A differential Feedback Scheme Exploiting the Temporal and Spectral Correlation

Mingxin Zhou, Leiming Zhang, Lingyang Song, Mérouane Debbah

To cite this version:

Mingxin Zhou, Leiming Zhang, Lingyang Song, Mérouane Debbah. A differential Feedback Scheme Exploiting the Temporal and Spectral Correlation. IEEE Transactions on Vehicular Technology, Institute of Electrical and Electronics Engineers, 2013, pp.4701-4707. hal-00932863

HAL Id: hal-00932863

https://hal-supelec.archives-ouvertes.fr/hal-00932863

Submitted on 17 Jan 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
A Differential Feedback Scheme Exploiting the Temporal and Spectral Correlation

Minxin Zhou, Leiming Zhang, Member, IEEE, Lingyang Song, Senior Member, IEEE, and Merouane Debbah, Senior Member, IEEE,

Abstract—Channel state information (CSI) provided by limited feedback channel can be utilized to increase the system throughput. However, in multiple input multiple output (MIMO) systems, the signaling overhead realizing this CSI feedback can be quite large, while the capacity of the uplink feedback channel is typically limited. Hence, it is crucial to reduce the amount of feedback bits. Prior work on limited feedback compression commonly adopted the block fading channel model where only temporal or spectral correlation in wireless channel is considered. In this paper, we propose a differential feedback scheme with full use of the temporal and spectral correlations to reduce the feedback load. Then, the minimal differential feedback rate over MIMO time-frequency (or doubly) selective fading channel is investigated. Finally, the analysis is verified by simulation results.

Index Terms—Differential feedback, correlation, MIMO

I. INTRODUCTION

In multiple input and multiple output (MIMO) systems, channel adaptive techniques (e.g., water-filling, interference alignment, beamforming, etc.) can enhance the spectral efficiency or the capacity of the system. However, these channel adaptive techniques require accurate channel conditions, often referred to channel state information (CSI). Oftentimes, in a Frequency-Division Duplex (FDD) setting, CSI is estimated at the receiver and conveyed to the transmitter via a feedback channel. In recent years, CSI feedback problems have been intensively studied, due to its potential benefits to the MIMO systems [1], [2]. It is significant to explore how to reduce the feedback load, due to the uplink feedback channel limitation.

In [3], four feedback rate reduction approaches were reviewed, where the lossy compression using the properties of the fading process was considered best. When the wireless channel experiences temporal-correlated fading, modeled as a finite-state Markov chain, the amount of CSI feedback bits can be reduced by ignoring the states occurring with small probabilities [4]–[8]. The feedback rate in frequency-selective fading channels was studied in [9], [10], by exploiting the frequency correlation.

In summary, all the above works mainly focus on feedback rate compression considering either temporal correlation or spectral correlation. However, doubly selective fading channels are more frequently encountered in wireless communications as the desired data rate and mobility grow simultaneously. To the best knowledge of the authors, the scheme of making full use of the two-dimensional correlations is not yet well studied. Using both of the orthogonal dimensional correlations in a cooperated way, the feedback overhead can be further reduced in the doubly selective fading channels. Thus, in this paper, we derive the minimal feedback rate using both the temporal and spectral correlations.

The main contributions of the present paper can be briefly summarized as: 1) We discuss the minimal feedback rate without differential feedback. 2) We propose a differential feedback scheme by exploiting the temporal and spectral correlations, and 3) We derive the minimal differential feedback rate expression over MIMO doubly selective fading channel.

The rest of the paper is organized as follows: In Section II, we describe the differential feedback model as well as the statistics of the doubly selective fading channel. In Section III, we propose a differential feedback scheme by exploiting the two-dimensional correlations and derive the minimal feedback rate. In Section IV, we provide some simulation results showing the performance of the proposed scheme.

II. SYSTEM MODEL

In this paper, we assume that the down-link channel is a mobile wireless channel which is always correlated in time and frequency domains, while the up-link channel is a limited feedback channel.

