A differential Feedback Scheme Exploiting the Temporal and Spectral Correlation

Mingxin Zhou, Leiming Zhang, Lingyang Song, Mérouane Debbah

A Differential Feedback Scheme Exploiting the Temporal and Spectral Correlation

Mingxin 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 [11]

$$\begin{aligned} r_H(\Delta t, \Delta f) &= \mathbb{E}\left\{H(t+\Delta t, f+\Delta f) H^*(t,f)\right\} \\ &= \sigma_H^2 r_t(\Delta t) r_f(\Delta f), \end{aligned} \quad (1)$$

Manuscript received November 25, 2012; revised February 20, 2013; accepted May 4, 2013. The associate editor coordinating the review of this paper and approving it for publication was Prof. Xianbing Wang. This work was partially supported by the National 973 project under grant 2013CB336700, National Nature Science Foundation of China under grant number 61222104 and 61061130561, the Ph.D. Programs Foundation of Ministry of Education of China under grant number 20110001110102, and the Opening Project of Key Laboratory of Cognitive Radio and Information Processing (Guilin University of Electronic Technology).

M. Zhou, and L. Song are with the State Key Laboratory of Advanced Optical Communication Systems and Networks, School of Electronics Engineering and Computer Science, Peking University, Beijing, P. R. China, 100871. (e-mail: {mingxin.zhou, lingyang.song}@pku.edu.cn).

L. Zhang is with Huawei Technol., Beijing, P. R. China, 100095. (e-mail: leiming.zhang@pku.edu.cn)

M. Debbah is with SUPELEC, Alcatel-Lucent Chair in Flexible Radio, 3 rue Joliot-Curie, FR-91192 Gif Sur Yvette, France. (e-mail: merouane.debbah@supelec.fr)

where $\mathbb{E}\{\cdot\}$ denotes expectation function, the superscript $(\cdot)^*$ denotes complex conjugate. σ_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

$$r_H[\Delta m, \Delta n] = \sigma_H^2 r_t[\Delta m] r_f[\Delta n], \tag{2}$$

where $r_t[\Delta m] = r_t(\Delta m t_s)$ and $r_f[\Delta n] = r(\Delta n f_s)$.

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

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

where $H_{m,n}$ denotes the channel coefficient of the mth symbol interval and the nth subchannel, W_t is a complex white noise variable, which is independent of $H_{m-1,n}$, with variance σ_H^2. The parameter α_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]

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

where W_f is a complex white noise variable, which is independent of $H_{m,n-1}$, with variance σ_H^2. The parameter α_f determines the correlation between the subchannels, which is given by $\alpha_f = r_f[1] = \frac{1}{\sqrt{1+(2\pi f_s \Delta)^2}}$, where Δ 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.

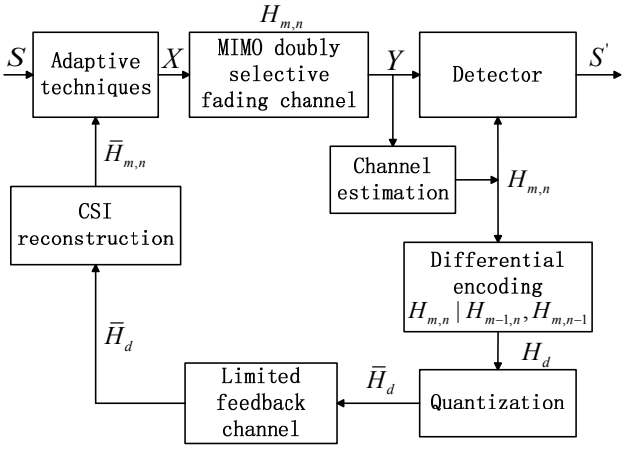

Fig. 1. System model of the differential feedback over MIMO doubly selective fading channel

