MIMO Relay with Finite-Rate Feedback

LGE Mobile Research
San Diego, CA 92131

***Abstract*—In this paper, the effects of finite-rate CQI feedback on multiple antenna relay systems are discussed. We start from multi-antenna beamforming transmission using uniform random codebook and formulate average SNR loss due to channel quantization error and limited-sized precoding codebook. After this, consider MIMO relay with multiple relay nodes and discussion how the relay channel capacity suffers from this finite-rate feedback and as well as codebook design.**

I. Introduction

Multiple antenna systems have received much attention for the past decade or so, due to their promise of higher spectrum efficiency without increasing transmit power. It is well-known that the performance and complexity of multiple-input multiple-output systems can be improved by making the channel state information (CSI) available at the transmitter. This is achieved through a special reverselink CSI feedback channel from the receiver. For example, there are HS-DPCCH and R-CQI channels for 3GPP HSDPA Rel. 6 and 3GPP2 UMB, respectively. However, the CSI received by the transmitter is not perfect and suffers from various impairments, including round-trip delay, signal processing errors, transmission loss, etc.

II. System Model And Problem Description

Consider a three-terminal one-hop MIMO relay channel as shown in Figure 1, where a relay is used to assist the transmission from source to destination. All terminals in the relay model are equipped with multiple antennas. The source transmits (broadcasts) to the relay and destination in channels 0 and 1, individually noted by $\mathbf{H}_0$ and $\mathbf{H}_1$, and the relay transmits to the destination in channel 2, noted by $\mathbf{H}_2$. We assume Channels 1 and 2 are orthogonal to each other. In practice, the two channels should be further divided in time, frequency or code domain and interference between Channels 1 and 2 can be modelled as colored additive noise if it is necessary. For fast data transmission and convenient implementation, division in frequency or code domain appears more advantageous than division in time. As long as the channel coherence time is larger than the reciprocal of the channel coherence bandwidth, all fading channels may be modelled as if they were frequency flat, through use of multiple narrow-band carriers (such as orthogonal frequency division multiplexing). This is the case for most practical environments. Therefore, we will assume that the MIMO channel responses between the source and the relay, the relay and the destination, and the source and the destination, are represented, respectively, by constant (as opposed to polynomial) matrices $\mathbf{H}_1$, $\mathbf{H}_2$, and $\mathbf{H}_0$. The transfer function of a non-regenerative relay is equivalent to a memoryless weighting matrix $\mathbf{W}$ that transforms the (baseband) waveform received at the relay to the (baseband) waveform transmitted from the relay. Furthermore, we assume that during the transmission of each packet of data, $\mathbf{H}_0$, $\mathbf{H}_1$, $\mathbf{H}_2$ and $\mathbf{W}$ remain constant (as opposed to time varying). The numbers of antennas equipped at the source, destination and relay are denoted as M, N and L, respectively, and thus we can write $\mathbf{H}_0 \in \mathbb{C}^{N\times M}$, $\mathbf{H}_1 \in \mathbb{C}^{L\times M}$, $\mathbf{H}_2 \in \mathbb{C}^{N\times L}$ and $\mathbf{W} \in \mathbb{C}^{L\times L}$. We assume that all L antennas at the relay can be used for both receiving and transmitting.

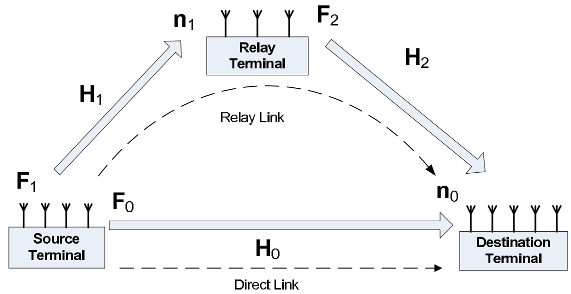

Fig. 1. A MIMO relay channel model

If there is no relay link between the source terminal and destination terminal the MIMO relay channel in Figure 1 becomes a MIMO channel and the received signals by the destination is

$$\mathbf{y}_0 = \mathbf{H}_0\mathbf{W}_0\mathbf{x} + \mathbf{n}_0 \tag{1}$$

where $\mathbf{x}$ is the transmitted $M \times 1$ signal vector by the source with $\mathbf{R}_{\mathbf{x}} = \mathrm{E}\left\{\mathbf{x}\mathbf{x}^{\mathrm{H}}\right\} = \frac{P_0}{M}\mathbf{I}_M$, where $[*]^{\mathrm{H}}$ is the Hermitian conjugate operation and P_0 is the total transmission power, $\mathbf{W}_0$ is a $M \times M$ MIMO linear beamforming matrix, and $\mathbf{n} \sim \mathcal{CN}(0, \sigma^2\mathbf{I}_N)$ is a complex circular white Gaussian vector. It is well-known that the achievable MIMO channel capacity is

$$C_0 = W\log\left|\mathbf{I}_N + \frac{P_0}{M}\mathbf{H}_0\mathbf{W}_0\mathbf{W}_0^{\mathrm{H}}\mathbf{H}_0^{\mathrm{H}}\right| \tag{2}$$

where $| * |$ denotes the determinant of matrix $*$.

If we consider the relay link, there are two basic relay modes among this three-terminal MIMO relay system: 1) Non-Cooperative MIMO Relay, where the direct link from the source to the destination and the relay link via the relay terminal are independent to each other; 2) Cooperative MIMO Relay, where the direct link and the relay link are cooperatively correlated to each other. In the following section we give the mathematic models for these two cases, respectively.

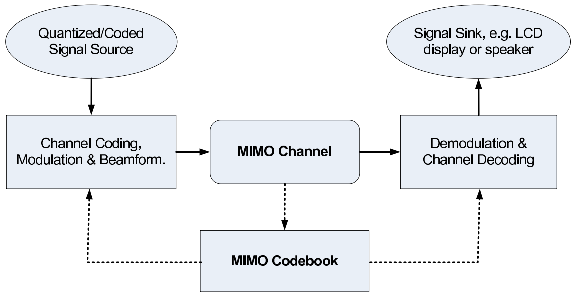

Fig. 2. A system model for MIMO with finite-rate feedback.

