where

$$\mathcal{M}_{\gamma_{\rm eq}}(s)=\sum_{l=1}^{N}\binom{N}{l}P_\lambda^l\bar{P}_\lambda^{N-l}l\sum_{m=0}^{l-1}\frac{(-1)^m\binom{l-1}{m}}{m+1}\frac{\mathcal{A}_m}{s+\mathcal{A}_m} \quad (26)$$

$$\mathcal{M}_{\gamma_{SD}}(s)=\int_0^\infty e^{-s\gamma}f_{\gamma_{SD}}(\gamma)d\gamma=\frac{1/\bar{\gamma}_{SD}}{s+1/\bar{\gamma}_{SD}} \quad (27)$$

with $f_{\gamma_{SD}}(\gamma)=(1/\bar{\gamma}_{SD})e^{-(\gamma/\bar{\gamma}_{SD})}$.

By performing inverse Laplace transform on $\mathcal{M}_{\gamma_{\rm total}}(s)$, the joint pdf of $\gamma_{\rm total}$ and $l$ can be derived as

$$f_{\gamma_{\rm total},l\neq 0}(\gamma,l\neq 0)=\sum_{l=1}^{N}\binom{N}{l}P_\lambda^l\bar{P}_\lambda^{N-l}l\sum_{m=0}^{l-1}\frac{(-1)^m\binom{l-1}{m}}{m+1} \times\frac{\mathcal{A}_m}{\mathcal{A}_m\bar{\gamma}_{SD}-1}\{e^{-\gamma/\bar{\gamma}_{SD}}-e^{-\mathcal{A}_m\gamma}\}. \quad (28)$$


## Acknowledgment

The authors would like to thank the anonymous reviewers for their critical comments, which greatly improved this paper.


## References


[1] Z. Zhang, W. Zhang, and C. Tellambura, "OFDMA uplink frequency offset estimation via cooperative relaying," *IEEE Trans. Wireless Commun.*, vol. 8, no. 9, pp. 4450–4456, Sep. 2009.

[2] R. U. Nabar, H. Bolcskei, and F. W. Kneubuhler, "Fading relay channels: Performance limits and space-time signal design," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1099–1109, Aug. 2004.

[3] W. Zhuang and M. Ismail, "Cooperation in wireless communication networks," *IEEE Wireless Commun.*, vol. 19, no. 2, pp. 10–20, Apr. 2012.

[4] Z. Zhang, W. Zhang, and C. Tellambura, "Cooperative OFDM channel estimation in the presence of frequency offsets," *IEEE Trans. Veh. Technol.*, vol. 58, no. 7, pp. 3447–3459, Sep. 2009.

[5] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[6] P. L. Yeoh, M. Elkashlan, Z. Chen, and I. B. Collings, "SER of multiple amplify-and-forward relays with selection diversity," *IEEE Trans. Commun.*, vol. 59, no. 8, pp. 2078–2083, Aug. 2011.

[7] Z. Zhang, C. Tellambura, and R. Schober, "Improved OFDMA uplink transmission via cooperative relaying in the presence of frequency offsets—Part I: Ergodic information rate analysis," *Eur. Trans. Telecommun.*, vol. 21, no. 3, pp. 224–240, Apr. 2010.

[8] M. Torabi, D. Haccoun, and J.-F. Frigon, "Impact of outdated relay selection on the capacity of AF opportunistic relaying systems with adaptive transmission over non-identically distributed links," *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3626–3631, Nov. 2011.

[9] K. B. Fredj and S. Aïssa, "Performance of amplify-and-forward systems with partial relay selection under spectrum-sharing constraints," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 500–504, Feb. 2012.

[10] D. S. Michalopoulos, H. A. Suraweera, G. K. Karagiannidis, and R. Schober, "Amplify-and-forward relay selection with outdated channel estimates," *IEEE Trans. Commun.*, vol. 60, no. 5, pp. 1278–1290, May 2012.

[11] D. Li and A. W. Long, "Outage probability of cognitive radio networks with relay selection," *IET Commun.*, vol. 5, no. 18, pp. 2730–2735, Dec. 2011.

[12] Z. Zhang, K. Long, and J. Wang, "Self-organization paradigms and optimization approaches for cognitive radio technologies: A survey," *IEEE Wireless Commun.*, vol. 20, no. 2, pp. 36–42, Apr. 2013.