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

$$\mathbf{y}_{m,n} = \mathbf{H}_{m,n}\mathbf{x}_{m,n} + \mathbf{n}_{m,n}. \tag{5}$$

In the above expression, $\mathbf{y}_{m,n}$ denotes the $N_r \times 1$ received vector at the mth symbol interval and the nth subchannel. $\mathbf{H}_{m,n}$, 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 σ_H^2. Different antennas have the same characteristic in temporal and spectral correlations, α_t and α_f, respectively. Besides, there is no spatial correlation between different antennas. $\mathbf{x}_{m,n}$ denotes the $N_t \times 1$ transmitter signal vector and is assumed to have unit variance. $\mathbf{n}_{m,n}$ is a $N_r \times 1$ additive white Gaussian noise (AWGN) vector with zero-mean and variance σ_0^2. Both $\mathbf{x}_{m,n}$ and $\mathbf{n}_{m,n}$ are independent for different m's and n's.

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

$$\mathbf{H}_{m,n} = \bar{\mathbf{H}}_{m,n} + \mathbf{E}_{m,n}, \tag{6}$$

where $\bar{\mathbf{H}}_{m,n}$ denotes the channel quantization matrix, and $\mathbf{E}_{m,n}$ is the independent additive quantization distortion matrix whose entries are zero-mean and with variance $\frac{D}{N_r N_t}$, 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 $\mathbf{H}_{m,n}$ at the transmitter

$$\hat{\mathbf{H}}_{m,n} = a_1 \mathbf{H}_{m-1,n} + a_2 \mathbf{H}_{m,n-1}, \tag{7}$$

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

$$\mathbf{H}_d = \mathrm{Diff}\left(\mathbf{H}_{m,n} | \mathbf{H}_{m-1,n}, \mathbf{H}_{m,n-1}\right), \tag{8}$$

where $\mathbf{H}_d$ represents the differential CSI which obviously is the prediction error, and $Diff(\cdot)$ is the differential function. Then through limited feedback channel, $\mathbf{H}_d$ 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 $\bar{\mathbf{H}}_{m,n} = \bar{\mathbf{U}}\bar{\mathbf{\Sigma}}\bar{\mathbf{V}}^+$ using singular value decomposition (SVD) at the transmitter. $\bar{\mathbf{U}}$ and $\bar{\mathbf{V}}$ are unitary matrixes, and $\bar{\mathbf{\Sigma}}$ is a non-negative diagonal matrix composed of eigenvalues of $\bar{\mathbf{H}}_{m,n}$.

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

$$C_{erg} = \mathbb{E}\left[\log\det\left(\mathbf{I}_{N_r} + \mathbf{J}\cdot\mathbf{J}^+\left(\mathbf{F}^{-1}\right)\right)\right], \tag{9}$$

where $\mathbf{J} = \bar{\mathbf{H}}_{m,n}\bar{\mathbf{V}}\bar{\mathbf{Z}}$, $\mathbf{J}_e = \mathbf{E}_{m,n}\bar{\mathbf{V}}\bar{\mathbf{Z}}$, and $\mathbf{F} = \frac{1}{A^2}\mathbf{I}_{N_r} + \mathbb{E}\left[\mathbf{J}_e\mathbf{J}_e^{+}|\mathbf{J}\right]$, where A represents the amplitude of signal symbol, and $\bar{\mathbf{Z}}$ denotes a diagonal matrix determined by water-filling [13]–[15]

$$\begin{cases} \bar{z}_i^2 = \begin{cases} \bar{\mu} - (\bar{\gamma}_{i,i}^2 A^2)^{-1}, & \bar{\gamma}_{i,i}^2 A^2 \geq \bar{\mu}^{-1} \\ 0, & \text{otherwise} \end{cases} \\ \sum\limits_{i=1}^{N_t} \bar{z}_i^2 A^2 = N_t A^2, \qquad \text{power constraint} \end{cases}, \tag{10}$$

where $\bar{\gamma}_{i,i}, i = 1, 2, ..., N_t$ are the entries of $\bar{\Sigma}$, $\bar{\mu}$ denotes a cut-off value chosen to meet the power constraint.

