# Limited Feedback for Temporally-Correlated Channels: Feedback Rate and Delay

Kaibin Huang, Bishwarup Mondal, Robert W. Heath, Jr and Jeffrey G. Andrews

Wireless Networking and Communication Group

Dept. of Electrical & Computer Engineering

University of Texas at Austin

1 University Station C0803

Austin, TX 78712-0240

Phone: +1-512-921-8582

{khuang, mondal, rheath, jandrews}@ece.utexas.edu

March 28, 2006

**Index Terms**

Limited Feedback, Temporally-Correlated Channels, Markov Chain, Feedback Delay, Beamforming Systems

**Abstract**

Limited feedback systems send quantized channel state information (CSI) from the receiver to the transmitter for improving either the throughput or the link reliability. This paper proposes a framework for analyzing multi-antenna limited feedback systems with temporally-correlated flat-fading channels. Developed on a Markov chain modeling quantized CSI, the proposed framework consists of analytical results concerning the feedback rate, the feedback compression, the capacity with feedback delay, and the decay rate of the capacity gain due to feedback delay. These results constitute a set of useful tools for designing a wide-range of limited feedback system. The application of the proposed framework is illustrated for a multiple-input-single-output (MISO) beamforming system, where the feedback consists of the quantized channel gain and direction. Feedback rates are shown to increase linearly with the Doppler frequency in a Rayleigh fading channel with Clarke's spectrum. Furthermore, capacity gains are found to decrease exponentially with feedback delay. From theses results, useful design guidelines can be derived for choosing system parameters ranging from the reception vehicular speed to the feedback channel rate.

## I. Introduction

For a multi-input-multiple-output (MIMO) wireless downlink, the benefits of transmit (TX) channel state information (CSI), such as improvements on the data throughput and the link reliability, can be

realized through intelligent CSI quantization techniques, called *limited feedback*. In most prior work on limited feedback, *block fading* is assumed for the downlink channel. Thereby, different channel realizations are assumed independent and potential temporal correlation is ignored [1], [2]. Under this assumption, the design of a limited feedback system reduces to a vector quantization problem [3]. Different methods for quantizing CSI have been developed based on line packing [4]–[6], combined parametrization and scalar quantization [7], interpolation of quantized vectors [8], [9], frame theory [10], random vector quantization [11], [12] and the Lloyd algorithm [13], [14]. Furthermore, different types of limited feedback systems have been investigated for beamforming [4], [5], [15], precoded orthogonal space-time block codes [16], [17], precoded spatial multiplexing [7], [18]–[21], and TX covariance optimization [22], [23].

In practical systems, the block fading assumption for the downlink channel is pessimistic since channel temporal correlation often exists and can be used for reducing feedback rates. The topic of feedback compression by exploiting channel temporal correlation is addressed to some extent in [19], [24] where 1-bit feedback algorithms are proposed for CSI feedback and channel tracking, respectively. In [19], a MIMO channel is parameterized and the feedback of each parameter is compressed to be one bit. Nevertheless, the multiplicity of the channel parameters compromises the feedback efficiency. In [25], the CSI feedback is compressed to be one bit but requires the transmitter to broadcast channel subspace matrices. Other than these algorithms, few results exist on how the feedback rate scales with the channel coherence time despite the importance of such results for determining the feedback channel rate. On the other hand, the block fading assumption hinders the analysis of feedback delay, which exists in a practical system due to sources such as signal processing, propagation and protocol delays. The inevitable existence of feedback delay calls for evaluation of its impacts on the system performance. Albeit not in the context of limited feedback, two related studies are conducted in [26], [27], which address respectively the effect of channel predication error on adaptive modulation and the capacity with delayed CSI feedback for a finite-state single-input-single-output (SISO) channel.

This paper proposes a framework for addressing the feedback rate and delay in a limited feedback

system. The novelty of the framework lies in the modeling of quantized CSI using a Markov chain and its use in deriving a set of analytical results. Specifically, the feedback rate, the feedback compression method, the downlink capacity with feedback delay, and the exponential decay rate of the capacity gain caused by feedback delay are analyzed. The proposed framework is useful for computing the required feedback rate and evaluating the impact of feedback delay. This framework is applicable for any typical limited feedback system having a temporally-correlated channel and employing adaptive transmission such as beamforming, precoding and adaptive modulation [4], [18], [28]. The application of the framework follows a two-step procedure: (1) construction of a Markov chain for a given temporally-correlated channel and a chosen CSI quantization method and (2) quantification of the analytical results using the Markov chain.

The proposed framework is applied to analyze two types of limited feedback: the *gain* (Frobenius norm) and the *shape* (unitary directional vector) of an i.i.d. downlink channel in a multiple-input-single-output (MISO) beamforming system[1]. The key of this application is the construction of two Markov chains for modeling the gain and the shape, respectively. Using the Markov chains, the feedback rates for both the gain and the shape are shown (partially using numerical results) to increase *linearly* with the Doppler frequency in the uniform scattering environment [29], [30]. For both cases, the downlink capacity gains are found to decrease exponentially with the feedback delay, which indicates the impact of feedback delay on the downlink capacity. Furthermore, we show that the feedback compression by truncating the Markov-chain transition probabilities results in one-bit gain feedback and reduces the rate of the shape feedback effectively as demonstrated by numerical results. These results are useful for designing a limited feedback system with either the gain or the shape feedback, specifically, for determining the tolerable vehicular speed and feedback delay, and the required feedback bandwidth and CSI quantization resolution.

The remainder of this paper is organized as follows. In Section II, a limited feedback system for a

---

[1]This application does not limit the applicability of the framework for a spatially correlated downlink channel as well as other types of limited feedback systems.

temporally-correlated channel is described. Based on this system, problem formulations are presented in Section III. An analytical framework is developed in Section IV for solving the formulated problems. This analytical framework is applied in Section V-VII to analyze a MISO beamforming system. The MISO system model is described in Section V, in which two types of feedback (*channel gain* and *channel shape*) are identified and two corresponding limited feedback systems are defined. The gain feedback system is analyzed in Section VI and the shape feedback system in Section VII. A design example and other numerical results are presented in Section VIII followed by concluding remarks in Section IX.

**Notation:** The matrix power in $\mathbf{P}^A$ represents $(A-1)$-time multiplications of the matrix $\mathbf{P}$ with itself. The $(i,j)$th element of the matrix $\mathbf{P}$ is denoted as $[\mathbf{P}]_{i,j}$. Similarly, the $i$th element of a vector $\mathbf{p}$ is written as $[\mathbf{p}]_i$. We denote the conjugate transpose of a matrix $\mathbf{P}$ as $\mathbf{P}^H$. The cardinality of a set $\mathbb{X}$ is represented using $|\mathbb{X}|$. The *variance distance* between two matrixes $\mathbf{P}$ and $\mathbf{Q}$ is defined as [31]

$$\text{(variance distance)} \quad \|\mathbf{P} - \mathbf{Q}\|_v = \sum_i \sum_j |[\mathbf{P}]_{i,j} - [\mathbf{Q}]_{i,j}| . \tag{1}$$

## II. SYSTEM DESCRIPTION

The objective of this section is to introduce the generalized limited feedback system for a temporally-correlated channel as illustrated in Fig. 1. We consider a discrete-time system model where continuous-time signals are sampled at the symbol rate $1/T_s$ with $T_s$ being the symbol duration. Consequently, each signal is represented by a sequence of samples with the subscript $n$ denoting the sample index. Furthermore, the signals and functions in this system are defined abstractly and denoted as follows:

$\mathbf{h}_n$      receiver (RX) CSI, $\mathbf{h}_n \in \mathbb{h}$ where $\mathbb{h}$ is the channel space;

$J_n$      channel state, $J_n \in \mathbb{J} = \{1, \cdots, N\}$;

$\mathcal{H}$      codebook containing $N$ transmission profiles[2];

$\mathbf{f}_i$      $i$th member of the codebook $\mathcal{H}$ where $1 \leq i \leq N$;

---

[2]A transmission profile is defined as an adaptable format of the transmitted signal. Examples of transmission profiles include beamforming vectors, linear precoders, and modulation orders.

$K_n$      transmitter (TX) channel state, $K_n \in \mathbb{J}$; and

$\mathbf{w}_n$      transmission profile, $\mathbf{w}_n \in \mathcal{H}$.

Note that for the case of a matrix channel, the RX CSI vector $\mathbf{h}_n$ is formed by stacking the matrix columns.

The RX CSI represents the downlink channel as the result of the following assumption (AS):

**AS 1:** *The mobile has perfect knowledge of the downlink channel.*

By making this assumption, we ignore channel estimation error at the mobile. The RX CSI $\{\mathbf{h}_n\}$ contains potentially correlated vectors, whose coefficients result from the use of multiple TX/RX antennas for the downlink.

The process of CSI feedback is illustrated in Fig. 1 and is described as follows. First, the *CSI quantizer* maps the RX CSI $\mathbf{h}_n$ onto a member of the codebook $\mathcal{H}$ [2] and outputs its index $J_n$, called the *channel state*. The codebook $\mathcal{H}$ is designed by partitioning the channel space into $N$ regions, called *Voronoi cells* and denoted as $\{\mathbb{V}_i\}$ [3]. The codebook members and the Voronoi cells are one-to-one mapped. Therefore, the operation of the CSI quantizer implies that

$$J_n = i \quad \text{if} \quad \mathbf{h}_n \in \mathbb{V}_i, \tag{2}$$

where $1 \leq i \leq N$. Second, the time instants for sending back $J_n$ are determined by the *aperiodic feedback scheme*, which initiates a feedback whenever the channel state changes ($J_n \neq J_{n-1}$). Third, the feedback channel conveys the channel state to the base-station and is specified by the following assumption:

**AS 2:** *The feedback channel is free of error but has a delay of $D$ samples.*

The error-free assumption in AS 2 is justifiable as the feedback channel is usually protected using error-correction coding and hence has a very low error probability. Given an error-free feedback channel with delay, the TX channel state lags behind the channel state by $D$ samples: $K_n = J_{n-D}$. Last, the *TX profile generator* simply performs the codebook lookup function using the TX channel state $K_n$, where the codebook is also $\mathcal{H}$. Mathematically, $\mathbf{w}_n = \mathbf{f}_{K_n}$ where $\mathbf{f}_i$ is the $i$th member of the codebook $\mathcal{H}$.

## III. PROBLEM FORMULATION

In this section, the problems to be solved in this paper are formulated mathematically. Consider the limited feedback system in Fig. 1. The average feedback rate is defined as the average bit rate for transmission over the feedback channel. Given a $N$-region partition of the channel space, each feedback requires

$$B = \log_2 N \quad \text{bits} \tag{3}$$

for identifying the current channel state. Note that the temporal correlation of the channel can be exploited to compress the feedback bits to be smaller than $B$ in (3). A detailed treatment of this topic is outside the scope of this paper but a simple method of feedback compression is discussed in Section IV-B.3. With the number of bits per feedback command fixed in (3) and the feedback frequency determined by the aperiodic feedback scheme, the time-average feedback rate for a duration of $MT_s$ can be expressed as

$$\bar{R}(M) = \frac{B}{MT_s} \sum_{n=1}^{M} 1\{J_n \neq J_{n+1}\} \quad \text{bits/s}, \tag{4}$$

where $B$ comes from (3), and $1\{\cdot\}$ is the indicator function being 1 if the channel state changes and 0 otherwise.

Consider the time-average downlink capacity in the interval $[0, MT_s]$. The instantaneous downlink capacity depends on not only the current channel sample $\mathbf{h}_n$ but also the delayed one $\mathbf{h}_{n-D}$, which is the source for the TX CSI given a feedback delay of $D$ samples. Therefore, the instantaneous downlink capacity can be represented as $C(\mathbf{h}_n, \mathbf{h}_{n-D})^3$ and hence the time-average downlink capacity can be written as

$$\bar{C}(M, D) = \frac{1}{M} \sum_{n=1}^{M} C(\mathbf{h}_n, \mathbf{h}_{n-D}). \tag{5}$$

The downlink temporally-correlated channel considered in this paper is a stochastic process specified by a PDF $f_{\mathbf{h}}$ and a joint PDF $f_{\mathbf{hh}}$ of two different channel samples. A set of natural questions to ask are:

---

[3] See Section VI-C and Section VII-C for specific examples.