III. MIMO Relay with Perfect CSI

If all signals are simultaneously transmitted and relayed using the same channel resource, the MIMO relay channel can be taken as a hybrid of a multi-antenna broadcast channel (BC) and a multi-antenna multi-access channel (MAC). With the assumption that the relay works in full-duplex mode, dirty paper coding (DPC) and successive interference cancellation (SIC) are capacity-achieving. The discussion of the achievable channel capacity for Gaussian MIMO relay channel and as well as the Rayleigh fading case can be found in [1]. In practices, the signals are usually precoded and transmitted using TDMA, FDMA or CDMA, etc. to separate the relay link and direct link and possibly the receiving and transmission at the relay terminal. In the this case, the signals through the direct link and relay link are orthogonally separated. The signals received directly through the direct link is the same as (1) but the relayed signals through the relay terminal is

$$\begin{aligned} \mathbf{y}_1 &= \mathbf{H}_2\mathbf{W}_2\left(\mathbf{H}_1\mathbf{W}_1\mathbf{x}+\mathbf{n}_1\right)+\mathbf{n}_0 \\ &= \mathbf{H}_2\mathbf{W}_2\mathbf{H}_1\mathbf{W}_1\mathbf{x}+\tilde{\mathbf{n}}_0 \end{aligned} \tag{3}$$

where $\mathbf{W}_1$ is the beamforming matrix at the source terminal and $\mathbf{W}_2$ is the $L \times L$ beamforming matrix at the relay terminal. $\mathbf{n}_1 \sim \boldsymbol{\mathcal{CN}}(0, \sigma^2\mathbf{I}_L)$ denotes the complex circular white Gaussian vector received by the relay terminal and the $\tilde{\mathbf{n}}_0$ denoted the composite spatial color noise vector received by the destination terminal

$$\tilde{\mathbf{n}}_0 = \mathbf{H}_2\mathbf{W}_2\mathbf{n}_1+\mathbf{n}_0 \tag{4}$$

with the correlation matrix

$$\mathbf{R}_{\tilde{\mathbf{n}}_0} = \sigma_1^2\mathbf{H}_2\mathbf{W}_2\mathbf{W}_2^{\mathrm{H}}\mathbf{H}_2^{\mathrm{H}}+\sigma_0^2\mathbf{I} \tag{5}$$

The MIMO channel capacity here is

$$\begin{aligned} C &= (1-\alpha)\,C_0+\alpha C_1 \\ &= (1-\alpha)\,W\log\left|\mathbf{I}_N+\tfrac{P_0}{M}\mathbf{H}_0\mathbf{W}_0\mathbf{W}_0^{\mathrm{H}}\mathbf{H}_0^{\mathrm{H}}\right|+ \\ &\alpha W\log\left|\mathbf{I}_N+\tfrac{P_0}{M}\mathbf{R}_{\tilde{\mathbf{n}}_0}^{+}\mathbf{H}_2\mathbf{W}_2\mathbf{H}_1\mathbf{W}_1\mathbf{W}_1^{\mathrm{H}}\mathbf{H}_1^{\mathrm{H}}\mathbf{W}_2^{\mathrm{H}}\mathbf{H}_2^{\mathrm{H}}\right| \end{aligned} \tag{6}$$

where C_1 denote the channel capacity of the relay link if all resource is allocated for it and it can be written by

$$\begin{aligned} C_1 &= \\ &W\log\left|\mathbf{I}_N+\tfrac{P_0}{M}\mathbf{R}_{\tilde{\mathbf{n}}_0}^{+}\mathbf{H}_2\mathbf{W}_2\mathbf{H}_1\mathbf{W}_1\mathbf{W}_1^{\mathrm{H}}\mathbf{H}_1^{\mathrm{H}}\mathbf{W}_2^{\mathrm{H}}\mathbf{H}_2^{\mathrm{H}}\right| , \end{aligned} \tag{7}$$

$\alpha \in [0\ 1]$ denotes the channel resource partition ratio between the direct link and the total resource in terms of degree of freedom (the symbol period multiple the total bandwidth) and $[*]^{+}$ denote the pseudo-inverse of the matrix $*$.

In (6), the MIMO relay channel capacity is a balance between direct MIMO and non-regenerative MIMO relay. When $\alpha = 0$, only the direct link exists and the MIMO relay channel becomes MIMO channel with $C = C_0$. When $\alpha = 1$, only the relay link works and the MIMO relay channel becomes non-regenerative MIMO relay channel with $C = C_1$.

IV. MIMO Channel: SINR Degradation with Finite-Rate Feedback

A. Codebook and Channel Quantization

In a finite-rate feedback model, the receiver estimates channel responses from each receiving antenna and decides the best beamforming vector(s) within a MIMO precoding codebook shared between the receiver and transmitter. This is also called channel quantization. The receiver then feeds back the chosen precoding index(es) to the transmitter for the next transmission. The MIMO beamforming codebook $\boldsymbol{\mathcal{W}}$ of the size 2^B consists of M-dimensional normalized vectors $\{\mathbf{w}_1, ..., \mathbf{w}_{2^B}\}$. It usually takes the receiver B feedback bits for each beamforming stream. The codebook design usually uses vector quantization (VQ) [2] to quantize channel responses with certain distortion measures. This procedure is also called Grassmannian line packing [3], which is similar to spherical packing on unit sphere $\boldsymbol{\mathcal{S}}_n\,(1)$, where $\boldsymbol{\mathcal{S}}_n(R) = \{\mathbf{v} : \|\mathbf{v}\| = R\}$ denotes $(n-1)$-sphere. The Grassmannian space is the space of all K-dimensional subspaces of an M-dimensional vector space denoted

$\mathcal{G}_{k,n}$. The Grassmannian space $\mathcal{G}_{1,n}$ used in Grassmannian line packing is the space of lines through the origin. Grassmanian can usually be taken a generalization of projective space. Obviously $(n-1)$-unit sphere $\mathcal{S}_n(1)$ can be one-to-one mapped onto $\mathcal{G}_{1,n}$. The following conclusion is very straightforward and it shows that the sphere packing and the Grassmanian line packing are the same.

Lemma 1. *For Grassmannian line pack using the metric*

$$d_{\mathrm{G}}(\mathbf{v}_1,\ \mathbf{v}_2) \quad = \quad \sqrt{1-\left(\mathbf{v}_1^{\mathrm{H}}\mathbf{v}_2\right)^2}\ , \tag{8}$$

where $d_{\mathrm{G}}(\mathbf{v}_1,\ \mathbf{v}_2) \in [0,\ 1]$*, is the same to sphere packing using a squared Euclid distance*

$$d_{\mathrm{S}}(\mathbf{v}_1,\ \mathbf{v}_2) \quad = \quad \|\mathbf{v}_1-\mathbf{v}_2\|_2\ , \tag{9}$$

where $d_{\mathrm{S}}(\mathbf{v}_1,\ \mathbf{v}_2) \in \left[0,\ \sqrt{2}\right]$.