It is obvious from (9) that the closed-loop ergodic capacity is determined by $\mathbf{H}_{m,n}$ and $\bar{\mathbf{H}}_{m,n}$, and the loss of capacity is mainly caused by the quantization error. Therefore, given 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 σ^2 is represented as [16]

$$h(X) = \frac{1}{2}\log 2\pi e\sigma^2. \tag{11}$$

Thus, the feedback load has positive relation with σ_H^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]

$$R = \inf\left\{I\left(\mathbf{H}_{m,n}; \bar{\mathbf{H}}_{m,n}\right) : \mathbb{E}\left[d\left(\mathbf{H}_{m,n}; \bar{\mathbf{H}}_{m,n}\right)\right] \leq D\right\}, \tag{12}$$

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

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

$$R = \inf\left\{N_t N_r I\left(H_{m,n}; \bar{H}_{m,n}\right) : \mathbb{E}[d(H_{m,n}, \bar{H}_{m,n})] \leq d\right\}, \tag{13}$$

where $d = \frac{D}{N_t N_r}$ is the one-dimensional average channel quantization distortion. $H_{m,n}$ and $\bar{H}_{m,n}$ represent the entries of $\mathbf{H}_{m,n}$, $\bar{\mathbf{H}}_{m,n}$, respectively. Also, from (6) the one-dimensional channel quantization is written as

$$H_{m,n} = \bar{H}_{m,n} + E_{m,n}. \tag{14}$$

The mutual information can be written as

$$I\left(H_{m,n}; \bar{H}_{m,n}\right) = h\left(H_{m,n}\right) - h\left(H_{m,n}\left|\bar{H}_{m,n},\right.\right). \tag{15}$$

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

$$I\left(H_{m,n}; \bar{H}_{m,n}\right) \geq h\left(H_{m,n}\right) - h\left(E_{m,n}\right). \tag{16}$$

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

$$R = N_r N_t \log\left(\frac{\sigma_H^2}{d}\right). \tag{17}$$

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

$$\mathrm{MSE}(a_1, a_2) = \mathbb{E}\left|\hat{\mathbf{H}}_{m,n} - \mathbf{H}_{m,n}\right|^2. \tag{18}$$

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

$$\mathbf{H}_{m,n} = \hat{\mathbf{H}}_{m,n} + \mathbf{H}_d = a_1\mathbf{H}_{m-1,n} + a_2\mathbf{H}_{m,n-1} + \mathbf{H}_d, \tag{19}$$

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

$$\begin{cases} \mathbb{E}\left[\left(\mathbf{H}_{m,n} - a_1\mathbf{H}_{m-1,n} - a_2\mathbf{H}_{m,n-1}\right)\mathbf{H}_{m-1,n}\right] = 0 \\ \mathbb{E}\left[\left(\mathbf{H}_{m,n} - a_1\mathbf{H}_{m-1,n} - a_2\mathbf{H}_{m,n-1}\right)\mathbf{H}_{m,n-1}\right] = 0 \end{cases}. \tag{20}$$

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

$$\begin{cases} \mathbb{E}\left[\left(H_{m,n} - a_1H_{m-1,n} - a_2H_{m,n-1}\right)H_{m-1,n}\right] = 0 \\ \mathbb{E}\left[\left(H_{m,n} - a_1H_{m-1,n} - a_2H_{m,n-1}\right)H_{m,n-1}\right] = 0 \end{cases}. \tag{21}$$

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

$$H_{m,n} = \hat{H}_{m,n} + H_d = a_1H_{m-1,n} + a_2H_{m,n-1} + H_d, \tag{22}$$

where $H_{m,n}$, $\hat{H}_{m,n}$, $H_{m-1,n}$, $H_{m,n-1}$ and H_d represent the corresponding entries.

