On MIMO Relay Network with Finite-Rate Feedback

Byung K. Yi, Shu Wang and Soon Y. Kwon
LG Electronics Mobile Research
San Diego, CA 92131
Email: {byungkyi, swang, sykwon}@lge.com

***Abstract*—Though feedback plays an important role in wireless relay networking, the information it carries is inherently imperfect and limited. In this paper, the imperfection of feedback and its impact on MIMO relay network are analyzed and presented. At first, we show that the imperfection of feedback is the result of both inaccurate estimation and unreliable transmission. For this, the achievable rate-distortion region and rate-reliability region for feedback are calculated. These two regions not only reveal what is the achievable performance of feedback itself but also help us understand how much feedback is necessary for MIMO relay. After that, we investigate how imperfect feedback affects the achievable throughput and networking protocol design for MIMO relay network. We start from a simple two-hop MIMO relay network and formulate the achievable throughput of MIMO relay networks for the cases with or without interference cancellation. We then compare the results and discuss the effect of finite-rate feedback and the scaling laws of throughput by codebook size and interference. Thereafter, we extend our analysis to a general multi-hop relay network and examine how the end-to-end throughput is limited by network size, interference, codebook design, etc. Through our discussion and analysis, the tradeoffs for MIMO relay network design are revealed and some considerations for MIMO relaying protocol design are presented. Some numerical results are also presented for demonstrating our conclusions.**

I. Introduction

Multi-antenna systems have received much attention over the last decades, due to their promise of higher spectrum efficiency with no transmit power increase. Combining multi-antenna transceiver with relay network is essential not only to provide comprehensive coverage but also to help relieve co-channel interference in existing wireless systems in a cost-effective fashion. For multiple-input multiple-output (MIMO) transmission, it is well-known that their performance and complexity can be improved by making channel state information (CSI) available at the transmitter side. This is usually achieved through a reverselink CSI feedback channel from receiver, e.g., there is a reverslink channel quality indicator channel (R-CQICH) for CSI feedback in UMB (Ultra Mobile Broadband), a 3.5G mobile network standard recently developed by 3GPP2. In practice, CSI received by transmitters is not perfect and suffers from various impairments and limitations that include round-trip delay, channel estimation error, codebook limitation, etc. Therefore the actual link throughput is degraded. This kind of degradation becomes more serious if the end-to-end capacity is considered for a multi-hop MIMO relay network.

MIMO beamforming with quantized feedback has intensively been investigated since 1990s [1]. MIMO channel quantization as well as codebook design in general is a NP-hard Voronoi decomposition problem. The Voronoi region for a uniform random codebook is known to be upper-bounded by the disk-covering problem solution and lower-bounded by the sphere-packing problem solution. These two problems themselves are still open. MISO/MIMO beamforming systems with perfect CQI Lloyd vector quantization (VQ) [2], different channel model [3] or different performance metrics [4], [5] have intensively been investigated. It is linked to Grassmannian line packing problem [6]. However, most of existing work is done without considering pilot design, channel estimation and the reliability of feedback, even though they are among the most important components of actual multi-antenna systems. In reality, MIMO CSI is estimated with forwardlink common pilot channels sent from each transmitter antenna. An overview of pilot-assisted transmission (PAT) including pilot placement and channel estimation can be found in [7]. In most multi-antenna systems, pilot channels are designed to be orthogonal to other channels and periodically sent by transmitter. Nonorthogonal pilot design like superimposed pilots (SIP) has recently received much attention for channel estimation too [8]. Optimal pilot placement was investigated in [9]. Besides pilot design, the feedback capacity and reliability have intensively been investigated over decades too. Though feedback doesn't increase the capacity of memoryless channels [10], [11], a feedback coding scheme with the decoding error probability decreasing more rapidly than the exponential of any order is achievable [12]. Since CSI feedback plays such a critical role in MIMO transmission, it is desired to understand how MIMO pilot and codebook design affects system behavior, what are the tradeoffs, etc. And these problems become more critical when a multi-hop MIMO relay network.

The feedback and sharing of CSI and/or network status information (NSI) help wireless relay network achieve high throughput and reliability with a little overhead increase. For example, CSI feedback helps nodes realize distributed cooperation for increasing the throughput and reliability of wireless relay networks. Cooperation diversity for wireless relay network has heavily been investigated in the past several years. The concept of distributed cooperation diversity is knowingly pioneered by Sendonaris et al. [13], where the transmitters cooperate with each other by repeating symbols

of others. It shows that a higher rate is achievable with this cooperation. Almost at the same time, this concept is also developed through other techniques such as code combining [14], coherent soft combining [15], power control [16] and later opportunistic routing [17]. Most of them are implemented with CSI feedback in the assumption. Besides this, from a network perspective, it is known that NSI feedback can also assist each source terminal or relay terminal to shape the dynamic behavior of the network and increase network agility through proper resource allocation [18]. Due to the limitation of the measuring, link capacity and network resource in reality, however, most CSI or NSI sent back by receivers is neither perfect nor sufficient in nature. It is interesting and important to understand the effect of imperfect feedback on wireless link and network, which are still not clear from many perspectives.

To the best of our knowledge, no systematic discussion on the effect of imperfect feedback on MIMO relay network has been found so far, though it has thoroughly been studied regarding the performance of MIMO channels by many literatures [3]–[5]. In this paper, the effects of finite-rate feedback on MIMO relay networks are investigated. At first, the distortion and reliability of MIMO feedback are analyzed. For this, the forwardlink is modelled as a Gaussian channel with multiplexed pilot and data signals for channel estimation. The Shannon rate-distortion lower bound is derived for the forward link and a heuristic sphere-packing upper bound is given for MIMO codebook. The achievable rate-distortion region is limited between these two bounds. In reverselink, the MIMO feedback channel is modelled as a Gaussian binary erasure channel with additive Gaussian noise and binary erasure. The binary erasure part is a simplified but popular model for the channel fading in reality. With the erasure mechanism, the reverselink is roughly simplified into a Gaussian channel model with forwardlink feedback. The achievable rate-reliability region of feedback channel is presented with Shannon reliability bounds. After that, the MIMO beamforming signal-to-interference/noise ratio (SINR) with imperfect feedback is formulated and analyzed. It shows that channel estimation error affects the accuracy of both MIMO beamforming precoding and power allocation. After this, we start from a simple two-hop relay network, where relays directly forward received signals to the destination, for studying the effect of finite-rate feedback on MIMO relay network. The achievable throughput for a two-hop relay network with or without interference cancellation is then formulated, respectively. With interference cancellation, the throughput of MIMO relay network mostly suffers from precoding loss and throughput essentially is power-limited. With no interference cancellation, the network throughput suffers from both inter-beam interference (IBI) and precoding loss. Because of IBI, the throughput of the MIMO relay networks actually becomes interference-limited, especially in high signal-to-noise ratio (SNR) region. After this, we extend our discussion to a multi-hop MIMO relay network, where each MIMO relay stream experiences the same number of hops. It shows that more relays can actually increase the IBI effect on the MIMO relay network. This means, in high SNR region, it may be more efficient for MIMO relay network to use more relays instead of SNR for achieving higher network throughput. This is different to the networks operating in low interference region, where SNR increase may be a better choice. Besides this, the scalability of MIMO relay throughput by codebook size, pilot size, relay network size, etc. is also discussed. Finally, some considerations on relaying protocol design are presented.