Proof. If we set the origin of the unit sphere $\mathcal{S}_n(1)$ to be the zero point, we may directly take the every point of $\mathcal{S}_n(1)$ as a vector in $\mathcal{G}_{1,n}$. And the relationship between $d_{\mathrm{G}}(\mathbf{v}_1,\ \mathbf{v}_2)$ and $d_{\mathrm{S}}(\mathbf{v}_1,\ \mathbf{v}_2)$ is

$$\begin{aligned} d_{\mathrm{G}}^2(\mathbf{v}_1,\ \mathbf{v}_2) &= 1-\left(\mathbf{v}_1^{\mathrm{H}}\mathbf{v}_2\right)^2 \\ &= 1-\left[1-\tfrac{1}{2}d_{\mathrm{S}}^2(\mathbf{v}_1,\ \mathbf{v}_2)\right]^2\ . \end{aligned} \tag{10}$$

Obviously the mapping of $d_{\mathrm{G}}^2(\mathbf{v}_1,\ \mathbf{v}_2)$ to $d_{\mathrm{S}}^2(\mathbf{v}_1,\ \mathbf{v}_2)$ is given by

$$y \quad = \quad f(x) \quad = \quad 1-\left(1-\tfrac{1}{2}x\right)^2\ . \tag{11}$$

Since the differential of $f(x)$ is

$$\tfrac{d}{dx}f(x) \quad = \quad 1-\tfrac{1}{2}x \quad > \quad 0 \tag{12}$$

for $x \in [0,\ 2]$, there exists one-to-one mapping between Grassmannian line packing and sphere packing. These two spaces are similar to each.

□

B. Received SNR Loss

At the transmitter side, the signal is precoded by the chosen codeword $\mathbf{w}_k$ following the feedback from receiver. The received signal then is similar to (1) but with the distortion δ_i due to the quantized precoding codeword. This distortion can be expressed by

$$\begin{aligned} \delta_i &= \min_{\mathbf{w}\in\mathcal{W}}\|\mathbf{H}\mathbf{v}_i-\mathbf{H}\mathbf{w}\| \\ &= \lambda_i-\max_{\mathbf{w}\in\mathcal{W}}\|\mathbf{H}\mathbf{w}\| \end{aligned} \tag{13}$$

where $\mathbf{v}_i$ denotes the original beamforming vector for eigen-mode i and λ_i is the corresponding eigen-value of $\mathbf{H}$, which denotes $\mathbf{H}_0$, $\mathbf{H}_1$ or $\mathbf{H}_2$. It is known that the degradation δ_i is one of the major sources limiting closed-loop MIMO throughput and the average SNR distortion decreases with increasing codebook size 2^B . From (13), the signal distortion δ_i is a random valuable which depends on $\mathbf{v}_i$, $\mathbf{w}_k$ and λ_i. The mean of δ_i is

$$\begin{aligned} \bar{\delta}_i &= \mathrm{E}\left\{\min_{\mathbf{w}\in\mathcal{W}}\|\mathbf{H}\mathbf{v}_i-\mathbf{H}\mathbf{w}_k\|\right\} \\ &= \mathrm{E}_{\mathbf{v}_i\in\mathcal{V}_k}\left\{\mathbf{H}\mathbf{v}_i-\mathbf{H}\mathbf{V}\mathbf{V}^{\mathrm{H}}\mathbf{w}_k\right\} \\ &= \mathrm{E}_{\mathbf{v}_i\in\mathcal{V}_k}\left\{\lambda_i\left(1-\mathbf{v}_i^{\mathrm{H}}\mathbf{w}_k\right)-\sum_{j\neq i}\lambda_j\mathbf{v}_j^{\mathrm{H}}\mathbf{w}_k\right\} \end{aligned} \tag{14}$$

where $\mathcal{V}_k$ denotes the Voronoi cell for $\mathbf{w}_k$. $\mathcal{V}_k$ is decided by the VQ criteria used in the codebook design. Different VQ criteria may lead to different Voronoi decomposition. For example, with Euclid distance criteria, $\mathcal{V}_k$ is defined by

$$\mathcal{V}_k \quad = \quad \left\{\mathbf{v}:\|\mathbf{v}\|=1,\ \tfrac{\|\mathbf{v}-\mathbf{w}_k\|}{\|\mathbf{v}-\mathbf{w}_j\|}<1,\ j\neq k\right\}\ . \tag{15}$$

In (14) it shows that there are two major sources causing SNR degradation in MIMO beamforming with imperfect feedback. The first major source denoted by

$$\begin{aligned} \Delta_P &= \mathrm{E}_{\mathbf{v}_i\in\mathcal{V}_k}\left|\lambda_i\left(1-\mathbf{v}_i^{\mathrm{H}}\mathbf{w}_k\right)\right| \\ &= \lambda_i\left(1-\mathrm{E}_{\mathbf{v}_i\in\mathcal{V}_k}\left|\mathbf{v}_i^{\mathrm{H}}\mathbf{w}_k\right|\right) \\ &= (1-\alpha_i)\,\lambda_i \end{aligned} \tag{16}$$

suggests there is a power degradation and the average power degradation ratio $\bar{\alpha}$ is

$$\begin{aligned} \bar{\alpha} &= \mathrm{E}\{\alpha_i\} \\ &= \mathrm{E}\left\{\mathbf{v}_i^{\mathrm{H}}\mathbf{w}_k\right\}\ , \end{aligned} \tag{17}$$

where $\alpha_i = \mathbf{v}_i^{\mathrm{H}}\mathbf{w}_k$. $\bar{\alpha}$ is irrelevant to the power distribution on the beams but depends on the codebook design. The second source is the inter-beam interference (IBI) due to the correlation between the precoding codeword $\mathbf{w}_k$ and other eigen-modes. It can be expressed by

$$I \quad = \quad \sum_{j\neq i}\lambda_j\mathbf{v}_j^{\mathrm{H}}\mathbf{w}_k \tag{18}$$