A. Statistics of the down-link channel

Since the channel corresponding to each antenna is independent and with the same statistics, we can describe the separation property of the channel frequency response \( H(t, f) \) at time \( t \) for an arbitrary transmit and receive antenna pair \( \{1, 2\} \)

\[
r_H(\Delta t, \Delta f) = \mathbb{E} \left\{ H(t + \Delta t, f + \Delta f) H^*(t, f) \right\} \\
= \sigma_H^2 r_f(\Delta t) r_f(\Delta f), \tag{1}
\]
where $\mathbb{E}\{\cdot\}$ denotes expectation function, the superscript $(\cdot)^*$ denotes complex conjugate. $\sigma_H^2$ is the power of the channel frequency response. $r_t(\Delta t)$ and $r_f(\Delta f)$ denotes the temporal and spectral correlation functions, respectively.

Assuming that the channel frequency response stays constant within the symbol period $t_s$ and the subchannel spacing $f_s$, the correlation function for different periods and subchannels is written as

$$r_H[\Delta m, \Delta n] = \sigma_H^2 r_f[\Delta m] r_f[\Delta n], \quad (2)$$

where $r_t[\Delta m] = r_t(\Delta mt_s)$ and $r_f[\Delta n] = r_f(\Delta nf_s)$.

Furthermore, if we just consider the time domain, the correlated channel can be modeled as a time-domain first-order autoregressive process (AR1) [4]

$$H_{m,n} = \alpha_t H_{m-1,n} + \sqrt{1 - \alpha_t^2} W_t, \quad (3)$$

where $H_{m,n}$ denotes the channel coefficient of the $m$th symbol interval and the $n$th subchannel, $W_t$ is a complex white noise variable, which is independent of $H_{m-1,n}$, with variance $\sigma_H^2$. The parameter $\alpha_t$ is the time autocorrelation coefficient, which is given by the zero-order Bessel function of first kind $\alpha_t = r_t[1] = J_0(2\pi f_d t_s)$, where $f_d$ is the Doppler frequency [12].

Similarly, if we just consider the frequency domain, the correlated channel can also be represented as a frequency-domain AR1 [9]

$$H_{m,n} = \alpha_f H_{m,n-1} + \sqrt{1 - \alpha_f^2} W_f, \quad (4)$$

where $W_f$ is a complex white noise variable, which is independent of $H_{m,n-1}$, with variance $\sigma_H^2$. The parameter $\alpha_f$ determines the correlation between the subchannels, which is given by $\alpha_f = r_f[1] = \frac{1}{\sqrt{1 + (2\pi f_d \Delta)^2}}$, where $\Delta$ is the root mean square delay spread [12].

**B. Differential Feedback Model**

The system model with differential feedback is illustrated in Fig. 1. By using differential feedback scheme, the receiver just feeds back the differential CSI.

![Fig. 1. System model of the differential feedback over MIMO doubly selective fading channel](image)

We suppose that there are $N_t$ and $N_r$ antennas at the transmitter and receiver, respectively. The received signal vector at the $n$th symbol interval and the $n$th subchannel is given by

$$y_{m,n} = H_{m,n} x_{m,n} + n_{m,n}. \quad (5)$$

In the above expression, $y_{m,n}$ denotes the $N_r \times 1$ received vector at the $n$th symbol interval and the $n$th subchannel. $H_{m,n}$, a $N_r \times N_t$ channel fading matrix, is the frequency response of the channel. The entries are assumed independent and identically distributed (i.i.d.) obeying a complex Gaussian distribution with zero-mean and variance $\sigma_H^2$. Different antennas have the same characteristic in temporal and spectral correlations, $\alpha_t$ and $\alpha_f$, respectively. Besides, there is no spatial correlation between different antennas. $x_{m,n}$ denotes the $N_t \times 1$ transmitter signal vector and is assumed to have unit variance. $n_{m,n}$ is a $N_r \times 1$ additive white Gaussian noise (AWGN) vector with zero-mean and variance $\sigma_e^2$. Both $x_{m,n}$ and $n_{m,n}$ are independent for different $m$'s and $n$'s.

Through CSI quantization, the feedback channel output is written as [13]–[15]

$$H_{m,n} = \bar{H}_{m,n} + E_{m,n}, \quad (6)$$

where $\bar{H}_{m,n}$ denotes the channel quantization matrix, and $E_{m,n}$ is the independent additive quantization distortion matrix whose entries are zero-mean and with variance $\frac{D_{N_x N_r}}{N_x N_r}$, where $D$ represents the channel quantization distortion constraint.