Direct calculation shows that (21) is equivalent to

$$\begin{cases} r_H[1,0] - a_1r_H[0,0] - a_2r_H[1,1] = 0 \\ r_H[0,1] - a_1r_H[1,1] - a_2r_H[0,0] = 0 \end{cases}. \tag{23}$$

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

$$\begin{cases} a_1\sigma_H^2 + a_2\alpha_t\alpha_f\sigma_H^2 - \alpha_t\sigma_H^2 = 0 \\ a_1\alpha_t\alpha_f\sigma_H^2 + a_2\sigma_H^2 - \alpha_f\sigma_H^2 = 0 \end{cases}. \tag{24}$$

From (24), a_1, a_2 are given by

$$\begin{cases} a_1 = \frac{\alpha_t\left(1-\alpha_f^2\right)}{1-\alpha_t^2\alpha_f^2} \\ a_2 = \frac{\alpha_f\left(1-\alpha_t^2\right)}{1-\alpha_t^2\alpha_f^2} \end{cases}. \tag{25}$$

Combing (25) and (22), the one-dimensional MSE of the channel estimator is

$$\mathrm{MSE} = \mathrm{Var}\left(H_d\right) = \sigma_H^2\left(1 - a_1^2 - a_2^2 - 2a_1a_2\alpha_t\alpha_f\right). \tag{26}$$

Finally, the channel estimator $\hat{\mathbf{H}}_{m,n}$ is given by

$$\hat{\mathbf{H}}_{m,n} = \frac{\alpha_t\left(1-\alpha_f^2\right)}{1-\alpha_t^2\alpha_f^2}\mathbf{H}_{m-1,n} + \frac{\alpha_f\left(1-\alpha_t^2\right)}{1-\alpha_t^2\alpha_f^2}\mathbf{H}_{m,n-1}. \tag{27}$$

And combining (19) and (27), $\mathbf{H}_{m,n}$ is given by

$$\mathbf{H}_{m,n} = \frac{\alpha_t\left(1-\alpha_f^2\right)}{1-\alpha_t^2\alpha_f^2}\mathbf{H}_{m-1,n} + \frac{\alpha_f\left(1-\alpha_t^2\right)}{1-\alpha_t^2\alpha_f^2}\mathbf{H}_{m,n-1} + \mathbf{H}_d. \tag{28}$$

Then, through the feedback channel, the error of the channel predictor $\mathbf{H}_d$ can be fed back from the transmitter to the receiver. Similarly, from (11), the feedback load is positive related with $\mathrm{Var}(H_d) = \sigma_H^2(1 - a_1^2 - a_2^2 - 2a_1a_2\alpha_t\alpha_f)$. Because $\frac{\partial \mathrm{MSE}}{\partial \alpha_t} < 0$, $\frac{\partial \mathrm{MSE}}{\partial \alpha_f} < 0$, the feedback load can be much smaller than σ_H^2, the non-differential one, especially when the channel is highly correlated. For example, given $\alpha_t > 0.75$, $\alpha_f > 0.75$, then $\mathrm{MSE}|_{\alpha_t>0.75,\ \alpha_f>0.75} < \mathrm{MSE}|_{\alpha_t=0.75,\ \alpha_f=0.75} = 0.28\sigma_H^2$.

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.

$$R = N_rN_t \log\left\{a_1^2 + a_2^2 + \frac{2a_1a_2\alpha_t\alpha_f d}{\sigma_H^2} + \frac{Var\left(H_d\right)}{d}\right\}, \tag{29}$$

where the channel predictor coefficients a_1, a_2 are determined by $a_1 = \frac{\alpha_t\left(1-\alpha_f^2\right)}{1-\alpha_t^2\alpha_f^2}$ and $a_2 = \frac{\alpha_f\left(1-\alpha_t^2\right)}{1-\alpha_t^2\alpha_f^2}$. The average power of H_d is $Var\left(H_d\right) = \sigma_H^2\left(1 - a_1^2 - a_2^2 - 2a_1a_2\alpha_t\alpha_f\right)$. 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 α_t, α_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 $N_r = N_t = 2$, and $\sigma_H^2 = 1$. 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 α_t or α_f comes to one, the MSE tends to zero.