Since we use the uniform random codebook and assume the MIMO channel response $\mathbf{H}$ is uniform random too, the eigenvector $\mathbf{v}_j$ is independent to the eigenvale λ. The correlation ratio between $\mathbf{w}_k$ and the jth eigen-mode can be denoted by

$$\alpha_j \quad = \quad \mathrm{E}_{\mathbf{v}_i\in\mathcal{V}_k}\left|\mathbf{v}_j^{\mathrm{H}}\mathbf{w}_k\right|\ . \tag{19}$$

Proposition 1. *For the uniform random codebook* $\mathcal{W}$ *and isotropically distributed channel response matrix* $\mathbf{H}$*, there is the following relationship for the correlations defined in (17) and (19).*

$$\alpha+\textstyle\sum_{j\neq i}\alpha_j \quad = \quad \alpha+(M-1)\tilde{\alpha} \quad = \quad 1\ , \tag{20}$$

where $\mathbf{v}_i \in \mathcal{V}_k$ *and*

$$\tilde{\alpha} = E_{j\neq i}\left\{\alpha_j\right\} \tag{21}$$

denotes the average interference correlation ratio.

With Proposition 1, the average IBI to other beams in (18) can then be written by

$$\bar{I} = \mathrm{E}_{\mathbf{v}_i\in\mathcal{V}_k}\left\{I\right\} = \tfrac{1-\alpha}{M-1}\sum_{j\neq i}\lambda_j \ . \tag{22}$$

And the average received SNR $\bar{\rho}_i$ for the ith eigen-mode using codeword $\mathbf{w}_k$ can be written by

$$\begin{aligned}\bar{\rho}_i &= \mathrm{E}_{\mathbf{v}_i\in\mathcal{V}_k}\left\{\frac{\lambda_i\left(\mathbf{v}_i^{\mathrm{H}}\mathbf{w}_k\right)^2 P_i}{n_0^2+\lambda_i\sum\limits_{l\neq i}\left(\mathbf{v}_i^{\mathrm{H}}\mathbf{w}_l\right)^2 P_l}\right\}\\ &= \mathrm{E}_{\mathbf{v}_i\in\mathcal{V}_k}\left\{\frac{\lambda_i\alpha_i^2 P_i}{n_0^2+\lambda_i\sum\limits_{l\neq i}\alpha_l^2 P_l}\right\}\\ &= \mathrm{E}_{\mathbf{v}_i\in\mathcal{V}_k}\left\{\frac{\alpha_i^2\rho_i}{1+\lambda_i\sum\limits_{l\neq i}\frac{\alpha_l^2}{\lambda_l}\rho_l}\right\} ,\end{aligned} \tag{23}$$

where $K \leq \min\left\{M,\ N\right\}$ denotes the number of beams in use and ρ_i, given by

$$\rho_i = \tfrac{\lambda_i P_i}{\sigma_0^2} \ , \tag{24}$$

denotes the maximum achievable SNR

With (17) the correlation ratio α can be decided if we know the Voronoi tessellation. However, the closed-form solution to the boundary of a Voronoi cell, which is a polytope, generally is unknown, even for a uniform Voronoi tessellation. This makes it difficult to accurately calculate the average $\bar{\rho}_i$ for the ith eigen-mode. Here we suggest an approach to approximate the Voronoi cell boundary. Instead of deciding the exact boundary for $\boldsymbol{\mathcal{V}}_i$, we use a hypercircle to approximate the bound with the constraint that the interior of the hypercircle has the same area as the Voronoi cell. In the following, we details this procedure as well as the theory principles to decide this hyper-circle. Before doing that, we present two propositions for calculating the cell area of uniform Voronoi tessellation and the surface area of hyperspherical cap.

Proposition 2. *The area of a Voronoi cell from a uniform random codebook of size* 2^B *in* M*-dimensional Euclid space is given by*

$$A\left(\boldsymbol{\mathcal{V}}_k\right) = \tfrac{2\pi^M}{2^B\Gamma(M)} \tag{25}$$

where $\Gamma\left(*\right)$ *is the gamma function.*

Proof. The surface area of a real n-ball (the $(n-1)$-sphere) in Euclid space is

$$S_n\left(R\right) = \tfrac{n}{R}V_n = \tfrac{2\pi^{\frac{n}{2}}R^{n-1}}{\Gamma\left(\frac{n}{2}\right)} \tag{26}$$

where V_n is the volume and R is the radius of the n-ball. A M-dimension complex ball in Euclid space with the radius R is defined by

$$\begin{aligned}\boldsymbol{\mathcal{S}}_M^{\mathrm{c}}\left(R\right) &= \left\{\mathbf{v}:\ \left\|\mathbf{v}\right\|_2 = R\right\} = \\ &\left\{\left[x_m+iy_m\right]_{M\times 1}:\ \sum_{m=1}^{M}\left(x_m^2+y_m^2\right) = R^2\right\} .\end{aligned} \tag{27}$$

Therefore, the interior area for the M-dimensional complex ball defined by (27) is

$$S_{2M}\left(R\right) = \tfrac{2\pi^M R^{2M-1}}{\Gamma(M)} \ . \tag{28}$$

With uniform Voronoi decomposition on $\boldsymbol{\mathcal{S}}_M^{\mathrm{c}}$, we have

$$S_{2M}\left(1\right) = \sum_{k=1}^{2^B}A\left(\boldsymbol{\mathcal{V}}_k\right) = 2^B A\left(\boldsymbol{\mathcal{V}}_k\right) \ . \tag{29}$$

□

After the area of $\boldsymbol{\mathcal{V}}_i$ is decided by (25), we define a $(m-1)$-complex sphere cap $\boldsymbol{\mathcal{C}}_m^{\mathrm{c}}\left(R,\ \psi,\ \mathbf{w}\right)$,

$$\boldsymbol{\mathcal{C}}_m^{\mathrm{c}}\left(\psi,\ \mathbf{w}\right) = \left\{\mathbf{v}:\ \left\|\mathbf{v}\right\| = \left\|\mathbf{w}\right\| = R,\ \angle\left(\mathbf{v},\ \mathbf{w}\right)\leq\psi\right\} , \tag{30}$$

where $\mathbf{w}$ is the center of the cap. The solution to calculate the area of $\boldsymbol{\mathcal{C}}_m^{\mathrm{c}}\left(\theta,\ \mathbf{w}\right)$ is presented in the following proposition.