The differential feedback is under consideration as shown in Fig. 1. We can use the previous CSI to forecast the present CSI $H_{m,n}$ at the transmitter

$$\hat{H}_{m,n} = a_1 H_{m-1,n} + a_2 H_{m,n-1}, \quad (7)$$

where $a_1$ and $a_2$ are the coefficients of the channel predictor which will be calculated by using the minimum mean square error (MMSE) principle in the next section. Meanwhile, the receiver calculates the differential CSI, given the previous ones. The differential CSI can be formulated as

$$H_d = \text{Diff}(H_{m,n}|H_{m-1,n}, H_{m,n-1}), \quad (8)$$

where $H_d$ represents the differential CSI which obviously is the prediction error, and $\text{Diff}(\cdot)$ is the differential function. Then through limited feedback channel, $H_d$ should be quantized and fed back.

Finally, The CSI reconstructed by combining the differential one and the channel prediction is utilized by the channel adaptive techniques. In this paper, we adopt the water-filling precoder, however, the analysis and conclusions given in this paper are also valid for other adaptive techniques.

The channel quantization matrix is decomposed as $\bar{H}_{m,n} = \bar{U} \Sigma \bar{V}^\dagger$ using singular value decomposition (SVD) at the transmitter. $\bar{U}$ and $\bar{V}$ are unitary matrices, and $\Sigma$ is a non-negative diagonal matrix composed of eigenvalues of $\bar{H}_{m,n}$.

With the water-filling precoder, the closed-loop capacity can be obtained as [13]–[15]

$$C_{erg} = \mathbb{E}\left[\log \det \left( I_{N_r} + J \cdot \bar{J}^\dagger \left( \bar{F}^{-1} \right) \right) \right], \quad (9)$$
where \( \mathbf{J} = \hat{\mathbf{H}}_{m,n} \mathbf{V} \mathbf{Z} \), \( \mathbf{J}_c = \mathbf{E}_{m,n} \mathbf{V} \mathbf{Z} \), and \( \mathbf{F} = \frac{1}{A} \mathbf{I}_{N_r} + \mathbb{E} \{ \mathbf{J}_c \mathbf{J}_c^\dagger | \mathbf{J}_c \} \), where \( A \) represents the amplitude of signal symbol, and \( \mathbf{Z} \) denotes a diagonal matrix determined by water-filling [13]–[15].

Thus, the feedback load has positive relation with consideration, the feedback rate is determined by the rate \( \bar{\gamma} \), the limited feedback channel, the capacity can be enhanced by exploiting the channel correlations to reduce the quantization error.

### III. Minimal Differential Feedback Rate

In this section, exploiting the temporal and spectral correlations, we study the minimal feedback rate that denotes the minimal feedback bits required per block to preserve the given channel quantization distortion.

We first describe the feedback rate using normal quantization. Without differential feedback scheme, the receiver feeds back \( \mathbf{H}_{m,n} \) to the transmitter. The information entropy of a Gaussian variable \( X \) with variance \( \sigma^2 \) is represented as [16]

\[
h(X) = \frac{1}{2} \log 2\pi e \sigma^2.
\]

Thus, the feedback load has positive relation with \( \sigma^2 \).

Furthermore, taking quantization of the channel matrix into consideration, the feedback rate is determined by the rate distortion theory of continuous-amplitude sources [16],

\[
R = \inf \{ I(\mathbf{H}_{m,n} ; \hat{\mathbf{H}}_{m,n}) : \mathbb{E} \{ d(\mathbf{H}_{m,n} ; \hat{\mathbf{H}}_{m,n}) \} \leq D \},
\]

where \( \inf \{ \cdot \} \) denotes the infimum function, \( I(\mathbf{H}_{m,n} ; \hat{\mathbf{H}}_{m,n}) \) denotes the mutual information between \( \mathbf{H}_{m,n} \) and \( \hat{\mathbf{H}}_{m,n} \), and \( d(\mathbf{H}_{m,n} ; \hat{\mathbf{H}}_{m,n}) = \| \mathbf{H}_{m,n} - \hat{\mathbf{H}}_{m,n} \|^2 \) denotes the channel quantization distortion which is constrained by \( D \).