1) *How can the time-average quantities, $\bar{R}$ in (4) and $\bar{C}$ in (5), be expressed as functions of the channel PDF's?*

2) *How can the feedback rate be reduced by exploiting the channel temporal correlation, which is characterized by the channel joint PDF?*

3) *How can the average downlink capacity $\bar{C}$ be expressed as a function of the feedback delay $D$?*

4) *How fast will the gain in capacity due to CSI feedback diminish with the feedback delay?*

To answer these questions, we propose an analytical approach based on ergodic theory and theory of Markov chains. This approach is captured in the analytical framework developed in the next section.

## IV. ANALYTICAL FRAMEWORK

In this section, based on ergodic theory and the theory of Markov chains, an analytical framework is developed, in which questions raised in the last section are answered. The downlink temporally-correlated channel is modeled using a Markov chain as discussed in Section IV-A. In Section IV-B and IV-C.1, the average feedback rate and the average downlink capacity are shown using ergodic theory to be functions of the Markov chain parameters. For reducing the feedback rate, a simple method of feedback compression is discussed in Section IV-B.3, which is based on truncation of the Markov chain transition probabilities. In Section IV-C.2, the downlink capacity gain due to CSI feedback are shown to decrease at least exponentially with the feedback delay.

### A. Channel State Markov Chain

We model the channel state $J_n$ (cf. Fig. 1) using a finite-state and discrete-time Markov chain by making the following assumption:

**AS 3:** *The channel state $J_n$ is a homogeneous (stationary) Markov chain [31].*

This assumption is reasonable for a time-varying downlink channel that varies slowly with respect to (w.r.t.) the symbol rate. The stationary and transition probabilities of the Markov chain are defined as

$$\text{(stationary probability)} \quad P_i \;=\; \Pr\{J_n = i\}, \tag{6}$$

$$\text{(transition probability)} \quad P_{i,j} \;=\; \Pr\{J_n = i \mid J_{n-1} = j\}, \tag{7}$$

where $1 \leq i, j \leq N$. From (2), they can be written in terms of the channel PDF's as

$$P_i \;=\; \Pr\{\mathbf{h}_n \in \mathbb{V}_i\} = \int_{\mathbb{V}_i} f_{\mathbf{h}}(\mathbf{h})d\mathbf{h}, \tag{8}$$

$$P_{i,j} \;=\; \Pr\{\mathbf{h}_n \in \mathbb{V}_j \mid \mathbf{h}_{n-1} \in \mathbb{V}_i\} = \frac{1}{P_i} \int_{\mathbb{V}_i} \int_{\mathbb{V}_j} f_{\mathbf{hh}}(\mathbf{h}_{n-1}, \mathbf{h}_n) d\mathbf{h}_{n-1} d\mathbf{h}_n, \tag{9}$$

where $\mathbb{V}_i$ is the $i$th Voronoi cell. Last, let $\mathbf{P}$ denote the *transition probability matrix* with $[\mathbf{P}]_{i,j} = P_{i,j} \; \forall \, 1 \leq i, j \leq N$ and $\boldsymbol{\pi}$ the *stationary probability vector* with $[\boldsymbol{\pi}]_i = P_i \; \forall \, 1 \leq i \leq N$.

To analyze the effect of feedback delay, we are interested in knowing transition probabilities between two channel state samples spaced $D$ samples apart: $J_n$ and $J_{n-D}$. To this end, the channel state Markov chain provides the following useful property [31],

$$\Pr(J_n = i \mid J_{n-D} = j) = \left[\mathbf{P}^D\right]_{i,j}, \quad \forall \, 1 \leq i, j \leq N. \tag{10}$$

This property is instrumental in deriving the downlink ergodic capacity with delayed CSI feedback in Section IV-C.1 and the exponential rate at which the feedback delay reduces the capacity gain with respect to no CSI feedback in Section IV-C.2.

### B. Average Feedback Rate

In Section IV-B.2, the average feedback rate in (4) is related to the stationary and transition probabilities of the channel state Markov chain in a theorem proved using ergodic theory. Auxiliary results needed in the proof of the theorem are presented in Section IV-B.1. In addition, a simple method for feedback compression is discussed in Section IV-B.3.

*1) Auxiliary Results:* The assumption and lemmas needed for proving the theorem in the next section are presented as follows. First, a Markov chain is *irreducible* if a transition can occur between any two states [31]. We make the following assumption:

**AS 4:** *The channel state Markov chain is irreducible.*

Define the area of a Voronoi cell $\mathbb{V}_i$ as

$$\text{(Voronoi cell area)} \quad \mathcal{A}_i = \int_{\mathbb{V}_i} d\mathbf{h}. \tag{11}$$

The assumption in AS 4 is justified for the i.i.d. complex Gaussian channel in the following lemma:

**Lemma 1:** *For the i.i.d. complex Gaussian downlink channel $\mathbf{h}_n$ whose space is partitioned into $N$ Voronoi cells $\{\mathbb{V}_i\}$ with nonzero volume: $\mathcal{A}_i > 0 \ \forall \ 1 \leq i \leq N$, the channel state Markov chain is irreducible.*

**Proof:** See Appendix A. $\qquad\square$

Second, a Markov chain is *positive recurrent* if it leaves and returns to each state within a finite interval [31]. The lemma given below follows from [32, Theorem 3.3].

**Lemma 2:** *An irreducible channel state Markov chain with a finite state space ($N < \infty$) is positive recurrent.*

Third, the specific ergodic theorem used for proving the theorem in the next section is rewritten from [32, Corolary 4.1] as the following lemma.

**Lemma 3 (Ergodic Theorem):** *Given that the channel state Markov chain is irreducible and positive recurrent and a function $z : \mathbb{J} \times \mathbb{J} \to \mathbb{R}$ that satisfies*

$$\sum_{i=1}^{N} \sum_{j=1}^{N} |z(i,j)| P_i P_{i,j} < \infty, \tag{12}$$

*the following a.e. (almost everywhere) convergence holds*

$$\lim_{M \to \infty} \frac{1}{M} \sum_{n=1}^{M} z(J_n, J_{n-1}) = \sum_{i=1}^{N} \sum_{j=1}^{N} z(i,j) P_i P_{i,j}. \tag{13}$$

*2) Main Result:* The first main result in the framework is stated in the following theorem. Its proof relies on the auxiliary results in the last section.

**Theorem 1:** *The average feedback rate in (4) converges (a.e.) with time as follows*

$$\lim_{M \to \infty} R(M) = \frac{B}{T_s} \sum_{i=1}^{N} P_i (1 - P_{i,i}), \tag{14}$$

*where $\{P_{i,i}\}$ and $\{P_i\}$ are the transition and stationary probabilities of the channel state Markov chain.*

**Proof:** See Appendix B. □

A few remarks are in order:

- Roughly speaking, (14) equates the time space (left side) and the probability space (right side).

- From (14), the average feedback rate increases linearly with $B$, which is related to the quantization resolution of RX CSI.

- The term $(1 - P_{i,i})$ in (14) is the total transition probability from state $i$ to other states. The values of $\{1 - P_{i,i}\}$ tend to be larger for smaller coherence time and vise versa.

- Since $P_i$ and $P_{i,i}$ are functions of the downlink channel PDF's (cf. (8) and (9)), Theorem 1 provides an answer to the first question asked in Section III from the aspect of average feedback rate.

*3) A Simple Method for Feedback Compression:* Modeled by a Markov chain, the downlink temporally-correlated channel can be treated as a Markov source and hence compressed using data compression techniques [33]–[36]. As the result, fewer bits than $B$ in (3) are required for each feedback. Rather than pursuing optimal compression [33]–[36], we introduce a sub-optimal method for feedback compression based on truncating the Markov chain transition probabilities.

We define the $\epsilon$-neighborhood of the Markov state $i$ as

$$\mathcal{N}_i(\epsilon) = \{1 \leq j \leq N \mid P_{i,j} \geq \epsilon\}, \quad i = 1, \cdots, N, \tag{15}$$

where $P_{i,j}$ is the transition probability and $\epsilon$ a small positive number. Therefore, the required number of feedback bits for identifying the current channel state, say $J_n = i$, is defined as

$$B_i = \lceil \log_2 |\mathcal{N}_i(\epsilon)| \rceil \text{ bits}, \quad i = 1, \cdots, N, \tag{16}$$

where $\lceil a \rceil$ is the smallest integer larger than $a$. It follows from (15) that $|\mathcal{N}_i(\epsilon)| < N$ and hence $B_i < B$ where $B$ in (3) corresponds to the case of no compression. Therefore, with feedback compression, we can rewrite the average feedback rate in (14) as

$$\text{(feedback compression)} \quad \mathcal{R} = \frac{1}{T_s} \sum_{i=1}^{N} B_i P_i (1 - P_{i,i}), \tag{17}$$

where $B_i$ is in (16). Comparison of this average feedback rate with that for optimal compression is a subject under ongoing investigation. Note that the above feedback compression method incurs an additional cost – storage of tables defining $\epsilon$-neighborhoods at both the mobile and the base-station.

### C. Effect of Feedback Delay

Define the *downlink ergodic capacity* as

$$\text{(downlink ergodic capacity)} \quad \mathcal{C}(D) = \lim_{M \to \infty} \bar{C}(D, M), \tag{18}$$

where $\bar{C}$ is the average downlink capacity in (5). In Section IV-C.1, the downlink ergodic capacity will be derived as a function of the feedback delay as well as parameters of the channel state Markov chain. CSI feedback provides a gain in the downlink ergodic capacity. How fast this gain diminishes with feedback delay is analyzed in Section IV-C.2.

*1) Downlink Ergodic Capacity:* The second main result in the framework is given in the following theorem.

**Theorem 2:** *The downlink ergodic capacity is given as,*

$$\mathcal{C}(D) = \sum_{i=1}^{N} \sum_{j=1}^{N} \mathcal{C}_{i,j} \left[ \mathbf{P}^D \right]_{i,j} P_i, \tag{19}$$

*where $\mathbf{P}$ and $P_i$ are the transition probability matrix and the stationary probability of the channel state Markov chain. Moreover,*

$$\mathcal{C}_{i,j} = \frac{1}{P_{i,j} P_i} \int_{\mathbb{V}_i} \int_{\mathbb{V}_j} C(\mathbf{h}_n, \mathbf{h}_{n-D}) f_{\mathbf{h}\mathbf{h}}(\mathbf{h}_n, \mathbf{h}_{n-D}) d\mathbf{h}_n d\mathbf{h}_{n-D}, \tag{20}$$

where $C(\mathbf{h}_n, \mathbf{h}_{n-D})$ is the instantaneous capacity (cf. Section III).

**Proof:** See Appendix C. □

A few remarks are in order:

- Similar to (14), (19) equates the time space (left side) and the probability space (right side).

- Theorem 2 answers partially the first question raised in Section III, namely how to derive the time-average feedback rate $\bar{R}$ in (4) as a function of the channel PDF's[4].

- The term $\mathcal{C}_{i,j}$ in (20) is the expected downlink capacity conditional on the event $(J_n = i, J_{n-D} = j)$. The other term $\left( \left[ \mathbf{P}^D \right]_{i,j} P_i \right)$ in (20) is the probability of this event.

Using the following asymptotic result related to high-resolution quantization [37]:

$$N \to \infty \text{ and } \mathbf{h} \in \mathbb{V}_i \;\Rightarrow\; \mathbf{h} \to \mathbf{f_i}, \quad 1 \le i \le N, \tag{21}$$

where $\mathbf{f_i}$ is the $i$th member of the codebook $\mathcal{H}$, we have the following corollary of Theorem 2.

