDD massive MIMO systems.

However, the channel reciprocity does not exist in FDD systems, so accurate channel feedback from users to the BS is required. Traditional codebooks for channel feedback such as Grassmannian codebook

[6], [7] and random vector quantization (RVQ) based codebook [8], [9] have been extensively investigated. For these codebooks, there is an important conclusion that, to maintain a constant capacity degradation due to channel quantization error, the required number of feedback bits approximately scales linearly with the number of BS antennas [9]. Thus, for massive MIMO with a large number of BS antennas, the channel feedback overhead will be overwhelming. Therefore, several new codebooks have been proposed to reduce the channel feedback overhead. Based on the assumption of high correlation among BS antennas, an antenna-grouping-based feedback scheme was proposed in [10]. Another promising solution is the channel-statistics-based codebook [11] that is designed by multiplying each vector of an original codebook (e.g., RVQ-based codebook) by the channel correlation matrix. It has been verified through extensive simulations that the number of feedback bits required by this channel-statistics-based codebook only scales linearly with the rank of the channel correlation matrix [12]. As the rank of the channel correlation matrix is much smaller than the number of BS antennas in massive MIMO [13], the feedback overhead can be reduced. However, Clerckx *et al.* [12] only provided a proof of this attractive conclusion under a particular channel assumption, where the singular values of the square root of the channel correlation matrix are the same except for the first one, whereas the rigorous proof under a general channel assumption has not been provided in the literature.

In this paper, we quantitatively analyze the performance of the channel-statistics-based codebook under a general channel assumption.[1] Specifically, we first introduce the rate gap between the ideal case of perfect CSIT and the practical case of limited channel feedback. We find that the rate gap depends on the quantization error of the codebook. Although the quantization error is difficult to obtain, we are able to derive an upper bound of the quantization error by magnifying and shrinking the involved inequation. After that, by substituting the upper bound of quantization error into the rate gap expression, we can obtain the upper bound of the rate gap. Finally, we prove that the number of feedback bits required to ensure a constant rate gap scales linearly with the rank of the channel correlation matrix. To the best of our knowledge, this work is the first one to provide a rigorous proof of this conclusion under a general channel assumption.

Notation: Lower case and upper case boldface letters denote vectors and matrices, respectively, $(\cdot)^H$ and $(\cdot)^{-1}$ denote the conjugate transpose and inverse of a matrix, respectively, $\mathbf{I}_K$ denotes the identity matrix of size $K \times K$, $\mathrm{E}[\cdot]$ denotes the expectation operator, and $\Pr\{A\}$ denotes the probability of A.

II. SYSTEM MODEL

In this section, we first introduce the massive MIMO channel model. Then, we present the limited channel feedback. Finally, we review the per-user rate.

A. Massive MIMO Channel Model

In this paper, we consider a massive MIMO system with M antennas at the BS and K single-antenna users ($M \gg K$). The downlink channel vector $\mathbf{h}_k \in \mathbb{C}^{M \times 1}$ for the kth user can be described as [12]

$$\mathbf{h}_k = \mathbf{R}_k^{1/2} \mathbf{h}_{w,k} \qquad (1)$$

---

[1]Simulation codes are provided to reproduce the results presented in this paper: http://oa.ee.tsinghua.edu.cn/dailinglong/publications/publications.html.

where $\mathbf{R}_k^{1/2} \in \mathbb{C}^{M \times M}$ is the square root of the kth user's channel correlation matrix[2] $\mathbf{R}_k \in \mathbb{C}^{M \times M}$, and $\mathbf{h}_{w,k} \in \mathbb{C}^{M \times 1}$ is a vector whose elements are independent identically distributed complex Gaussian distributed with zero mean and unit variance. Furthermore, $\mathbf{R}_k^{1/2}$ can be decomposed as $\mathbf{U}_k \mathbf{\Lambda}_k^{1/2} \mathbf{U}_k^H$, where $\mathbf{U}_k$ is a unitary matrix, and $\mathbf{\Lambda}_k^{1/2}$ is a diagonal matrix with the diagonal elements denoted by $\{\sigma_1, \sigma_2, \ldots, \sigma_r, 0, \ldots\}$, where r is the rank of correlation matrix. The concatenation of K channel vectors can be denoted by $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_K] \in \mathbb{C}^{M \times K}$.

B. Limited Channel Feedback

Although the training overhead to obtain the downlink channel vector at the user side is increased in massive MIMO systems, there are some recently proposed effective downlink training methods [13],[14] proposed with reduced training overhead. Thus, in this paper, each user is assumed to know its channel vector $\mathbf{h}_k$. Such information is required at the BS via limited feedback channel. Each user quantizes its channel to B bits and then feeds them back to the BS. Quantization is realized by using a quantization codebook, which is known to the BS and users.