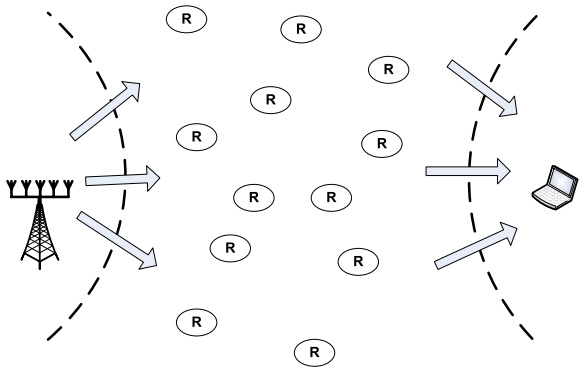


Fig. 1. A MIMO Relay Network with K relay terminals and one pair of multi-antenna source-destination terminals.

II. Network Model And Problem Description

We consider a wireless MIMO relay network consisting of $K+2$ terminals with a designated multi-antenna source and destination pair and L relay terminals $\mathcal{R}$ located randomly and independently in a domain of fixed area. This is shown in Fig. 1. Both the source terminal $\mathcal{S}$ and the destination terminal $\mathcal{D}$ are equipped with M antennas. And we assume there is no direct link between $\mathcal{S}$ and $\mathcal{D}$ and each $\mathcal{R}$ is equipped with N antennas. In the following discussion, we assume all channels experience independent frequency-flat block fading with the same block length. Therefore, a MIMO link between $\mathcal{S}$ and $\mathcal{R}$ consisting of a transmitter with M transmit antennas, a receiver with N receive antenna and a MIMO channel represented by the $N \times M$ matrix $\mathbf{H} = [\mathbf{h}_1\ \mathbf{h}_2\ \dots\ \mathbf{h}_N]^{\mathrm{T}}$ with $\mathbf{h}_n = [h_{n,1}\ h_{n,2}\ \dots\ h_{n,M}]^{\mathrm{T}}$. The $N \times 1$ received signal $\mathbf{y}$ is

$$\mathbf{y} \;=\; [y_1\ y_2\ \dots\ y_N]^{\mathrm{T}} \;=\; \mathbf{HWx} + \mathbf{n} \qquad (1)$$

where $\mathbf{x} = [x_1\ x_2\ \dots\ x_M]^{\mathrm{T}}$ is the $M \times 1$ signal vector transmitted by the source with $\mathbf{R}_{\mathbf{x}} = \mathrm{E}\left\{\mathbf{x}\mathbf{x}^{\mathrm{H}}\right\} = \frac{P}{M}\mathbf{I}_M$, P is the total transmit power, $\mathbf{W} = [\mathbf{w}_1\ \mathbf{w}_2\ \dots\ \mathbf{w}_M]$ is a $M \times M$ MIMO beamforming precoding matrix with $\|\mathbf{w}_m\|_2 = 1$, $\mathbf{n} \sim \mathcal{CN}(0, \sigma^2\mathbf{I}_N)$ is a complex circular white Gaussian vector, $[*]^{\mathrm{T}}$ and $[*]^{\mathrm{H}}$ denotes the transpose operator and Hermitian conjugate operator, respectively. It is known that the normalized MIMO channel achievable throughput or MIMO channel spectral efficiency is

$$\eta = \log\left|\mathbf{I}_N + \tfrac{P}{M}\mathbf{HWW}^{\mathrm{H}}\mathbf{H}^{\mathrm{H}}\right| = \sum_{n=1}^{N}\log\left(1+\tfrac{\rho_n}{M}\right) \qquad (2)$$

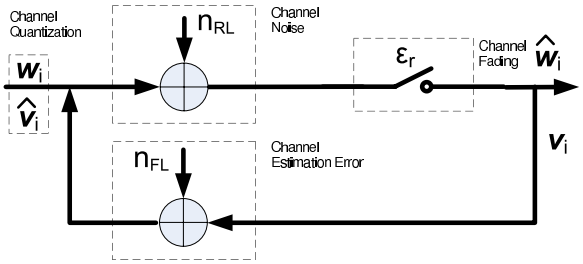


Fig. 2. Noisy Gaussian binary erasure feedback channel with channel quantization.

where ρ_n denotes the received signal-to-noise ratio (SNR) of the nth beam, which is given by

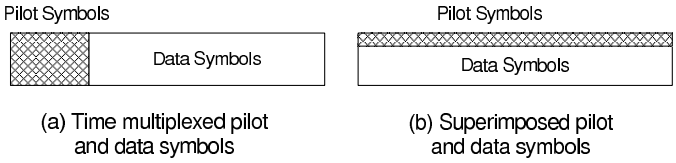
$$\rho_n \quad = \quad \frac{\text{var}\{\mathbf{H}\mathbf{w}_n x_n\}}{\sigma^2} \quad = \quad \frac{\lambda_n P_n}{\sigma^2} \tag{3}$$

with λ_n denoting the antenna gain of the nth beam and also the nth eigenvalue of the MIMO channel autocorrelation matrix $\mathbf{R_H}$ defined by

$$\mathbf{R_H} = \mathbf{H}^{\mathrm{H}}\mathbf{H} = \sum_{m=1}^{M} \mathbf{h}_m \mathbf{h}_m^{\mathrm{H}} = \sum_{i=1}^{M} \lambda_i \mathbf{v}_i \mathbf{v}_i^{\mathrm{H}} \ , \tag{4}$$

$\mathbf{v}_i$ denotes the ith eigenvector, the operator $|*|$ denotes the determinant of matrix $*$, and $\text{var}\{*\}$ denotes the variance of random variable $*$. (2) and (3) are the results from the assumption of perfect CSI available at transmitter. This may not be consistent with practical applications.