[13] Z. Zhang, K. Long, J. Wang, and F. Dressler, "On swarm intelligence inspired self-organized networking: Its bionic mechanisms, designing principles and optimization approaches," *IEEE Commun. Surveys Tuts.*, to appear.

[14] S. I. Hussian, M.-S. Alouini, M. Hasna, and K. Qarage, "Partial relay selection in underlay cognitive networks with fixed gain relays," in *Proc. IEEE VTC*, May 2012, pp. 1–5.

[15] Z. Zhang, C. Tellambura, and R. Schober, "Improved OFDMA uplink transmission via cooperative relaying in the presence of frequency offsets—Part II: Outage information rate analysis," *Eur. Trans. Telecommun.*, vol. 21, no. 3, pp. 241–250, Apr. 2010.

[16] M. O. Hasna and M.-S. Alouini, "A performance study of dual-hop transmissions with fixe gain relays," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 1963–1968, Nov. 2004.

[17] J. G. Proakis and M. Salehi, *Digital Communications*. New York, NY, USA: McGraw-Hill, 2008.

[18] Y. Zhao, R. Adve, and T. J. Lim, "Symbol error rate of selection amplify-and-forward relay systems," *IEEE Commun. Lett.*, vol. 10, no. 11, pp. 757–759, Nov. 2006.

[19] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. New York, NY, USA: Academic, 2007.


# 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 a limited feedback channel can be utilized to increase system throughput. However, in multiple-input–multiple-output (MIMO) systems, the signaling overhead realizing this CSI feedback can be quite large, whereas 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 a 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 a MIMO time–frequency (or doubly)-selective fading channel is investigated. Finally, the analysis is verified by simulation results.**



***Index Terms*—Correlation, differential feedback, multiple-input multiple-output (MIMO).**


## I. Introduction

In multiple-input–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


Manuscript received November 25, 2012; revised March 18, 2013; accepted May 3, 2013. Date of publication June 5, 2013; date of current version November 6, 2013. This work was supported in part by the National 973 project under Grant 2013CB336700, by the National Natural Science Foundation of China under Grant 61222104 and Grant 61061130561, by the Ph.D. Programs Foundation of the Ministry of Education of China under Grant 20110001110102, and by the Opening Project of the Key Laboratory of Cognitive Radio and Information Processing (Guilin University of Electronic Technology). The review of this paper was coordinated by Prof. X. Wang.



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 100871, China (e-mail: mingxin.zhou@pku.edu.cn; lingyang.song@pku.edu.cn).

L. Zhang is with Huawei Technologies Company Ltd., Beijing 100095, China (e-mail: leiming.zhang@pku.edu.cn).

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



Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2013.2266379

system. However, these channel adaptive techniques require accurate channel conditions, which are often referred to channel state information (CSI). Oftentimes, in a frequency-division-duplex 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 lossy compression using the properties of the fading process was considered best. When the wireless channel experiences temporal-correlated fading, which is 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] and [10] by exploiting the frequency correlation.

In summary, all the given 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 simultaneously grow. To the best of the authors' knowledge, the scheme of making full use of the 2-D correlations has yet to be 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 this paper can be briefly summarized as follows: 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 a MIMO doubly selective fading channel.

The rest of this paper is organized as follows: In Section II, we describe the differential feedback model and the statistics of the doubly selective fading channel. In Section III, we propose a differential feedback scheme by exploiting the 2-D 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 downlink channel is a mobile wireless channel, which is always correlated in time and frequency domains, whereas the uplink channel is a limited feedback channel.

### A. Statistics of the Downlink Channel

Since the channel corresponding to each antenna is independent and with the same statistics, we can describe the separation property of channel frequency response $H(t,f)$ at time $t$ for an arbitrary transmit and receive antenna pair [11], i.e.,

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

where $\mathbb{E}\{\cdot\}$ denotes the expectation function, and superscript $(\cdot)^*$ denotes the 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 symbol period $t_s$ and subchannel spacing $f_s$, the correlation function for different periods and subchannels is written as

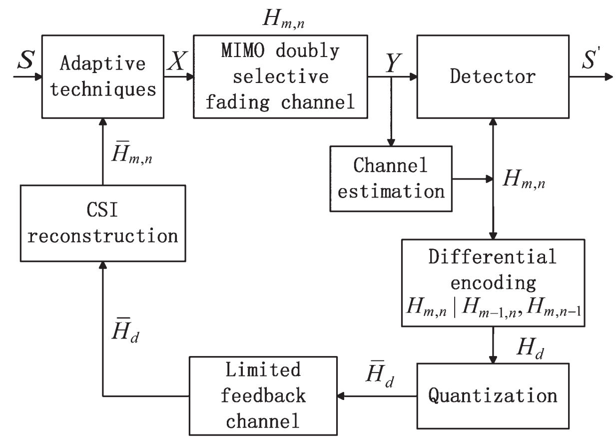


Fig. 1. System model of the differential feedback over MIMO doubly selective fading channels.