For the traditional RVQ-based codebook [9] $\mathcal{W} = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{2^B}\}$, the unit-norm column vector $\mathbf{w}_i \in \mathbb{C}^{M \times 1}$ is randomly generated by selecting vector independently from the uniform distribution on the complex unit sphere. The required number of feedback bits scales linearly with the number of BS antennas to ensure a constant capacity degradation [9]. Thus, the channel feedback overhead becomes overwhelming for massive MIMO with a large number of BS antennas. In order to reduce the channel feedback overhead, a more effective codebook has been designed based on the channel statistics [11]. Specifically, the quantization vector $\mathbf{c}_{k,i} \in \mathbb{C}^{M \times 1}$ in the kth users' codebook $\mathcal{C}_k$ can be obtained by multiplying the vector $\mathbf{w}_i$ by the square root of the channel correlation matrix $\mathbf{R}_k^{1/2}$, i.e., $\mathbf{c}_{k,i} = \mathbf{R}_k^{1/2} \mathbf{w}_i$. Note that to ensure the unit-norm vector requirement of $\mathbf{c}_{k,i} \in \mathbb{C}^{M \times 1}$, $\mathbf{c}_{k,i}$ should be normalized as

$$\mathbf{c}_{k,i} = \frac{\mathbf{R}_k^{1/2} \mathbf{w}_i}{\|\mathbf{R}_k^{1/2} \mathbf{w}_i\|}. \tag{2}$$

The distribution of the resulted quantization vector $\mathbf{c}_{k,i}$ is closer to the distribution of actual channel vector $\mathbf{h}_k$ in (1), thus, the channel-statistics-based codebook has better quantization performance [10]–[12]. Accordingly, we mainly consider the channel-statistics-based codebook in this paper.

The kth user quantizes its own channel $\mathbf{h}_k$ to a quantization vector $\mathbf{c}_{k,F_k}$ that is closest to $\mathbf{h}_k$, where "closeness" is measured by the angle between two vectors. Thus, user k computes the quantization index F_k according to

$$F_k = \underset{i \in [1,2^B]}{\arg\min} \sin^2(\angle(\mathbf{h}_k, \mathbf{c}_{k,i})) = \underset{i \in [1,2^B]}{\arg\max} |\tilde{\mathbf{h}}_k^H \mathbf{c}_{k,i}|^2 \tag{3}$$

where the normalized channel vector $\tilde{\mathbf{h}}_k = \frac{\mathbf{h}_k}{\|\mathbf{h}_k\|}$ denotes the direction of channel vector. Note that only the direction of channel vector is quantized, whereas the channel magnitude $\|\mathbf{h}_k\|$ is not quantized by using codebook $\mathbf{C}$. Magnitude information is just a scalar value that can be easily fed back. In this paper, we focus on the quantization of channel direction. After that, with the received index F_k, the BS can obtain the feedback channel vector $\hat{\mathbf{h}}_k = \|\mathbf{h}_k\| \mathbf{c}_{k,F_k}$. The concatenation of the feedback channel vectors can be denoted as $\hat{\mathbf{H}} = [\hat{\mathbf{h}}_1, \hat{\mathbf{h}}_2, \ldots, \hat{\mathbf{h}}_K] \in \mathbb{C}^{M \times K}$.

---

[2]The channel correlation matrix $\mathbf{R}_k$ is closely related to the antenna spacing at the BS and the user location, which is the second-order channel statistics that can be usually assumed to be static due to the moderate user velocities [11].

C. Downlink Precoding and Per-User Rate

We can utilize the widely used zero-forcing (ZF) precoding at the BS based on the feedback channel matrix $\hat{\mathbf{H}}$ to eliminate interferences among multiple users. The transmitted signal $\mathbf{x} \in \mathbb{C}^{M \times 1}$ after precoding at the BS is given by

$$\mathbf{x} = \sqrt{\frac{\gamma}{K}} \mathbf{V} \mathbf{s} \tag{4}$$

where γ is the transmit power, $\mathbf{s} = [s_1, s_2, \ldots, s_K] \in \mathbb{C}^{K \times 1}$ is the symbol vector intended for K users with the normalized power $\mathrm{E}[|s_i|^2] = 1$, and $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_K] \in \mathbb{C}^{M \times K}$ is the precoding matrix consisting of K different M-dimensional unit-norm precoding vector $\mathbf{v}_i \in \mathbb{C}^{M \times 1}$. We denote $\mathbf{U} = \hat{\mathbf{H}}(\hat{\mathbf{H}}^H \hat{\mathbf{H}})^{-1}$, then the precoding vectors $\mathbf{v}_i$ can be obtained as the normalized ith column of $\mathbf{U}$, i.e., $\mathbf{v}_i = \frac{\mathbf{U}(:,i)}{\|\mathbf{U}(:,i)\|}$.

