Precoding Performance with Codebook Feedback in a MIMO-OFDM System

Min Zhang *, Mansoor Shafi †, Peter J. Smith ‡, and Pawel A. Dmochowski §

* Alcatel-Lucent, New Zealand
† Telecom New Zealand, New Zealand
‡ Department of Electrical and Computer Engineering, University of Canterbury, New Zealand
§ School of Engineering and Computer Science, Victoria University of Wellington, New Zealand

Email: min.hs.zhang@alcatel-lucent, mansoor.shafi@telecom.co.nz, p.smith@elec.canterbury.ac.nz, pdmochowski@ieee.org

Abstract—Limited feedback precoding is part of the LTE standard. Despite standardization, important fundamental questions, especially relating to the performance due to the use of codebooks and receiver processing techniques, remain to be explored. In order to understand these questions, we consider a single user in a single cell employing single or multi-stream transmission using a variety of codebooks and a choice of different receiver types. We derive expressions for capacity loss relative to perfect feedback due to the limited size of codebook. When multi-stream transmission is deployed, we show that codebook feedback manifests itself as inter-stream interference resulting in a capacity loss for all receiver types. In the case of SVD receivers, this interference results in a capacity floor. We define a precoding matrix index (PMI) coherence time and bandwidth and show how these parameters are respectively related to channel coherence time and bandwidth. The PMI coherence parameters shed new light on how often feedback is required, both in time and frequency domains, and help us to determine the capacity penalty if feedback is delayed beyond the coherence parameters. Finally, we show the distribution of PMI and also show that this is environment dependent. This supports the need for codebooks that are adaptable to different environments and are not strictly tied to i.i.d. channels.

I. INTRODUCTION

Precoding at the transmitter (Tx) requires knowledge of the multiple-input-multiple-output (MIMO) channel. Transmitting full channel state information (CSI) back to the Tx involves a significant overhead, thus necessitating limited feedback. Indeed, limited feedback precoding is now a topic of large interest [1], [2] and forms a part of the LTE standard [3]. The receiver (Rx) sends a precoding matrix index (PMI) to the Tx. Both Tx and Rx share a common look up table (codebook) and use the PMI to identify the precoding matrix (PM) [3], [4]. The selected PMs are generally limited feedback versions of the right singular vectors of the MIMO channel. There are many criteria for the design of codebooks [5] and for the selection criteria of PM. In this paper we have chosen two kinds of codebooks: a standardized LTE codebook with 16 PMs [3], and a Grassmannian codebook with 16 or 64 PMs [6]. The PMs are selected according to various types of selection criteria, e.g. minimal subspace angle [1], [2] (often referred to as minimum Chordal distance), maximum capacity or simply random selection.

We use the ITU standardized M.2135 channel model [7] for the Urban and Rural Macro environments to simulate the necessary channel profiles. The urban area has large delay spread but low mobility, whereas the rural area has low delay spread but high mobility. Therefore, these two environments represent two ends of the multipath spectrum in terms of channel coherence bandwidth and time.

In order to perfectly construct the transmit streams (referred to as layers in LTE) exact knowledge of CSI is required at the Tx. Regardless of the codebook type used, the limited number of PMs in a codebook results in imperfect CSI. This leads to leakage between the transmit streams causing inter-stream interference for a given user. Consequently, the constraint of limited feedback leads to a capacity loss (CL). In this paper we address the following questions:

• What is the CL and how is it sensitive to codebook design, selection criteria and receiver type?
• Is there a PMI coherence time and coherence bandwidth and how often should the PMI be updated in time and in frequency so that the CL is not further degraded?
• Are the PMI coherence parameters related to channel parameters such as delay spread and Doppler frequency?
• Are all PMs selected uniformly or do some have a higher likelihood of selection than others. What insights does this distribution provide into codebook design?