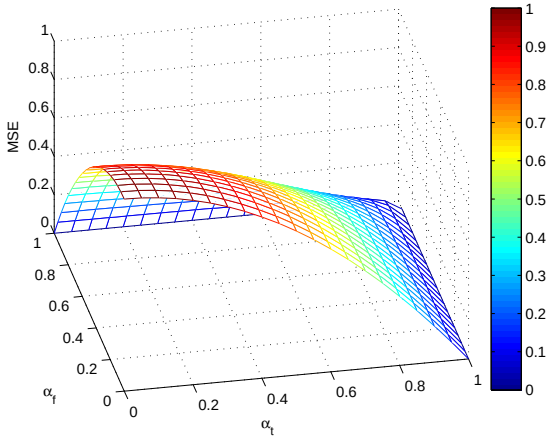

Fig. 2. The MSE of the predictor at the transmiter, for $N_r = 2, N_t = 2, \sigma_H^2 = 1$ and $D = 0.1$.

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.

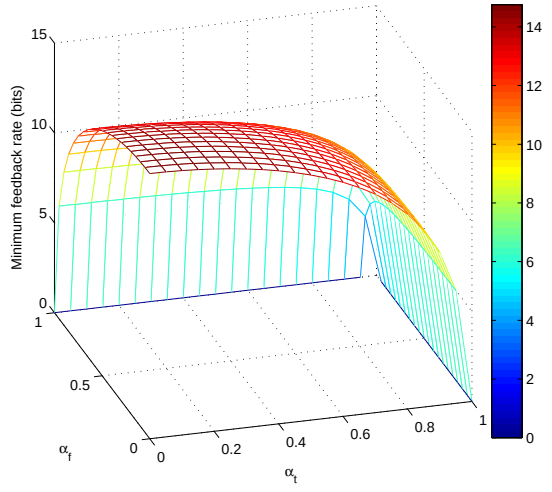

Fig. 3. The minimal differential feedback rate, for $N_r = 2, N_t = 2, \sigma_H^2 = 1$ and $D = 0.1$.

Additionally, because α_t and α_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 $\alpha_t = \alpha_f = 0.95$.

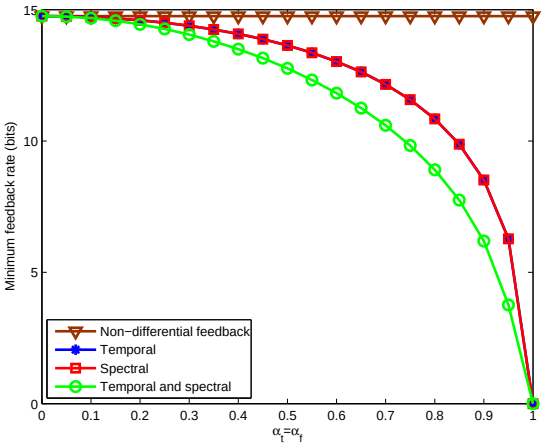

Fig. 4. The relationship between the minimal feedback rate and temporal and spectral correlations, when they are equal, for $N_r = 2, N_t = 2, \sigma_H^2 = 1$ and $D = 0.1$.