The received signal y_k at the kth user can be described as

$$y_k = \mathbf{h}_k^H \mathbf{x} + n_k$$

$$= \sqrt{\frac{\gamma}{K}} \mathbf{h}_k^H \mathbf{v}_k s_k + \sqrt{\frac{\gamma}{K}} \sum_{i=1, i \neq k}^{K} \mathbf{h}_k^H \mathbf{v}_i s_i + n_k \tag{5}$$

where n_k is the complex Gaussian noise at the kth user with zero mean and unit variance. Thus, the signal-to-interference-plus-noise ratio (SINR) at the kth user is [9]

$$\mathrm{SINR}_k = \frac{\frac{\gamma}{K} |\mathbf{h}_k^H \mathbf{v}_k|^2}{1 + \frac{\gamma}{K} \sum_{i=1, i \neq k}^{K} |\mathbf{h}_k^H \mathbf{v}_i|^2}. \tag{6}$$

Accordingly, the per-user rate R is

$$R = \mathrm{E}\left[\log_2\left(1 + \frac{\frac{\gamma}{K} |\mathbf{h}_k^H \mathbf{v}_k|^2}{1 + \frac{\gamma}{K} \sum_{i=1, i \neq k}^{K} |\mathbf{h}_k^H \mathbf{v}_i|^2}\right)\right]. \tag{7}$$

Clearly, the per-user rate depends on the precoding matrix $\mathbf{V}$, which is significantly affected by the quality of feedback channel $\hat{\mathbf{H}}$. In the next section, we will analyze the per-user rate when the channel-statistics-based codebook is applied.

III. Performance Analysis

In this section, we first compute the rate gap. Then, we analyze the quantization error of the channel-statistics-based codebook. Finally, we derive an upper bound of the required feedback bits to ensure a constant rate gap.

A. Rate Gap

Providing the ideal case of perfect CSIT at the BS, i.e., $\hat{\mathbf{H}} = \mathbf{H}$, the ZF precoding vector $\mathbf{v}_{\text{ideal},i}$ is obtained as the normalized ith column of $\mathbf{H}(\mathbf{H}^H \mathbf{H})^{-1}$. Therefore, we can obtain the ideal per-user rate as

$$R_{\text{ideal}} = \mathrm{E}\left[\log_2\left(1 + \frac{\gamma}{K} |\mathbf{h}_k^H \mathbf{v}_{\text{ideal},k}|^2\right)\right]. \tag{8}$$

However, in the practical case of limited channel feedback, the BS can only obtain the feedback channel $\hat{\mathbf{H}}$ using the channel-statistics-based codebook. The ZF precoding is performed based on $\hat{\mathbf{H}}$, and the precoding vector $\mathbf{v}_i$ is obtained as the normalized ith column of $\hat{\mathbf{H}}(\hat{\mathbf{H}}^H \hat{\mathbf{H}})^{-1}$. Thus, the interuser interference $|\mathbf{h}_k^H \mathbf{v}_i| \neq 0$, which degrades the per-user rate as

$$R_{\text{practical}} = \mathrm{E}\left[\log_2\left(1 + \frac{\frac{\gamma}{K} |\mathbf{h}_k^H \mathbf{v}_k|^2}{1 + \frac{\gamma}{K} \sum_{i=1, i \neq k}^{K} |\mathbf{h}_k^H \mathbf{v}_i|^2}\right)\right]. \tag{9}$$

We define the rate gap $\Delta R(\gamma)$ as the difference between per-user rate achieved by ideal CSIT and limited channel feedback using the channel-statistics-based codebook

$$\Delta R(\gamma) = R_{\text{ideal}} - R_{\text{practical}}. \tag{10}$$

Following the results from [9] and[10] and using Jensen's inequality, the rate gap $\Delta R(\gamma)$ can be upper bounded as

$$\Delta R(\gamma) \leq \log_2 \left(1 + \frac{\gamma}{K}(K-1)\text{E}\left[|\mathbf{h}_k^H \mathbf{v}_i|^2\right]\right) \tag{11}$$

where the multiuser interference $\text{E}[|\mathbf{h}_k^H \mathbf{v}_i|^2]$ can be upper bounded in *Lemma 1*.

Lemma 1: The upper bound of multiuser interference $\text{E}[|\mathbf{h}_k^H \mathbf{v}_i|^2]$ depends on the channel quantization error $\text{E}[\sin^2(\angle(\tilde{\mathbf{h}}_k, \hat{\mathbf{h}}_k))]$, i.e.

$$\text{E}\left[|\mathbf{h}_k^H \mathbf{v}_i|^2\right] \leq \text{E}\left[\|\mathbf{h}_k\|^2\right] \text{E}\left[\sin^2(\angle(\tilde{\mathbf{h}}_k, \hat{\mathbf{h}}_k))\right]. \tag{12}$$

Proof: Denote the quantization error $X = \sin^2(\angle(\tilde{\mathbf{h}}_k, \hat{\mathbf{h}}_k)) = 1 - |\tilde{\mathbf{h}}_k^H \mathbf{c}_{k,F_k}|^2$. Since the size of the codebook is limited, $X \neq 0$. Thus, the normalized channel vector $\tilde{\mathbf{h}}_k$ can be decomposed along two orthogonal direction, one is the direction of quantization vector $\mathbf{c}_{k,F_k}$, and the other is in the null space of $\mathbf{c}_{k,F_k}$. Mathematically

$$\tilde{\mathbf{h}}_k = \sqrt{1-X}\mathbf{c}_{k,F_k} + \sqrt{X}\mathbf{s} \tag{13}$$

where $\mathbf{s}$ is a unit vector distributed in the null space of $\mathbf{c}_{k,F_k}$ [9]. Combining $\mathbf{h}_k = \|\mathbf{h}_k\|\tilde{\mathbf{h}}_k$ and (13), we have

$$|\mathbf{h}_k^H \mathbf{v}_i|^2 = \|\mathbf{h}_k\|^2 \left((1-X)|\mathbf{c}_{k,F_k}^H \mathbf{v}_i|^2 + X|\mathbf{s}^H \mathbf{v}_i|^2\right). \tag{14}$$