MIMO beamforming with finite-rate feedback is modelled as a noisy Gaussian binary erasure feedback channel depicted in Fig. 2. In reality, the receivers estimate $\mathbf{H}$ or $\mathbf{R_H}$ with the pilots sent by transmitter. Accuracy of the channel estimation depends on both forwardlink design and receiver design. The pilot transmission is important for receiver to efficiently estimate CSI. There are two popular pilot patterns, TMP and SIP, receiving much attention for MIMO CSI estimation. They are shown in Fig. 3. TMP is a typical example of orthogonal pilot design where pilot symbols and data symbols are separated in time and/or frequency domain, which makes them orthogonal to each other. With orthogonal pilots, the CSI estimation and data demodulation can be done separately which may lead to simple receiver design [9]. SIP does the opposite. In SIP design, pilots and data nonorthogonally share the same time period and frequency band. In this case, joint channel estimation/demodulation and the demodulation with pilot interference cancellation are among the most popular receiver design techniques [8]. After channel estimation, the receiver chooses a beamforming vector from a shared MIMO precoding codebook. This is called channel quantization. It means the receiver actually feeds back the chosen precoding index(es) to transmitter(s) instead of channel response for MIMO precoding. With a MIMO codebook $\boldsymbol{\mathcal{W}}$ of the size 2^R consisting of M-dimensional normalized vectors $\{\mathbf{w}_1, ..., \mathbf{w}_{2^R}\}$, it usually takes the receiver to feedback R bits for precoding each beam stream. The codebook is designed to quantize channel responses with certain distortion measures [2]. It is related to Grassmannian line packing,



(a) Time multiplexed pilot and data symbols

(b) Superimposed pilot and data symbols

Fig. 3. Pilot patterns for channel estimation.

the spherical packing on unit sphere $\boldsymbol{\mathcal{S}}_n(1)$ [6] etc., where $\boldsymbol{\mathcal{S}}_n(r) = \{\mathbf{v} : \|\mathbf{v}\| = r\}$ denotes $(n-1)$-sphere of radius r. When a codeword $\mathbf{w}_k$ is chosen to precode signals for ith beam $\mathbf{v}_i$ at the transmitter side, degradation will happen on the signals received by receiver because of the imperfect channel estimation and finite-rate feedback. The degradation on received signals can be expressed by

$$\begin{aligned} \delta_i \quad &= \quad \min_{\mathbf{w}\in\boldsymbol{\mathcal{W}}} \|\mathbf{H}\mathbf{v}_i - \mathbf{H}\mathbf{w}\| \\ &= \quad \lambda_i \left(1 - \mathbf{v}_i^{\mathrm{H}}\mathbf{w}_k\right) - \sum_{j\neq i} \lambda_j \mathbf{v}_j^{\mathrm{H}}\mathbf{w}_k \end{aligned} \tag{5}$$

where $\mathbf{v}_i$ is the ideal precoding vector for ith eigen-mode of $\mathbf{H}$. It is known that δ_i, which depends on $\mathbf{v}_i$, $\mathbf{w}_k$ and λ_i, etc., is one of the major factors limiting closed-loop MIMO beamforming throughput. The reverselink channel is modelled by concatenating a Gaussian channel and a binary erasure channel due to feedback redundancy and the employed erasure mechanism, which drops those CSI packets with low SNR. On the other hand, there are two components in (5). The first term is

$$\Delta \quad = \quad \lambda_i \left(1 - \mathbf{v}_i^{\mathrm{H}}\mathbf{w}_k\right) \quad = \quad (1-\alpha_i)\,\lambda_i \ , \tag{6}$$

which denotes the degradation of antenna gain with

$$\alpha_i \quad = \quad \mathbf{v}_i^{\mathrm{H}}\mathbf{w}_k \ . \tag{7}$$

The second term is

$$I \quad = \quad \sum_{j\neq i} \lambda_j \mathbf{v}_j^{\mathrm{H}}\mathbf{w}_k \quad = \quad \sum_{j\neq i} \alpha_j \mathbf{w}_k \tag{8}$$

denoting the IBI due to the correlation between $\mathbf{w}_k$ and other eigen-modes $\{\mathbf{v}_j : j \neq i\}$. The average received SINR for the ith eigen-mode with the precoding codeword $\mathbf{w}_k$ can be written by

$$\begin{aligned} \bar{\rho}_i \quad &= \quad \mathrm{E}_{\mathbf{v}_i\in\mathcal{V}_k} \left\{ \frac{\lambda_i \left(\mathbf{v}_i^{\mathrm{H}}\mathbf{w}_k\right)^2 P_i}{n_0^2 + \lambda_i \sum\limits_{l\neq i}^{K} \left(\mathbf{v}_i^{\mathrm{H}}\mathbf{w}_l\right)^2 P_l} \right\} \\ &= \quad \mathrm{E}_{\mathbf{v}_i\in\mathcal{V}_k} \left\{ \frac{\alpha_i^2 \rho_i}{1 + \lambda_i \sum\limits_{l\neq i}^{K} \frac{\alpha_l^2}{\lambda_l} \rho_l} \right\} \ , \end{aligned} \tag{9}$$

where $K \leq \min\{M,\ N\}$ denotes the spatial multiplexity order.

III. The Imperfection of Feedback

For feedback channel design, the rate-distortion function and the rate-reliability function are of the most important tools.

With the rate-distortion function, the region of achievable rate distortion pair $(R,\ D)$ with the squared error distortion

$$D\left(R\right) \quad = \quad \max_{i} \sum_{m=1}^{M} \left(w_{i,m} - v_{i,m}\right)^2 \tag{10}$$

is of the most interests for understanding the effects of channel estimation and quantization. On the other hand, the achievable rate reliability pair $(R,\ \epsilon)$, with the error exponent ϵ denoting how fast the reverselink bit-error rate (BER) P_e decays for the transmit rate R and given by

$$\epsilon \quad = \quad \lim_{n\to\infty} \sup \epsilon_n \quad = \quad \lim_{n\to\infty} \sup -\tfrac{1}{n} \ln P_e \ , \tag{11}$$

where $\epsilon_n = -\frac{1}{n}\ln P_e$ denotes the reliability of n-bits block transmission and n is the minimum block-length needed in order to operate at rate R with the error probability P_e, is of the most importance to understand the performance of feedback design. In the following section, we present how the imperfect channel estimation limits the achievable rate-distortion region and rate-reliability region of MIMO feedback channel.

