Influence of CSI Feedback Delay on Capacity of Linear Multi-User MIMO Systems

Bartosz Mielczarek and Witold A. Krzymień
TRLabs/University of Alberta
Edmonton, Alberta
Canada
Email: {bmielcza, wak}@trlabs.ca

Abstract—In this paper, we discuss the influence of feedback channel delay on throughput of linear multi-user multiple-input multiple-output (MIMO) systems employing vector quantization (VQ) algorithms for encoding channel state information (CSI).

We consider an approach where the mobile receiver estimates the downlink channel and chooses one of the predefined channel characterization codewords whose indices are transmitted back to the base station transmitter. In practice, the feedback channel will introduce delay in transmission of the indices and we evaluate the resulting channel throughput loss for different system setups. Moreover, we define two new parameters of VQ MIMO systems: the channel eigenmode coherence time and singular value coherence time, and discuss their influence on system design.

We demonstrate the performance of two types of systems, one using time division multiplexing of individual mobile users and another transmitting to multiple users at the same time. The simulation results show that feedback delay sensitivity of the multiple-user system throughput is much higher than in the TDM approach. Based on those results we propose a simple method allowing a selection of the VQ resolution for the given channel characteristics and system setup.

Index Terms—Multi-user MIMO systems, channel state information feedback, multi-user diversity.

I. INTRODUCTION

One of the current trends in designing future high capacity wireless systems is the use of multiple antennas [1] to increase spectral efficiency on fading channels. The fundamental issue in multiple-input multiple-output (MIMO) systems is the availability of the channel state information (CSI) at transmitters and receivers. Full CSI at the transmitter (CSIT) allows the use of such techniques as nonlinear or linear precoding, which significantly improve the transmission quality either by increasing the average signal-to-interference-plus-noise ratio (SINR) or by enabling spatial multiplexing.

There are basically two types of methods allowing the transmitter to obtain CSI. The first one exploits reciprocity of the radio channel and assumes that the downlink and uplink channels are almost identical so that the transmitter can use the uplink signal to estimate the downlink channel. The second method uses the CSI feedback link to transmit the information about the downlink channel from the downlink receiver back to the downlink transmitter. In practice, the channel reciprocity method is limited to time-division duplex (TDD) systems and is subject to the ‘ping-pong’ problem, i.e., the downlink transmission must always closely follow the uplink transmission in order to fit within the channel coherence time. If this condition is not met, the downlink transmitter will use outdated CSI, which will lead to performance loss. Moreover, even though mobile receivers may use multiple antennas to receive the signal, they will typically use only one to transmit the signal. This means that channel reciprocity cannot be used at all in such setups and, even though some hybrid solutions may be possible, the explicit CSI feedback approach must be implemented. The problem has attracted attention of the scientific community and papers like [2]–[7] provide solutions for different kinds of systems requiring explicit CSI feedback.

In this paper, we consider only the second approach with the assumption of perfect CSI at the receivers and delayed quantized CSIT. We analyze the system using different values of normalized Doppler frequencies and delays in the feedback channel and evaluate the performance of both time-division multiplexing (TDM) and multiple-user scheduling systems. We also propose two new parameters, the eigenmode coherence time and singular value coherence time, which can be used to quantify the system’s robustness to feedback channel delays.

In Section II, we present the assumed system model and discuss its operation. Section III briefly discusses CSI VQ encoding methods for TDM and multi-user scheduling systems and Section IV analyzes the impact of delayed feedback on performance of closed-loop MIMO systems. We present the performance of the proposed algorithms in Section V, and conclude the paper, together with a discussion of possible directions of future research work, in Section VI.

II. SYSTEM ARCHITECTURE

A. Signal model

We assume that the communication system consists of a transmitter equipped with \( n_T \) antennas and \( K \geq n_T \) mobile receivers with \( n_R(k) \) antennas, where \( k = 1, 2, \ldots, K \). The mobile user channels are modeled by a set of i.i.d. complex Gaussian channel matrices \( \mathbf{H}_k \) of dimension \( n_R(k) \times n_T \).

The received signal of the \( k \)th user is then given by the \( n_R(k) \)-dimensional vector \( \mathbf{y}_k \) defined as

\[
\mathbf{y}_k = \mathbf{H}_k \mathbf{x} + \mathbf{n}_k \tag{1}
\]

Throughout the paper we use the upper-case bold letters to denote matrices and lower-case bold letters to denote column vectors.
where $x$ is the $n_T$-dimensional vector of the transmitted signal and $n_k$ is the $n_R(k)$-dimensional vector containing independent circularly symmetric complex Gaussian entries with zero means and unit variances. All entries in $H_k$ are time-correlated according to the Clark’s model [8]. The frame duration is given as $T_{\text{frame}}$ and we normalize the Doppler frequency as $f_D T_{\text{frame}}$, where $f_D$ is the maximum Doppler shift.