Since the precoding vector $\mathbf{v}_i$ is obtained as the normalized ith column of $\hat{\mathbf{H}}(\hat{\mathbf{H}}^H \hat{\mathbf{H}})^{-1}$, $\mathbf{v}_i$ is orthogonal to the kth user's feedback channel vector $\hat{\mathbf{h}}_k = \|\mathbf{h}_k\|\mathbf{c}_{k,F_k}$, i.e., $\mathbf{c}_{k,F_k}^H \mathbf{v}_i = 0$. Thus, we have $|\mathbf{h}_k^H \mathbf{v}_i|^2 = \|\mathbf{h}_k\|^2 X|\mathbf{s}^H \mathbf{v}_i|^2$. Since $|\mathbf{s}^H \mathbf{v}_i|^2 = \cos^2(\angle(\mathbf{s}, \mathbf{v}_i)) \leq 1$, we have

$$|\mathbf{h}_k^H \mathbf{v}_i|^2 \leq \|\mathbf{h}_k\|^2 X. \tag{15}$$

As the norm of a vector is independent of its direction [9], $\|\mathbf{h}_k\|^2$ is independent of X. Therefore, we obtain (12). ∎

Combining (11) and *Lemma 1*, we can further obtain

$$\Delta R(\gamma) \leq \log_2 \left(1 + \frac{\gamma}{K}(K-1)\text{E}[\|\mathbf{h}_k\|^2] \text{E}\left[\sin^2(\angle(\tilde{\mathbf{h}}_k, \hat{\mathbf{h}}_k))\right]\right). \tag{16}$$

B. Quantization Error

In this section, we discuss the quantization error $\text{E}[\sin^2(\angle(\tilde{\mathbf{h}}_k, \hat{\mathbf{h}}_k))]$ in (16) when the channel-statistics-based codebook is considered. For the rest of this paper, we omit the subscript k for simplicity but without loss of generality.

Lemma 2: The quantization error $\text{E}[\sin^2(\angle(\tilde{\mathbf{h}}, \hat{\mathbf{h}}))]$ of $\tilde{\mathbf{h}}$ can be upper bounded as

$$\text{E}\left[\sin^2(\angle(\tilde{\mathbf{h}}, \hat{\mathbf{h}}))\right] < 2^{-\frac{B}{r-1}}. \tag{17}$$

Proof: Denote that $Z = \cos^2(\angle(\tilde{\mathbf{h}}, \hat{\mathbf{h}})) = |\tilde{\mathbf{h}}^H \mathbf{c}_F|^2$. Since $F = \arg\max_{i \in [1, 2^B]} |\tilde{\mathbf{h}}^H \mathbf{c}_i|^2$, we have $\Pr\{|\tilde{\mathbf{h}}^H \mathbf{c}_F|^2 < z\} = \Pr\{|\tilde{\mathbf{h}}^H \mathbf{c}_1|^2 < z, |\tilde{\mathbf{h}}^H \mathbf{c}_2|^2 < z, \ldots |\tilde{\mathbf{h}}^H \mathbf{c}_{2^B}|^2 < z\} \overset{(a)}{=} \Pr\{|\tilde{\mathbf{h}}^H \mathbf{c}_i|^2 < z\}^{2^B}$, where (a) is true due to the independence among $\{|\tilde{\mathbf{h}}^H \mathbf{c}_i|^2\}_{i=1}^{2^B}$ [9]. Therefore

$$\Pr\{Z < z\} = \Pr\left\{|\tilde{\mathbf{h}}^H \mathbf{c}_i|^2 < z\right\}^{2^B}. \tag{18}$$

Using (1) and (2), we can rewrite $|\tilde{\mathbf{h}}^H \mathbf{c}_i|^2$ as

$$|\tilde{\mathbf{h}}^H \mathbf{c}_i|^2 = \frac{|(\mathbf{R}^{1/2}\mathbf{h}_w)^H \mathbf{R}^{1/2}\mathbf{w}_i|^2}{\|\mathbf{R}^{1/2}\mathbf{h}_w\|^2 \|\mathbf{R}^{1/2}\mathbf{w}_i\|^2}$$

$$\overset{(a)}{=} \frac{|\mathbf{h}_w^H \mathbf{U}(\boldsymbol{\Lambda}^{1/2})^H \boldsymbol{\Lambda}^{1/2}\mathbf{U}^H \mathbf{w}_i|^2}{\|\mathbf{R}^{1/2}\mathbf{h}_w\|^2 \|\mathbf{R}^{1/2}\mathbf{w}_i\|^2} \tag{19}$$

where (a) is obtained by using $\mathbf{U}^H \mathbf{U} = \mathbf{I}_M$. Since the unitary matrix $\mathbf{U}$ does not change the distribution of an isotropically distributed vector, $\mathbf{h}_w^H \overset{d}{=} \mathbf{h}_w^H \mathbf{U}$ and $\mathbf{w}_i \overset{d}{=} \mathbf{U}^H \mathbf{w}_i$, where $\overset{d}{=}$ denotes the equality in terms of distribution. Thus