**Corollary 1:** *The downlink ergodic capacity converges with the number of Voronoi cells in the partition of the channel space as follows*

$$N \to \infty \quad \Rightarrow \quad \mathcal{C}(D) \to \sum_{i=1}^{N} \sum_{j=1}^{N} C(\mathbf{f}_i, \mathbf{f}_j) \left[ \mathbf{P}^D \right]_{i,j} P_j, \tag{22}$$

*where $\mathbf{f}_i$ is the $i$th member of the codebook $\mathcal{H}$.*

The above corollary is useful as it avoids the evaluation of the integration in (20), which is usually complicated.

It is also interesting to consider the extreme cases of zero delay ($D = 0$) and infinite delay ($D \to \infty$). Related results are presented in the following lemmas.

**Lemma 4:** *For zero feedback delay $D = 0$, the downlink ergodic capacity is given as*

$$\mathcal{C}(0) = \sum_{i=1}^{N} \mathcal{C}_i P_i, \tag{23}$$

*where*

$$\mathcal{C}_i = \frac{1}{P_i} \int_{\mathbb{V}_i} C(\mathbf{h}, \mathbf{h}) f_{\mathbf{h}}(\mathbf{h}) d\mathbf{h}. \tag{24}$$

---

[4]In (19) the channel PDF's are absorbed into the stationary and transition probabilities of the channel-state Markov chain.

**Proof:** See Appendix D. □

**Lemma 5:** *For infinite feedback delay $D \to \infty$, the downlink ergodic capacity is given as*

$$\mathcal{C}(\infty) = \sum_{i=1}^{N} \sum_{j=1}^{N} \mathcal{C}_{i,j} P_i P_j, \tag{25}$$

*where $\mathcal{C}_{i,j}$ is defined in (20).*

**Proof:** For $D \to \infty$, the positive recurrent (cf. Lemma 2) channel state Markov chain has the following property [31]:

$$D \to \infty \Rightarrow \mathbf{P}^D \to [\boldsymbol{\pi}, \cdots, \boldsymbol{\pi}]. \tag{26}$$

Substitution of (26) into (19) leads to (25). □

Two observations can be made:

- Contrary to the case of feedback delay in (19), the downlink ergodic capacity without feedback delay no longer depends on the transition probabilities of the channel state Markov chain, as observed from (23). Nevertheless, the transition probabilities still influence the feedback rate.

- The downlink ergodic capacity for infinite feedback delay does not depend on the transition probabilities as shown by (25). The reason is that the TX channel state is independent of the channel state as feedback delay goes to infinity. In another word, the CSI feedback is as good as no feedback.

*2) Reduction Rate of Feedback Capacity Gain:* We define the *feedback capacity gain* as the difference in downlink ergodic capacity between the cases of finite and infinite feedback delay

$$\text{(feedback capacity gain)} \quad \Delta\mathcal{C}(D) = \mathcal{C}(D) - \mathcal{C}(\infty), \tag{27}$$

where $\mathcal{C}(D)$ and $\mathcal{C}(\infty)$ are given in (19) and (25), respectively. It is by system design that

$$\mathcal{C}(\infty) < \mathcal{C}(D) < \mathcal{C}(0). \tag{28}$$

It follows from (27) and (28) that the feedback capacity gain is positive: $\Delta\mathcal{C}(D) > 0$.

We will now analyze how fast the feedback capacity gain $\Delta\mathcal{C}(D)$ converges to 0 as the feedback delay $D$ goes to infinity. From (26) and (27),

$$\mathbf{P}^D \to [\boldsymbol{\pi}, \cdots, \boldsymbol{\pi}] \Rightarrow \Delta\mathcal{C}(D) \to 0. \tag{29}$$

As shown in Lemma 6 rewritten from [38, Proposition 1], the columns of $\mathbf{P}^D$ converge to $\boldsymbol{\pi}$ at least exponentially with the delay $D$. Given (29), the feedback capacity gain is also expected to decreases at least exponentially with $D$. This result is formally stated as Theorem 3.

**Lemma 6:** *Consider the channel state Markov chain. The variance distance in (1) between* $\mathbf{p}_i(D)$, *the $i$th column of $\mathbf{P}^D$, and the stationary probability vector $\boldsymbol{\pi}$ is bounded as*

$$\|\mathbf{p}_i(D) - \boldsymbol{\pi}\|_v \leq \mu_i \zeta^D, \tag{30}$$

*where $\zeta$ is the second largest absolute eigenvalue of $\mathbf{P}$ and*

$$\mu_i = \tfrac{1}{2}\sqrt{(1 - [\boldsymbol{\pi}]_i)^2/[\boldsymbol{\pi}]_i + N - 1}. \tag{31}$$

**Theorem 3:** *The feedback capacity gain $\Delta\mathcal{C}(D)$ defined in (27) decreases with the feedback delay $D$ in samples at least exponentially. Specifically, $\Delta\mathcal{C}(D)$ is bounded as:*

$$0 \leq \Delta\mathcal{C}(D) \leq \alpha\beta^D, \tag{32}$$

*where*

$$\alpha = \sum_{j=1}^{N}\sqrt{\sum_{i=1}^{N}\mathcal{C}_{i,j}}\sqrt{\sum_{n=1}^{N}\mu_n}, \quad \beta = \sqrt{\zeta} \tag{33}$$

*with $\mathcal{C}_{i,j}$ given in (20) and $\mu_n$ and $\zeta$ in Lemma 6.*

**Proof:** See Appendix E. □

Numerical results are provided in Section VIII-C show that the decreasing rate of the feedback capacity gain with the feedback delay is indeed exponential.

## V. APPLICATION: MISO BEAMFORMING SYSTEM

In previous sections, a general analytical framework is developed based on an abstract channel state Markov chain. To quantify its analytical results, the framework is applied in Section V-VII to analyze a multiple-input-single-output (MISO) system with beamforming and power control. In this section, the MISO system is described and two types of limited feedback are defined.

Consider the discrete-time MISO system in Fig. 2 where the *shape* and the *gain* of the downlink channel are defined shortly. The received signal is given as

$$y_n = \sqrt{\eta_n} \mathbf{w}_n^H \mathbf{h}_n x_n + \nu_n, \tag{34}$$

where

$L$      number of TX antennas;

$\mathbf{h}_n$      ($L \times 1$ vector) downlink channel;

$\eta_n$      TX power;

$\mathbf{w}_n$      ($L \times 1$ vector) beamforming vector with $\|\mathbf{w}_n\|^2 = 1$;

$x_n$      transmitted symbol with $\|x_n\| = 1$;

$y_n$      received symbol; and

$\nu_n$      AWGN sample with $\nu_n \in \mathcal{CN}(0, 1)$.

We make the following assumption about the downlink channel.[5]

**AS 5:** *The downlink channel* $\mathbf{h}_n$ *is an i.i.d. vector whose coefficients are* $\mathcal{CN}(0, 1)$.

This channel model is reasonable for the scenario where the TX antennas are sufficiently separated such that different channel coefficients are independent [39].

To define two types of CSI feedback, we decompose $\mathbf{h}_n$ into two random processes $g_n$ and $\mathbf{u}_n$

$$\mathbf{h}_n = g_n \mathbf{u}_n, \tag{35}$$

where $g_n = \|\mathbf{h}_n\|$ and $\mathbf{u}_n = \mathbf{h}_n / \|\mathbf{h}_n\|$. Adopting the terminologies in [40], $g_n$ and $\mathbf{u}_n$ are named the *gain* and the *shape* of the channel vector $\mathbf{h}_n$, respectively. Under AS 5, $g_n$ and $\mathbf{u}_n$ are independently distributed: the former follows the chi distribution with the PDF given as [41]

$$f_g(g) = 2g^{2L-1} \exp(-g^2)/(L-1)!, \quad g \geq 0, \tag{36}$$

and the latter is uniformly distributed on the surface of a unit hyper-sphere [5].

---

[5]Note that our analysis can be extended to the case of spatially correlated channels, which mainly changes the properties of the Markov chains modeling the RX CSI.

The channel gain and the channel shape are used at the transmitter for different purposes: controlling the TX power and choosing the beamforming vector, respectively. Separating the analysis of the gain and the shape feedback not only simplifies our analysis but also yields more insights into the required feedback rate and the effect of feedback delay for each feedback. To this end, we define two types of MISO limited feedback systems similar to that in Fig. 1:

1) (*Gain Feedback System*) The gain feedback system is the MISO system with ideal beamforming ($\mathbf{w}_n = \mathbf{u}_n$) [42] and limited feedback of the channel gain.

2) (*Shape Feedback System*) The shape feedback system is the MISO system with constant TX power ($\eta_n = \rho$) and limited feedback of the channel shape.

Despite above simplifying assumptions, our analysis can be extended to analyze the joint feedback of the gain and the shape of the downlink channel.

## VI. APPLICATION: FEEDBACK OF CHANNEL GAIN

In this section, the analytical framework developed in Section IV is used to analyze the gain feedback system defined in Section V. First, a Markov chain is constructed in Section VI-A for modeling the quantized channel gain. Second, the average feedback rate and the downlink ergodic capacity for the gain feedback system are derived in Section VI-B and Section VI-C, respectively. Note that superscripts $^{(g)}$ are added to variables in the gain feedback system to avoid confusion.

### A. Gain Markov Chain

The state of the channel gain generated by the CSI quantizer (cf. Fig. 1) is modeled by a Markov chain, called the *gain Markov chain*. In Section VI-A.1, the state space of the gain Markov chain is constructed. The transition and stationary probabilities of the gain Markov chain are derived in Section VI-A.3 and VI-A.2, respectively.

*1) State Space:* The states of the gain Markov chain are one-to-one mapped to the Voronoi cells in the channel gain space ($\mathbb{R}^+ \cup 0$). The Voronoi cells are obtained using the *scalar quantization* method

with the square-error distortion function [3]. In brief, the application of scalar quantization generates a codebook containing possible quantized gain values: $\mathcal{H}^{(g)} = \{f_1^{(g)}, \cdots, f_N^{(g)}\}$. In terms of the codebook members, the Voronoi cells are written as

$$
\mathbb{V}_i^{(g)} = \begin{cases} \left\{ g \in \mathbb{R}^+ \cup 0 \mid \sqrt{\theta_{i-1}} \leq g \leq \sqrt{\theta_i} \right\}, & 1 < i < N, \\ \left\{ g \in \mathbb{R}^+ \cup 0 \mid 0 \leq g \leq \sqrt{\theta_1} \right\}, & i = 1, \\ \left\{ g \in \mathbb{R}^+ \cup 0 \mid \sqrt{\theta_{i-1}} \leq g \leq \infty \right\}, & i = N, \end{cases} \tag{37}
$$

where $\theta_i$ is defined as

$$
\theta_i = \frac{1}{4} \left( f_i^{(g)} + f_{i+1}^{(g)} \right)^2, \quad \forall \, 1 \leq i \leq N - 1. \tag{38}
$$

*2) Stationary Probabilities:* The stationary probabilities of the gain Markov chain are defined as

$$
P_i^{(g)} = \Pr \left\{ J_n^{(g)} = i \right\}, \quad i = 1, \cdots, N. \tag{39}
$$

Equivalently, from (2),

$$
P_i^{(g)} = \Pr \left\{ g_n \in \mathbb{V}_i^{(g)} \right\}, \quad i = 1, \cdots, N, \tag{40}
$$

where $\mathbb{V}_i^{(g)}$ is the $i$th Voronoi cell given in (37). The expression of $P_i^{(g)}$ is obtained in the following lemma.