The above questions are fundamental to determine the performance of limited feedback precoding whether it is applied to single or multiple users in a single or multiple cells [8]. In order to fully understand these questions it is best to consider the simple case of a single user in a single cell with single or multi-stream (multi-layer) transmission.

This paper is organized as follows. Section II describes the system model and gives an analytical framework for the simulations to follow. The expressions for CL due to imperfect feedback for various receiver types are also derived here. Section III presents the simulation results and key findings of the paper. Finally, conclusions are given in Section IV.

II. MIMO CHANNEL WITH LIMITED FEEDBACK

We consider a single-user MIMO-OFDM feedback system where the transmitter uses an orthogonal precoding matrix to assist with the performance of a Minimum Mean Square Error (MMSE) receiver, a Zero-Forcing (ZF) receiver or with a Singular Vector Decomposition (SVD) receiver. These receivers
are being considered in the LTE-Advanced standard. In this section, we define the system, and derive expressions for the CL due to the utilization of a codebook.

A. MIMO-OFDM System Model

Consider a system with \(N_T\) transmit antennas, \(N_R\) receive antennas, and employing \(N\) sub-carriers. For notational convenience we describe the system for a single sub-carrier and omit any subscripts which identify the particular sub-carrier. Hence, for a generic sub-carrier we have a flat fading channel matrix, \(\mathbf{H}\) with dimension \(N_R \times N_T\). For ease of exposition we assume \(N_T \geq N_R\). The singular value decomposition of \(\mathbf{H}\) gives \(\mathbf{H} = \mathbf{UDV}^\dagger\) where \(\mathbf{U}, \mathbf{V}\) are unitary, the diagonal entries of \(\mathbf{D}\) are denoted by \(\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_{N_T}\) and \(\lambda_i = \sigma_i^2, i = 1, 2, \ldots, N_T\) are the eigenvalues of \(\mathbf{H}^\dagger\mathbf{H}\) where \(\dagger\) denotes the Hermitian transpose. The system equation is

\[
\mathbf{r} = \mathbf{Hs} + \mathbf{n},
\]

where \(\mathbf{r}\) is the \(N_R \times 1\) received signal vector, \(\mathbf{s}\) is the \(N_T \times 1\) transmitted signal and \(\mathbf{n}\) is an \(N_R \times 1\) vector of independent and identically distributed (i.i.d.) \(CN(0, \sigma^2)\) noise terms. If \(L \leq N_T\) streams of data are used then the \(N_T \times L\) precoding matrix, \(\mathbf{P}\), is applied to the original \(L \times 1\) data vector, \(\mathbf{b}\), so that \(\mathbf{s} = \mathbf{Pb}\) and equation (1) becomes

\[
\mathbf{r} = \mathbf{HPb} + \mathbf{n}.
\]

We assume unitary precoding matrices so that \(\mathbf{P}^\dagger\mathbf{P} = \mathbf{I}_L\). The total transmit power is \(P_T = \mathbb{E}(\mathbf{s}\mathbf{s}^\dagger) = \mathbb{E}(\mathbf{b}\mathbf{b}^\dagger)\) so that \(\mathbb{E}(|\mathbf{s}_i|^2) = P_T/L\) and we assume a zero mean i.i.d. structure for \(\mathbf{b}\). The noise power is \(\mathbb{E}(|\mathbf{n}_i|^2) = \sigma^2\) so the link SNR is \(\rho = P_T/\sigma^2\).

We assume that \(\mathbf{v}(i)\) denotes the \(i\)th column of \(\mathbf{V}\), and \(\mathbf{V}_1 = [\mathbf{v}(1) \ \mathbf{v}(2) \ \ldots \ \mathbf{v}(L)]\) is a sub-matrix of \(\mathbf{V}\) containing the first \(L\) columns. The column vectors \(\mathbf{p}(i)\) and \(\mathbf{v}_c(i)\) are defined as the \(i\)th column of \(\mathbf{P}\) and \(\mathbf{V}_c\), respectively, where the \(N_T \times L\) matrix \(\mathbf{V}_c\) is a quantized version of \(\mathbf{V}_1\).