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 $\Delta = 8\mu$s, 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 $\text{Diff}\left(\mathbf{H}_{m,n}|\mathbf{H}_{m-1,n}\mathbf{H}_{m,n-1}\right) = \mathbf{H}_{m,n} - a_1\mathbf{H}_{m-1,n} - a_2\mathbf{H}_{m,n-1}$ as a differential function, where $a_1 = \frac{\alpha_t\left(1-\alpha_f^2\right)}{1-\alpha_t^2\alpha_f^2}$, $a_2 = \frac{\alpha_f\left(1-\alpha_t^2\right)}{1-\alpha_t^2\alpha_f^2}$ in the two-dimensional differential feedback compression and $a_1 = \alpha_t$, $a_2 = 0$ 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 $\mathbf{H}_d$. Thirdly, the differential CSI is quantized to the optimal coodbook value $\bar{\mathbf{H}}_d$ according to the Euclidean distance. Finally, the transmitter reconstructs the channel quantization matrix by $\mathbf{H}_{m,n} = a_1\mathbf{H}_{m-1,n} + a_2\mathbf{H}_{m,n-1} + \bar{\mathbf{H}}_d$.

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.

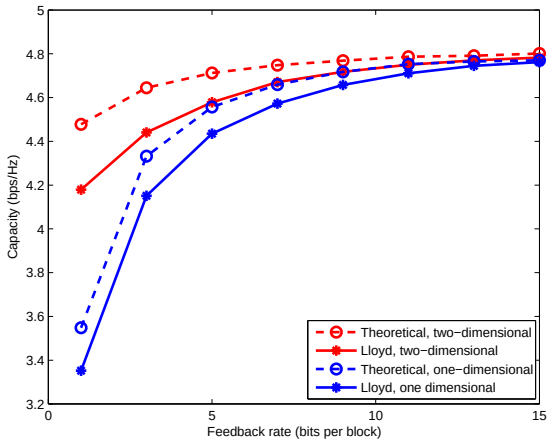

Fig. 5. The relationship between the ergodic capacity and feedback rate with Lloyd's algorithm in AR1 model for $N_r = 2, N_t = 2, \sigma_H^2 = 1$ and SNR $= 5$dB.

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 $\mathbf{H}$ 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 $\bar{\mathbf{H}}_{m-1,n}$ and $\bar{\mathbf{H}}_{m,n-1}$ at the transmitter, the differential feedback rate can be represented as

$$R = \inf\left\{I\left(\mathbf{H}_{m,n};\bar{\mathbf{H}}_{m,n}|\bar{\mathbf{H}}_{m-1,n},\bar{\mathbf{H}}_{m,n-1}\right):\mathbb{E}\left[d\left(\mathbf{H}_{m,n};\bar{\mathbf{H}}_{m,n}\right)\right]\leq D\right\}. \quad (30)$$

Since the entries are i.i.d. complex Gaussian variables, (30) can be written as

$$R = \inf\left\{I\left(H_{m,n};\bar{H}_{m,n}|\bar{H}_{m-1,n},\bar{H}_{m,n-1}\right):\mathbb{E}\left[d\left(H_{m,n};\bar{H}_{m,n}\right)\right]\leq D\right\}. \quad (31)$$

The one-dimensional channel quantization equality can be written as

$$\begin{aligned} H_{m-1,n} &= \bar{H}_{m-1,n} + E_{m-1,n} \\ H_{m,n-1} &= \bar{H}_{m,n-1} + E_{m,n-1}. \end{aligned} \tag{32}$$

Similarly, (28) yields

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

where $a_1 = \frac{\alpha_t\left(1-\alpha_f^2\right)}{1-\alpha_t^2\alpha_f^2}$, $a_2 = \frac{\alpha_f\left(1-\alpha_t^2\right)}{1-\alpha_t^2\alpha_f^2}$. The conditional mutual information $I\big(H_{m,n};\bar{H}_{m,n}|\bar{H}_{m-1,n},\bar{H}_{m,n-1}\big)$ can be written as

$$\begin{aligned} I\big(H_{m,n};\bar{H}_{m,n}|\bar{H}_{m-1,n},\bar{H}_{m,n-1}\big) &= h\big(H_{m,n}|\bar{H}_{m-1,n},\bar{H}_{m,n-1}\big) \\ &\quad - h\big(H_{m,n}|\bar{H}_{m,n},\bar{H}_{m-1,n},\bar{H}_{m,n-1}\big). \end{aligned} \tag{34}$$