Since the entries of \( \mathbf{H} \) and \( \hat{\mathbf{H}} \) are i.i.d. complex Gaussian variables, the feedback rate can be written as

\[
R = \inf \{ N_r N_r I(\mathbf{H}_{m,n} ; \hat{\mathbf{H}}_{m,n}) : \mathbb{E} \{ d(\mathbf{H}_{m,n} , \hat{\mathbf{H}}_{m,n}) \} \leq d \},
\]

where \( d = \frac{D}{N_r N_r} \) is the one-dimensional average channel quantization distortion. \( \mathbf{H}_{m,n} \) and \( \hat{\mathbf{H}}_{m,n} \) represent the entries of \( \mathbf{H}_{m,n}, \hat{\mathbf{H}}_{m,n} \), respectively. Also, from (9) the one-dimensional channel quantization is written as

\[
\mathbf{H}_{m,n} = \hat{\mathbf{H}}_{m,n} + \mathbf{E}_{m,n}.
\]

The mutual information can be written as

\[
I(\mathbf{H}_{m,n} ; \hat{\mathbf{H}}_{m,n}) = h(\mathbf{H}_{m,n}) - h(\mathbf{E}_{m,n}).
\]

Combining (14), (15) can be rewritten as

\[
I(\mathbf{H}_{m,n} ; \hat{\mathbf{H}}_{m,n}) \geq h(\mathbf{H}_{m,n}) - h(\mathbf{E}_{m,n}).
\]

Substituting (11) and (16) into (13), we obtain

\[
R = N_r N_r \log \left(\frac{\sigma^2}{d} \right).
\]

From (17), the feedback rate required for the non-differential feedback is very large. Nevertheless, by employing the temporal and spectral correlations, we can use the differential feedback scheme to reduce the feedback bits significantly. The transmitter can predict the present CSI \( \mathbf{H}_{m,n} \) depending on the previous ones in time domain \( \mathbf{H}_{m-1,n} \) and frequency domain \( \mathbf{H}_{m,1-n} \). Then, the receiver quantizes \( \mathbf{H}_d \) or equivalently, the error of the channel prediction, and feeds back to the transmitter. Finally, the transmitter reconstructs the CSI by both the channel prediction and the differential CSI. It is obvious that the more accurate the channel is predicted, the less bits is fed back from the receiver. As \( \mathbf{H}_{m-1,n}, \mathbf{H}_{m,n-1} \) and \( \mathbf{H}_{m,n} \) are correlated, an MMSE channel predictor can be constructed as (7), where the coefficients \( a_1 \) and \( a_2 \) are selected to minimize

\[
\text{MSE}(a_1, a_2) = \mathbb{E} \{ \hat{\mathbf{H}}_{m,n} - \mathbf{H}_{m,n} \}^2.
\]

The MSE represents the statistical difference between the predicted value and the true one. We can obtain the minimized quantization bits by minimizing the MSE.

We can rewrite \( \mathbf{H}_{m,n} \) as

\[
\mathbf{H}_{m,n} = \hat{\mathbf{H}}_{m,n} + \mathbf{H}_d = a_1 \mathbf{H}_{m-1,n} + a_2 \mathbf{H}_{m,n-1} + \mathbf{H}_d,
\]

where \( \mathbf{H}_d \) is the differential feedback load to minimize. By the orthogonality principle [17], \( a_1, a_2 \) are determined by

\[
\begin{align*}
\mathbb{E} \{ (\mathbf{H}_{m,n} - a_1 \mathbf{H}_{m-1,n} - a_2 \mathbf{H}_{m,n-1}) \mathbf{H}_{m-1,n} \} &= 0, \\
\mathbb{E} \{ (\mathbf{H}_{m,n} - a_1 \mathbf{H}_{m-1,n} - a_2 \mathbf{H}_{m,n-1}) \mathbf{H}_{m,n-1} \} &= 0.
\end{align*}
\]

Since the entries of \( \mathbf{H}_{m,n}, \hat{\mathbf{H}}_{m-1,n}, \mathbf{H}_{m,n-1} \) are i.i.d. complex Gaussian variables, the orthogonality principle can be rewritten as

\[
\begin{align*}
\mathbb{E} \{ (\mathbf{H}_{m,n} - a_1 \mathbf{H}_{m-1,n} - a_2 \mathbf{H}_{m,n-1}) \mathbf{H}_{m-1,n} \} &= 0, \\
\mathbb{E} \{ (\mathbf{H}_{m,n} - a_1 \mathbf{H}_{m-1,n} - a_2 \mathbf{H}_{m,n-1}) \mathbf{H}_{m,n-1} \} &= 0.
\end{align*}
\]