**Lemma 7:** *Given the i.i.d. downlink channel in AS 5, the stationary probabilities for the gain Markov chain are given as*

$$
P_i^{(g)} = \sum_{k=0}^{L-1} \frac{\exp(-\theta_{i-1})\theta_{i-1}^k - \exp(-\theta_i)\theta_i^k}{k!}, \quad i = 1, \cdots, N, \tag{41}
$$

*where $\theta_i$ is defined in (38).*

**Proof:** See Appendix F. □

The stationary probability $\{P_i^{(g)}\}$ is the probability for the channel gain $g$ to lie in the $i$th region (Voronoi cell) on the nonnegative real axis. Since these regions are separated by $\{\theta_i\}$ in (38), the probability $P_i^{(g)}$ is a function of $\theta_i$ and $\theta_{i-1}$ as shown in (41).

*3) Transition Probabilities:* The transition probabilities of the gain Markov chain are defined as

$$P_{i,j}^{(g)} = \Pr\left\{ J_n^{(g)} = i \mid J_{n-1}^{(g)} = j \right\}, \quad i,j = 1, \cdots, N. \tag{42}$$

or equivalently,

$$P_{i,j}^{(g)} = \Pr\left\{ g_n \in \mathbb{V}_i^{(g)} \mid g_{n-1} \in \mathbb{V}_j^{(g)} \right\}, \quad i,j = 1, \cdots, N. \tag{43}$$

It is proposed in [43] that the transition probabilities of a channel Markov chain, similar to the gain Markov chain but for a SISO channel, be derived from the level crossing rate (LCR) of the channel gain. The LCR can be defined for the continuous-time channel gain $g(t)$ as the average frequency at which $g(t)$ crosses a given level in the same direction. This LCR-based method is applied shortly for deriving the transition probabilities of the gain Markov chain.

Without loss of generality, let the channel state values $1, \cdots, N$ be mapped to the Voronoi cells in (37) according to their ascending order on the real axis. Following [43], we make the following assumption that is reasonable for a slowly time-varying channel w.r.t. the symbol rate:

**AS 6:** *The transitions in the gain Markov chain occur only between adjacent states, namely $J_n = J_{n-1} \pm 1$.*

The above assumption results in truncations of transition probabilities between non-adjacent Markov states and hence implicit application of the feedback compression method in Section IV-B.3. Given this assumption, the gain Markov chain can be represented by the graph of a one-dimensional chain as in Fig. 3.

Following the interpretation in [43], the term $\left( P_{i,j}^{(g)} P_i^{(g)} \right) / T_s$ is equal to the LCR of the channel gain $g(t)$ at $\theta_i$ in (38), the boundary between the Voronoi cells $\mathbb{V}_i^{(g)}$ and $\mathbb{V}_j^{(g)}$. Such interpretation leads to the following lemma.

**Lemma 8:** *Under the assumption AS 6, the transition probabilities of the gain Markov chain are given as*

$$P_{i,i+1}^{(g)} = P_{i+1,i}^{(g)} = \frac{T_s}{P_i^{(g)}} \beta\left(\theta_i\right), \quad i = 1, \cdots, N-1, \tag{44}$$

$$P_{i,i}^{(g)} = \begin{cases} 1 - P_{i,i-1}^{(g)} - P_{i,i+1}^{(g)}, & 1 < i < N, \\ 1 - P_{i,i+1}^{(g)}, & i = 1, \\ 1 - P_{i,i-1}^{(g)}, & i = N, \end{cases}$$ (45)

where $\beta(\theta)$ denotes the LCR of $g(t)$ at the level $\theta$ and $P_i^{(g)}$ is the stationary provability in (41).

This lemma shows that the transition probabilities is proportional to the channel gain LCR. Consequently, the transition probabilities increase with the variation speed of the channel gain and hence inversely with the channel coherence time. As shown in [44], in the uniform scattering environment [29], [30], the LCR is actually linearly proportional to the Doppler frequency $f_D$, which is inversely proportional to the channel coherence time,

$$\beta(\theta) = \frac{\sqrt{2\pi} f_D \theta^{(L-1/2)}}{(L-1)! \exp(\theta)}.$$ (46)

### B. Average Feedback Rate

In this section, the average feedback rate for the gain feedback system is derived using Theorem 1. Recall that the feedback compression method in Section IV-B.3 is implicitly applied by making the assumption AS 6. To be specific, the $\epsilon$-neighborhood of each Markov state is comprised of two neighboring states. As the result, the feedback is compressed to be 1-bit for indicating either a leftward or a rightward transition in the gain Markov chain. Given 1-bit feedback, by applying Theorem 1 together with Lemma 7 and Lemma 8, we have the following proposition.

**Proposition 1:** *Under assumptions AS 5 and AS 6, the average feedback rate for the gain feedback system is given as*

$$R^{(g)} = \frac{1}{T_s} \sum_{i=1}^{N} \left(1 - P_{i,i}^{(g)}\right) P_i^{(g)} \quad bit/s,$$ (47)

where $P_{i,i}^{(g)}$ is the probability of no transition in (45) and $P_i^{(g)}$ the stationary probability in (41).

In summary, the average feedback rate for the gain feedback system can be computed by first constructing the gain Markov chain following the procedure in Section VI-A and second applying (47). The following corollary of Proposition 1 follows from Lemma 8 and (46).

**Corrollary 2:** In the uniform scattering environment [29], [30], the average feedback rate of the gain feedback system increases linearly with the Doppler frequency: $R^{(g)} \propto f_D$.

## C. Downlink Ergodic Capacity

The average downlink capacity for the gain feedback system is obtained as follows. Consider a transmission duration of $MT_s$ and a feedback delay of $D$ samples. From (34) substituted with the ideal beamforming vector $\mathbf{w}_n = \mathbf{u}$, we can write the downlink average capacity for the gain feedback system as

$$\bar{C}^{(g)}(M, D) = \frac{1}{M} \sum_{n=1}^{M} \log_2(1 + g_n \eta(g_{n-D})), \tag{48}$$

where $\eta(g_{n-D})$ is the TX power given feedback delay. Due to feedback quantization, the base-station treats the channel gain as having a finite set of values and hence "water fill" [45], [46] the TX power onto these values:

$$\text{(TX power)} \quad \eta(g_n) = \left( \mu - \frac{1}{\rho \mathcal{Q}(g_n)} \right)^+, \tag{49}$$

where $\rho$ is the average TX power. The scalar $\mu$ in (49) is chosen for satisfying the following power constraint

$$\frac{1}{M} \sum_{n=1}^{M} \eta_n(D) = \rho. \tag{50}$$

Moreover, $\mathcal{Q}$ in (49) is the scalar quantization function (cf. Section VI-A.1) defined as

$$\mathcal{Q}(g_n) = \arg \min_{1 \le i \le N} \left( g_n - f_i^{(g)} \right)^2, \tag{51}$$

where $f_i^{(g)}$ is the $i$th codebook member.

The downlink ergodic capacity is derived as follows. To avoid evaluation of the integration in (20) that is usually complicated, we use Corollary 1 rather than Theorem 2 to obtain an approximation of the downlink ergodic capacity as given in the following proposition.

**Proposition 2:** *Under the assumption AS 6, the downlink ergodic capacity for the gain feedback system is*

$$\mathcal{C}^{(g)}(D) \approx \sum_{i=1}^{N} \sum_{j=1}^{N} C_{i,j}^{(g)} \left[ \left( \mathbf{P}^{(g)} \right)^D \right]_{i,j} P_j^{(g)}, \quad \text{for } N \gg 1, \tag{52}$$

*where the elements of the transition probability matrix $\mathbf{P}^{(g)}$ are in (44) and (45) and the stationary probability $P_i^{(g)}$ is obtained in (41). Moreover,*

$$C_{i,j}^{(g)} = \log_2(1 + f_i^{(g)} \cdot \eta(f_j^{(g)})), \tag{53}$$

*where $\eta(f_j)$ is given in (49).*

Note that the condition $N \gg 1$ is usually ensured to keep the quantization error, the difference between the channel gain and its quantized value, small. Using Proposition 2, we can compute an approximation of the downlink ergodic capacity with delayed CSI feedback by constructing the gain Markov chain and applying (52).

The downlink ergodic capacity for the case of infinite delay can be obtained by applying Lemma 5,

$$\mathcal{C}^{(g)}(\infty) \approx \sum_{i=1}^{N} \sum_{j=1}^{N} C_{i,j}^{(g)} P_i^{(g)} P_j^{(g)}, \tag{54}$$

where $C_{i,j}^{(g)}$ is defined in (53) and $P_i^{(g)}$ in (40). Following the definition in (27), we define the feedback capacity gain for the gain feedback system as

$$\Delta \mathcal{C}^{(g)} = \mathcal{C}^{(g)}(D) - \mathcal{C}^{(g)}(\infty), \tag{55}$$

where $\mathcal{C}^{(g)}(D)$ and $\mathcal{C}^{(g)}(\infty)$ are given in (52) and (54), respectively. Numerical results in Section VIII-C show that the feedback capacity gain decreases exponentially with the feedback delay in the uniform scattering environment [29].

## VII. APPLICATION: FEEDBACK OF CHANNEL SHAPE

In this section, the analytical framework developed in Section IV is used to analyze the shape feedback system defined in Section V. First, a Markov chain is constructed in VII-A for modeling the quantized channel shape. Second, the average feedback rate and the downlink ergodic capacity for the shape feedback

system are derived in Section VII-B and Section VII-C, respectively. Note that superscripts $(s)$ are added to variables in the shape feedback system to avoid confusion.

### A. Shape Markov Chain

The quantized shape of the downlink channel is modeled by a Markov chain, called the *shape Markov chain*. In Section VII-A.1, the state space of the shape Markov chain is constructed. The transition and stationary probabilities of the shape Markov chain are derived in Section VII-A.3 and VII-A.2, respectively.

*1) State Space:* The states of the shape Markov chain are one-to-one mapped to the Voronoi cells in the channel shape space – the unit hyper sphere denoted as $\mathbb{S}$. The Voronoi cells are obtained using the modified Lloyd algorithm in [28] with the following distortion function [4], [28]

$$d^{(s)}(\mathbf{u}, \mathbf{u}') = 1 - |\mathbf{u}^H \mathbf{u}'|^2. \tag{56}$$

A detailed description of this algorithm can be found in [28]. In brief, the partitioning of the channel shape space generates a codebook containing a set of possible beamforming vectors: $\mathcal{H}^{(s)} = \{\mathbf{f}_i^{(s)}, \cdots, \mathbf{f}_N^{(s)}\}$. In terms of the codebook members, the Voronoi cells can be defined as

$$\mathbb{V}_i^{(s)} = \left\{\mathbf{u} \in \mathbb{S} \mid |\mathbf{u}^H \mathbf{f}_i^{(s)}| > |\mathbf{u}^H \mathbf{f}_j^{(s)}| \; \forall \; j \neq i \right\}, \quad i = 1, \cdots, N. \tag{57}$$

From (2) and (57), the channel-state $J_n^{(s)}$ can be computed as

$$J_n^{(s)} = \arg \max_{1 \leq i \leq N} \left| \mathbf{u}_n^H \mathbf{f}_i^{(s)} \right|. \tag{58}$$

Moreover, the beamforming vector at the base-station is chosen as [47]

$$\mathbf{w}_n = \mathcal{Q}(\mathbf{u}_{n-D}) = \arg \max_{\mathbf{f} \in \mathcal{H}^{(s)}} \left| \mathbf{f}^H \mathbf{u}_{n-D} \right|, \tag{59}$$

where the function $\mathcal{Q}$ represents the limited feedback process and $D$ is the feedback delay in samples.