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

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], i.e.,

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

where $H_{m,n}$ denotes the channel coefficient of the $m$th symbol interval and the $n$th subchannel, and $W_t$ is a complex white noise variable, which is independent of $H_{m-1,n}$, with variance $\sigma_H^2$. Parameter $\alpha_t$ is the time autocorrelation coefficient, which is given by the zero-order Bessel function of the 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], i.e.,

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

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

### B. Differential Feedback Model

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

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

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

In the given expression, $\mathbf{y}_{m,n}$ denotes the $N_r \times 1$ received vector at the $m$th symbol interval and the $n$th subchannel. $\mathbf{H}_{m,n}$, which is an $N_r \times N_t$ 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, i.e., $\alpha_t$ and $\alpha_f$, respectively.

Moreover, 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 an $N_r \times 1$ additive white Gaussian noise vector with zero mean and variance $\sigma_0^2$. 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 variance $D/N_r N_t$, 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, i.e.,

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

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 following section. Meanwhile, the receiver calculates the differential CSI, given the previous CSI. 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 is obviously the prediction error, and $Diff(\cdot)$ is the differential function. Then, through the limited feedback channel, $\mathbf{H}_d$ should be quantized and fed back.

Finally, the CSI reconstructed by combining the differential CSI 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 at the transmitter. $\bar{\mathbf{U}}$ and $\bar{\mathbf{V}}$ are unitary matrices, and $\bar{\mathbf{\Sigma}}$ is a nonnegative 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]

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

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 $\mathbf{F} = (1/A^2)\mathbf{I}_{N_r} + \mathbb{E}[\mathbf{J}_e\mathbf{J}_e^+|\mathbf{J}]$, in which $A$ represents the amplitude of the signal symbol, and $\bar{\mathbf{Z}}$ denotes a diagonal matrix determined by water-filling [13]–[15], i.e.,

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

where $\bar{\gamma}_{i,i}$, $i = 1, 2, \ldots, N_t$ are the entries of $\bar{\Sigma}$, and $\bar{\mu}$ denotes a cutoff 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

Here, 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 the differential feedback scheme, the receiver feeds back $\mathbf{H}_{m,n}$ to the transmitter. The information entropy of Gaussian variable $X$ with variance $\sigma^2$ is represented as [16]

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

Thus, the feedback load has a positive relation with $\sigma_H^2$.

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

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

where $\inf\{\cdot\}$ denotes the infimum function, $I(\mathbf{H}_{m,n}; \bar{\mathbf{H}}_{m,n})$ denotes the mutual information between $\bar{\mathbf{H}}_{m,n}$ and $\mathbf{H}_{m,n}$, and $d(\mathbf{H}_{m,n}; \bar{\mathbf{H}}_{m,n}) = \|\mathbf{H}_{m,n} - \bar{\mathbf{H}}_{m,n}\|^2$ 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

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

where $d = D/(N_t N_r)$ is the 1-D average channel quantization distortion. $H_{m,n}$ and $\bar{H}_{m,n}$ represent the entries of $\mathbf{H}_{m,n}$ and $\bar{\mathbf{H}}_{m,n}$respectively. In addition, from (6), the 1-D channel quantization is written as

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

The mutual information can be written as

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

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

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

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 nondifferential 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 fewer bits that will be 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

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

The mean square error (MSE) represents the statistical difference between the predicted value and the true value. 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], $a_1$ and $a_2$ are determined by

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

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

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

Moreover, the 1-D 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), $a_1$ and $a_2$ 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 1-D MSE of the channel estimator is

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

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

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

Moreover, 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 $\partial\mathrm{MSE}/\partial\alpha_t < 0$, $\partial\mathrm{MSE}/\partial\alpha_f < 0$, the feedback load can be much smaller than $\sigma_H^2$, which is the nondifferential feedback, particularly when the channel is highly correlated. For example, given $\alpha_t > 0.75$, $\alpha_f > 0.75$, then $\mathrm{MSE}|_{\alpha_t>0.75,\alpha_f>0.75} < \mathrm{MSE}|_{\alpha_t=0.75,\alpha_f=0.75} = 0.28\sigma_H^2$.