Moreover, the one-dimensional frequency response of the channel can be represented as

\[
\mathbf{H}_{m,n} = \hat{\mathbf{H}}_{m,n} + \mathbf{H}_d = a_1 \mathbf{H}_{m-1,n} + a_2 \mathbf{H}_{m,n-1} + \mathbf{H}_d,
\]

where \( \mathbf{H}_{m,n}, \hat{\mathbf{H}}_{m,n}, \mathbf{H}_{m-1,n}, \mathbf{H}_{m,n-1} \) and \( \mathbf{H}_d \) represent the corresponding entries.

Direct calculation shows that (21) is equivalent to

\[
\begin{align*}
\mathbf{r}_H [1, 0] - a_1 \mathbf{r}_H [0, 0] - a_2 \mathbf{r}_H [1, 1] &= 0, \\
\mathbf{r}_H [0, 1] - a_1 \mathbf{r}_H [1, 1] - a_2 \mathbf{r}_H [0, 0] &= 0.
\end{align*}
\]

With the separation property of the correlations of the channel frequency response (2), and combining \( r_1[0] = r_f[0] = 1 \) and \( r_1[1] = \alpha_1, r_f[1] = \alpha_f \), (23) can be simplified by

\[
\begin{align*}
& a_1 \sigma^2_H + a_2 \alpha_1 \sigma^2_f - \alpha_1 \sigma^2_f = 0, \\
& a_1 \alpha_1 \sigma^2_f + a_2 \sigma^2_H - \alpha_f \sigma^2_f = 0.
\end{align*}
\]
From (24), \( a_1, a_2 \) are given by
\[
\begin{align*}
  a_1 &= \frac{\alpha_i(1-\alpha_f^2)}{1-\alpha_i^2\alpha_f^2}, \\
  a_2 &= \frac{\alpha_f(1-\alpha_i^2)}{1-\alpha_i^2\alpha_f^2},
\end{align*}
\] (25)

Combining (25) and (22), the one-dimensional MSE of the channel estimator is
\[
\text{MSE} = \text{Var}(H_d) = \sigma_H^2 (1-a_i^2-a_f^2 - 2a_1 a_2 \alpha_i \alpha_f). 
\] (26)

Finally, the channel estimator \( \hat{H}_{m,n} \) is given by
\[
\hat{H}_{m,n} = \frac{\alpha_t (1-\alpha_f^2)}{1-\alpha_t^2\alpha_f^2} H_{m-1,n} + \frac{\alpha_f (1-\alpha_t^2)}{1-\alpha_t^2\alpha_f^2} H_{m,n-1} + H_d. 
\] (27)

And combining (19) and (27), \( H_{m,n} \) is given by
\[
H_{m,n} = \frac{\alpha_t (1-\alpha_f^2)}{1-\alpha_t^2\alpha_f^2} H_{m-1,n} + \frac{\alpha_f (1-\alpha_t^2)}{1-\alpha_t^2\alpha_f^2} H_{m,n-1} + H_d. 
\] (28)

Then, through the feedback channel, the error of the channel predictor \( H_d \) can be fed back from the transmitter to the receiver. Similarly, from (11), the feedback load is positive related with \( \text{Var}(H_d) = \sigma_H^2 (1-a_i^2-a_f^2 - 2a_1 a_2 \alpha_i \alpha_f) \). Because \( \frac{\partial \text{MSE}}{\partial \alpha_i} < 0 \), \( \frac{\partial \text{MSE}}{\partial \alpha_f} < 0 \), the feedback load can be much smaller than \( \sigma_H^2 \), the non-differential one, especially when the channel is highly correlated. For example, given \( \alpha_i > 0.75 \), \( \alpha_f > 0.75 \), then \( \text{MSE}_{|\alpha_i>0.75, \alpha_f>0.75} < \text{MSE}_{|\alpha_i=0.75, \alpha_f=0.75} = 0.28 \sigma_H^2 \).