*2) Stationary Probabilities:* The stationary probabilities of the shape Markov chain are defined as

$$P_i^{(s)} = \Pr\left\{J_n^{(s)} = i\right\}, \quad i = 1, \cdots, N. \tag{60}$$

Equivalently, from (2),

$$P_i^{(s)} = \Pr\left\{\mathbf{u}_n \in \mathbb{V}_i^{(s)}\right\}, \quad i = 1, \cdots, N. \tag{61}$$

Since $\mathbf{u}_n$ is uniformly distributed on the unit hyper-sphere, the stationary probability of each Markov state is proportional to the area of the corresponding Voronoi cell [5]

$$P_i^{(s)} = \frac{\mathcal{A}_i^{(s)}}{\mathcal{A}^{(s)}}, \quad i = 1, \cdots, N, \tag{62}$$

where $\mathcal{A}_i^{(s)}$ defined in (11) is the area of the $i$th Voronoi cell and $\mathcal{A}^{(s)}$ the surface area of the hyper unit sphere given as [5],

$$\mathcal{A}^{(s)} = \frac{4\pi^L}{(L-1)!}. \tag{63}$$

Next, we will show that the stationary probabilities in (62) are equal. The well-known Gersho's conjecture [37] states that the optimal quantizer for a random vector uniformly distributed on some convex set partitions the set into congruent regions. Since the channel shape is isotropic, based on the Gersho's conjecture, we make the following assumption.

**AS 7:** *The Voronoi cells $\{\mathbb{V}_i^{(s)}\}$ defined in (57) are congruent.*

The above assumption ensures equal area for the Voronoi cells: $\mathcal{A}_1 = \cdots = \mathcal{A}_N$, where $\mathcal{A}_i$ is defined in (11). Therefore, together with (62) and the fact that $\sum_{i=1}^{N} P_i^{(s)} = 1$, AS 7 results in the following lemma.

**Lemma 9:** *Under the assumptions AS 5 and AS 7, the stationary probabilities of the shape Markov chain are homogeneous, hence*

$$P_i^{(s)} = \frac{1}{N} \quad \forall\, 1 \leq i \leq N. \tag{64}$$

Compared with their counterparts in (41) for the gain feedback system, the stationary probabilities for the shape feedback system have a simpler expression due to the uniform distribution of the channel shape. Last, from the congruency of the Voronoi cells and the fact that the channel shape is isotropic, we can infer that the graph of the shape Markov chain is a "mesh" network on the hyper unit sphere as illustrated in Fig. 4.

*3) Transition Probabilities:* The transition probabilities of the shape Markov chain are defined as

$$P_{i,j}^{(s)} = \Pr\left\{ J_n^{(s)} = i \mid J_{n-1}^{(s)} = j \right\}, \tag{65}$$

where $1 \le i, j \le N$. Equivalently, from (2),

$$P_{i,j}^{(s)} = \Pr\left\{ \mathbf{u}_n \in \mathbb{V}_i^{(s)} \mid \mathbf{u}_{n-1} \in \mathbb{V}_j^{(s)} \right\} = \int_{\mathbb{V}_i^{(s)}} \int_{\mathbb{V}_j^{(s)}} f_{\mathbf{uu}}(\mathbf{u}_n, \mathbf{u}_{n-1}) d\mathbf{u}_n d\mathbf{u}_{n-1}, \tag{66}$$

where $f_{\mathbf{uu}}$ is the joint PDF of $\mathbf{u}_n$ and $\mathbf{u}_{n-1}$. To the best of our knowledge, no closed-form expression for $f_{\mathbf{uu}}$ exists in the literature. Neither is there the concept of LCR for the channel shape and hence the LCR method is not applicable here. Recall that this method is used in Section VI-A.2 for deriving the transition probabilities of the gain Markov chain. Therefore, we resolve to the Monte Carlo method for computing the transition probabilities of the shape Markov chain. Related details are provided in Section VIII.

Last, we provide a lemma useful for deriving the average feedback rate in the next section.

**Lemma 10:** *For the shape Markov chain, the probabilities of no transition are homogeneous, hence*

$$P_{11}^{(s)} = \cdots = P_{NN}^{(s)}. \tag{67}$$

**Proof:** See Appendix G. □

Lemma 10 shows that like the stationary probabilities the probabilities of no transition $\{P_{ii}^{(s)}\}$ also exhibit homogeneity given that the channel shape is isotropic. Despite its being an interesting topic, we do not investigate in this paper whether other transition probabilities $\{P_{i,j}^{(s)}\}$ for $i \ne j$ possess homogeneity.

*B. Average Feedback Rate*

In this section, the average feedback rate for the shape feedback system will be derived using the shape Markov chain constructed Section VII-A. By applying Theorem 1 and Lemma 9 and 10, we have the following proposition.

**Proposition 3:** *Under the assumption AS 5 and AS 7, the average feedback rate for the shape feedback system is given as*

$$R^{(s)} = \frac{B}{T_s}\left(1 - P_{11}^{(s)}\right), \tag{68}$$

where $P_{11}^{(s)}$ follows the definition in (66).

A few remarks are in order:

- The feedback compression method in Section IV-B.3 is not accounted for in Proposition 3 but related numerical results are presented in Section VIII.

- As shown by numerical results in Section VIII, the average feedback rate increases linearly with the Doppler frequency in the uniform scattering environment [29]. Note that the same result exists for the gain feedback system according to Corollary 2.

- Due to the homogeneity of the shape Markov chain shown in Lemma 9 and Lemma 10, the average feedback rate in (68) has a simpler expression than that in (47) for the gain feedback system.

### C. Downlink Ergodic Capacity

The downlink ergodic capacity for the shape feedback system will be derived as a function of the feedback delay as well as the parameters of the shape Markov chain constructed in Section VII-A. From (34) and given constant TX power $\eta_n = \rho$, the downlink average capacity for the shape feedback system can be written as

$$\bar{C}^{(s)}(M) = \frac{1}{M} \sum_{n=1}^{M} \log_2(1 + \rho g |\mathbf{w}_n^H \mathbf{u}_n|^2), \tag{69}$$

where $\rho$ the TX power and $\mathbf{w}_n$ is the beamforming vector in (59). An approximation of the downlink ergodic capacity is obtained using Corollary 1 in the following proposition.

**Proposition 4:** *Under the assumption AS 5 and AS 7, the downlink ergodic capacity for the shape feedback system is given as*

$$\mathcal{C}^{(s)}(D) \approx \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} C_{i,j}^{(s)} \left[ \left( \mathbf{P}^{(s)} \right)^D \right]_{i,j}, \quad \text{for } N \gg 1, \tag{70}$$

*where $\mathbf{P}^{(s)}$ is the transition probability matrix and*

$$C_{i,j}^{(s)} = \int_0^\infty \log_2 \left( 1 + \rho g \left| \mathbf{f}_i^{(s)H} \mathbf{f}_j^{(s)} \right|^2 \right) f_g(g) dg. \tag{71}$$

*where $f_g$ is the PDF of the channel gain in (36).*

Note that the condition $N \gg 1$ is usually ensured to keep the feedback information on the channel shape accurate. Using Proposition 4, we can compute an approximation of the downlink ergodic capacity with delayed CSI feedback by constructing the shape Markov chain and applying (70).

The downlink ergodic capacity for infinite feedback delay can be obtained by applying Lemma 5 and Lemma 9,

$$\mathcal{C}^{(s)}(\infty) \approx \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} C_{i,j}^{(s)}, \quad N \gg 1, \tag{72}$$

where $C_{i,j}^{(s)}$ is defined in (71). Following the definition in (27), we define the feedback capacity gain for the shape feedback system as

$$\Delta \mathcal{C}^{(s)} = \mathcal{C}^{(s)}(D) - \mathcal{C}^{(s)}(\infty), \tag{73}$$

where $\mathcal{C}^{(s)}(D)$ and $\mathcal{C}^{(s)}(\infty)$ are given in (70) and (72), respectively. Numerical results in Section VIII-C show that the feedback capacity gain for the shape feedback system decreases exponentially with the feedback delay. Recall that the same observation is also made for the gain feedback system.

## VIII. NUMERICAL RESULTS

For demonstrating the application of analytical tools developed in this paper, a design example for the shape feedback system is provided in Section VIII-A. Other numerical results and discussion concerning the feedback rates, the feedback capacity gains and the channel-state Markov chains of the gain and the shape feedback systems are presented in Section VIII-B.

For all numerical results, the following assumption is made about the downlink channel:

**AS 8:** *The scattering environment is uniform [29], [30].*

The direct result of the above assumption is that the temporal correlation of the downlink channel coefficients are specified by Clarke's function [30]

$$\mathrm{E}[\mathbf{h}_n^H \mathbf{h}_{n-1}] = 2L J_0(2\pi f_D T_s), \tag{74}$$

where $J_0$ is the Bessel function of 0th order, and $f_D$ the Doppler frequency.

*A. Design Example*

The shape feedback system analyzed in Section VII will be designed to achieve a feedback capacity gain of 1 bit/s/Hz with respect to the case of no CSI feedback. We assume that the number of TX antennas is $L = 2$, the carrier frequency is $f_c = 2$ GHz, and the symbol rate is $1/T_s = 1$ MHz. The average feedback rate is constrained as $\mathcal{R} \leq 20$ kbits/s. Furthermore, the feedback CSI accuracy, represented by the signal-to-quantization-noise ratio (SQNR), is required to be $\text{SQNR}^{(s)} \geq 21$ dB. For the shape feedback system, the $\text{SQNR}^{(s)}$ is defined as:

$$\text{SQNR}^{(s)} = \sum_{i=1}^{N} \text{E}_{\text{V}_i^{(s)}} \left[ \frac{1}{1 - |\mathbf{u}^H \mathcal{Q}(\mathbf{u})|^2} \right], \tag{75}$$

where the quantization function $\mathcal{Q}$ is given in (59). The design objective is to determine a set of system parameters including the vehicular speed, the feedback delay and the CSI quantization complexity by using the analytical tools developed in this paper.

The curves useful for designing the system parameters are presented. We plot the curves of *average feedback rate vs. Doppler frequency* in Fig. 5 for different codebook sizes $N$. The dashed curves correspond to the case of no feedback compression, where average feedback rates are computed using (68) with $P_{11}^{(s)}$ obtained using the Monte Carlo method[6]. The solid curves correspond to the feedback compression method in Section IV-B.3 with $\epsilon = 10^{-6}$ , where average feedback rates are computed using (17). Also shown in the figure are $\text{SQNR}^{(s)}$ and SNR Loss[7] defined as

$$\text{SNR Loss} = -10 \log_{10} \sum_{i=1}^{N} \text{E}_{\text{V}_i^{(s)}} \left[ |\mathbf{u}^H \mathbf{f}_i^{(s)}|^2 \right] \quad \text{dB}. \tag{76}$$

where $\mathbf{f}_i$ is the $i$th member of the codebook $\mathcal{H}^{(s)}$. We also plot the *feedback capacity gain vs. the feedback delay* curves in Fig. 6. The feedback capacity gain is computed using (73). Different values of normalized

---

[6] A sequence of the temporally-correlated channel shape $\{\mathbf{u}_n\}$ is generated and converted into the channel-sate sequence $\{J_n^{(s)}\}$ using (58). The pairs of consecutive 1's and the number of 1's in the channel-state sequence are counted. The ratio between these two counted values is used for approximating the transition probability $P_{11}^{(s)}$.

[7] The SNR Loss is defined for evaluating the effect of quantizing the channel shape on the SNR and hence the effect of feedback delay is omitted by setting $D = 0$. For the shape feedback system, the instantaneous SNR Loss with respect to ideal beamforming ($\mathbf{w}_n = \mathbf{u}_n$) is $|\mathbf{w}_n^H \mathbf{u}_n|^2$ where the beamforming vector $\mathbf{w}_n$ is given in (59). By taking expectation of $|\mathbf{w}_n^H \mathbf{u}_n|^2$ and substituting (59) and (57), (76) follows.

Doppler frequencies are considered.