Proposition 3. *The area of* $(m-1)$*-complex sphere cap* $\boldsymbol{\mathcal{C}}_m^{\mathrm{c}}\left(\psi,\ \mathbf{w}\right)$ *with* $\mathbf{w} = R$ *is*

$$A\left(\boldsymbol{\mathcal{C}}_m^{\mathrm{c}}\left(\psi,\ \mathbf{w}\right)\right) = \Phi_m\left(\psi\right)S_{2m}\left(R\right) \ , \tag{31}$$

where $\Phi_m\left(\psi\right)$ *is defined by*

$$\Phi_m\left(\psi\right) = 1-\cos^{(2m-2)}\left(\psi\right) \ . \tag{32}$$

And it can be verified that

$$S_{2m}\left(\left\|\mathbf{w}\right\|\right) = A\left(\boldsymbol{\mathcal{C}}_m^{\mathrm{c}}\left(\pi,\ \mathbf{w}\right)\right) \ . \tag{33}$$

Proof. At first, a complex hypercircle $\boldsymbol{\mathcal{O}}_m^{\mathrm{c}}\left(r,\ \mathbf{c}\right)$ on $(m-1)$-complex sphere with radius R, the central point $\mathbf{w}$ is defined by

$$\begin{aligned}\boldsymbol{\mathcal{O}}_m^{\mathrm{c}}\left(r,\ \mathbf{c}\right) &= \boldsymbol{\mathcal{S}}_m^{\mathrm{c}}\left(R\right)\bigcap\boldsymbol{\mathcal{P}}_m\left(\mathbf{c}\right)\\ &= \left\{\mathbf{v}:\ \left\|\mathbf{v}\right\| = R,\ \mathbf{c}^{\mathrm{H}}\left(\mathbf{v}-\mathbf{c}\right) = 0\ \right\} ,\end{aligned} \tag{34}$$

where $R = \sqrt{r^2+\left\|\mathbf{c}\right\|_2^2}$. Obviously $\boldsymbol{\mathcal{O}}_m^{\mathrm{c}}\left(r,\ \mathbf{w}\right)$ can be taken as a $(m-2)$-complex sphere and its area is

$$A\left(\boldsymbol{\mathcal{O}}_m^{\mathrm{c}}\left(r,\ \mathbf{w}\right)\right) = \tfrac{2\pi^{m-1}r^{2m-2}}{\Gamma(m)} \ . \tag{35}$$

The area of the complex sphere cap then is

$$\begin{aligned}A\left(\boldsymbol{\mathcal{C}}_m^{\mathrm{c}}\left(\psi,\ \mathbf{w}\right)\right) &= \int_0^{\psi}A\left(\boldsymbol{\mathcal{O}}_m^{\mathrm{c}}\left(R\sin\phi,\ \mathbf{w}\cos\phi\right)\right)d\phi\\ &= \left[1-\sin^{(2m-2)}\left(\psi\right)\right]S_{2m}\left(R\right)\\ &= \Phi_m\left(\psi\right)S_{2m}\left(R\right)\end{aligned}$$

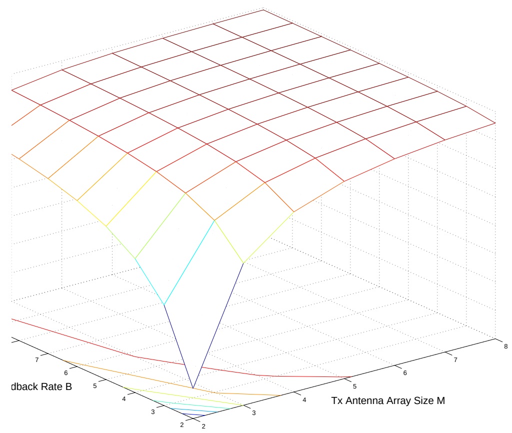

Fig. 3. Approximately maximum beamforming distortion α with various feedback rate B and Tx antenna array size M.

(36)

$\square$

Now the boundary of a Voronoi cell can be approximated by a hypershpere or a closed space curve decided in the following lemma.

Lemma 2. *The boundary of the uniform complex Voronoi cell $\boldsymbol{\mathcal{V}}_k$ can be approximated by a $(M-1)$-unit complex sphere or a closed complex space curve.*

$$\begin{aligned} \mathbf{B}\left(\boldsymbol{\mathcal{V}}_k\right) &\approx \boldsymbol{\mathcal{S}}_M^{\mathrm{c}}(1) \bigcap \boldsymbol{\mathcal{L}}_M^{\mathrm{c}}(\mathbf{w}_k,\ \cos(\theta)) \\ &= \{\mathbf{v}:\ \|\mathbf{v}\|=1,\ \angle\left(\mathbf{v},\ \mathbf{w}_k\right)=\theta\}\ , \end{aligned} \tag{37}$$

where $\boldsymbol{\mathcal{L}}_M^{\mathrm{c}}\left(\mathbf{w}_k,\ \cos(\theta)\right) = \left\{\mathbf{v}:\ \mathbf{v}^{\mathrm{H}}\mathbf{w}_k = \cos(\theta)\right\}$ denotes a complex space curve and θ is

$$\theta = \arccos\left(\left(\frac{2^B-1}{2^B}\right)^{\frac{1}{2M-2}}\right)\ . \tag{38}$$

From Lemma 2, the maximum beamforming power loss is

$$\alpha_0 = \left(\frac{2^B-1}{2^B}\right)^{\frac{1}{2M-2}}\ . \tag{39}$$

After the Voronoi boundary is approximately decided, we need find the probability distribution for beamforming distortion. In the following lemma, we outline the results for it.