$$|\tilde{\mathbf{h}}^H \mathbf{c}_i|^2 \overset{d}{=} \frac{|\mathbf{h}_w^H (\boldsymbol{\Lambda}_k^{1/2})^H \boldsymbol{\Lambda}_k^{1/2}\mathbf{w}_i|^2}{\|\boldsymbol{\Lambda}_k^{1/2}\mathbf{h}_w\|^2 \|\boldsymbol{\Lambda}_k^{1/2}\mathbf{w}_i\|^2} \tag{20}$$

where the denominator is accordingly changed to ensure the unit-norm requirement of $\tilde{\mathbf{h}}$ and $\mathbf{c}_i$. Therefore, we have

$$\Pr\left\{|\tilde{\mathbf{h}}^H \mathbf{c}_i|^2 < z\right\} = \Pr\left\{\frac{|\mathbf{h}_w^H (\boldsymbol{\Lambda}^{1/2})^H \boldsymbol{\Lambda}^{1/2}\mathbf{w}_i|^2}{\|\boldsymbol{\Lambda}^{1/2}\mathbf{h}_w\|^2 \|\boldsymbol{\Lambda}^{1/2}\mathbf{w}_i\|^2} < z\right\}. \tag{21}$$

Since $\boldsymbol{\Lambda}^{1/2}$ is a diagonal matrix with r nonzero diagonal elements $\{\sigma_1, \sigma_2, \ldots, \sigma_r\}$, we rewrite (21) as

$$\Pr\left\{|\tilde{\mathbf{h}}^H \mathbf{c}_i|^2 < z\right\} = \Pr\left\{\frac{|\mathbf{g}^H \boldsymbol{\Gamma}^H \boldsymbol{\Gamma}\mathbf{v}|^2}{\|\boldsymbol{\Gamma}\mathbf{g}\|^2 \|\boldsymbol{\Gamma}\mathbf{v}\|^2} < z\right\} \tag{22}$$

where $\mathbf{g} \in \mathbb{C}^{r \times 1}$ with $\mathbf{g}(j) = \mathbf{h}_w(j)$, $\mathbf{v} \in \mathbb{C}^{r \times 1}$ with $\mathbf{v}(j) = \mathbf{w}_i(j)$, and $\boldsymbol{\Gamma} \in \mathbb{C}^{r \times r}$ with $\boldsymbol{\Gamma}(j,j) = \boldsymbol{\Lambda}^{1/2}(j,j), j = 1, 2, \ldots, r$. The nondiagonal elements $\boldsymbol{\Gamma}(i,j) = 0$. Actually, $\mathbf{g}$ and $\mathbf{v}$ are random vectors in the r-dimensional hypersphere, whereas $\boldsymbol{\Gamma}\mathbf{g}$ and $\boldsymbol{\Gamma}\mathbf{v}$ are randomly distributed in the r-dimensional hyperellipse, which is obtained by stretching the hypersphere according to the diagonal elements of $\boldsymbol{\Gamma}$. As we know, $\frac{|\mathbf{g}^H \mathbf{v}|^2}{\|\mathbf{g}\|^2 \|\mathbf{v}\|^2}$ is the squared cosine of the angle between two vectors $\mathbf{g}$ and $\mathbf{v}$ in hypersphere, whose cumulative distribution function (CDF) is given by $\Pr\{\frac{|\mathbf{g}^H \mathbf{v}|^2}{\|\mathbf{g}\|^2 \|\mathbf{v}\|^2} \leq z\} = 1 - (1-z)^{r-1}$ for $z \in [0,1]$ [8]. In (22), $\frac{|\mathbf{g}^H \boldsymbol{\Gamma}^H \boldsymbol{\Gamma}\mathbf{v}|^2}{\|\boldsymbol{\Gamma}\mathbf{g}\|^2 \|\boldsymbol{\Gamma}\mathbf{v}\|^2}$ is the squared cosine of the angle between two vectors $\boldsymbol{\Gamma}\mathbf{g}$ and $\boldsymbol{\Gamma}\mathbf{v}$ in the hyperellipse, whose CDF is very difficult be obtained. However, we can prove in the *Appendix* that $\Pr\{\frac{|\mathbf{g}^H \boldsymbol{\Gamma}^H \boldsymbol{\Gamma}\mathbf{v}|^2}{\|\boldsymbol{\Gamma}\mathbf{g}\|^2 \|\boldsymbol{\Gamma}\mathbf{v}\|^2} \leq z\} < \Pr\{\frac{|\mathbf{g}^H \mathbf{v}|^2}{\|\mathbf{g}\|^2 \|\mathbf{v}\|^2} \leq z\}$. Thus, we have

$$\Pr\left\{\frac{|\mathbf{g}^H \boldsymbol{\Gamma}^H \boldsymbol{\Gamma}\mathbf{v}|^2}{\|\boldsymbol{\Gamma}\mathbf{g}\|^2 \|\boldsymbol{\Gamma}\mathbf{v}\|^2} < z\right\} < 1 - (1-z)^{r-1}. \tag{23}$$