Consider the case of no feedback compression. Under the constraints of $\mathcal{R} \leq 20$ kbits/s and SQNR$^{(s)} \geq$ 21 dB, the minimum codebook size is determined from Fig. 5 to be $N = 16$. It follows that the corresponding range of the Doppler frequency $f_D$ is given as $0 \leq f_D \leq 600$ Hz. Using the relationship $V = f_D V_c / f_c$ where $V_c$ denote the speed of light, we can compute the range of the vehicular speed to be $0 \leq V \leq 324$ km/h. Next, to ensure the feedback capacity gain of 1 bit/s/Hz is achieved for all Doppler frequencies in the range of $0 \leq f_D \leq 600$ Hz, we determine from Fig. 6 the feedback delay $T_D = DT_s$ must be constrained as $0 \leq T_D \leq 120$ $\mu$s. Last, from (59), the quantization of the channel shape $\mathbf{u}$ involves an exhaustive search over the codebook $\mathcal{H}^{(s)}$ and hence requires $N$-times computation of the vector product $|\mathbf{u}^H \mathbf{f}|$ where the $\mathbf{f}$ is a codebook member. The system parameters designed above are summarized in the first column of Table I.

The system parameters for the case of feedback compression (cf. Section IV-B.3) are obtained following a similar procedure and shown in the second column of Table I. Note that the feedback compression method does not require an exhaustive search over the codebook $\mathcal{H}^{(s)}$ and hence has a low quantization complexity compared with the case of no feedback compression. From Table I, we can observe that compared with no feedback compression the feedback compression method allows a larger range of vehicular speed but a smaller range of feedback delay and requires less computation for quantizing the channel shape.

*B. Additional Results and Discussions*

*1) Average Feedback Gain:* For the gain feedback system with four TX antennas, the curves of *average feedback rate vs. Doppler frequency* are plotted in Fig 7 for different codebook size $N$, where the average feedback rate is computed using (47). Also shown in the figure are the signal-to-quantization-noise-ratio (SQNR) defined as

$$\text{SQNR}^{(g)} = \sum_{i=1}^{N} \text{E}_{\text{V}_i^{(g)}} \left[ \frac{L}{(f_i^{(s)} - g)^2} \right]. \tag{77}$$

where $f_i^{(g)}$ is the $i$th member of the codebook $\mathcal{H}^{(g)}$. The condition $N \gg 1$ in Proposition 1 is satisfied for SQNR$^{(g)}$ being sufficiently large, say $> 10$ dB.

It can be observed from Fig. 7 and Fig. 5 that for both the gain and the shape feedback systems *the average feedback rates increase linearly with the Doppler frequency*. Moreover, for the shape feedback system, feedback compression significantly reduces the average feedback rates especially for larger codebook sizes ($N$).

Feedback rate curves similar to those in Fig. 7 and Fig. 5 are useful for designing a multi-antenna limited-feedback system in three aspects as illustrated in Section VIII-A. First, for a required SQNR, we can use the feedback-rate curves to determine the feedback rates for different Doppler frequencies. Second, for a given Doppler frequency, we can obtain from these curves SQNR's achieved by various feedback rates. Third, for a fixed feedback rate, we can use these curves to evaluate the impact of Doppler frequency on the SQNR.

*C. Feedback Capacity Gain*

Consider the gain feedback system with two TX antennas: $L = 2$. We plot the *feedback capacity gain vs. the feedback delay* curves in Fig. 8. Different values of normalized Doppler frequencies are considered. The feedback capacity gain is computed using (55).

Two key observations can be made from Fig. 8 and Fig. 6. First, *the feedback capacity gain decreases exponentially with the feedback delay*, which agrees with Theorem 3 in Section IV-C.2. Second, *the decreasing rate, namely* the slope of the feedback capacity gain vs. feedback delay curve*, increases with the normalized Doppler frequency.*

Feedback rate curves similar to those in Fig. 8 and Fig. 6 are useful for designing a multi-antenna limited-feedback system as demonstrated in Section VIII-A. First, for a given Doppler frequency, we can use these curves to determine the tolerable feedback delay to achieve a particular feedback capacity gain. Second, for a given feedback delay, the maximum Doppler frequency, hence the limit of the mobile speed, can be obtained from these curves.

*D. Channel-State Markov Chains*

For verifying Lemma 7 and Lemma 8, analytical and simulated values for the stationary and transition probabilities of the channel-state Markov chain for the gin feedback system are plotted in Fig. 9. The "analytical" curve for the stationary probabilities is generated using (41) and that for the transition probabilities using (44) and (46). It can be observed that the analytical and simulated results agree with each other very well.

To illustrate Lemma 9 and Lemma 10, the stationary and transition probabilities $\{P_{ii}^{(s)}\}$ of the channel-state Markov chain for the shape feedback system are plotted in Fig. 10. Codebooks $\mathcal{H}^{(s)}$ used for generating the plots are constructed using the modified Lloyd algorithm [28] with the training sequence length fixed at 20000 (cf. Section VII-A.1). According to Lemma 9 and Lemma 10, all curves in Fig. 10 should be horizontal lines, but fluctuations of the curves are observed for $N \geq 16$ albeit in a small range. The explanation we offer is that with the training sequence length fixed the optimality of the codebooks constructed numerically degrades with the codebook size $N$. The degradation contributes to the fluctuations of the curves for large $N$ values. Nevertheless, given the small ranges of the fluctuations, Lemma 9 and Lemma 10 are verified to a large extent.

## IX. CONCLUSION

Limited systems with temporally-correlated channels have been analyzed using ergodic theory and theory of Markov chains. We have modeled the downlink temporally-correlated channels using Markov chains whereby feedback rates and downlink capacities have been derived as functions of the Markov chain parameters. By analyzing Markov chain convergence rates, we have shown that downlink capacity gains w.r.t. the case of no feedback decrease at least exponentially with feedback delay. Furthermore, we have shown that modeling downlink channels using Markov chains facilitates feedback compression by truncating transition probabilities of the Markov chains. Specific results have been presented on limited feedback of the channel gain and the channel shape in a MISO beamforming system. We have constructed

two Markov chains modeling the channel gain and shape. Using them, average feedback rates have been shown to increase linearly with Doppler frequencies in the uniform scattering environment. Moreover, the predicted exponential reduction rate of feedback capacity gains has been confirmed using numerical results. We have also shown that feedback compression effectively reduces feedback rates.

This paper opens several interesting questions for future research. First, how can the feedback bits be optimally compressed by exploiting the channel temporal correlation? Second, how to construct the channel Markov chain at the packet level rather than at the present physical level [48]? Such Markov chain is useful not only for limited feedback analysis but also for queueing analysis. Third, can the channel coherence time be defined rigorously using the relationship between feedback capacity gain and feedback delay addressed in this paper? Answers to these questions will be useful for minimizing limited feedback rates and designing a wireless network with limited feedback.

## APPENDIX

### A. Proof of Lemma 1

First, we will prove that the stationary probabilities are nonzero. From (8) and AS 5,

$$P_i = \prod_{l=1}^{L} \int_{a_{i,l}^-}^{a_{i,l}^+} f_\nu(h_{l,\alpha}) dh_{l,\alpha} \int_{b_{i,l}^-}^{b_{i,l}^+} f_\nu(h_{l,\beta}) dh_{l,\beta}, \tag{78}$$

where $f_\nu$ is the PDF of $\mathcal{N}(0, \frac{1}{2})$, $h_{l,\alpha}$ and $h_{l,\beta}$ are the real and imaginary parts of the $l$th coefficient of the downlink channel $\mathbf{h}_n$, the real values $\{a_{i,1}^-, a_{i,1}^+, b_{i,1}^-, b_{i,1}^+, \cdots, a_{i,L}^-, a_{i,L}^+, b_{i,L}^-, b_{i,L}^+\}$ determines the volume of the $i$ Voronoi cell in (11)

$$\mathcal{A}_i = \prod_{l=1}^{L} (a_{i,l}^+ - a_{i,l}^-)(b_{i,l}^+ - b_{i,l}^-). \tag{79}$$

Since $\mathcal{A}_i > 0 \ \forall \ 1 \leq i \leq N$,

$$a_{i,l}^+ > a_{i,l}^- \quad \text{and} \quad b_{i,l}^+ > b_{i,l}^-, \quad \forall \ 1 \leq i \leq N \text{ and } 1 \leq l \leq L. \tag{80}$$

It follows that

$$\int_{a_{i,l}^-}^{a_{i,l}^+} f_\nu(h) dh > 0 \text{ and } \int_{b_{i,l}^-}^{b_{i,l}^+} f_\nu(h) dh > 0, \quad \forall \ 1 \leq i \leq N \text{ and } 1 \leq l \leq L. \tag{81}$$

By substituting (81) into (78), we prove that $P_i > 0 \ \forall \ 1 \leq i \leq N$.

Given that $P_i > 0$, to prove $P_{i,j} > 0$, from (9), it is sufficient to show that

$$\int_{\mathbb{V}_i} \int_{\mathbb{V}_j} f_{\mathbf{hh}}(\mathbf{h}_n, \mathbf{h}_{n-1}) d\mathbf{h}_n d\mathbf{h}_{n-1} > 0. \tag{82}$$

The procedure for proving the above inequality is similar to that for proving $P_i > 0$ and hence omitted.

*B. Proof of Theorem 1*

We say a channel state transition occurs at the $n$th sample if $J_n \neq J_{n-1}$. Now consider a duration of $MT_s$. We denote the number of transitions as $K(M)$ and it can be defined as

$$K(M) = \frac{1}{M} \sum_{n=2}^{M} 1\{J_n \neq J_{n-1}\}. \tag{83}$$

By substituting (83) into (4),

$$R(M) = \frac{K(M)B}{T_s}. \tag{84}$$

Define the function $z(i,j) = 1\{i \neq j\}$. The condition in (12) is satisfied since $\sum_{i,j} 1\{i \neq j\} P_i P_{i,j} < 1$. Hence by applying Lemma 1-3 on (83),

$$\lim_{M \to \infty} K(M) = \sum_{i=1}^{N} \sum_{j=1}^{N} z(i,j) P_i P_{i,j} = \sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} P_i P_{i,j}. \tag{85}$$

By combining (85) and (84) and using $\sum_{j=1}^{N} P_{i,j} = 1$, (14) follows.

*C. Proof of Theorem 2*

By applying the Ergodic Theorem [49] on (5), the downlink ergodic capacity is given as

$$\mathcal{C}(D) = \lim_{M \to \infty} \bar{C}(D, M) = \int \int C(\mathbf{h}_n, \mathbf{h}_{n-D}) f_{\mathbf{hh}}(\mathbf{h}_n, \mathbf{h}_{n-D}) d\mathbf{h}_n d\mathbf{h}_{n-D}. \tag{86}$$

By partitioning the channel space into Voronoi cells $\{\mathbb{V}_i\}$, (86) can be re-written as

$$\mathcal{C}(D) = \sum_{i=1}^{N} \sum_{j=1}^{N} \int_{\mathbb{V}_i} \int_{\mathbb{V}_j} C(\mathbf{h}_n, \mathbf{h}_{n-D}) f_{\mathbf{hh}}(\mathbf{h}_n, \mathbf{h}_{n-D}) d\mathbf{h}_n d\mathbf{h}_{n-D}. \tag{87}$$

By substituting (20) into (87),

$$
\mathcal{C}(D) = \sum_{i=1}^{N}\sum_{j=1}^{N} \bar{C}_{i,j} \Pr\{\mathbf{h}_n \in \mathbb{V}_i, \mathbf{h}_{n-D} \in \mathbb{V}_j\}, \tag{88}
$$

$$
= \sum_{i=1}^{N}\sum_{j=1}^{N} \bar{C}_{i,j} \Pr\{\mathbf{h}_n \in \mathbb{V}_i | \mathbf{h}_{n-D} \in \mathbb{V}_j\} \Pr\{\mathbf{h}_{n-D} \in \mathbb{V}_j\}. \tag{89}
$$

By substituting (2), (89) is equivalently written as

$$
\mathcal{C}(D) = \sum_{i=1}^{N}\sum_{j=1}^{N} \bar{C}_{i,j} \Pr\{J_n = i | J_{n-D} = j\} \Pr\{J_{n-D} = j\}. \tag{90}
$$

Using the property of the channel Markov chain in (10) and the definition of $P_i$ in (6), (19) follows from (89).