Lemma 3. *The probability density function (PDF) $p\left(\alpha\right)$ and the cumulated density function (CDF) $P\left(\alpha\right)$ of α are*

$$\begin{aligned} p\left(\alpha\right) &= \mathit{Prob}\left\{x=\alpha\right\} \\ &= \begin{cases} 0 & 0 \le \alpha < \alpha_0 \\ \frac{(2M-2)\alpha\left(1-\alpha^2\right)^{M-2}}{\left(1-\alpha_0^2\right)^{M-1}} & \alpha_0 \le \alpha \le 1 \end{cases} \end{aligned} \tag{40}$$

and

$$\begin{aligned} P\left(\alpha\right) &= \mathit{Prob}\left\{x \le \alpha\right\} \\ &= \begin{cases} 0 & 0 \le \alpha < \alpha_0 \\ 1-\left(\frac{1-\alpha^2}{1-\alpha_0^2}\right)^{M-1} & \alpha_0 \le \alpha \le 1 \end{cases}\ . \end{aligned} \tag{41}$$

Proof. From Proposition 3, we have

$$\mathrm{Prob}\left\{\angle(\mathbf{v},\ \mathbf{w}) \le \psi\right\} = \Phi_M\left(\psi\right)\ . \tag{42}$$

Since α is defined by $\alpha = \cos\left(\angle(\mathbf{v},\ \mathbf{w})\right)$ and $\alpha \in [\alpha_0,\ 1]$, (42) can be re-written by

$$\mathrm{Prob}\left\{x \ge \alpha\right\} = \left(\frac{1-\alpha^2}{1-\alpha_0^2}\right)^{M-1}\ . \tag{43}$$

And we also have

$$\begin{aligned} \mathrm{Prob}\left\{x=\alpha\right\} &= \frac{d}{d\alpha}\mathrm{Prob}\left\{x \le \alpha\right\} \\ &= \begin{cases} 0 & 0 \le \alpha < \alpha_0 \\ \frac{(2M-2)\alpha\left(1-\alpha^2\right)^{M-2}}{\left(1-\alpha_0^2\right)^{M-1}} & \alpha_0 \le \alpha \le 1 \end{cases}\ . \end{aligned} \tag{44}$$

$\square$

Theorem 1. *The average received SINR for MIMO beamforming with B-bit feedback is*

$$\bar{\rho} \ge \frac{\bar{\alpha}^2 P_0}{\frac{\sigma_0^2}{\lambda}+\left(\frac{1-\bar{\alpha}}{M-1}\right)^2(P-P_0)}\ , \tag{45}$$

where λ is the idea channel gain, $K \in [1,\ \min\{M,\ N\}]$ denotes the number of selected beams, P_0 is the transmit power for this beam and P is the total transmit power for all K beams and

$$\begin{aligned} \bar{\alpha} &= E_{\mathbf{v}_i \in \mathcal{V}_k}\left\{\alpha_i\right\} \\ &\approx \int_1^{\alpha_0} \frac{(2M-2)\alpha^2\left(1-\alpha^2\right)^{M-2}}{\left(1-\alpha_0^2\right)^{M-1}} d\alpha\ . \end{aligned} \tag{46}$$

Proof.

$$\begin{aligned} \hat{\rho}_i &= \frac{\lambda_i \alpha_i^2 P_i}{n_0^2+\lambda_i \sum\limits_{l \ne i} \alpha_l^2 P_l} \\ &\approx \frac{\alpha_i^2 \rho_i}{1+\frac{1}{M-1}(1-\alpha_i)^2 \rho_i} \\ &= \Theta\left(\alpha_i\right)\ , \end{aligned} \tag{47}$$

where $\Theta\left(x\right)$ is defined by

$$\Theta\left(x\right) = \frac{x^2 \rho_i}{1+\frac{1}{M-1}(1-x)^2 \rho_i} \tag{48}$$

with $x \in [\alpha_0,\ 1]$. It can prove that $\Theta\left(x\right)$ when $x \in [\alpha_0,\ 1]$. Following Jensen's inequality, we have

$$\begin{aligned} \bar{\rho} &= \mathrm{E}\left\{\hat{\rho}_i\right\} \\ &= \mathrm{E}\left\{\Theta\left(\alpha_i\right)\right\} \\ &\ge \Theta\left(\bar{\alpha}_i\right) \\ &= \frac{\bar{\alpha}_i^2 \rho_i}{1+\frac{1}{M-1}(1-\bar{\alpha}_i)^2 \rho_i}\ , \end{aligned} \tag{49}$$

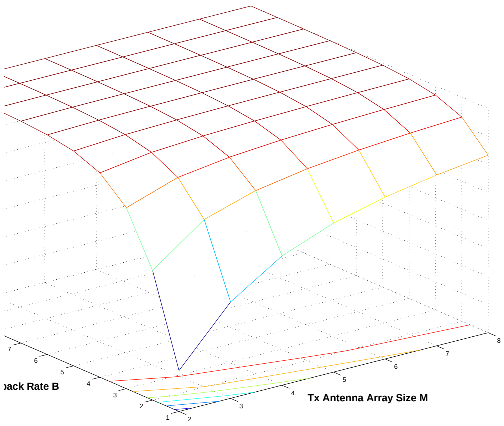

Fig. 4. A lower bound for averaged SINR $\bar{\rho}_i$ with $\rho_i = 10$dB with various feedback rate B and Tx antenna array size M.

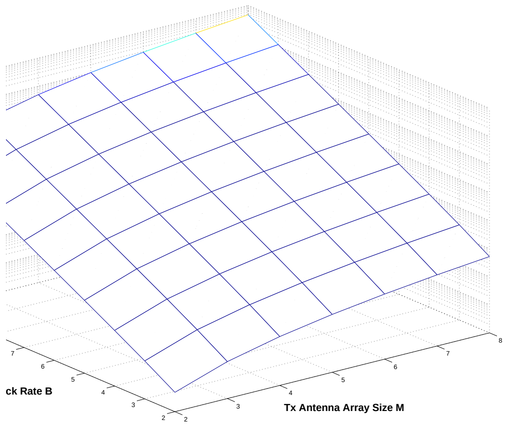

Fig. 5. SIR in high-SNR region with various feedback rate B and Tx antenna array size M.

□

It is difficult to give a closed-form solution to $\bar{\alpha}$ in (46).