From (28), taking quantization impact into consideration, the minimal differential feedback rate over doubly selective fading channels can be calculated by the rate distortion theory of continuous-amplitude sources in a similar way.
\[
R = N_r N_t \log \left\{ a_i^2 + a_f^2 + \frac{2a_1 a_2 \alpha_i \alpha_f d}{\sigma_H^2} + \frac{\text{Var}(H_d)}{d} \right\}, 
\] (29)

where the channel predictor coefficients \( a_1, a_2 \) are determined by \( a_1 = \frac{\alpha_i(1-\alpha_f^2)}{1-\alpha_i^2\alpha_f^2} \) and \( a_2 = \frac{\alpha_f(1-\alpha_i^2)}{1-\alpha_i^2\alpha_f^2} \). The average power of \( H_d \) is \( \text{Var}(H_d) = \sigma_H^2 (1-a_i^2-a_f^2 - 2a_1 a_2 \alpha_i \alpha_f) \). The detailed derivation is given in Appendix A.

The above expression gives the minimal differential feedback rate simultaneously utilizing the temporal and spectral correlations. From (29), the minimal differential feedback rate is a function of \( \alpha_i, \alpha_f \) and the channel quantization distortion \( d \), and much smaller than that of the non-differential one (17).

IV. SIMULATION RESULTS AND DISCUSSION

In this section, we first provide the relationship between the MSE of the predictor and the two-dimensional correlations in Fig. 2. The minimal differential feedback rate over MIMO doubly selective fading channels is given in Fig. 3. Then, a longitudinal section of Fig. 3 is presented, where we assume the temporal correlation and spectral correlation is equal. Finally, we verify our theoretical results by a practical differential feedback system with water-filling precoder and Lloyd’s quantization algorithm [18].

A. MSE of the predictor and Minimal Differential Feedback Rate

For simplicity and without loss of generality, we consider \( N_r = N_t = 2, \sigma_H^2 = 1 \). Fig. 2 presents the MSE between the predicted value and the true value. As the temporal or spectral correlation increases, the MSE decreases. Furthermore, when either \( \alpha_i \) or \( \alpha_f \) comes to one, the MSE tends to zero.

Fig. 3 plots the relationship between the minimal differential feedback rate and the two-dimensional correlations with the channel quantization distortion \( D = 0.1 \). It is very similar to the MSE shown in Fig. 2 because it presents the minimal bits required to quantize the differential CSI.

Additionally, because \( \alpha_i \) and \( \alpha_f \) could be any value, we provide one of the longitudinal section of Fig. 3 where the temporal correlation is equal to the spectral correlation in Fig. 4. For comparison, the differential feedback compression only using one-dimensional correlation and the non-differential
feedback scheme are also included in Fig. 4. It is observed from Fig. 4 that the scheme using both temporal and spectral correlations is always better than the scheme using only one-dimensional correlation. As the correlations increase, the two-dimensional differential feedback compression exhibits a significant improvement compared to one-dimensional one. This performance advantage even reaches up to 67% with \( \alpha_t = \alpha_f = 0.95 \).

![Fig. 4. The relationship between the minimal feedback rate and temporal and spectral correlations, when they are equal, for \( N_r = 2, N_t = 2, \sigma_H^2 = 1 \) and \( D = 0.1 \).](image)

**B. Differential Feedback System with Lloyd’s Algorithm**

In this subsection, we consider the temporal correlation \( \alpha_t = 0.9 \), with carrier frequency 2 GHz, the normalized Doppler shift \( f_d = 100 \) Hz and spectral correlation \( \alpha_f = 0.9 \), with \( \Delta = 8 \) ms, which is a reasonable assumption [12]. We design a differential feedback system using Lloyd’s quantization algorithm to verify our theoretical results [18]. We use
\[
\text{Diff}(H_{m,n},H_{m-1,n},H_{m,n-1}) = H_{m,n} - \alpha_1 H_{m-1,n} - \alpha_2 H_{m,n-1}
\]
as a differential function, where \( \alpha_1 = \frac{\alpha_t(1-\alpha_f^2)}{1-\alpha_t^2} \) and \( \alpha_2 = \frac{\alpha_t(1-\alpha_f^2)}{1-\alpha_f^2} \) in the two-dimensional differential feedback compression and \( \alpha_1 = \alpha_t, \alpha_2 = 0 \) in the one-dimensional one.