All receivers are assumed to have identical statistical properties, i.e., their path and shadowing losses are identical and known to the transmitter. Since the mean received power at all receivers is identical, we can omit power losses from the large-scale effects, assume that the total transmit and receive power at each transmission instant is equal to $P$ and normalize all entries in the channel matrix to unit variance. The system model covers a wide class of wireless systems and can easily be further expanded to include orthogonal frequency division multiplexing (OFDM) on frequency-selective channels or users with different received powers.

B. System operation

Prior to every transmission epoch, all channel matrices $H_k$ are estimated by receivers. After singular value decomposition (SVD) of $H_k$ at each receiver, the VQ index selector finds the best match of the actual matrix characteristics with the precomputed quantized codewords [7]. In this paper, the measured channel characteristics are the eigenmode vectors and the distribution of singular values. However, any other channel parameters can be easily used instead.

Once the codebook entries are found, their indices are sent back to the transmitter, which uses them to choose from two independent linear modulation codebooks $\hat{B}$ and $\hat{S}$ [7]. The first codebook consists of power-independent $n_T \times n_T$ matrices of eigenmodes, mutually orthogonal in case of single-user transmission and non-orthogonal in case of multi-user transmission. The second codebook is a set of power-dependent non-negative diagonal $n_T \times n_T$ matrices describing power allocation over the eigenmodes and constrained by $\text{Tr} (\hat{S}) = 1$.

The choice of actual modulation matrices $\hat{B}(I_B)$ and $\hat{S}(I_S)$ depends on the indices $I_B$ and $I_S$, which are generated by indexers at the transmitter. In this paper, the indexers are designed to maximize the sum-rate of the system but any other choice of their design criteria is possible.

Once the indexers have received the information from all users and have chosen the optimum modulation format, the transmitted signal $x$ is given as

$$x = \sqrt{P} \hat{B}(I_B) \hat{S}(I_S)^{1/2} \hat{x}$$

(2)

where $\hat{x}$ is the data signal with $E[\hat{x}\hat{x}^H] = I_{n_T}$. The modulated signal is then sent to $n_T$ downlink antennas and transmitted with the total power of $P$.

C. The feedback channel

Transmission of the indices of the CSI codebooks to the transmitter will inevitably suffer from channel disturbances and delays. The first problem is treated in more detail in [9]. In this paper, we concentrate only on the feedback delay effects.

Delays in transmission will occur due to the processing latency at the receivers and finite propagation time of the feedback signal. If the channel changes very fast, the time span over which the chosen CSI codebook indices are valid is very short and the transmitter may be forced to use outdated CSIT. Such a situation results in lower throughput of the system, since the base station may choose suboptimal modulation formats and/or transmit data to users, which already moved to a position characterized by poor propagation conditions.

In this paper, we measure the feedback delay $\Delta$ in number of transmission epochs (frames) between channel estimation at the receiver and using the received VQ indices at the base station. In practical implementations, the transmitter will be able to estimate both the delay and the normalized Doppler frequency, which can be used by channel prediction algorithms and similar solutions not discussed in this paper.

D. Multi-user implementations

In case of multiple user systems, multi-user diversity may be exploited by a simple time-division multiplexing mode (when only one user at a time is given the full bandwidth of the channel) or scheduling the transmission to multiple users at a time. In this paper, we will analyze sensitivity of both approaches to feedback delays.

III. CSI ENCODING IN MIMO SYSTEMS

A. TDM systems

In TDM systems, if the accurate matrix $H_k$ is available at the transmitter, it can adjust the signaling vector $x_k$ to achieve the closed-loop system capacity

$$C = \log_2 \det[I_{n_R(k)} + H_k Q_k H_k^H],$$

(3)

using covariance matrix $Q_k = E[x_k x_k^H] = PB_k S_k B_k^H$ [6], [10]. With such an approach, the transmitter directs all $n_T$ streams of information to one selected user $k$.

As shown in [6], the straightforward approach in TDM MIMO is to let a receiver directly request quantized matrices $\hat{B}(I_B)$ and $\hat{S}(I_S)$ based on estimated $H_k$. Having received indices from all $K$ users, the base station estimates the capacity of their respective channels and, in the next transmission epoch, activates the user with the highest one.