Then, by combining (18), (22), and (23), we can obtain the upper bound of Z's CDF as

$$\Pr\{Z < z\} < (1 - (1-z)^{r-1})^{2^B}. \tag{24}$$

Utilizing the fact that $\text{E}[Z] = \int_0^1 \Pr\{Z \geq z\}$, the expectation $\text{E}[Z]$ can be expressed as

$$\text{E}[Z] = \int_0^1 (1 - \Pr\{Z < z\})dz = 1 - \int_0^1 \Pr\{Z < z\}\,dz. \tag{25}$$

Combining (24) and (25), we have

$$\text{E}[Z] > 1 - \int_0^1 (1 - (1-z)^{r-1})^{2^B}\,dz$$

$$\overset{(a)}{=} 1 - \int_0^1 (1 - s^{r-1})^{2^B}\,ds$$

$$\overset{(b)}{=} 1 - 2^B \beta\left(2^B, \frac{r}{r-1}\right) \overset{(c)}{\geq} 1 - 2^{-\frac{B}{r-1}} \tag{26}$$
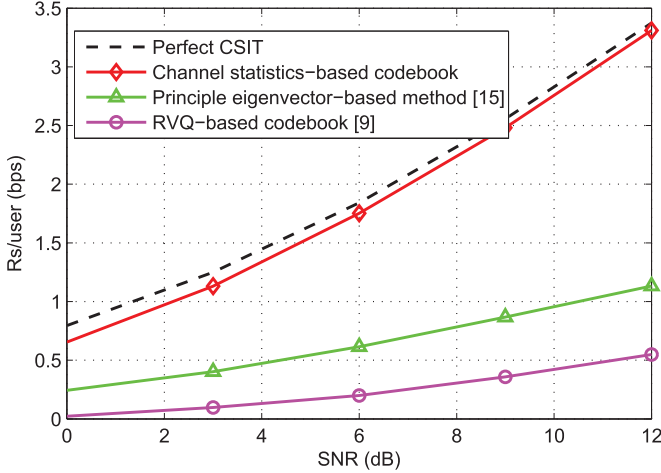
Fig. 1.   Per-user rate against the SNR at the receiver.

Fig. 2.   Number of feedback bits against the rank of correlation matrix.

where (a) is obtained by setting $s = 1 - z$, and (b) and (c) are obtained from [9, Appendixes I and II], respectively. Finally, we can obtain the upper bound of quantization error as

$$\mathrm{E}\left[\sin^2(\angle(\tilde{\mathbf{h}}, \hat{\mathbf{h}}))\right] = 1 - \mathrm{E}\left[Z\right] < 2^{-\frac{B}{r-1}}. \qquad (27)$$

∎

C. Feedback Bits

In this section, we discuss the required number of feedback bits B to ensure a constant rate gap $\Delta R(\gamma)$. By combining (16) and (27), we can easily obtain

$$\Delta R(\gamma) \leq \log_2\left(1 + \frac{\gamma}{K}(K-1)\mathrm{E}\left[\|\mathbf{h}\|^2\right] 2^{-\frac{B}{r-1}}\right). \qquad (28)$$

To ensure the rate gap $\Delta R(\gamma) \leq \log_2(b)$, based on (28) and letting $\log_2(1 + \frac{\gamma}{K}(K-1)\mathrm{E}[\|\mathbf{h}\|^2]2^{-\frac{B}{r-1}}) \leq \log_2(b)$ bps/Hz, we have

$$B \geq \frac{r-1}{3}\mathrm{SNR} + (r-1)\log_2\left(\frac{K-1}{b-1}\right) \qquad (29)$$

where the signal-to-noise ratio (SNR) at the receiver is defined as $\mathrm{SNR} = 10\log_{10}\frac{\gamma}{K}\mathrm{E}[\|\mathbf{h}\|^2]$. We can observe that the required number of feedback bits only scales linearly with the rank of correlation matrix r when SNR increases.

IV. SIMULATION VERIFICATION

In this section, simulation results are provided to verify the derived theoretical result. The simulation setup is as follows: The number of BS antennas, the number of users, and the rank of the channel correlation matrix is set as $(M, K, r) = (64, 10, 4)$; the square root of the channel correlation matrix is $\mathbf{R}_k^{1/2} = \mathbf{U}_k\mathbf{\Lambda}_k^{1/2}\mathbf{U}_k^H$, where $\mathbf{U}_k$ is a random unitary matrix, and $\mathbf{\Lambda}_k^{1/2}$ is a diagonal matrix with r nonzero diagonal elements [12]; the feedback bits $B = \frac{r-1}{3}\mathrm{SNR} + 3.17$ for Fig. 1 and SNR = 6 dB for Fig. 2.