B. Receiver Structures

All three receiver types perform linear combining where the output of the combiner is \(\bar{\mathbf{r}} = \mathbf{W}^\dagger \mathbf{r}\) and the SINR of the \(k\)th stream is denoted \(\text{SINR}_k\). Using standard results from the literature we have:

ZF Receiver

\[
\mathbf{W}_{ZF} = \mathbf{H}\mathbf{P}^\dagger \left(\mathbf{P}^\dagger\mathbf{H}\mathbf{P}\right)^{-1}
\]

\[
\text{SINR}_k = \frac{\rho}{L[\mathbf{P}^\dagger\mathbf{H}\mathbf{P}]^{-1})_{kk}}
\]

MMSE Receiver

\[
\mathbf{W}_{MMSE} = \mathbf{H}\left(\frac{L}{\rho} \mathbf{I} + \mathbf{P}^\dagger\mathbf{H}\mathbf{P}\right)^{-1}
\]

\[
\text{SINR}_k = \frac{\lambda_i}{L[\mathbf{P}^\dagger\mathbf{H}\mathbf{P}]^{-1})_{kk}}
\]

where \(\lambda_i\) denotes the desired channel, is the \(k\)th column of \(\mathbf{H}\) and \(\mathbf{H}_f\), the interfering matrix, is the matrix \(\mathbf{HP}\) with the \(k\)th column removed.

SVD Receiver

\[
\mathbf{W}_{SVD} = \mathbf{U}
\]

\[
\text{SINR}_k = \frac{|\mathbf{v}(k)^\dagger\mathbf{p}(k)|^2}{\sum_{j=1,j\neq k}^{L}|\mathbf{v}(k)^\dagger\mathbf{p}(j)|^2 + \frac{L}{\rho\lambda_k} \text{noise}}
\]

The capacity per subcarrier for the three receiving techniques under inter-stream interference is given by:

\[
C = \sum_{k=1}^{L} \log_2 (1 + \text{SINR}_k)
\]

for a MIMO system using \(L\) stream (layer) transmission and SNR \(\rho\). For reasons of space, the proof that \(\text{SINR}_k = \lambda_k\rho/L\) for ZF, MMSE and SVD receivers using \(\mathbf{P} = \mathbf{V}_1\) is omitted.

With limited feedback, \(\mathbf{P} = \mathbf{V}_1\) is not possible in reality. An alternative PM, \(\mathbf{V}_c\), is used where \(\mathbf{V}_c\) is one of the candidate PMs in a codebook. Therefore, only the PMI of \(\mathbf{V}_c\) in such a codebook is fed back to the Tx. By substituting \(\mathbf{P} = \mathbf{V}_c\) into (4), (6) and (8) numerical results for the SINR are obtained and hence the capacity follows from (II-B).

We define the CL, \(C_{LOSS}\), as the capacity difference between the capacity in (II-B) with perfect feedback, \(\mathbf{P} = \mathbf{V}_1\), and the capacity with quantized PM feedback, \(\mathbf{P} = \mathbf{V}_c \approx \mathbf{V}_1\).

Hence,

\[
C_{LOSS} = C|\mathbf{P} = \mathbf{V}_1 - C|\mathbf{P} = \mathbf{V}_c
\]