B. Multi-user scheduling systems

When multiple users are scheduled simultaneously, we assume that the base station treats each user as if it was equipped with only one antenna, regardless of the actual number of antennas it may have. With such an approach, the transmitter directs $n_T$ streams of information to $n_T$ selected users in a set $S$. While suboptimal, such a method allows any type of a receiver to work with any base station and greatly reduces the feedback rate burden [7].

To construct the system, we follow the approach of [11], where each user performs singular value decomposition of
$H_k = U_k S_k V_k^H$ and converts its respective $H_k$ to a $n_T$-dimensional vector $h_k = s_k^{\text{max}} v_k^H$ where $s_k^{\text{max}}$ is the largest singular value of $S_k$ and $v_k$ is its corresponding vector from the unitary matrix $V_k$. Based on quantized vectors $h_k$, the base station employs linear block diagonalization approach, attempting to minimize MUI by choosing the proper modulation $B(I_B)$ and power allocation $S(I_S)$ matrices. Having received indices from all $K$ users, the base station estimates the sum-rate of all combinations of $n_T$ users out of $K$ and, in the next transmission epoch, activates the set $S$ providing the highest throughput [7].

IV. ANALYSIS OF DELAYED FEEDBACK EFFECTS

The general effect of the transmitter having delayed channel information is a mismatch between the actual channel characteristics and the modulation matrices used. The extent of such mismatch is dependent on the rate of change of time-correlated matrices $H_k$, but also on the resolution of the VQ codebooks at the receiver. In this paper, we consider a dual codebook system with two independent indices characterizing channel eigenmodes and singular values, represented by the numbers of feedback bits $N_v$ and $N_s$. Hence, there are two parameters that may be used to quantify the severity of the feedback channel delay.

A. Eigenmode coherence time

As mentioned in Section II, the modulation matrices $B$ in TDM systems consist of unit-length vectors, which are mutually orthogonal. The changes of actual channel matrices $B$ will be directly reflected by $I_B$ indices fed to the transmitter from receivers and their characterization is necessary to analyze systems with feedback delay. Fig. 1(a) shows a trajectory in a simple 2x2 MIMO system with time-correlated channel matrices where a pair of orthogonal eigenmodes draw a complicated figure on the surface of unit-radius hemisphere. In case of multiple-user scheduling, similar trajectory behavior can be observed, albeit for only one eigenmode at a time (see Section III). Fig. 1(b) presents an example of such trajectory. In both cases, the extent of eigenmode vector changes will depend on the rate of change of the channel, which is directly influenced by the normalized Doppler frequency.

To fully quantify the effects of feedback delay on VQ CSI systems it is, however, necessary to relate the eigenmode trajectories to the resolution of the quantizer expressed as the number of feedback bits $N_v$. With channel changes caused by fading processes, actual eigenmodes will wander around the reported centroid eigenmode before they leave its assigned Voronoi region and a new index has to be reported to the base station. With larger Voronoi regions (lower resolution), the system will be less susceptible to the feedback delays since the actual eigenmode will stay in the same region longer.

To quantify the system dependence on its resolution and channel change rate, we introduce a parameter called eigenmode coherence time $\tau_{\text{eig}}(N_v)$ defined as the average time it takes an eigenmode vector to leave a multi-dimensional Voronoi region for a given normalized Doppler frequency. Its dependence on the number of VQ bits is monotonically decreasing - with smaller regions, it will take less time before the vector leaves its current region. This can be seen in Fig. 2 where a 2D VQ tessellation of an eigenmode hemisphere surface is shown for two resolutions of a VQ quantizer. If the identical channel eigenmode trajectory is projected over the surface of both quantizers, one can see that with higher VQ resolution, the changes of centroid indices will be frequent as the eigenmodes will wander into different Voronoi regions faster. On the other hand, with increasing resolution, the distances between the adjacent centroids will become smaller, and the losses of performance due to a mismatch of indices will diminish. Thus, a VQ MIMO system design must include a trade-off between a theoretical system performance and its robustness to feedback channel delays. As shown in [7], there exist special techniques, which alleviate the problem, but we do not discuss them in this paper.
B. Singular value coherence time

Similarly to eigenmode coherence time, we define singular value coherence time, \( \tau_{\sin}(N_s) \) as the average time it takes a singular value representation to leave a one-dimensional Voronoi region for a given normalized Doppler frequency. The rationale behind separating these two related parameters lies in the fact that if \( \tau_{\text{eig}}(N_v) \) and \( \tau_{\sin}(N_s) \) differ, the frequency of change of indices \( I_B \) and \( I_S \) will not be the same and hence, design of the feedback link will have to take it into consideration.