Fig. 1 shows the per-user rate against SNR. We can observe from Fig. 1 that the rate gap between the ideal case of perfect CSIT and the practical case using the channel-statistics-based codebook to realize CSIT remains constant when SNR increases. Such simulated result is consistent with our theoretical analysis in Section III. On the other hand, we can also find that the rate gap between the ideal case of
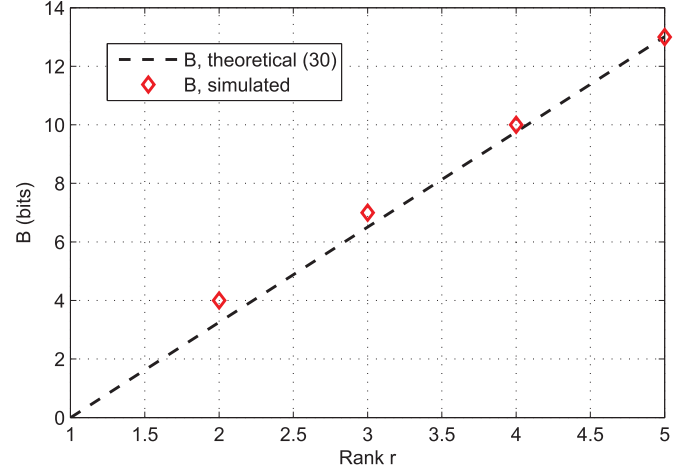
perfect CSIT and the practical case using RVQ-based codebook to realize CSIT becomes greater when SNR increases since the required number of feedback bits for RVQ-based codebook should scale linearly with the number of BS antennas, which is much larger than the rank of the channel correlation matrix. Moreover, we also show the per-user rate of the principle eigenvector-based method [15] for comparison, which is overperformed by the channel-statistics-based codebook.

Fig. 2 shows the required number of feedback bits to limit the rate gap between the ideal case of perfect CSIT and the practical case of using the channel-statistics-based codebook within 0.09 bps/Hz. We can observe that the required number of feedback bits B scales linearly with the rank of the channel correlation matrix r, which is consistent with the theoretical result (29) shown by the dash curve in Fig. 2.

V. CONCLUSION

It has been shown that the number of feedback bits scales linearly with the rank of the channel correlation matrix for the channel-statistics-based codebook. However, this attractive conclusion was only proved under a particular channel assumption in the literature. To this end, in this paper we provided a rigorous proof under a general channel assumption for the first time. At first, we found that the rate gap between the ideal case of perfect CSIT and using the channel stastics-based codebook depends on the channel quantization error. Then, we derived an upper bound of the quantization error, based on which we proved that the required number of feedback bits scales linearly with the rank of the channel correlation matrix. Moreover, the proof techniques in this paper can be used to analyze the performance of other codebooks.

APPENDIX

The CDF of the squared cosine of the angle between two vectors $\mathbf{\Gamma g}$ and $\mathbf{\Gamma v}$ in the hyperellipse is upper bounded as

$$\Pr\left\{\frac{|\mathbf{g}^H\mathbf{\Gamma}^H\mathbf{\Gamma v}|^2}{\|\mathbf{\Gamma g}\|^2\|\mathbf{\Gamma v}\|^2} \leq z\right\} < \Pr\left\{\frac{|\mathbf{g}^H\mathbf{v}|^2}{\|\mathbf{g}\|^2\|\mathbf{v}\|^2} \leq z\right\}. \qquad (30)$$

Proof: First, we consider the squared cosine of the angle $\frac{|\mathbf{g}^H\mathbf{v}|^2}{\|\mathbf{g}\|^2\|\mathbf{v}\|^2}$ between two isotropic vectors $\mathbf{g}$ and $\mathbf{v}$ uniformly distributed in a hypersphere. For $z \in [0, 1]$, the CDF of $\frac{|\mathbf{g}^H\mathbf{v}|^2}{\|\mathbf{g}\|^2\|\mathbf{v}\|^2}$ is given by [8]

$$\Pr\left\{\frac{|\mathbf{g}^H\mathbf{v}|^2}{\|\mathbf{g}\|^2\|\mathbf{v}\|^2} \leq z\right\} = 1 - (1 - z)^{r-1}. \qquad (31)$$

For $\frac{|\mathbf{g}^H \boldsymbol{\Gamma}^H \boldsymbol{\Gamma} \mathbf{v}|^2}{\|\boldsymbol{\Gamma} \mathbf{g}\|^2 \|\boldsymbol{\Gamma} \mathbf{v}\|^2}$ in (30), the elements of $\mathbf{g}$ and $\mathbf{v}$ are multiplied by the diagonal elements $\{\sigma_1, \sigma_2, \ldots, \sigma_r\}$ of $\boldsymbol{\Gamma}$. The uniformly distributed vectors $\mathbf{g}$ and $\mathbf{v}$ in hypersphere are projected onto nonuniformly distributed vectors $\boldsymbol{\Gamma} \mathbf{g}$ and $\boldsymbol{\Gamma} \mathbf{v}$ in hyperellipse, which is generated by stretching the hypersphere. That means the angle between two vectors in hyperellipse tends to be smaller, i.e., the squared cosine of the angle between two vectors in hyperellipse tends to be larger. Thus, the CDF of $\frac{|\mathbf{g}^H \boldsymbol{\Gamma}^H \boldsymbol{\Gamma} \mathbf{v}|^2}{\|\boldsymbol{\Gamma} \mathbf{g}\|^2 \|\boldsymbol{\Gamma} \mathbf{v}\|^2}$ is smaller than that of $\frac{|\mathbf{g}^H \mathbf{v}|^2}{\|\mathbf{g}\|^2 \|\mathbf{v}\|^2}$, i.e., we have (30).