An approximate expression for capacity loss at high SNR for an SVD receiver can be derived as:

\[
C_{LOSS, SVD} \approx \sum_{k=1}^{L} \log_2 \left(\frac{\rho\lambda_k/L + 1}{\rho\lambda_k + |\mathbf{v}(k)^\dagger\mathbf{v}(c)|^2}\right)
\]

where \(\rho\lambda_k = \sum_{j=1,j\neq k}^{L}|\mathbf{v}(j)^\dagger\mathbf{v}(c)|^2\) is defined as the inter-stream interference to layer \(k\) caused by the mismatch between \(\mathbf{V}_c\) and \(\mathbf{V}_1\). Due to space limitation, we omit the derivation. For single layer transmission with \(L = 1\), there is no inter-stream interference so that \(\rho\lambda_k = 0\). Therefore, the CL of an SVD receiver can be approximated by

\[
C_{LOSS, SVD} \approx \log_2 |\mathbf{v}(1)^\dagger\mathbf{v}(c)|^2,
\]

which is always non-negative. For multiple layer transmission, \(L > 1\), the inter-stream interference is the dominant limiting factor so that the CL increases at higher SNR. Consequently, the CL for \(L\) stream transmission can be approximated from
\( C_{LOSS, SVD} \approx L \log_2(p), \) \hfill (14)
which clearly shows the increased CL at high SNR.

At high SNR, the performance of ZF and MMSE receivers is the same and consequently both receiving techniques lead to a similar CL. To investigate this loss, we substitute (4) into (11) for large SNR with \( N_T \) layer transmission and obtain the following upper bound:

\[
C_{LOSS,ZF/MMSE} \leq N_T \times \\
\log_2 \sum_{k=1}^{N_T} \left( \sum_{j=1}^{N_T} \left| \frac{\lambda_k}{\lambda_j} v_c(j)^\dagger v(k) \right|^2 - \log_2 N_T \right)
\]

(15)

Details of the derivation are omitted because of limited space. With perfect codebook feedback, \( V_c = V_1 = V \), the upper bound in (15) equals zero. The upper bound also shows that the CL of ZF and MMSE receivers is independent of SNR. However, the upper bound is related to both specific codebook and MIMO channel characteristics. Taking the expectation of (15) leads to an upper bound on the ergodic capacity loss (ECL). Simulations shown in Section III indicate that this ECL is also an upper bound in \( L \leq N_T \) layer transmission. However at present this remains a conjecture.

\section{C. Codebook Selection}

In this paper, we consider LTE [3] and Grassmannian codebooks [6], and three standard methods of selection criteria.

1) Minimum Subspace Angle (MSA): \( V_c \) is chosen as the best \( \text{PM} \) that satisfies following condition [2]:

\[
\min_{V_c \in W} \left\{ L - \sum_{k=1}^{L} |v(k)^\dagger v_c(k)|^2 \right\}
\]

(16)

where \( W \) is the codebook.

2) Maximum Capacity (MC): \( V_c \) is chosen as the best \( \text{PM} \) that maximizes equation (II-B) over all possible PMs in the codebook \( W \). This is the optimal approach but it requires more calculations than MSA since, for each candidate matrix, all \( L \) SINRs must be computed from (4), (6) and (8). Hence, for a codebook of size \( M \), \( LM \) matrix inversions are required in both ZF and MMSE receivers.

3) Random (RND): \( V_c \) is chosen randomly so that a lower bound on codebook performance can be obtained.

\section{III. Simulation Results}

We present simulation results here for the 2.6 GHz band using the M.2135 channel model, which is designed for the evaluation of radio interface technologies for IMT-Advanced [7] by the ITU-R. Channel profiles for the urban macrocell (UMa) and rural macrocell (RMa) environments are simulated using this model. Details of the simulation parameter settings are given in Table I. Each drop of the M.2135 channel is converted into the frequency domain by FFT. The bandwidth of each subcarrier in the OFDM system is 15 KHz as specified in the LTE and LTE-Advanced standards. The total power of the MIMO channels for all subcarriers and OFDM symbols in one drop is normalized such that the average power equals \( N_T \times N_R \) over both time and frequency.