A. Rate-Distortion Region

1) Bemaforming Mismatch Lower Bound: In general, the channel quantization distortion is decided by channel quality, channel estimation and codebook design. Given σ^2_{FL} the mean-squared error (MSE) of forwardlink channel estimation, the minimum rate at distortion $D_{\mathbf{h}}$ is given by

$$\begin{aligned} R\left(D_{\mathbf{h}}\right) \quad &= \quad \min_{E\|\mathbf{h}-\hat{\mathbf{w}}\|_2^2 \le D_{\mathbf{h}}} I\left(\mathbf{h}_i;\ \hat{\mathbf{h}}_i\right) \\ &= \quad \min_{E\|\mathbf{h}_i-\hat{\mathbf{w}}_i\|_2^2 \le D_{\mathbf{h}}} \left\{H\left(\mathbf{h}\right) - H\left(\mathbf{h}|\hat{\mathbf{h}}\right)\right\} \\ &\ge \quad M\log_2\left(2\pi e\sigma^2_{\text{FL}}\right) - M\log_2\left(2\pi e E\|h_i - \hat{w}_i\|_2^2\right) \\ &= \quad M\log_2\left(\frac{\sigma_h^2}{\frac{1}{M}D_{\mathbf{h}}+\sigma^2_{\text{FL}}}\right) \end{aligned} \tag{12}$$

with $D_{\mathbf{h}}$ the channel quantization distortion defined by

$$D_{\mathbf{h}} \quad = \quad \max \mathrm{E}\left\|\mathbf{h} - \hat{\mathbf{h}}\right\|_2^2 \tag{13}$$

and $0 \le \sigma_h^2 \le \frac{1}{M}D_{\mathbf{h}} + \sigma^2_{\text{FL}}$; otherwise, $R\left(D_{\mathbf{h}}\right) = 0$. Meanwhile, the lower bound to the MSE of unbiased channel estimates of $\mathbf{h}$ is given by Cramer-Rao lower bound, which is defined as the inverse of the Fisher information matrix [7].

In most existing MIMO beamforming systems, the receiver tracks the channel norm information for link adaptation purpose and the phase information for beamforming precoding. In this case, (13) can be rewritten by

$$\begin{aligned} D_{\mathbf{h}} \quad &= \quad \max \mathrm{E}\left\|\mathbf{h} - r\mathbf{w}_\phi\right\|_2^2 \\ &= \quad \max \mathrm{E}\left\|r\left(\mathbf{v} - \mathbf{w}_\phi\right)\right\|_2^2 \\ &= \quad 2M\sigma_h^2\left(1 - \sqrt{1-D_\phi}\right) \end{aligned} \tag{14}$$

with

$$D_\phi \quad = \quad \mathrm{var}\left\{\sin\angle\left(\mathbf{v},\ \mathbf{w}\right)\right\} \tag{15}$$

denoting the phase quantization deviation. With (12), a lower bound for distortion rate of channel quantization for MIMO beamforming is

$$\begin{aligned} R(D_\phi) \quad &= \quad \min_{\mathrm{E}\|\mathbf{v}-\hat{\mathbf{v}}\|_2^2 \le D_\phi} I\left(\mathbf{v};\ \hat{\mathbf{v}}\right) \\ &\ge \quad \max\left\{-M\log_2\left(2 - 2\sqrt{1-D_\phi} + \tfrac{\sigma^2_{\text{FL}}}{\sigma_h^2}\right),\ 0\right\} \end{aligned} \tag{16}$$

for each beam. With the feedback rate of R, (16) also tells us that the minimum precoding mismatch for forwardlink MIMO beamforming is

$$\begin{aligned} D_\phi \quad &= \quad \max \mathrm{E}\left\|\hat{\mathbf{v}} - \mathbf{v}\right\|_2^2 \\ &\le \quad 1 - \left(1 - 2^{-\frac{R}{M}-1} + \tfrac{\sigma^2_{\text{FL}}}{2\sigma_h^2}\right)^2 \ . \end{aligned} \tag{17}$$

2) Bemaforming Mismatch Upper Bound: Here we discuss the beamforming mismatch due to codebook design. With proper channel quantization, the maximum beamforming mismatch can be determined by the largest radius of the codebook Voronoi cell $\left\{\boldsymbol{\mathcal{V}}_i :\ 1 \le i \le 2^R\right\}$, which in general is the solution to the disk-covering problem that still is open. Instead of the exact boundary for the Voronoi cell $\boldsymbol{\mathcal{V}}_i$, we suggest a heuristic approach using sphere-packing bound and sphere cap to approximate the actual polytope boundary. It is an approximate of the sphere packing solution, in which all spheres are supposed to be non-overlappedly placed. With our approach, sphere caps are overlapped with each other in space but the interior of them has the same area as the Voronoi cell. The border of this sphere cap is termed sphere-packing boundary. The relationship between sphere-packing boundary and Voronoi cell is shown in Fig. 4. For an uniform random codebook of size 2^R in M-dimensional Euclid space, the area of a Voronoi cell is given by

$$A\left(\boldsymbol{\mathcal{V}}_k\right) \quad = \quad \tfrac{2\pi^M}{2^R\Gamma(M)} \tag{18}$$

where $\Gamma\left(*\right)$ denotes the gamma function. On the other hand, the area of $(m-1)$-complex sphere cap $\boldsymbol{\mathcal{C}}^{\text{c}}_m\left(\psi,\ R\right)$ with contact angle ψ and radius r is

$$A\left(\boldsymbol{\mathcal{C}}^{\text{c}}_m\left(\psi,\ r\right)\right) \quad = \quad \left[1 - \cos^{2(m-1)}\left(\psi\right)\right] S^{\text{c}}_m\left(r\right), \tag{19}$$

where $\boldsymbol{\mathcal{S}}^{\text{c}}_m\left(r\right)$ denotes a M-dimension complex ball in Euclid space with the radius r. The relationship between $\boldsymbol{\mathcal{S}}^{\text{c}}_m\left(r\right)$ and $\boldsymbol{\mathcal{C}}^{\text{c}}_m\left(\psi,\ r\right)$ can be shown in Fig. 5 and it can be verified that

$$A\left(S^{c}_m\left(r\right)\right) \quad = \quad A\left(\boldsymbol{\mathcal{C}}^{\text{c}}_m\left(\pi,\ r\right)\right) \ . \tag{20}$$

With matching the sum area of the sphere-cap with the whole sphere area, the boundary of a Voronoi cell can be

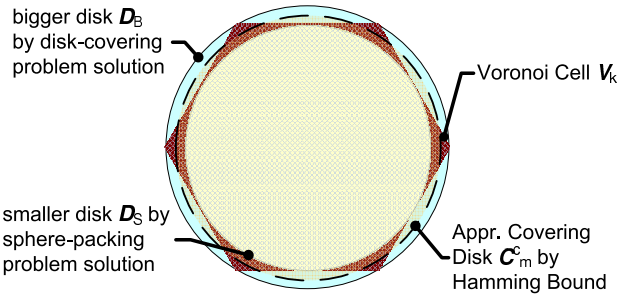


Fig. 4. Voronoi cell and various bounds

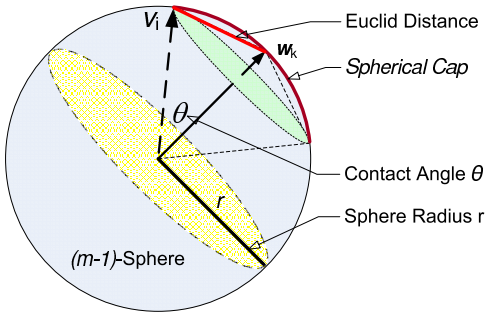


Fig. 5. Hypersphere and hypersphere Cap.

approximated by a hypershpere or a closed space curve defined in the following proposition.