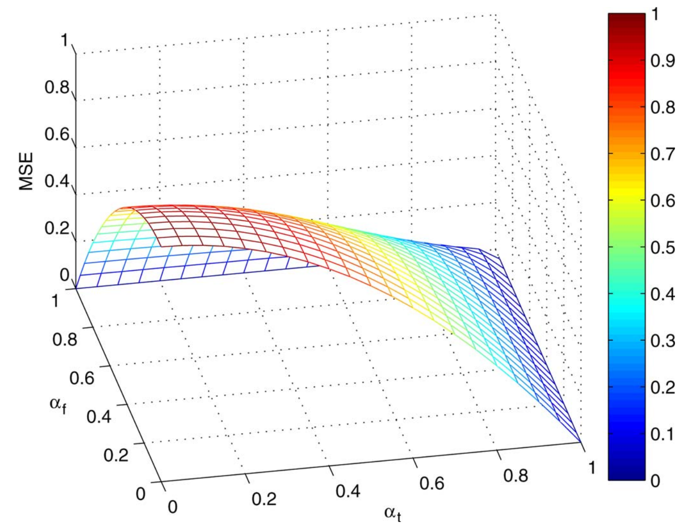


Fig. 2. MSE of the predictor at the transmitter for $N_r = 2$, $N_t = 2$, $\sigma_H^2 = 1$, and $D = 0.1$.

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, i.e.,

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

where channel predictor coefficients $a_1$ and $a_2$ are determined by $a_1 = (\alpha_t(1-\alpha_f^2))/(1-\alpha_t^2\alpha_f^2)$ and $a_2 = (\alpha_f(1-\alpha_t^2))/(1-\alpha_t^2\alpha_f^2)$. The average power of $H_d$ is $\mathrm{Var}(H_d) = \sigma_H^2(1 - a_1^2 - a_2^2 - 2a_1a_2\alpha_t\alpha_f)$. The detailed derivation is given in the Appendix.

The given 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 $\alpha_t$, $\alpha_f$ and channel quantization distortion $d$ and much smaller than that of the nondifferential feedback (17).

## IV. Simulation Results and Discussion

Here, we first provide the relationship between the MSE of the predictor and the 2-D correlations in Fig. 2. The minimal differential feedback rate over MIMO doubly selective fading channels is given in Fig. 3. Then, a longitudinal section in Fig. 3 is presented, where we assume that the temporal correlation and the spectral correlation are equal. Finally, we verify our theoretical results by a practical differential feedback system with the 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 $\alpha_t$ or $\alpha_f$ comes to one, the MSE tends to zero.

Fig. 3 plots the relationship between the minimal differential feedback rate and the 2-D correlations with 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.

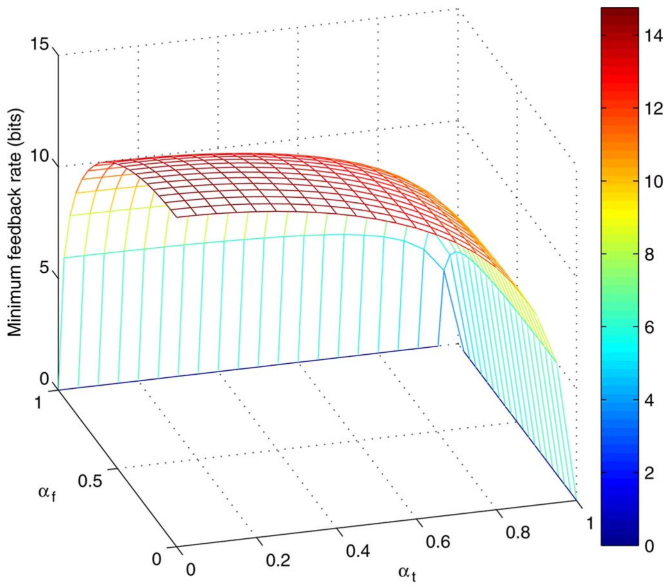


Fig. 3. Minimal differential feedback rate for $N_r = 2$, $N_t = 2$, $\sigma_H^2 = 1$, and $D = 0.1$.