During simulation, the transmission power is fixed at the Tx for all multi-stream transmissions. The MIMO CSI is assumed to be perfectly known by the MS and only the PMI is fed back to the Tx. As a comparison, the ideal case of perfect feedback is also considered here where the PM, \( P = V_1 \), is fed back to the Tx. Then the ergodic capacity (EC) is obtained by averaging instantaneous capacity from equation (II-B) over a large number M.2135 channel drops. The notation \( L1, L2, L3 \) and \( L4 \) refers to one, two, three and four stream transmissions with a constant transmission power. For simplicity, no power control or water filling techniques have been used.

\subsection{A. Capacity Loss}

Figure 1 shows the performance of ZF, MMSE and SVD receivers. Layer 1 (L1) corresponds to a single stream for each subcarrier and OFDM symbol, and layer 4 (L4) corresponds to 4 streams. It is shown in Fig. 1 that the LTE codebook leads to a CL, compared to perfect feedback. The CL is significant for L4, but nominal for L1. The MMSE receiver is always the best of the three receiving methods. Moreover, it is also shown in Fig. 1 that L1 transmission performs much better than L4 transmission within certain SNR ranges.
The break point between L1 and L4 performance will be highly related to the specific receiving strategy and channel characteristics. High layer transmission leads to destructive inter-stream interference, resulting in a significant CL. For example, at 20 dB SNR the EC with L4 transmission and the LTE codebook can only achieve 67%, 47% and 17% of the EC with perfect feedback for MMSE, ZF and SVD receivers respectively. The performance of MMSE and ZF receivers merge at high SNR, but this convergence can only be seen for the extremely high SNR range.

Figure 2 compares MIMO EC for different numbers of layers in the high SNR range. Although the values of high SNR are not practical, Fig. 2 shows that the EC with the LTE codebook is eventually ordered by the number of layers. For example, L4 transmission outperforms L3 transmission but only when the SNR is greater than 60dB. Similarly, L4 outperforms L2 transmission and L1 transmission when the SNR is greater than 40dB and 15dB respectively. These extremely large SNR values required to obtain the benefit of multi stream/layer transmission indicate a low opportunity of using high layer transmission in practice and suggests that single layer transmission will be a more common strategy.

Moreover, it is also shown in Fig. 2 that the ECL of the MMSE receiver is roughly constant and ordered by the number of layers for high SNR. For example, the ECL of the MMSE receiver at high SNR for layer 1 to 4 transmissions are roughly 0 bps/Hz, 2 bps/Hz, 6 bps/Hz and 12 bps/Hz respectively.

Figure 3 compares different codebook selection criteria using the LTE codebook and MMSE receiver. The EC difference between MC and MSA is marginal, especially for L1. This is an important result since MSA involves much less computation than the optimal MC approach. However, the criterion of RND, which is the worst selection method, shows a large degradation. It is also interesting to observe that L1 transmission suffers more than L4 transmission when RND is used. A possible explanation is that the performance of high layer transmission is mostly limited by inter-stream interference. Hence, the L4 CL is less sensitive to the "right" PM decision since all PMs from a codebook might lead to similar but dominant inter-stream interference.

In Fig. 4, we wish to determine the impact of size of codebook and also codebook design on the EC. LTE codebook is only defined with 16 PMs for 4 antennas [3]. The Grassmannian codebooks proposed in [6], [9] with 16 and 64 PMs are compared in Fig. 4. Quadrupling the size of the codebook represents a 50% increase in the feedback rate. Intuitively this should improve the performance [10]. However, in Fig. 4 increasing the size of the codebook shows a negligible performance gain for L2 transmission with ZF and MMSE decoding. On the other hand there is a small improvement for SVD decoding whilst increasing size of Grassmannian code from 16 to 64 since the EC is severely limited by inter-stream interference here. This is because the compared codebook design here is possibly for an I.I.D. channel and does not account for channel correlation. The later has dominant impact on the EC than increasing the number of feedback bits. It is also interesting to see that the Grassmannian codebook with
16 PMs is worse than the existing LTE codebook if the SVD receiver is applied. Note the EC values for LTE with ZF and MMSE are indistinguishable from Grassmannian equivalents.