*D. Proof of Lemma 4*

Consider the case of zero feedback delay ($D = 0$), hence $K_n = J_n$. By applying the Ergodic Theorem [49] on (5), the ergodic performance is given as

$$
\mathcal{C} = \lim_{M \to \infty} \bar{C}(0, M) = \int C(\mathbf{h}, \mathbf{h}) f_{\mathbf{h}}(\mathbf{h}) d\mathbf{h}. \tag{91}
$$

By partitioning the channel space into $N$ Voronoi cells, (91) is rewritten as

$$
\mathcal{C} = \sum_{n=1}^{N} \int_{\mathbb{V}_n} C(\mathbf{h}, \mathbf{h}) f_{\mathbf{h}}(\mathbf{h}) d\mathbf{h}. \tag{92}
$$

By substituting (24) into (92), (23) follows.

*E. Proof of Theorem 3*

From (27), (25) and (19),

$$
\Delta\mathcal{C}(D) = \sum_{i=1}^{N}\sum_{j=1}^{N} \mathcal{C}_{i,j} \left( \left[ \mathbf{P}^D \right]_{i,j} P_i - P_j P_i \right), \tag{93}
$$

$$
\leq \sum_{i=1}^{N}\sum_{j=1}^{N} \mathcal{C}_{i,j} \left| \left[ \mathbf{P}^D \right]_{i,j} P_i - P_j P_i \right|. \tag{94}
$$

By applying Schwarz inequality on (94),

$$
\Delta\mathcal{C}(D) \leq \sum_{i=1}^{N} \sqrt{\sum_{j=1}^{N} \mathcal{C}_{i,j} \sum_{n=1}^{N} \left| \left[ \mathbf{P}^D \right]_{i,n} P_i - P_n P_i \right|}. \tag{95}
$$

Substitute (30) into (95) and we have the upper-bound in (32). The lower bound is by system design.

*F. Proof of Lemma 7*

From (37) and (40),

$$P_i^{(g)} = \int_{\sqrt{\theta_{i-1}}}^{\sqrt{\theta_i}} f_g(g)dg, \quad i = 1, \cdots, N,$$ (96)

where $\theta_i$ is defined in (38) and $f_g(g)$ is the PDF of $g$ in (36). From (96), we can show that

$$P_i^{(g)} = \frac{\Gamma(L, \theta_{i-1}) - \Gamma(L, \theta_i)}{(L-1)!},$$ (97)

where $\Gamma(L, x)$ is the incomplete Gamma function defined as [50]

$$\Gamma(L, x) = \int_x^{\infty} \tau^{L-1} \exp(-\tau)d\tau.$$ (98)

Since $L$ is a integer, the above incomplete Gamma function has a closed-form expression [50].

*G. Proof of Lemma 10*

From (66),

$$P_{ii}^{(s)} = \int_{\mathbb{V}_i^{(s)}} \int_{\mathbb{V}_i^{(s)}} f_{\mathbf{uu}}(\mathbf{u}_n, \mathbf{u}_{n-1})d\mathbf{u}_n d\mathbf{u}_{n-1}.$$ (99)

Since the channel shape $\mathbf{u}_n$ is isotropic, the joint PDF $f_{\mathbf{uu}}$ remains unchanged to the rotation of the coordinate system. From AS 7, every two Voronoi cells $\mathbb{V}_i^{(s)}$ and $\mathbb{V}_j^{(s)}$ for $i \neq j$ are congruent. Hence, one is the rotated version of the other. By applying the rotation invariant property of $f_{\mathbf{uu}}$ and the congruency of Voronoi cells $\{\mathbb{V}_i^{(s)}\}$ on (99), (67) follows.

## REFERENCES

[1] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Trans. on Info. Theory*, vol. 45, pp. 1468–89, July 1999.

[2] D. J. Love, R. W. Heath Jr., W. Santipach, and M. L. Honig, "What is the value of limited feedback for MIMO channels?," vol. 42, pp. 54–59, Oct. 2004.

[3] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Kluwer Academic Press, 1992.

[4] D. J. Love, R. W. Heath Jr., and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. on Info. Theory*, vol. 49, pp. 2735–47, Oct. 2003.

[5] K. K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, "On beamforming with finite rate feedback in multiple antenna systems," *IEEE Trans. on Info. Theory*, vol. 49, pp. 2562–79, Oct. 2003.

[6] J. H. Conway, R. H. Hardin, and N. J. A. Sloane, "Packing lines, planes, etc.: Packings in Grassmannian spaces," *Experimental Mathematics*, vol. 5, pp. 139–159, 1996.

[7] J. C. Roh and B. D. Rao, "Channel feedback quantization methods for MISO and MIMO systems," in *Proc., IEEE PIMRC*, (Barcelona, Spain), pp. 805–9, Sept. 2004.

[8] J. Choi and R. W. Heath Jr., "Interpolation based transmit beamforming for MIMO-OFDM with limited feedback," vol. 1, pp. 249 – 253, June 2004.

[9] J. Choi and R. W. Heath Jr., "Interpolation based unitary precoding for spatial multiplexing MIMO-OFDM with limited feedback," vol. 1, pp. 214–218, Nov. 2004.

[10] B. Mondal, R. Samanta, and R. W. Heath Jr., "Frame theoretic quantization for limited feedback MIMO beamforming systems," in *Proc, Intl. Conf. Wireless Networks, Comm. and Mobile Computing*, vol. 2, pp. 1065–1070, June 2005.

[11] C.-K. A. Yeung and D. J. Love, "Performance analysis of random vector quantization limited feedback beamforming," in *Proc., IEEE Asilomar*, pp. 408–412, Oct. 2005.

[12] W. Santipach and M. L. Honig, "Signature optimization for DS-CDMA with limited feedback," in *Proc. 7th Int. Symp. Spread Spect. Tech. Appl.*, vol. 1, pp. 180–184, Sept. 2002.

[13] J. C. Roh and B. D. Rao, "Vector quantization techniques for multiple-antenna channel information feedback," in *Proc., Int. Conf. on Sig. Proc. and Comms.*, pp. 11–14, Dec. 2004.

[14] P. Xia, S. Zhou, and G. B. Giannakis, "Achieving the Welch bound with difference sets," *IEEE Trans. on Info. Theory*, vol. 51, pp. 1900–07, May 2005.

[15] V. K.-N. Lau, Y. Liu, and T.-A. Chen, "Optimal partial feedback design for MIMO block fading channels with feedback capacity constraint," in *Proc., IEEE Intl. Symposium on Information Theory*, p. 65, June 2003.

[16] E. G. Larsson, G. Ganesan, P. Stoica, and W.-H. Wong, "On the performance of orthogonal space-time block coding with quantized feedback," *IEEE Communications Letters*, vol. 11, pp. 487–89, Nov. 2002.

[17] D. J. Love and R. W. Heath Jr., "Limited feedback unitary precoding for orthogonal space-time block codes," *IEEE Trans. on Signal Processing*, vol. 53, pp. 64–73, Jan. 2005.

[18] D. J. Love and R. W. Heath Jr., "Limited feedback unitary precoding for spatial multiplexing systems," *IEEE Trans. on Info. Theory*, vol. 51, pp. 1967–76, Aug. 2005.

[19] J. C. Roh and B. D. Rao, "An efficient feedback method for MIMO systems with slowly time-varying channels," in *Proc., IEEE Wireless Communications and Networking Conf.*, (Atlanta, GA), pp. 11–14, Mar. 2004.

[20] J. C. Roh and B. D. Rao, "MIMO spatial multiplexing systems with limited feedback," in *Proc., IEEE Intl. Conf. on Communications*, pp. 777–82, May 2005.

[21] M. A. Sadrabadi, A. K. Khandani, and F. Lahouti, "A new method of channel feedback quantization for high data rate MIMO systems,"

in *Proc., IEEE Globecom*, pp. 91–95, Nov. 2004.

[22] R. S. Blum, "MIMO with limited feedback of channel state information," pp. 89–92, Apr. 2003.

[23] V. K. N. Lau, Y. Liu, and T.-A. Chen, "On the design of MIMO block-fading channels with feedback-link capacity constraint," *IEEE Trans. on Communications*, vol. 52, pp. 62–70, Jan. 2004.

[24] B. C. Banister and J. R. Zeidler, "Feedback assisted transmission subspace tracking for MIMO systems," *IEEE Journal on Sel. Areas in Communications*, vol. 21, no. 3, pp. 452–63, 2003.

[25] B. C. Banister and J. R. Zeidler, "A simple gradient sign algorithm for transmit antenna weight adaptation with feedback," vol. 51, pp. 1156 – 1171, May 2003.

[26] S. Zhou and G. B. Giannakis, "How accurate channel prediction needs to be for adaptive modulation in Rayleigh MIMO channels?," *IEEE Trans. on Communications*, vol. 3, pp. 1285–94, July 2004.

[27] H. Viswanathan, "Capacity of Markov channels with receiver CSI and delayed feedback," *IEEE Trans. on Info. Theory*, vol. 45, pp. 761–771, Mar. 1999.

[28] P. Xia, S. Zhou, and G. B. Giannakis, "Multiantenna adaptive modulation with beamforming based on bandwidth-constrained feedback," vol. 53, pp. 526–36, Mar. 2005.

[29] R. H. Clarke, "A statistical theory of mobile radio reception," *Bell Syst. Tech. J.*, pp. 957–1000, 1974.

[30] J. W. C. Jakes, *Microwave mobile comunications*. New York: Willey, 1974.

[31] D. W. Stroock, *An Introduction to Markov Processes*. Springer, 2005.

[32] P. Bremaud, *Markov Chains*. Springer, 1999.

[33] S. A. Savari and R. G. Gallager, "Generalized Tunstall codes for sources with memory," *IEEE Trans. on Info. Theory*, vol. 43, pp. 658–68, Mar. 1997.

[34] S. A. Savari, "Renewal theory and source coding," *Proceedings of the IEEE*, vol. 88, pp. 1692–1702, Nov. 2000.

[35] I. Tabus and J. Rissanen, "Asymptotics of greedy algorithms for variable-to-fixed length coding of Markov sources," *IEEE Trans. on Info. Theory*, vol. 48, pp. 2022–35, July 2002.

[36] T. Tjalkens and F. Willems, "Variable to fixed-length codes for Markov sources," *IEEE Trans. on Info. Theory*, vol. 33, pp. 246–57, Mar. 1987.

[37] A. Gersho, "Asymptotically optimal block quantization," vol. 25, pp. 373–380, Jul. 1979.

[38] J. Rosenthal, "Markov chain convergence: From finite to infinite," *Stochastic Processes and their Applications*, vol. 62, pp. 55–72, 1996.

[39] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. on Telecomm.*, vol. 10, no. 6, pp. 585–595, 1999.

[40] M. J. Sabin and R. M. Gray, "Product code vector quantizers for waveform and voice coding," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. ASSP-32, pp. 474–88, June 1984.

[41] T. K. Y. Lo, "Maximum ratio transmission," *IEEE Trans. on Communications*, vol. 47, pp. 1458–61, Oct. 1999.