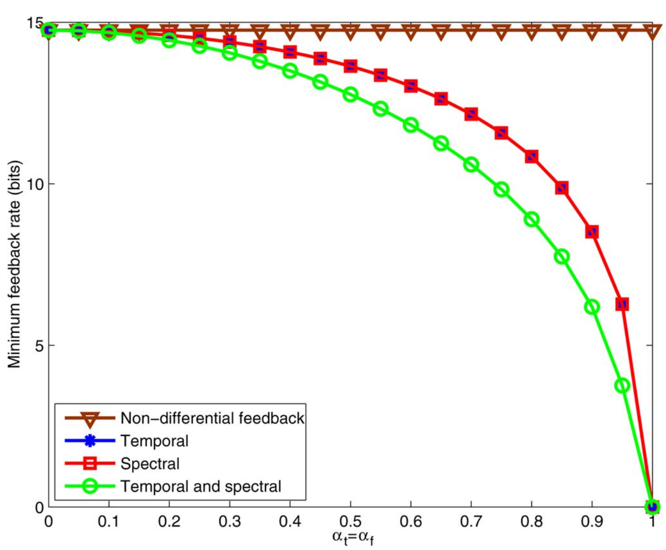


Fig. 4. Relationship between the minimal feedback rate and temporal and spectral correlations, when they are equal, for $N_r = 2$, $N_t = 2$, $\sigma_H^2 = 1$, and $D = 0.1$.

Additionally, because $\alpha_t$ and $\alpha_f$ could be any value, we provide one of the longitudinal section in Fig. 3 where the temporal correlation is equal to the spectral correlation in Fig. 4. For comparison, the differential feedback compression only using the 1-D correlation and the nondifferential feedback scheme are also included in Fig. 4. It is observed in Fig. 4 that the scheme using both temporal and spectral correlations is always better than the scheme using only the 1-D correlation. As the correlations increase, the 2-D differential feedback compression exhibits a significant improvement compared with 1-D compression. This performance advantage even reaches up to 67% with $\alpha_t = \alpha_f = 0.95$.

### *B. Differential Feedback System With Lloyd's Algorithm*

Here, we consider the temporal correlation $\alpha_t = 0.9$ with 2-GHz carrier frequency, 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 $\text{Diff}(\mathbf{H}_{m,n}|\mathbf{H}_{m-1,n}\mathbf{H}_{m,n-1}) = \mathbf{H}_{m,n} - a_1\mathbf{H}_{m-1,n} - a_2\mathbf{H}_{m,n-1}$ as a differential function, where $a_1 = \alpha_t(1-\alpha_f^2)/1 - \alpha_t^2\alpha_f^2$ and $a_2 = (\alpha_f(1-\alpha_t^2))/(1-\alpha_t^2\alpha_f^2)$ in the 2-D differential feedback compression and $a_1 = \alpha_t$ and $a_2 = 0$ in the 1-D compression.

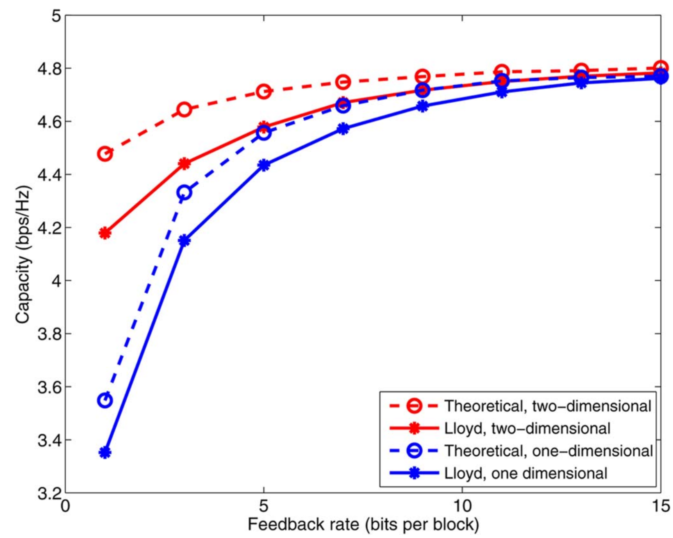


Fig. 5. Relationship between the ergodic capacity and the feedback rate with Lloyd's algorithm in AR1 model for $N_r = 2$, $N_t = 2$, $\sigma_H^2 = 1$, and SNR = 5 dB.