B. PMI Coherence Time and Bandwidth

The PMI must be independently updated in time and frequency domain. We wish to know if this updating should be done at least as often as channel coherence time and bandwidth respectively or the nearest values constrained by simulations. We also would like to know the penalty on EC if the updating period and bandwidth exceed corresponding channel values.

Given a certain selection criterion, the PMI is determined by the instantaneous MIMO channel realization for every subcarrier, and also every OFDM symbol. The PMI is then fed back to the Tx. We also need to define PMI coherence parameters in time and frequency domains and see if there are related to channel equivalents. The PMI coherence time is defined as the total simulation time $T$ divided by the number of PMI updates required in the time domain ($T$ is chosen as long time horizon). The PMI coherence bandwidth is defined as the total bandwidth divided by the number of PMI updates required in the frequency domain.

From Fig. 5, 6 and Table I we observe that channels with higher channel coherence time and bandwidth also have higher PMI coherence time and bandwidth respectively. However there is no simple relationship to determine the PMI coherence values given the corresponding channel coherence parameters. We also observe that PM coherence parameters are layer-dependent. Figures 5 and 6 show a decrease in PMI coherence time and PMI coherence bandwidth with an increasing number of transmission layers. Comparing the MC and MSA selection criteria, we observe that MC requires a higher feedback rate for both time and frequency domains, especially for L4 transmission. The increased PMI selectivity of MC may be due to the fact that the MC selects a PM, $V_c$, on the basis of maximizing instantaneous capacity. From the exact formulation in (6) and (II-B) and, more easily, from (15) we see that all the inter-stream interference terms $|v_c(j)^\dagger v(k)|$ for $k, j \in [1, \cdots, N_T]$ and $j \neq k$ play an important role in MC. Hence, MC is roughly equivalent to minimizing inter-stream interference, whilst MSA only deals with the projection terms $\sum_{k=1}^{N_T} |v_c(k)^\dagger v(k)|^2$ on a column-by-column basis.

In order to observe the capacity penalty when reducing feedback rate, the PMI update is fixed for a period of time or a particular bandwidth. In Fig. 7, each PMI is based on the instantaneous CSI from the first OFDM symbol, and then remains static for a time period as specified on the x-axis (PMI feedback time interval). The PMI is still updated for each individual subcarrier with 15kHz frequency interval. The ECs are obtained by varying PMI feedback time interval, and then normalized relative to the largest EC, which represents the finest PMI update within each individual subcarrier and individual OFDM symbol. Similarly, simulations in Fig. 8 vary PMI feedback bandwidth interval in the frequency domain, whilst the PMI is updated for each individual OFDM symbol. Both Fig. 7 and Fig. 8 show the capacity penalty as a result of either reducing the feedback rate in time, equivalent to increasing PMI feedback time interval, or reducing the feed-
This paper has discussed the performance of limited feedback precoding using various codebooks, codebook selection criteria and receiver types. The results indicate that imperfect PMs result in a significant CL that is more prominent when multi layer transmission is invoked. The CL is present for all decoder types. We have derived explicit expressions for the CL. In the case of SVD receivers, we have demonstrated a capacity floor that arises due to irreducible interference between the different layers. We show that increasing the codebook size from 16 to 64 does not have a noticeable impact on EC as the underlying codebooks are not designed for correlated channels and instead designed for an i.i.d. channel. This is further re-enforced by the observation that only some of PMs in the codebooks are selected with high probability. We simulate various PMI feedback times and bandwidth intervals resulting in a capacity penalty, which is useful in determining the rate of feedback in time and frequency. Even though our results are for a single user case, we believe they provide key insights into the performance of limited feedback precoding for the multi-user case also where each user is given single or multiple streams.

IV. CONCLUSIONS

Figure 9 clearly shows that such a design does not work very well for a correlated MIMO channel.

REFERENCES