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

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

Substituting (35) into (34), we obtain

$$I = h(a_1E_{m-1,n} + a_2E_{m,n-1} + H_d) - h\big(E_{m,n}|\bar{H}_{m-1,n},\bar{H}_{m,n-1}\big). \tag{36}$$

Considering inequality $h\big(E_{m,n}|\bar{H}_{m-1,n},\bar{H}_{m,n-1}\big) \le h(E_{m,n})$ (36) can be written as

$$I \ge h\left(a_1E_{m-1,n} + a_2E_{m,n-1} + H_d\right) - h\left(E_{m,n}\right). \tag{37}$$

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 σ^2 is $h\left(X\right) = \frac{1}{2}\log 2\pi e\sigma^2$, we calculate the variance of $\left(a_1E_{m-1,n} + a_2E_{m,n-1} + H_d\right)$

$$\begin{aligned} Var\left(a_1E_{m-1,n} + a_2E_{m,n-1} + H_d\right) &= a_1^2d + a_2^2d \\ &\quad + Var\left(H_d{}^2\right) + 2a_1a_2r\left(E_{m-1,n}, E_{m,n-1}\right). \end{aligned} \tag{38}$$

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

$$\begin{aligned} E_{m-1,n} &= \tfrac{\sigma_H^2-\sigma_{\bar{H}}^2}{\sigma_H^2}H_{m-1,n} + \psi_{m-1,n} \\ E_{m,n-1} &= \tfrac{\sigma_H^2-\sigma_{\bar{H}}^2}{\sigma_H^2}H_{m,n-1} + \psi_{m,n-1} \end{aligned}, \tag{39}$$

where

$$\begin{aligned} \psi_{m,n-1} &= \bar{H}_{m,n-1} - \tfrac{\sigma_{\bar{H}}^2}{\sigma_H^2}H_{m,n-1} \\ \psi_{m-1,n} &= \bar{H}_{m-1,n} - \tfrac{\sigma_{\bar{H}}^2}{\sigma_H^2}H_{m-1,n} \end{aligned}, \tag{40}$$

ψ is a Gaussian variable with zero-mean and variance $\frac{\sigma_{\bar{H}}^2\left(\sigma_H^2-\sigma_{\bar{H}}^2\right)}{\sigma_H^2}$, independent with H.

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

$$r\left(E_{m-1,n}, E_{m,n-1}\right) = \frac{\left(\sigma_H^2-\sigma_{\bar{H}}^2\right)^2}{\sigma_H^2}\alpha_t\alpha_f = \frac{d^2}{\sigma_H^2}\alpha_t\alpha_f. \tag{41}$$

Substituting (41) into (38), we obtain

$$\begin{aligned} Var\left(a_1E_{m-1,n} + a_2E_{m,n-1} + H_d\right) &= a_1^2d + a_2^2d \\ &\quad + \sigma_{H_d}^2 + 2a_1a_2\frac{d^2}{\sigma_H^2}\alpha_t\alpha_f. \end{aligned} \tag{42}$$

From (31), (37) and (42), it yields that

$$R = N_rN_t\log\left\{a_1^2 + a_2^2 + \frac{2a_1a_2\alpha_t\alpha_fd}{\sigma_H^2} + \frac{Var\left(H_d\right)}{d}\right\}, \tag{43}$$

where $a_1 = \frac{\alpha_t\left(1-\alpha_f^2\right)}{1-\alpha_t^2\alpha_f^2}$, $a_2 = \frac{\alpha_f\left(1-\alpha_t^2\right)}{1-\alpha_t^2\alpha_f^2}$ and $Var\left(H_d\right) = \sigma_H^2\left(1 - a_1^2 - a_2^2 - 2a_1a_2\alpha_t\alpha_f\right)$.