The feedback steps can be summarized as follows. First, based on Lloyd's quantization algorithm, the channel codebook can be generated according to the statistics of the corresponding differential feedback load at both the transmitter and the receiver. Second, the receiver calculates the current differential CSI $\mathbf{H}_d$. Third, the differential CSI is quantized to the optimal codebook 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 in Fig. 5 that the performance of the 2-D compression is always better than the 1-D compression, which verifies our theoretical analysis.

As shown in Fig. 5, with the increase in feedback rate $b$, the ergodic capacities rapidly increase 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. In addition, the capacities of Lloyd's quantization are lower than the theoretical results. 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 channel coefficient $\mathbf{H}$ is correlated in both temporal and spectral domains. 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. Conclusion

In this paper, we have designed a differential feedback scheme making full use of both the temporal and spectral correlations 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 with nondifferential feedback, as the channel is highly correlated in both temporal and spectral domains. Finally, we provide simulations to verify our analysis.

## APPENDIX
## DERIVATION OF THE MINIMAL DIFFERENTIAL FEEDBACK RATE USING TEMPORAL AND SPECTRAL CORRELATIONS

The minimal differential feedback rate over MIMO doubly selective fading channels 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

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

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

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

The 1-D channel quantization equality can be written as

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

Similarly, (28) yields

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

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

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

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

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

Substituting (35) into (34), we obtain

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

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

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

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

$$\mathrm{Var}(a_1 E_{m-1,n} + a_2 E_{m,n-1} + H_d) = a_1^2 d + a_2^2 d + \mathrm{Var}\left(H_d^2\right) + 2a_1 a_2 r(E_{m-1,n}, E_{m,n-1}). \quad (38)$$

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, i.e.,

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

where

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

and $\psi$ is a Gaussian variable with zero mean and variance $(\sigma_{\bar{H}}^2(\sigma_H^2 - \sigma_{\bar{H}}^2))/\sigma_H^2$, independent of $H$.

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

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

Substituting (41) into (38), we obtain

$$\mathrm{Var}(a_1 E_{m-1,n} + a_2 E_{m,n-1} + H_d) = a_1^2 d + a_2^2 d + \sigma_{H_d}^2 + 2a_1 a_2 \frac{d^2}{\sigma_H^2} \alpha_t \alpha_f. \quad (42)$$

From (31), (37), and (42), we obtain

$$R = N_r N_t \log \left\{ a_1^2 + a_2^2 + \frac{2a_1 a_2 \alpha_t \alpha_f d}{\sigma_H^2} + \frac{\mathrm{Var}(H_d)}{d} \right\} \quad (43)$$

where $a_1 = (\alpha_t(1-\alpha_f^2))/(1-\alpha_t^2\alpha_f^2)$, $a_2 = (\alpha_f(1-\alpha_t^2))/(1-\alpha_t^2\alpha_f^2)$, and $\mathrm{Var}(H_d) = \sigma_H^2(1 - a_1^2 - a_2^2 - 2a_1 a_2 \alpha_t \alpha_f)$.

## 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," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 54–59, Oct. 2004.

[2] D. J. Love, R. W. Heath, V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1341–1365, Oct. 2008.

[3] T. Eriksson and T. Ottosson, "Compression of feedback for adaptive transmission and scheduling," *Proc. IEEE*, vol. 95, no. 12, 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, no. 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 Commun. Mag.*, vol. 49, no. 4, pp. 44–52, Apr. 2011.

[6] K. Huang, B. Mondal, R. W. Heath., and J. G. Andrews, Jr., "Markov models for limited feedback MIMO systems," in *Proc. IEEE ICASSP*, 2006, pp. IV-9–IV-12.

[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 Netw. Mag.*, 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*, Jun. 2012, pp. 3741–3745.

[9] Y. Sun and M. L. Honing, "Asymptotic capacity of multicarrier transmission with frequency-selective 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 *Proc. IEEE VTC-Spring*, 2008, pp. 827–831.

[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. VT-21, no. 1, 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. IET ICC CCWMC*, Shanghai, China, Dec. 2009, pp. 547–550.

[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, Syst. Signal Process.*, vol. 27, no. 6, pp. 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 ed. Cambridge, U.K.: Cambridge Univ. Press, 2002.

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

[18] S. P. Lloyd, "Least-square quantization in PCM," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 2, pp. 129–137, Mar. 1982.