Proposition 1. *(Sphere-Packing Boundary) 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.*

$$\begin{aligned}\mathbf{B}\left(\boldsymbol{\mathcal{V}}_k\right) &\approx \boldsymbol{\mathcal{S}}^{\mathrm{c}}_M(1)\bigcap\boldsymbol{\mathcal{L}}^{\mathrm{c}}_M(\mathbf{w}_k,\ \cos(\theta)) \\ &= \left\{\mathbf{v}:\ \|\mathbf{v}\|=1,\ \angle(\mathbf{v},\ \mathbf{w}_k)=\theta\right\}\ ,\end{aligned} \tag{21}$$

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

$$\theta \quad = \quad \arccos\left(\alpha_0\right) \tag{22}$$

with

$$\alpha_0 \quad = \quad \left(\tfrac{2^R-1}{2^R}\right)^{\frac{1}{2M-2}}\ . \tag{23}$$

And, with Proposition 1 and the assumption of channel status being uniformly distributed over codebook, the standard deviation of precoding mismatch can be calculated with the following conclusion.

Lemma 1. *The standard deviation of α is given by*

$$\sigma^2_\alpha \quad = \quad \mathrm{E}\left\{\alpha^2\right\} \quad = \quad \tfrac{1}{M}+\tfrac{M-1}{M}\left(\tfrac{2^R-1}{2^R}\right)^{\frac{1}{M-1}}\ . \tag{24}$$

With Lemma 1, a heuristic upper bound of MIMO beamforming mismatch is given by

$$D_\phi \quad \geq \quad \tfrac{M-1}{M}-\tfrac{M-1}{M}\left(\tfrac{2^R-1}{2^R}\right)^{\frac{1}{M-1}}\ . \tag{25}$$

With (17) and (25), the rate-distortion region can heuristically be determined. One example of the rate-distortion region is shown in Fig. 6. In addition, with (17) and Lemma 1, we can have the following result, which will be used for understanding the scalability of MIMO relay network with finite-rate feedback.

Lemma 2. *The precoding loss σ_α due to finite-rate feedback satisfies*

$$1 \quad \geq \quad \sigma^2_\alpha \quad \geq \quad 1-2^{-\frac{R}{M-1}}+\tfrac{\sigma^2_{FL}}{2\sigma_h}\ , \tag{26}$$

where R is the codebook size, M is transmit antenna size and σ^2_{FL} denotes the MSE of the channel estimation and σ_h is the channel variation, respectively.

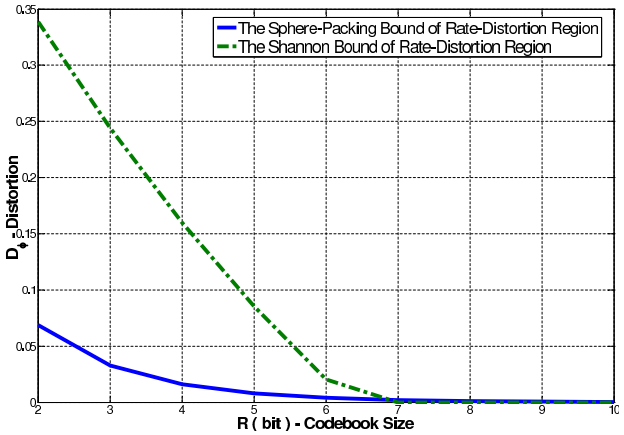


Fig. 6. The rate-distortion region with $M=4$ and $\frac{\sigma^2_h}{\sigma^2_{\mathrm{FL}}}=4.8$dB.

B. The Rate-Reliability Region

In general, the reliability of reverselink is controlled by both channel fading and received SNR. When the erasure rate ε_r is high, it means the amount of fading of reverselink is very high. Higher erasure rate also means it takes the forwardlink transmitter longer time to accurately filter out a proper forwardlink precoding word and it usually yields higher MIMO precoding mismatch given a certain channel coherent time. Since the unreliable symbols are erased based on their received SNR, the left symbols are more reliable and their reliability is mostly decided by γ_{RL}. In this case, the well-known sphere-packing upper bound of Gaussian channel reliability function is

$$\epsilon\left(R\right) \quad \leq \quad \epsilon_{\mathrm{sp}}\left(R\right) \tag{27}$$

with $\epsilon_{\mathrm{sp}}\left(R\right)=\frac{\gamma_{\mathrm{RL}}}{2}-\frac{\sqrt{\gamma_{\mathrm{RL}}}\eta\cos\theta}{2}-\ln\left(\eta\sin\theta\right)$ and the lower bound is

$$\epsilon\left(R\right)\geq\begin{cases}\frac{\gamma_{\mathrm{RL}}}{4}\left(1-\cos\theta\right) & 0\leq R\leq R_1\\ \frac{\gamma_{\mathrm{RL}}}{4}\left(1-\cos\theta_1\right)+R_1-R & R_1\leq R\leq R_2\\ \epsilon_{\mathrm{sp}}\left(R\right) & R_2\leq R\leq C\end{cases} \tag{28}$$

where $\gamma_{\mathrm{RL}}=\frac{P_{\mathrm{RL}}}{\sigma^2_{\mathrm{RL}}}$ denotes the reverselink SNR, $\theta=\arcsin e^{-R}$ denotes the sphere-packing angle,

$$\eta \quad = \quad \tfrac{1}{2}\left(\sqrt{\gamma_{\mathrm{RL}}}\cos\theta+\sqrt{\gamma_{\mathrm{RL}}\cos^2\theta+4}\right)\ , \tag{29}$$

$$R_1 \quad = \quad \tfrac{1}{2}\ln\left(\tfrac{1}{2}+\tfrac{1}{2}\sqrt{1+\tfrac{\gamma_{\mathrm{RL}}}{4}}\right)\ , \tag{30}$$

$$R_2 \quad = \quad \tfrac{1}{2}\ln\left(\tfrac{1}{2}+\tfrac{\gamma_{\mathrm{RL}}}{4}+\tfrac{1}{2}\sqrt{1+\tfrac{\gamma_{\mathrm{RL}}}{4}}\right) \tag{31}$$

and $\theta_1=\arcsin e^{-R_1}$. The rate-reliability region with sphere-packing bounds is shown in Fig. 7.