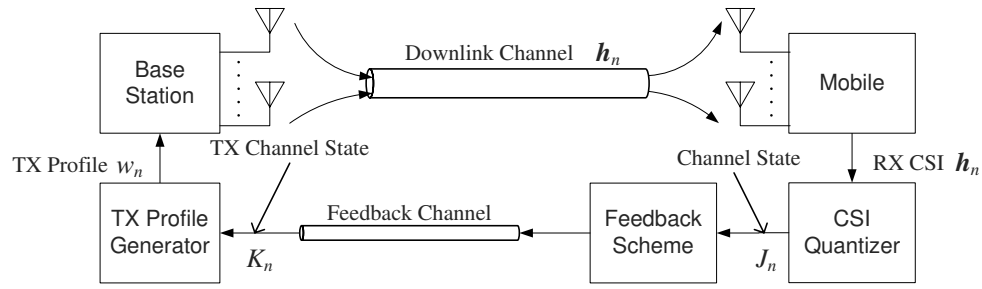[42] R. A. Monzingo and T. W. Miller, *Introduction to Adaptive Arrays*. SciTech Publishing, 2003.

Fig. 1. Discrete-time limited feedback system with a temporally-correlated channel (time is represented by the sample index $n$)

[43] H.-S. Wang and N. Moayeri, "Finite-state Markov channel – a useful model for radio communication channels," *IEEE Trans. on Veh. Technology*, vol. 44, pp. 163–71, Feb. 1995.

[44] Y.-C. Ko, A. Abdi, M.-S. Alouini, and M. Kaveh, "A general framework for the calculation of the average outage duration of diversity systems over generalized fading channels," *IEEE Trans. on Veh. Technology*, vol. 51, pp. 1672–80, Nov. 2002.

[45] A. J. Goldsmith and P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. on Info. Theory*, vol. 43, pp. 1986–92, Nov. 1997.

[46] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley-Interscience, 1991.

[47] D. J. Love, R. W. Heath Jr., and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," vol. 4, pp. 2618–2622, May 2003.

[48] Y.-Y. Kim and S.-Q. Li, "Modeling multipath fading channel dynamics for packet data performance analysis," *Wireless Networks*, vol. 6, pp. 481–92, June 2000.

[49] P. Billingsley, *Ergodic Theory and Information*. R. E. Krieger Pub. Co, 1978.

[50] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. New York: Dover: Dover Publications, 1965.

TABLE I

SUMMARY OF PARAMETERS FOR THE MISO BEAMFORMING SYSTEM IN THE DESIGN EXAMPLE

|  | w/o Feedback Compr. | w/ Feedback Compr. |
| --- | --- | --- |
| Codebook Size $N$ | 16 | 16 |
| Vehicular Speed $V$ (km/h) | $\leq 324$ | $\leq 432$ |
| Feedback Delay $T_D$ ($T_s$) | $\leq 120$ | $\leq 90$ |
| Av. Quant. Complexity [a] | 16 | 5.8 |

[a] Each unit refers to the computation of $|\mathbf{u}^H \mathbf{f}|$ where $\mathbf{u}$ is the channel shape and $\mathbf{f}$ a particular member of the codebook $\mathcal{H}^{(s)}$.
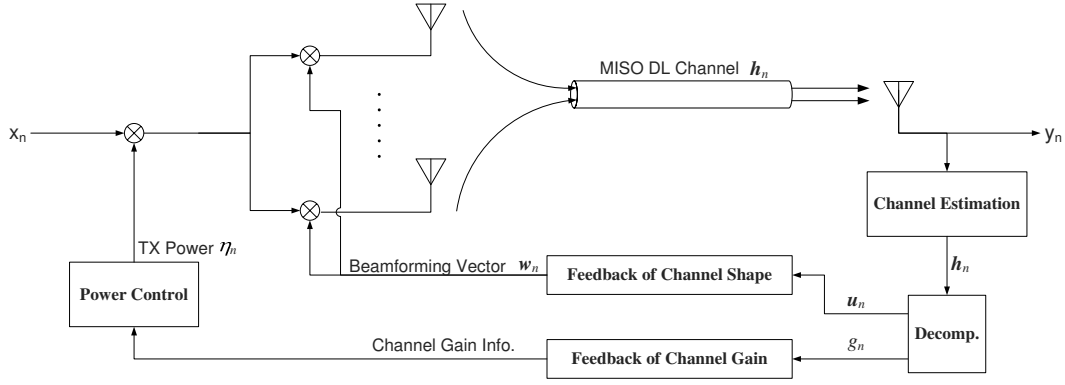
Fig. 2. The MISO system with transmit beamforming and power control



Fig. 3. Graph for the gain Markov chain. Circles: Markov states; Arrows: positive transition probabilities
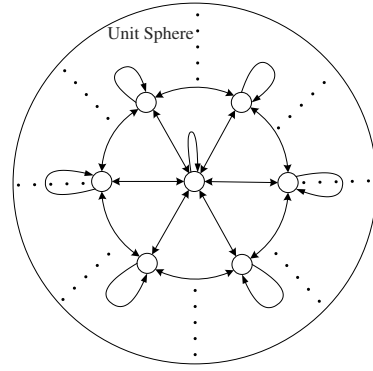


Fig. 4. Graph for the shape Markov chain; Circles: Markov states; Arrows: positive transition probabilities
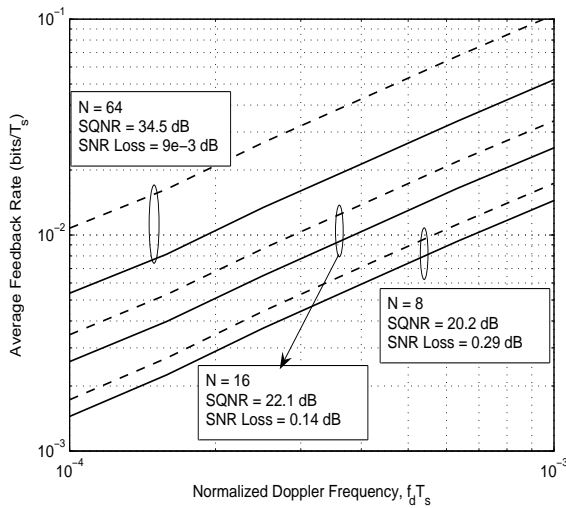


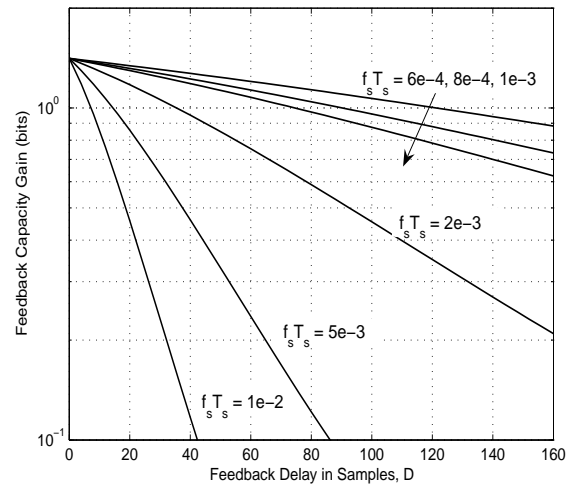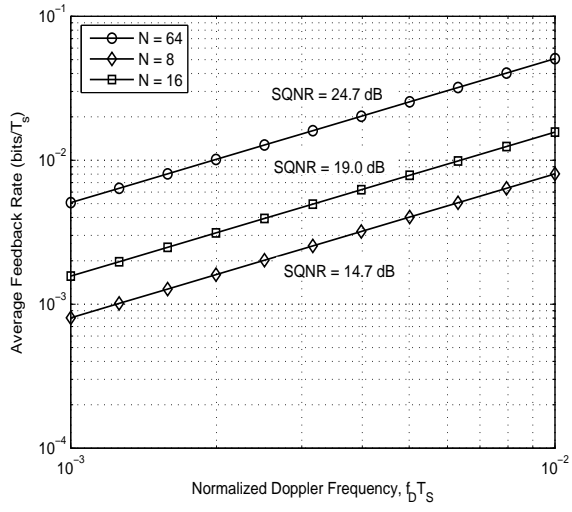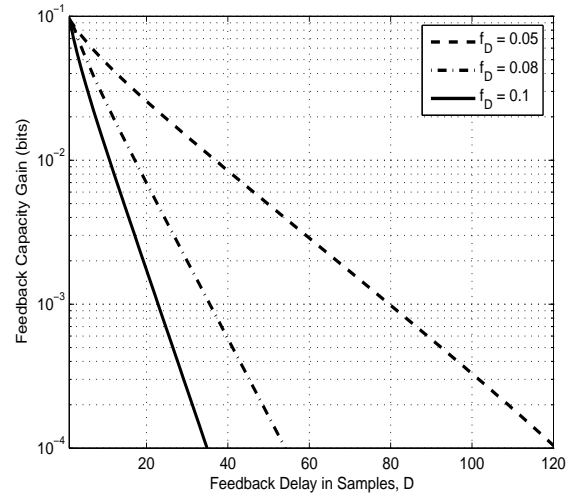Fig. 5. Average feedback rate for the shape feedback system for $L = 2$



Fig. 6. Effect of feedback delay on the feedback capacity gain for the shape feedback system: $L = 2$ and $N = 16$.

Fig. 7. Average feedback rate for the gain feedback system for $L = 4$



Fig. 8. Effect of feedback delay on the feedback capacity gain for the gain feedback system: $L = 2$ and $N = 16$.
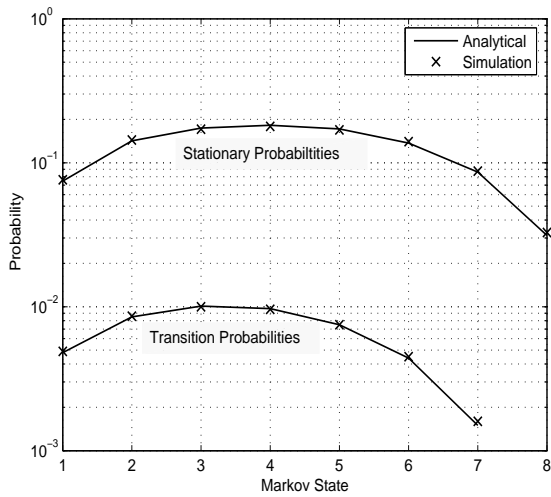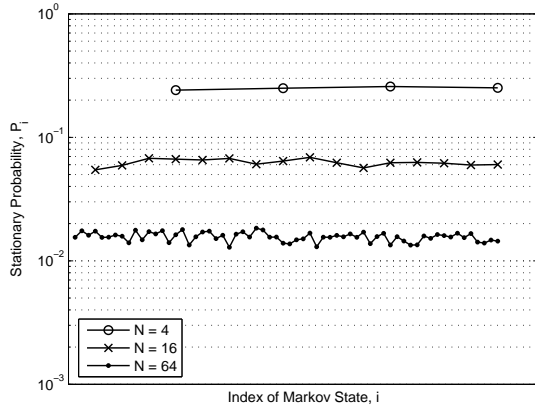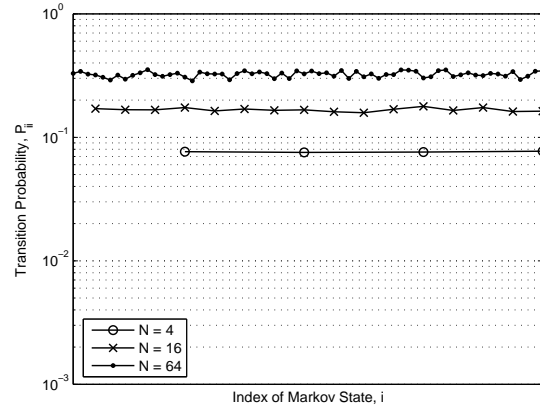


Fig. 9. Transition and stationary probabilities of the gain Markov chain ($L = 2$, $N = 8$, and $f_D T_s = 0.01$)

(a) Stationary Probabilities, $\left\{ P_i^{(s)} \right\}$          (b) Transition Probabilities, $\left\{ P_{ii}^{(s)} \right\}$

Fig. 10.   Stationary probabilities and probabilities of no transition for the shape Markov chain ($L = 2$ and $f_D T_s$=0.02)