The feedback steps can be summarized as follows. Firstly, based on Lloyd’s quantization algorithm, the channel codebook can be generated according to the statistics of the corresponding differential feedback load at both transmitter and receiver. Secondly, the receiver calculates the current differential CSI \( \bar{H}_d \). Thirdly, the differential CSI is quantized to the optimal codebook value \( \bar{H}_d \) according to the Euclidean distance. Finally, the transmitter reconstructs the channel quantization matrix by
\[
H_{m,n} = \alpha_1 H_{m-1,n} + \alpha_2 H_{m,n-1} + \bar{H}_d
\]
In Fig. 5 we give the simulation results of the ergodic capacity employing Lloyd’s algorithm. The theoretical capacity results are also provided in Fig. 5. We can see from Fig. 5 that the performance of the two-dimensional one are always better than the one-dimensional one, which verifies our theoretical analysis.

![Fig. 5. The relationship between the ergodic capacity and feedback rate with Lloyd’s algorithm in AR1 model for \( N_r = 2, N_t = 2, \sigma_H^2 = 1 \) and SNR = 5dB.](image)

As shown in Fig. 5, with the increase of feedback rate \( b \), the ergodic capacities increase rapidly when \( b \) is small, and then slow down in the large \( b \) region, because when \( b \) is large enough, the quantization errors tend to zero. Also, the capacities of Lloyd’s quantization are lower than the theoretical ones. The reasons are as follows. The Lloyd’s algorithm is optimal only in the sense of minimizing a variable’s quantization error, but not in data sequence compression while the channel coefficient \( \bar{H} \) is correlated in both temporal and spectral domain. However, the imperfection reduces as \( b \) increases, because the quantization errors of both Lloyd’s algorithm and theoretical results tend to zero with sufficient feedback bits \( b \).

**V. Conclusions**

In this paper, we have designed a differential feedback scheme making full use of both the temporal and spectral correlation and compared the performance with the scheme without differential feedback. We have derived the minimal differential feedback rate for our proposed scheme. The feedback rate to preserve the given channel quantization distortion is significantly small compared to non-differential one, as the channel is highly correlated in both temporal and spectral domain. Finally, we provide simulations to verify our analysis.

**APPENDIX A**

**Derivation of the Minimal Differential Feedback Rate Using Temporal and Spectral Correlations**

The minimal differential feedback rate over MIMO doubly selective fading channel can also be derived by the rate distortion theory. Given \( H_{m-1,n} \) and \( H_{m,n-1} \) at the transmitter, the differential feedback rate can be represented as
\[
R = \inf \{ I(H_{m,n};\bar{H}_{m,n}|H_{m-1,n},\bar{H}_{m,n-1});\mathbb{E}[d(H_{m,n};\bar{H}_{m,n})]\leq D \}.
\]
(30)
Since the entries are i.i.d. complex Gaussian variables, (30) can be written as
\[
R = \inf \{ I(H_{m,n};\bar{H}_{m,n}|H_{m-1,n},\bar{H}_{m,n-1});\mathbb{E}[d(H_{m,n};\bar{H}_{m,n})]\leq D \}.
\]
(31)
The one-dimensional channel quantization equality can be written as

\[ H_{m-1,n} = \bar{H}_{m-1,n} + E_{m-1,n} \]
\[ H_{m,n} = \bar{H}_{m,n} - E_{m,n-1} \tag{32} \]

Similarly, \[ (38) \] yields

\[ H_{m,n} = a_1 H_{m-1,n} + a_2 H_{m,n-1} + H_d, \tag{33} \]

where \( a_1 = \frac{\alpha_1(1-\alpha_2)}{1-\alpha_1\alpha_2}, \quad a_2 = \frac{\alpha_1(1-\alpha_2)}{1-\alpha_1\alpha_2} \). The conditional mutual information \( I(H_{m,n};H_{m,n}|H_{m-1,n},\bar{H}_{m-1,n}) \) can be written as

\[ I(H_{m,n};H_{m,n}|H_{m-1,n},\bar{H}_{m-1,n}) = h(H_{m,n}|H_{m-1,n},\bar{H}_{m-1,n}) \]
\[ -h(H_{m,n}|\bar{H}_{m,n},H_{m-1,n},\bar{H}_{m-1,n}). \tag{34} \]