Fig. 3 presents the curves of \( \tau_{\text{eig}}(N_v) \) and \( \tau_{\sin}(N_s) \) for different values of \( N_v \) and \( N_s \) and normalized Doppler frequencies \( f_D T_{\text{frame}} \). As one can see, both coherence times decrease rapidly with the increasing rate of feedback and normalized Doppler frequency. Keeping in mind that the typical feedback delay in current systems is in the range of 3-5 frames, it can be thus difficult to provide the base station with good quality channel estimates. Note that \( \tau_{\sin}(N_s) < \tau_{\text{eig}}(N_v) \) for \( N_v = N_s \). However, as shown in [6], [7], the resolution of a singular value quantizer can be an order of magnitude lower than the resolution of eigenmode quantizer. Hence, in a typical system, a rate of change of index \( I_S \) will be much lower than a rate of change of \( I_B \) and we can assume that a transmitter will have much better knowledge about singular value distribution of the receiver channels than about their eigenmodes.

V. SIMULATION RESULTS

We have implemented our system using a base station with \( n_T = 2 \) and a set of \( K = 2 \) or 5 mobile receivers with identical statistical properties and \( n_R(k) = n_T = 2 \). We have tested the systems for different numbers of feedback bits \( N_v \) and delays \( \Delta \) for a given normalized Doppler frequency \( f_D T_{\text{frame}} = 0.01 \). Moreover, we assumed that \( N_v \gg N_s \) resulting in \( \tau_{\sin}(N_s) \gg \tau_{\text{eig}}(N_v) \), which allowed us to only model the effects of eigenmode index delays (we assume that the base station has access to non-delayed singular value indices). Finally, each system setup has been simulated using 100,000 time-correlated channel realizations.

A. TDM systems

Fig. 4 shows the curves representing TDM system throughput with varying numbers of \( N_v \) feedback bits and delays \( \Delta \). We compare these results with the optimum full CSIT and no delay system (Opt). As one can see, only relatively large delays will cause some penalty in throughput performance but, in general, TDM systems are quite robust to feedback channel latency. Fig. 5 confirms this statement, showing that even for \( \Delta = 10 \), throughput losses are still marginal. For low values of \( \Delta \), increasing the number of feedback bits increases the throughput, while for \( \Delta = 50 \), there is a clear maximum where improved quality of modulation is offset by decreased eigenmode coherence time and rapid changes of indices effectively cancel the effects of higher VQ resolution.

B. Multi-user scheduling systems

Figs. 6 and 7 show similar type of analysis for multi-user systems, where we compare the VQ systems with the
feedback resolution brings minimal gains. Points of throughput curves, beyond which further increase of user systems, one can use the graphs in Fig. 7 to find saturation to select a number of feedback bits in different setups of multi-user systems, one can use the graphs in Fig. 7 to find saturation.

Losses in throughput due to lower eigenmode coherence times. Large feedback rates are shown to cause much more significant losses in throughput due to lower eigenmode coherence times. This can be clearly seen in Fig. 6, where the increasing delays with time is much lower than in the TDM systems (see Fig. 3). This is much lower than in the TDM systems (see Fig. 3).

VQ quantizers is used and the resulting eigenmode coherence delay is present and shown simple techniques allowing the suboptimum choice of VQ eigenmode codebook resolution. Our future work will be focused on the systematic approach to codebook design and development of eigenmode vector prediction algorithms allowing reduction of feedback resolution.

VI. CONCLUSIONS AND FUTURE WORK

We have considered multi-antenna, multi-user systems operating on a flat fading channel with imperfect CSIT. Based on earlier proposed vector-quantization approaches to CSI encoding, we investigated effects of the delayed feedback on overall system throughput. Moreover, we have proposed new tools for CSI VQ system analysis when the the feedback delay is present and shown simple techniques allowing the suboptimum choice of VQ eigenmode codebook resolution. Our future work will be focused on the systematic approach to codebook design and development of eigenmode vector prediction algorithms allowing reduction of feedback resolution.

ACKNOWLEDGMENT

Funding for this work has been provided by TRLabs, the Rohit Sharma Professorship and the Natural Sciences and Engineering Research Council (NSERC) of Canada.

REFERENCES