IV. The Capacity of MIMO Relay Network

A. The Capacity Upper Bound

The capacity of MIMO relay network was recently studied in [19], [20]. With perfect CSI, the capacity of the coherent MIMO relay network satisfies

$$\eta \quad \leq \quad \eta_{\max} \quad = \quad \tfrac{M}{2}\log\left(1+\tfrac{KN}{M}\bar{\rho}\right) \tag{32}$$

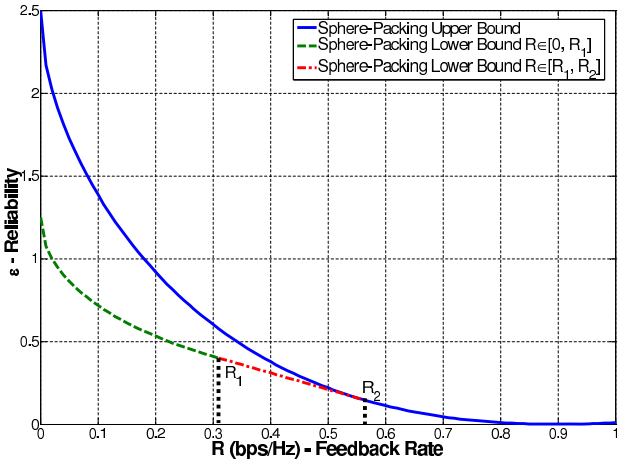


Fig. 7. The rate-reliability region with $\gamma_{\text{RL}} = 7\text{dB}$.

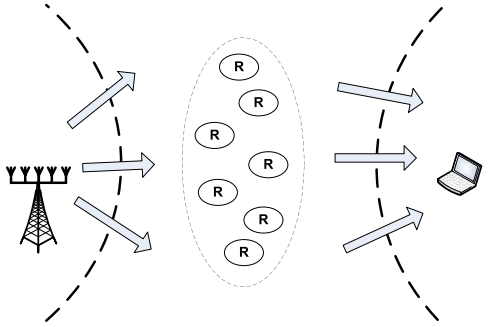


Fig. 8. A two-hop MIMO relay network.

where $\bar{\rho}$ denotes the average received SNR of K relays. The scalability of MIMO relay network throughput $\eta_{\max}$ can be expressed by

$$\eta_{\max}^{\infty} = \frac{M}{2}\log(K) + O(1). \tag{33}$$

With imperfect CSI, the received signals by each $\mathcal{R}$ and $\mathcal{D}$ suffer from the signal power degradation and IBI expressed in (5) and (9). With the consideration of imperfect CSI and the similar reasoning for (32), the upper bound on the capacity of coherent relay network with a codebook size R is the broadcast cut of a corresponding two-hop relay network illustrated in Fig. 8 and the following result can be obtained.

Lemma 3. *The capacity of the coherent MIMO relay network under two-hop relaying satisfies*

$$\eta \leq \eta_L = \frac{M}{2}\log\left[1 + \frac{KN}{M}\bar{\rho}\left(\frac{\sigma_\alpha^2}{1+\frac{(1-\sigma_\alpha)^2}{M-1}\bar{\rho}}\right)\right], \tag{34}$$

if there is no interference cancellation applied; Otherwise, the achievable throughput is

$$\eta \leq \eta_H = \frac{M}{2}\log\left(1 + \frac{KN}{M}\bar{\rho}\sigma_\alpha^2\right). \tag{35}$$

where σ_α denotes the signal degradation due to finite-rate feedback with

$$1 \geq \sigma_\alpha^2 \geq 1 - 2^{-\frac{R}{M}-1} + \frac{\sigma_{EL}^2}{2\sigma_h}, \tag{36}$$

if the channel status is of a uniformly random distribution and a random codebook is applied.

Comparing (34) and (35), the achievable throughput of wireless MIMO relay networks is degraded because the SINR of each MIMO stream decreases by $\alpha_i = \left[1 + \frac{(1-\sigma_\alpha)^2}{M-1}\bar{\rho}\right]$, if no interference cancellation is applied. Because the existing of $\bar{\rho}$ in α_i, the resulted interference becomes correlated with transmit signal power. Furthermore, increasing transmit signal power also leads to stronger interference I. And the throughput of MIMO relay network with finite-rate feedback may become interference-limited.

B. The Scalability in High SNR Region

In order to further understand how imperfect CSI affects the capacity of MIMO relay network, it is interesting to know the asymptotic behavior of the network in high SNR region where IBI becomes one of the dominant factor limiting the network throughput. In this case, the following conclusion can be derived from (17), (25) and Lemma 3.

Lemma 4. *In high SNR region, the capacity of the coherent MIMO relay network under two-hop relaying satisfies*

$$\begin{aligned}\eta_L^{\infty} &= \lim_{\bar{\rho}\to\infty}\eta_L \\ &= \frac{M}{2}\log\left(\frac{KN}{M}\right) + \frac{M}{2}\log\left[\frac{(M-1)\sigma_\alpha^2}{(1-\sigma_\alpha)^2}\right],\end{aligned} \tag{37}$$

if there is no interference cancellation applied; Otherwise, the achievable throughput is

$$\begin{aligned}\eta_H^{\infty} &= \lim_{\bar{\rho}\to\infty}\eta_H \\ &= \frac{M}{2}\log\left(\frac{KN}{M}\right) + \frac{M}{2}\log\left(\sigma_\alpha^2\right) + \frac{M}{2}\log\left(\bar{\rho}\right).\end{aligned} \tag{38}$$

Relaying in high SNR region usually is the result when a transmitter, $\mathcal{S}$ or $\mathcal{R}$, always chooses the closest $\mathcal{R}$ for relaying its packets, especially when their separation is small. If no interference cancellation by relays, the achievable throughput is bounded by η_{L}^{∞} in (37). And it shows that η_{L}^{∞} depends on codebook size R, antenna size M and network size K but the transmit signal power. This means the capacity relay network is interference-limited and the increase of signal power can't help the network for higher throughput. However, if interference cancellation is available, the achievable throughput η_{H}^{∞} can be scaled up by $\frac{M}{2}\log(\bar{\rho})$, which is expressed in (38).