Theorem 2. *The bound of the average received SINR satisfy the following inequality*

$$\bar{\rho}_i \quad \geq \quad \frac{\sigma_\alpha^2 \rho_i}{1+\frac{1}{M-1}(1-\sigma_\alpha)^2\rho_i}\,, \tag{50}$$

where σ_α denotes the standard deviation of α with σ_α^2 given by

$$\begin{aligned} \sigma_\alpha^2 &= E\left\{\alpha_i^2\right\} \\ &= \int_1^{\alpha_0} \frac{(2M-2)\alpha^3\left(1-\alpha^2\right)^{M-2}}{\left(1-\alpha_0^2\right)^{M-1}} d\alpha \\ &= \frac{1}{M} + \frac{M-1}{M}\alpha_0^2\,. \end{aligned} \tag{51}$$

C. High-SNR Region

Theorem 3. *In high-SNR region, MIMO beamforming becomes interference-limited and the average received SINR is*

$$\begin{aligned} \bar{\rho}_i &\geq \frac{\bar{\alpha}_i^2\rho_i}{1+\frac{1}{M-1}(1-\bar{\alpha}_i)^2\rho_i} \\ &\geq \frac{\left[1-\frac{M-1}{M}\left(1-\alpha_0^2\right)\right]^2\rho_i}{1+\frac{M-1}{M^2}\left(1-\alpha_0^2\right)^2\rho_i}\,. \end{aligned} \tag{52}$$

$$\bar{\rho}_i \quad \approx \quad \frac{\left[1-\frac{M-1}{M}\left(1-\alpha_0^2\right)\right]^2}{1+\frac{M-1}{M^2}\left(1-\alpha_0^2\right)^2}\,. \tag{53}$$

V. MIMO Codebook Design for Imperfect Channel Estimation

A. System Model and Problem Formulation

Consider M-input and L-output channels. The channels are assumed to be frequency-selective block fading, which means the random channels taps remain constant for some data packets and change to independent values for the next. The baseband received signal can be written as:

$$\begin{aligned} \mathbf{y}(t) &= \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_L(t) \end{bmatrix} \\ &= \begin{bmatrix} \sum_{m=1}^{M} h_{m,1}(t)\otimes s_m(t) \\ \sum_{m=1}^{M} h_{m,2}(t)\otimes s_m(t) \\ \vdots \\ \sum_{m=1}^{M} h_{m,L}(t)\otimes s_m(t) \end{bmatrix} + \mathbf{n} \\ &= \begin{bmatrix} \mathbf{h}_1^{\mathrm{T}}(t)\otimes\mathbf{s}(t) \\ \mathbf{h}_2^{\mathrm{T}}(t)\otimes\mathbf{s}(t) \\ \vdots \\ \mathbf{h}_L^{\mathrm{T}}(t)\otimes\mathbf{s}(t) \end{bmatrix} + \mathbf{n}\,, \end{aligned} \tag{54}$$

where $\mathbf{h}_l(t) = [h_{1,l}(t)\ h_{2,l}\ \ldots\ h_{M,l}(t)(t)]^{\mathrm{T}}$ for $t = 0,\ 1,\ \ldots\ L_c$ is the channel impulse response vector from the mth transmit antenna to the L receive antenna and L_c is the maximum channel order for all L subchannels. $s_m(t)$ denotes the transmitted symbol from the mth antenna and n represents the independent additive Gaussian white noise. Due to the commutativity of convolution, the stretched discrete received signal vector could as well be expressed by the following expression

$$\mathbf{y} \quad = \quad \mathbf{S}\mathbf{h} + \mathbf{n} \ , \tag{55}$$

where $\mathbf{h}$ is the stretched channel response vector defined by

$$\mathbf{h} \quad = \quad [\mathrm{vec}(\mathbf{H}_1)\ \mathrm{vec}(\mathbf{H}_2)\ \ldots\ \mathrm{vec}(\mathbf{H}_M)]^{\mathrm{T}} \tag{56}$$

with $\mathbf{H}_k = [\mathbf{h}_k(0)\ \mathbf{h}_k(1)\ \ldots\ \mathbf{h}_k(L_c-1)]$ and

$$\mathbf{S} \quad = \quad \mathrm{kron}\left([\mathbf{S}_1\ \mathbf{S}_2\ \ldots\ \mathbf{S}_M]^{\mathrm{T}}\ ,\ \mathbf{I}\right) \tag{57}$$

with

$$\mathbf{S}_k \quad = \quad \begin{bmatrix} s_k(Q) & \cdots & s_k(Q-L_c) \\ s_k(Q-1) & \cdots & s_k(Q-L_c-1) \\ \vdots & \ddots & \vdots \\ s_k(1) & \cdots & s_k(1-L_c) \end{bmatrix} . \tag{58}$$

Due to pilot design and channel impair, the

$$\mathbf{R} \quad = \quad \mathbf{H}\mathbf{H}^{\mathrm{H}} \ . \tag{59}$$

When there is an estimation error in $\mathbf{R}$, the eigen-decomposition of it becomes

$$\hat{\mathbf{R}} \quad = \quad \sum_{k=1}^{K} \hat{\lambda}_k \hat{\mathbf{v}}_k \hat{\mathbf{v}}_k^{\mathrm{H}} \ . \tag{60}$$

The first-order approximation of eigen-values and eigen-vectors is

$$\hat{\lambda}_k \quad = \quad \lambda_k + \hat{\mathbf{v}}_k^{\mathrm{H}} \mathbf{\Delta}_{\mathbf{R}} \hat{\mathbf{v}}_k \tag{61}$$

and

$$\hat{\mathbf{v}}_k = \mathbf{v}_k + \left(\hat{\mathbf{R}} - \hat{\lambda}_k \mathbf{I}\right)^{+} \left[\Delta_{\lambda_k}\mathbf{I} - \mathbf{\Delta}_{\mathbf{R}}\right] \mathbf{v}_k \tag{62}$$

where $\Delta_{\lambda_k} = \hat{\lambda}_k - \lambda_k$ and $\mathbf{\Delta}_{\mathbf{R}} = \hat{\mathbf{R}} - \mathbf{R}$.

Proposition 4.

$$\left[\mathbf{\Delta}_{\mathbf{R}} - \Delta_{\lambda_k}\mathbf{I}\right]\mathbf{v}_k \quad \approx \quad \sum_{i=1}^{K} \lambda_i \mathbf{v}_i \mathbf{\Delta}_{\mathbf{v}_i}^{\mathrm{H}} \mathbf{v}_k + \lambda_k \Delta_{\mathbf{v}_k} \tag{63}$$

$$\left(\hat{\mathbf{R}} - \hat{\lambda}_k \mathbf{I}\right)\Delta_{\mathbf{v}_k} \quad \approx \quad \sum_{i=1}^{K} \lambda_i \mathbf{v}_i \mathbf{\Delta}_{\mathbf{v}_i}^{\mathrm{H}} \mathbf{v}_k + \lambda_k \Delta_{\mathbf{v}_k} \tag{64}$$

$$\begin{aligned} \alpha_k \quad &= \quad -\hat{\mathbf{v}}_k^{\mathrm{H}} \Delta_{\mathbf{v}_k} \\ &\approx \quad -\hat{\mathbf{v}}_k^{\mathrm{H}} \left(\hat{\mathbf{R}} - \hat{\lambda}_k \mathbf{I} - \lambda_k \mathbf{I}\right)^{+} \sum_{i=1}^{K} \lambda_i \mathbf{v}_i \mathbf{\Delta}_{\mathbf{v}_i}^{\mathrm{H}} \mathbf{v}_k \\ &\approx \quad \frac{1}{\lambda_k} \hat{\mathbf{v}}_k^{\mathrm{H}} \lambda_i \mathbf{v}_i \mathbf{\Delta}_{\mathbf{v}_i}^{\mathrm{H}} \mathbf{v}_k \end{aligned} \tag{65}$$