REFERENCES

[1] D. J. Love, R. W. Heath Jr., W. Santipach, and M.L. Honig, "What is the value of limited feedback for MIMO channels," *lEEE Commun. Mag.*, vol. 42, no. 10. pp. 54-59, Oct. 2004

[2] D. J. Love, R. W. Heath, V. K. N. Lau, etc. "An Overview of limited feedback in wireless Communication Systems," *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 1341-1365, Oct. 2008.

[3] T. Eriksson and T. Ottosson, "Compression of feedback for adaptive transmission and scheduling," *Proc. IEEE*, pp. 2314-2321, Dec. 2007.

[4] K. E. Baddour and N. C. Beaulieu, "Autoregressive modeling for fading channel simulation," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 1650-1662. Jul. 2005.

[5] Y. Zhang, R. Yu, W. Yao, S. Xie, Y. Xiao and M. Guizani, "Home M2M Networks: Architectures, Standards, and QoS Improvement", *IEEE Communications Magazine*, vol.49, no.4, pp.44–52, Apr. 2011.

[6] K. Huang, B. Mondal, R. W. Heath. Jr., and J. G. Andrews, "Markov models for limited feedback MIMO systems," in *Proc. 2006 IEEE ICASSP*.

[7] Y. Zhang, R. Yu, M. Nekovee, Y. Liu, S. Xie, and S. Gjessing, "Cognitive Machine-to-Machine Communications: Visions and Potentials for the Smart Grid", *IEEE Network Magazine*, vol.26, no.3, pp.6–13, May/Jun. 2012.

[8] M. Zhou, L. Zhang, L. Song, M. Debbah, and B. Jiao, "Interference alignment with delayed differential feedback for time-correlated MIMO channels," in *Proc. IEEE ICC* ,pp.3741–3745, Jun. 2012.

[9] Y. Sun and M. L. Honing, "Asymptotic capacity of multicarrier transmission with frequency-selecive fading and limited feedback," *IEEE Trans. Inf. Theory*, vol. 54, no.7, pp. 2879-2902, Jul. 2008.

[10] W. H. Chin and C. Yuen, "Design of Differential Quantization for Low Bitrate Channel State Information Feedback in MIMO-OFDM Systems," in *IEEE VTC-Spring 2008*.

[11] Y. Li, L. J. Cimini, Jr., and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 902–915, Jul. 1998.

[12] M. J. Gans, "A power-spectral theory of propagation in the mobile-radio environment," *IEEE Trans. Veh. Technol.*, vol. 21, pp. 27-38, Feb. 1972.

[13] L. Zhang, L. Song; B. Jiao, and H. Guo, "On the minimum feedback rate of MIMO block-fading channels with time-correlation," in *Proc. CCWMC 2009, IET ICC*, pp.547–550, Shanghai, China, Dec. 2009.

[14] Z. Shi, S. Hong, J. Chen, K. Chen and Y. Sun, "Particle Filter Based Synchronization of Chaotic Colpitts Circuits Combating AWGN Channel Distortion", *Circuits, Systems and Signal Processing (Springer)*, Vol.27(6): 833–845, Dec. 2008.

[15] L. Zhang, L. Song, M. Ma, and B. Jiao, "On the minimum differential feedback rate for time-correlated MIMO Rayleigh block-fading channels," *IEEE Trans. Commun.*, vol. 60, no. 2, pp. 411–420, Feb. 2012.

[16] R. J. McEliece, *The Theory of Information and Coding*, 2nd edition. Cambridge University Press, 2002.

[17] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, 1991.

[18] S. P. Lloyd, "Least-square quantization in PCM," *IEEE Trans. Inf. Theory*, vol. 28, pp. 129-137, Mar. 1982.