C. The Amplify-and-Forward Capacity

In order to extend our results (34)-(38) to a more general case of MIMO relay network and understand the scalability of MIMO relay network throughput with finite-rate feedback, we consider a simple multi-hop amplify-and-forward MIMO relay network setting, which can be shown in Fig. 9. In this scenario, there are multiple relays grouping together and forming G virtual MIMO transceivers. Signals are relayed through amplify-and-forward by each virtual MIMO transceiver so that every MIMO stream experiences the same number of relays. In each relay group, there are at least N relays, which may internally exchange CSI for collaborated processing received signals. This means, for each group of

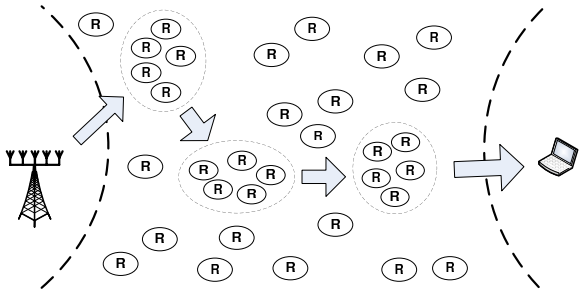


Fig. 9. A possible multi-hop MIMO relay scenario.

relays, they either either do joint interference cancellation on received signals before forwarding, which usually is optimal but more complicate, or just simple forward received signals separately. With straightforwardly extending Lemma 3 and 4, the end-to-end average throughput for a MIMO relay link of the length G is

$$\begin{aligned}\eta &\leq \eta_{\text{L}}(G) \\ &= \frac{M}{2}\log\left[1+\frac{KN}{M}\bar{\rho}\left(\frac{\sigma_\alpha^2}{1+\frac{(1-\sigma_\alpha)^2}{M-1}\bar{\rho}}\right)^G\right], \end{aligned} \quad (39)$$

if there is no interference cancellation applied; Otherwise, the achievable throughput is

$$\eta \leq \eta_{\text{H}}(G) = \frac{M}{2}\log\left(1+\frac{KN}{M}\bar{\rho}\sigma_\alpha^{2G}\right). \quad (40)$$

When the network is operating in high SNR region with $\bar{\rho}\to\infty$, (39) and (40) can, respectively, be simplified by

$$\eta_{\text{L}}^{\infty}(G) = \frac{M}{2}\log\left(\frac{KN}{M}\right)+\frac{MG}{2}\log\left[\frac{(M-1)\sigma_\alpha^2}{(1-\sigma_\alpha)^2}\right] \quad (41)$$

and

$$\eta_{\text{H}}^{\infty}(G) = \frac{M}{2}\log\left(\frac{KN}{M}\right)+\frac{MG}{2}\log\left(\sigma_\alpha^2\right)+\frac{M}{2}\log\left(\bar{\rho}\right). \quad (42)$$

An example of MIMO relay network throughput is shown in Fig. 10. With no interference cancellation, it shows that SIP is outperform TMP and the end-to-end throughput of both cases decreases with more hops. Therefore it may be more efficient to increase SNR instead of use more hops to increase network throughput. However, if the performance of multi-antenna relay network in high SNR region is concerned with IBI being the dominant factor affecting network throughput, the end-to-end throughput of MIMO relay network with finite-rate feedback may actually decrease in high SNR region even though the transmit power of each hop increases. This can be illustrated by Fig. 11, where higher SNR actually brings down the network throughput. Therefore it may be more efficient to use more hops instead of increase SNR in each hop to increase network throughput. This means that the network capacity becomes interference-limited. This is different to the previous low-IBI case.

V. The MIMO Relaying Protocol

From the previous study on the impact of finite-rate feedback on MIMO relay network, many valuable insights in protocol design are obtained. Following previous analysis, some considerations on wireless relaying protocol design are presented here.

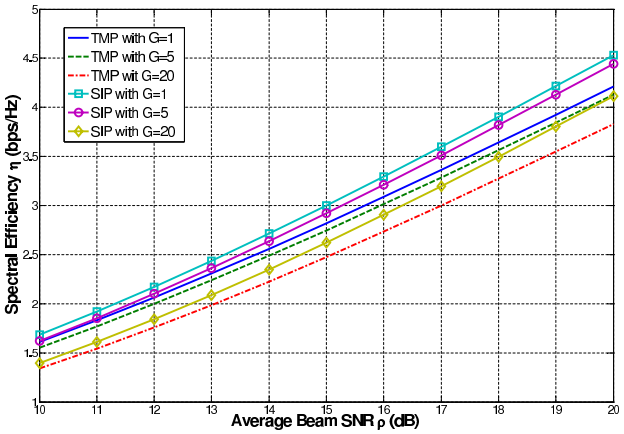


Fig. 10. MIMO relay throughput in low IBI region with $M=N=4$, $B=4$ and $\frac{Q}{L}=\frac{\sigma_p^2}{P}=10\%$.

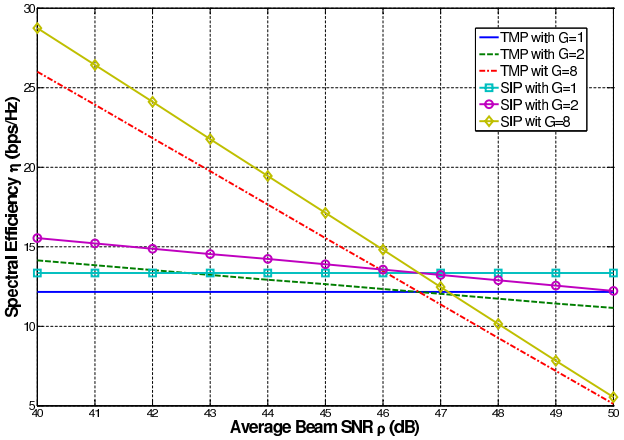


Fig. 11. MIMO relay throughput in high SNR region with $M=N=4$, $B=4$ and $\frac{Q}{L}=\frac{\sigma_p^2}{P}=10\%$.

A. CSI Structure

For facilitating the relay selection and routine for each transmission, it is helpful for relay to broadcast its CSI observed from each received receive pilot. The basic CSI structure is shown in Fig. 12. In addition to the preamble, which implicitly includes information like source identity (S-ID) and relay identity (R-ID), CSI includes the rank and CQI section and the interference-over-thermal (IoT) section. Preamble usually is modulated by spreading-spectrum signals so that it may be easily captured by receivers. And since it is scrambled by the random sequence, e.g., pseudo-noise sequence, generated with

Fig. 12. The basic structure of CSI .

R-ID and S-ID information, receivers can tell which relay it comes from upon processing the preamble. The rank and CQI section is for informing transmitter how many MIMO beams or layers the relay can accommodate, the corresponding precoding vector(s) $\{\mathbf{w}_n : \ n = 1, \ 2, \ \ldots, \ N_k\}$, where N_k denotes the rank, and the expect data rate $\{r_n : \ n = 1, \ 2, \ \ldots, \ N_k\}$ for each of the pecoding vectors. The rank and CQI parameters can be optional if the channel reciprocity can be fully exploited with TDD; Otherwise, it is necessary for relay to calculate and send them back to transmitter. The IoT section, which usually is of the length about one bit or two bits, is designed for informing other sources or relays the interference level it is experiencing. IoT usually comes from two sources. The first source is the concurrent transmission through the same relay from other transmitter(s). It is called multi-access interference (MAI). The other source is the unexpected signals arrived from other transmitters. It is called other-transmitter interference (OTI). High IoT means either MAI or OTI received by this relay is strong, even though the link between the transmitter and relay may be good. In this case, the relay basically suggests other transmitters that it may not be good idea to ask him to relay signals for them.