First, we calculate \( h(H_{m,n}|H_{m-1,n},\bar{H}_{m-1,n}) \). Substituting \[ (32) \] into \[ (33) \], it yields that

\[ H_{m,n} = a_1 (\bar{H}_{m-1,n} + E_{m-1,n}) + a_2 (\bar{H}_{m,n-1} + E_{m-1,n}) + H_d. \tag{35} \]

Substituting \[ (35) \] into \[ (34) \], we obtain

\[ I = h(a_1 E_{m-1,n} + a_2 E_{m-1,n} + H_d) - h(E_{m,n}|H_{m-1,n},\bar{H}_{m-1,n}). \tag{36} \]

Considering inequality \( h(E_{m,n}|\bar{H}_{m-1,n},H_{m-1,n}) \leq h(E_{m,n}) \) \[ (36) \] can be written as

\[ I \geq h(a_1 E_{m-1,n} + a_2 E_{m-1,n} + H_d) - h(E_{m,n}). \tag{37} \]

Since \( E_{m-1,n}, E_{m,n-1} \) and \( H_d \) are complex Gaussian variables, and the information entropy of a Gaussian variables with variance \( \sigma^2 \) is \( h(X) = 1/2 \log 2\pi e \sigma^2 \), we calculate the variance of \( (a_1 E_{m-1,n} + a_2 E_{m-1,n} + H_d) \)

\[ \text{Var} (a_1 E_{m-1,n} + a_2 E_{m-1,n} + H_d) = a_1^2 \sigma_1^2 d + a_2^2 \sigma_2^2 d \tag{38} \]
\[ + \text{Var} (H_d)^2 + 2a_1 a_2 \text{Var} (E_{m-1,n},E_{m,n-1}). \]

Now we give the derivation of the correlation function of two noise terms \( r(E_{m-1,n},E_{m,n-1}). \) From \[ (32) \], the quantization error can be decomposed into two parts

\[ E_{m-1,n} = \frac{\sigma_{m-1,n}^2 - \sigma_H^2}{\sigma_H^2} \bar{H}_{m-1,n} + \psi_{m-1,n}, \tag{39} \]
\[ E_{m,n-1} = \frac{\sigma_{m,n-1}^2 - \sigma_H^2}{\sigma_H^2} \bar{H}_{m,n-1} + \psi_{m,n-1}, \]

where

\[ \psi_{m,n} \equiv \bar{H}_{m,n} - \frac{\sigma_H^2}{\sigma_{m,n}^2} \bar{H}_{m,n-1} = H_{m,n-1} - \frac{\sigma_H^2}{\sigma_{m,n}^2} H_{m,n-1}, \tag{40} \]

\( \psi \) is a Gaussian variable with zero-mean and variance \( \frac{\sigma_H^2}{\sigma_{m,n}^2} \), independent with \( H \).

Then the correlation function of \( E_{m-1,n} \) and \( E_{m,n-1} \) can be calculated as

\[ r(E_{m-1,n},E_{m,n-1}) = \frac{(\sigma_{m-1,n}^2 - \sigma_H^2)^2}{\sigma_H^2} \alpha_t \alpha_f = \frac{d^2}{\sigma_H^2} \alpha_t \alpha_f. \tag{41} \]

Substituting \[ (41) \] into \[ (38) \], we obtain

\[ \text{Var} (a_1 E_{m-1,n} + a_2 E_{m-1,n} + H_d) = a_1^2 \sigma_1^2 d + a_2^2 \sigma_2^2 d \tag{42} \]
\[ + \sigma_H^2 + 2a_1 a_2 \frac{d^2}{\sigma_H^2} \alpha_t \alpha_f. \]

From \[ (31), (37) \] and \[ (42) \], it yields that

\[ R = N_r N_r \log \left( a_1^2 + a_2^2 + 2a_1 a_2 \frac{d^2}{\sigma_H^2} \alpha_t \alpha_f + \text{Var} (H_d) \right), \tag{43} \]

where \( a_1 = \frac{\alpha_1(1-\alpha_2)}{1-\alpha_1\alpha_2}, \quad a_2 = \frac{\alpha_1(1-\alpha_2)}{1-\alpha_1\alpha_2} \) and \( \text{Var} (H_d) = \sigma_H^2 (1 - a_1^2 - a_2^2 - 2a_1 a_2 \alpha_t \alpha_f) \).

**References**