Now we give an intuitive example of an extreme case where $\sigma_1 = \sigma_2 =, \ldots, = \sigma_{r-1}$ and $\sigma_r \ll \sigma_2$. The r-dimensional hyperellipse is generated by compressing the last dimension of a r-dimensional hypersphere. Thus, the vectors in hypersphere are projected onto a compressed hypersphere, i.e., the hyperellipse. The angle between two vectors in hyperellipse becomes smaller. Mathematically, the numerator of the squared cosine can be approximately expressed as

$$|\mathbf{g}^H \boldsymbol{\Gamma}^H \boldsymbol{\Gamma} \mathbf{v}|^2 \approx |\sigma_1|^4 |\Sigma_{i=1}^{r-1} g_i^* v_i|^2. \tag{32}$$

The dominator of the squared cosine can be approximately expressed as $\|\boldsymbol{\Gamma} \mathbf{g}\|^2 \|\boldsymbol{\Gamma} \mathbf{v}\|^2 \approx |\sigma_1|^4 (\Sigma_{i=1}^{r-1} |g_i|^2)(\Sigma_{i=1}^{r-1} |v_i|^2)$. Thus, we have $\frac{|\mathbf{g}^H \boldsymbol{\Gamma}^H \boldsymbol{\Gamma} \mathbf{v}|^2}{\|\boldsymbol{\Gamma} \mathbf{g}\|^2 \|\boldsymbol{\Gamma} \mathbf{v}\|^2} \approx \frac{|\Sigma_{i=1}^{r-1} g_i^* v_i|^2}{(\Sigma_{i=1}^{r-1} |g_i|^2)(\Sigma_{i=1}^{r-1} |v_i|^2)}$. Considering (31), we can get

$$\Pr \left\{ \frac{|\mathbf{g}^H \boldsymbol{\Gamma}^H \boldsymbol{\Gamma} \mathbf{v}|^2}{\|\boldsymbol{\Gamma} \mathbf{g}\|^2 \|\boldsymbol{\Gamma} \mathbf{v}\|^2} \leq z \right\} \approx 1 - (1-z)^{r-2}. \tag{33}$$

Utilizing the fact that $1 - (1-z)^{r-2} < 1 - (1-z)^{r-1}$, we can obtain (30). ∎

REFERENCES

[1] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[2] C. Qi, G. Yue, L. Wu, and A. Nallanathan, "Pilot design for sparse channel estimation in OFDM-based cognitive radio systems," *IEEE Trans. Veh. Technol.*, vol. 63, no. 2, pp. 982–987, Feb. 2014.

[3] C. Huang, L. Liu, C. Yuen, and S. Sun, "Iterative channel estimation using LSE and sparse message passing for mmWave MIMO systems," arXiv:1611.05653, 2016.

[4] X. Gao, L. Dai, S. Han, C.-L. I, and R. W. Heath, "Energy-efficient hybrid analog and digital precoding for mmWave MIMO systems with large antenna arrays," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 998–1009, Apr. 2016.

[5] H. Xie, F. Gao, S. Zhang, and S. Jin, "A unified transmission strategy for TDD/FDD massive MIMO systems with spatial basis expansion model," *IEEE Trans. Veh. Technol.*, to be published.

[6] D. J. Love, R. W. Heath, and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2735–2747, Oct. 2003.

[7] K. K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, "On beamforming with finite rate feedback in multiple-antenna systems," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2562–2579, Oct. 2003.

[8] C. K. Au-yeung and D. J. Love, "On the performance of random vector quantization limited feedback beamforming in a MISO system," *IEEE Trans. Wireless Commun.*, vol. 6, no. 2, pp. 458–462, Feb. 2007.

[9] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.

[10] B. Lee, J. Choi, J. Y. Seol, D. J. Love, and B. Shim, "Antenna grouping based feedback compression for FDD-based massive MIMO systems," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3261–3274, Sep. 2015.

[11] D. J. Love and R. W. Heath, "Limited feedback diversity techniques for correlated channels," *IEEE Trans. Veh. Technol.*, vol. 55, no. 2, pp. 718–722, Mar. 2006.

[12] B. Clerckx, G. Kim, and S. Kim, "MU-MIMO with channel statistics-based codebooks in spatially correlated channels," in *Proc. IEEE Global Commun. Conf.*, Nov. 2008, pp. 1–5.

[13] W. Shen, L. Dai, B. Shim, S. Mumtaz, and Z. Wang, "Joint CSIT acquisition based on low-rank matrix completion for FDD massive MIMO systems," *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2178–2181, Dec. 2015.

[14] Z. Gao, L. Dai, Z. Wang, and S. Chen, "Spatially common sparsity based adaptive channel estimation and feedback for FDD massive MIMO," *IEEE Trans. Signal Process.*, vol. 63, no. 23, pp. 6169–6183, Dec. 2015.

[15] S. Zhou and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel mean feedback," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2599–2613, Oct. 2002.