B. Routine Adaptation

With CSI's received from each relay, transmitter can make decision on which relays it can use for the next transmission and how to allocate power $\{P_k : \ k = 1, \ 2, \ \ldots, \ N\}$ for the N MIMO relay streams. At first, it chooses up to N relay based on its routine selection algorithm, which considers both the link condition $\{r_n : \ n = 1, \ 2, \ \ldots, \ N_k\}$ and IoT of each relay. The general rule here is to avoid the relays experiencing heavy interference in addition to choose the relays with better link condition. On the other hand, when the transmitter decides N the number of MIMO streams for the next transmission, it shall also consider R the MIMO codebook size. When R is large, the transmitter will transmit data streams through more MIMO relay streams than the case when R is small. With doing this, transmitters can help minimize IBI on the signals received by relays and therefore achieve better performance.

C. Link Power Allocation

There are many criteria for allocating power on relay links. When maximizing throughput is a major concern, the so-called water-filling strategy can be an option, in wich each beam's transmit power of each beam satisfies

$$P_k = \left(\mu - \frac{\sigma_n^2}{\lambda_k}\right)^+ = \begin{cases} \mu - \frac{\sigma_n^2}{\lambda_k} & \mu - \frac{\sigma_n^2}{\lambda_k} > 0 \\ 0 & \mu - \frac{\sigma_n^2}{\lambda_k} < 0 \end{cases} \tag{43}$$

with μ chosen so that the total power constraint is satisfied

$$\sum_{k=1}^{K} P_k \quad = \quad P \ . \tag{44}$$

If a delay-limited power allocation strategy is used, then

$$P_k \quad = \quad \frac{\mu}{\lambda_k} \tag{45}$$

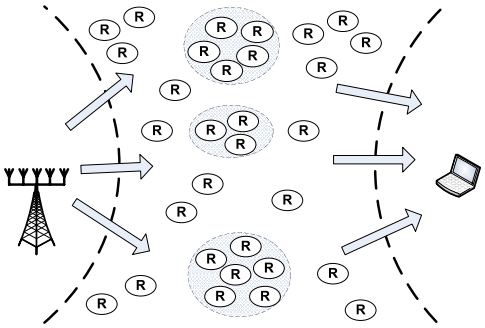


Fig. 13. A virtual MIMO grouping scenario.

with μ chosen to satisfy (44). It is also called channel inversion strategy. Another simple but popular strategy for allocating transmit power is to assign each beam the average power which is

$$P_k \quad = \quad \frac{P}{K} \ . \tag{46}$$

D. Virtual MIMO & Interference Cancel.

With (39) and (40), it shows that interference cancellation can help minimize IBI for higher network throughput at the cost of receiver complexity increas. However, if multiple relays have the capability for fast communicating with each other for their signal demodulation/decoding information, they can form a virtual MIMO group, in which multiple MIMO data streams can be jointly precessed by the relays within one group. This can be illustrated in Fig. 13. The discussion of virtual MIMO can also be found in [19].

VI. Conclusions

In this paper, the impact of finite-rate feedback on MIMO relay work is studied. The achievable throughput and the scalability of this MIMO relay network are formulated and analyzed. The tradeoffs and their impact on network relaying protocol design are discussed and presented.

References

[1] D. Gerlach and A. Paulraj. Spectrum reuse using transmitting antenna arrays with feedback. In *Proc. Int. Conf. Acoust., Speech, Signal Processing*, pages 97–100, Adelaide, Australia, April 1994.

[2] A. Narula, et al. 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] K. K. Mukkavilli, A. Sabharwal, E. Erkip and B. Aazhang. On beamforming with finite rate feedback in multiple-antenna systems. *IEEE Trans Info. Theo.*, 49:2562–2579, October 2003.

[4] P. Xia and G. B. Giannakis. Design and analysis of transmit-beamforming based on limited-rate feedback. In *Proc. IEEE VTC*, September 2004.

[5] J. C. Roh and B. D. Rao. Performance analysis of multiple antenna systems with vq-based feedback. In *Proc. Asilomar Conference 2004*, Pacific Grove, CA, November 2004.

[6] D. J. Love, R. W. Heath and T. Strohmer. Quantized maximal ratio transmission for multiple-input multiple-output wireless systems. In *Proc. Asilomar Conf.*, Pacific Grove, CA, Nov. 2002.

[7] L. Tong, B. M. Sadler and M. Dong. Pilot-assisted wireless transmissions: general model, design criteria, and signal processing. *IEEE Signal Processing Mag.*, 21(56):12–25, November 2004.

[8] M. Coldrey and P. Bohlin. Training-based mimo syetems: Part i/ii. *Technical Report (http://db.s2.chalmers.se/)*, (R032/033), June 2006.

[9] M. Dong and L. Tong. Optimal design and placement of pilot symbols for channel estimation. *IEEE Trans. on Signal Processing*, 50(12):3055–3069, December 2002.

[10] C. E. Shannon. The zero error capacity of a noisy channel. *IRE Trans. Inf. Theory*, 2(3):8–19, September 1956.

[11] Y. H. Kim. Feedback capacity of the first-order moving average gaussian channel. *IEEE Trans. on Inf. Theory*, 52(7):3063–3079, July 2006.

[12] A. J. Kramer. Improving communication reliability by use of an intermittent feedback channel. *IEEE Trans. Inf. Theory*, 15:52–60, January 1969.

[13] A. Sendonaris, E. Erkip and B. Aazhang. User cooperation diversity - part i/ii. *IEEE Trans. Commun.*, 51(11):1927–1948, Nov. 2003.

[14] T. E. Hunter and A. Nosratinia. Cooperative diversity through coding. In *Proc. IEEE int. Symp. Info. Theory*, page 220, 2002.

[15] J. N. Laneman. *Cooperative Diversity in Wireless Networks: Algorithm and Archiectures*. Ph.D. Thesis, MIT, Cambridge, MA, 2002.

[16] N. Ahmed, M. A. Khojastepour and B. Aazhang. Outage minimization and optimal power control for the fading relay channel. In *IEEE Information Theory Workshop 2004*, pages 458–462, Oct. 2004.

[17] C. K. Lo, R. W. Heath and S. Vishwanath. Opportunistic relay selection with limited feedback. In *Vech. Tech. Conf. 2007*, Dublin, Ireland, Apr. 2007.

[18] Special issue on networks and control. In *IEEE Control Systems Magazine*, February 2001.

[19] H. Blcskei, R. U. Nabar, . Oyman and A. J. Paulraj. Capacity scaling laws in mimo relay networks. *IEEE Trans. Wireless Communications*, pages 1433–1444, June 2006.

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