B. Pilot Design

1) Orthogonal Multiplexed Pilot:

2) Superimpose Pilot:

C. Cramer-rao Lower Bound

The lower bound to the mean-squared errors of unbiased estimates can be given by the Cramer-Rao bound (CRB), which is defined as the inverse of the Fisher Information Matrix (FIM). If we denotes $\boldsymbol{\vartheta} = \left[\mathbf{h}^{\mathrm{T}}\ \mathbf{s}_d^{\mathrm{T}}\right]^{\mathrm{T}}$, the complex FIM is given by

$$\begin{aligned} \mathrm{F}(\boldsymbol{\vartheta}) \quad &= \quad \mathrm{E}\left\{ \left[\frac{\partial \ln \Pr(\mathbf{y}|\vartheta)}{\partial \vartheta^*}\right] \left[\frac{\partial \ln \Pr(\mathbf{y}|\vartheta)}{\partial \vartheta^*}\right]^{\mathrm{H}} \right\} \\ = \quad &\sigma_0^{-2} \begin{bmatrix} \mathrm{E}\left\{\mathbf{S}^{\mathrm{H}}\mathbf{S}\right\} + \rho_h^2 \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathrm{E}\left\{\boldsymbol{\mathcal{H}}^{\mathrm{H}}\boldsymbol{\mathcal{H}}\right\} + \rho_{s_d}^2 \mathbf{I} \end{bmatrix} \end{aligned} , \tag{66}$$

where $\rho_h^2 = \mathrm{E}\left\{\left|\frac{\partial \ln \Pr(h)}{\partial h^*}\right|^2\right\}$ and $\rho_{s_d}^2 = \mathrm{E}\left\{\left|\frac{\partial \ln \Pr(s_d)}{\partial {s_d}^*}\right|^2\right\}$.

1) Orthogonal Multiplexed Pilot:

2) Superimpose Pilot:

$$\begin{aligned} \hat{r} \quad &= \quad \hat{\mathbf{u}}^{\mathrm{H}} \mathbf{H} \hat{\mathbf{w}} s \\ &= \quad [\mathbf{u} + (\hat{\mathbf{u}} - \mathbf{u})]^{\mathrm{H}} \mathbf{H} \left\{\mathbf{w} + [Q(\hat{\mathbf{v}}) - Q(\mathbf{v})]\right\} s \\ &\approx \quad \mathbf{u}^{\mathrm{H}} \mathbf{H} \mathbf{w} s + \left(\mathbf{\Delta}_{\mathbf{u}}^{\mathrm{H}} \mathbf{H} \mathbf{w} + \mathbf{u}^{\mathrm{H}} \mathbf{H} \mathbf{\Delta}_{\mathbf{w}}\right) s \end{aligned} \tag{67}$$

$$\begin{aligned} \mathbf{P}_{\hat{\mathbf{r}}} \quad &= \quad \mathbf{H}\hat{\mathbf{W}}\mathbf{R}_{\mathbf{s}}\hat{\mathbf{W}}^{\mathrm{H}}\mathbf{H}^{\mathrm{H}} \\ = \quad &\mathbf{H}\left[\mathbf{V} + \left(\hat{\mathbf{W}} - \mathbf{V}\right)\right]\mathbf{P}_{\mathbf{s}}\left[\mathbf{V} + \left(\hat{\mathbf{W}} - \mathbf{V}\right)\right]^{\mathrm{H}}\mathbf{H}^{\mathrm{H}} \\ = \quad &\frac{P}{K}\mathbf{H}\left[\mathbf{V} + \left(Q(\hat{\mathbf{V}}) - \mathbf{V}\right)\right]\left[\mathbf{V} + \left(Q(\hat{\mathbf{V}}) - \mathbf{V}\right)\right]^{\mathrm{H}}\mathbf{H}^{\mathrm{H}} \\ \approx \quad &\frac{P}{K}\mathbf{\Lambda} + \frac{P}{K}\mathbf{H}\left[Q(\hat{\mathbf{V}}) - \mathbf{V}\right]\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{U}^{\mathrm{H}} \\ &+\frac{P}{K}\mathbf{U}\mathbf{\Lambda}^{\frac{1}{2}}\left[Q(\hat{\mathbf{V}}) - \mathbf{V}\right]^{\mathrm{H}}\mathbf{H}^{\mathrm{H}} \end{aligned} \tag{68}$$

$$\begin{aligned} \mathbf{\Delta}_{\mathbf{v}} \quad &= \quad \hat{\mathbf{w}} - \mathbf{v} \\ &= \quad Q(\hat{\mathbf{v}}) - \mathbf{v} \\ &= \quad [Q(\hat{\mathbf{v}}) - \hat{\mathbf{v}}] + (\hat{\mathbf{v}} - \mathbf{v}) \\ &= \quad [Q(\hat{\mathbf{v}}) - Q(\mathbf{v})] + [Q(\mathbf{v}) - \mathbf{v}] \end{aligned} \tag{69}$$

VI. MIMO Beamforming with Interference Cancellation

VII. MIMO Relay with Finite-Rate Feedback

VIII. Conclusions

References

[1] Bo Wang, Junshan Zhang and Anders Host-Madsen. On the capacity of mimo relay channels. *IEEE Transactions on Information Theory*.

[2] A. Narula, M. J. Lopez, M. D. Trott and G. W. Wornell. Efficient use of side information in multiple-antenna data transmission over fading channels. *IEEE J. Select Areas in Communications*, 16(8):1423–1436, October 1998.

[3] J. Conway, R. Hardin, and N. Sloane. Packing lines, planes, etc.: Packings in grassmannian spaces, 1996.