FEEDBACK OF DIFFERENTIAL PRECODER FOR GEOMETRICAL MEAN DECOMPOSITION SYSTEMS

Hung-Chun Chen and Yuan-Pei Lin

Dept. Electrical Engr., National Chiao Tung Univ., Hsinchu, Taiwan

ABSTRACT

For a time-correlated channel, we consider the differential feedback of the geometrical mean decomposition (GMD) precoder, which is known to be optimal for a number of criteria. When the channel varies slowly, we can expect the optimal GMD precoders of consecutive channel uses to be close. We consider the feedback of the so-called differential precoder and show that it lies in a neighborhood of the identity matrix using matrix perturbation theory. Furthermore the radius of the neighborhood is proportional to a time-correlation parameter. Such a characterization is crucial for efficient quantization of the differential precoder. Simulations are given to demonstrate that, with a small feedback rate, the performance of the proposed differential GMD comes close to the case when perfect channel state information is available to the transmitter.

Index Terms— MIMO system, precoder, differential feedback, time-correlated channel, geometrical mean decomposition.

1. INTRODUCTION

Multiple-input multiple-output (MIMO) systems with limited feedback has been widely studied in recent years [1]-[6]. It has been demonstrated that limited feedback of channel state information improves the performance significantly. Various feedback schemes have been proposed. Feedback of precoder has been extensively investigated, e.g., [2]-[4]. When there is decision feedback at the receiver, the GMD precoder is shown to be optimal for minimizing bit error rate in [4] and codebook design is addressed therein. Feedback of bit loading is considered in [5][6]. The channel considered in these works is an independent fading channel, not time correlated.

In practical systems, there is substantial temporal correlation between consecutive channel uses. Exploitation of the channel correlation leads to more efficient feedback of channel information [7]-[11]. Channel diagonalizing precoders are parameterized in [7] and the differential changes of the parameters are fed back to the transmitter. With the assumption that the channel Gram matrices of consecutive time instants are along a geodesic curve, differential feedback of the channel Gram matrix is presented in [8]. The temporally correlated channel is modeled as a first-order Gauss-Markov process in [9]-[11] to further exploit the statistics of the channel. Differential feedback of precoder based on rotation matrices is proposed in [9]. In [10], the difference of consecutive channel matrices is fed back to transmitter and differential feedback of bit loading is considered in [11] using predictive quantization.

In this paper, we consider differential quantization of the GMD precoder for a time-correlated channel when the receiver has decision feedback. Differential feedback of precoder has also been addressed earlier [7]-[9], but the precoder considered therein are not optimized for error rate. When the terminal speed is low and the channel varies slowly, we can expect the optimal GMD precoders of consecutive channel uses to be very close. We define the so-called differential precoder between two consecutive optimal precoders and exploit properties of GMD to characterize the differential precoder. Using matrix perturbation theory, we show that the differential precoder lies in a small neighborhood around the identity matrix for a slowly varying Gauss-Markov channel. The radius of the neighborhood is quantified and is shown to be proportional to a time-correlation parameter. The radius can be used in the design of codebooks for efficient quantization of the differential precoder. Simulations are given to demonstrate that, with a small feedback rate, the performance of the proposed differential GMD is only 0.5 dB away from the case when the transmitter has full channel information for a moderate moving speed. The sections are organized as follows. In Sec. 2, we introduce the system model for the time-varying MIMO system. Statistical characterization of the differential precoder is presented in Sec. 3. Simulation examples are shown in Sec. 4 and a conclusion is given in Sec. 5.

This work was supported by Ministry of Science and Technology, Taiwan, R. O. C., under MOST 104-2221-E-009-080-MY2.

2. SYSTEM MODEL

Consider a MIMO communication system with M_t transmit antennas and M_r receive antennas. At time n, the channel is modeled by an $M_r \times M_t$ matrix $\mathbf{H}_n$ whose entries are independent and identically distributed circularly symmetric complex Gaussian random variables with zero mean and unit variance. A useful time correlated channel model is the first-order Gauss-Markov process [12]

$$\mathbf{H}_{n+1} = \sqrt{1-\epsilon^2}\mathbf{H}_n + \epsilon \mathbf{W}_{n+1}, \tag{1}$$

where $\mathbf{W}_{n+1}$ is independent of $\mathbf{H}_n$ and its entries have the same statistics as those of $\mathbf{H}_n$. Using Jake's model [13], $\epsilon = \sqrt{1-(J_0(2\pi f_d T_s))^2}$, where $J_0(\cdot)$ is the zeroth order Bessel function, f_d is the maximum Doppler frequency and T_s is the time interval between consecutive channel uses. The $M_r \times 1$ channel noise vector $\mathbf{q}_n$ is additive white Gaussian with zero mean and variance N_0. The precoder $\mathbf{F}_n$ is an $M_t \times M$ matrix, where M is the number of substreams with $M \leq \min(M_r, M_t)$. The input vector $\mathbf{s}_n$ is assumed to be uncorrelated, and zero mean with $E[\mathbf{s}_n \mathbf{s}_n^\dagger] = P_t/M \mathbf{I}_M$, where P_t is the total transmission power.

Fig. 1. A MIMO communication system.

Let the eigenvalue decomposition of $\mathbf{H}_n^\dagger \mathbf{H}_n$ be $\mathbf{V}_n \mathbf{\Lambda}_n \mathbf{V}_n^\dagger$, where $\mathbf{V}_n$ is an $M_t \times M_t$ unitary matrix and the diagonal matrix $\mathbf{\Lambda}_n$ contains the eigenvalues in nonincreasing order, i.e., $\lambda_{n,0} \geq \lambda_{n,1} \geq \cdots \geq \lambda_{n,M_t-1}$. Let the $M \times M$ leading principal

matrix of $\mathbf{\Lambda}_n$ be $[\mathbf{\Lambda}_n]_M$. Compute the geometric mean decomposition of $[\mathbf{\Lambda}_n]_M^{1/2}$ [14],

$$[\mathbf{\Lambda}_n]_M^{1/2} = \mathbf{Q}_n \mathbf{R}_n \mathbf{P}_n^\dagger, \tag{2}$$

where $\mathbf{Q}_n$ and $\mathbf{P}_n$ are $M \times M$ unitary matrices and $\mathbf{R}_n$ is upper triangular with $[\mathbf{R}_n]_{k,k} = \left(\prod_{j=0}^{M-1} \lambda_{n,j}\right)^{1/2M}$. When the zero-forcing decision feedback equalizer is used, the optimal precoder that minimizes the mean square error and average error rate have been shown to be [4][5]

$$\mathbf{F}_n = \mathbf{V}_{n,0} \mathbf{P}_n, \tag{3}$$

where $\mathbf{V}_{n,0}$ contains the first M columns of $\mathbf{V}_n$. The optimal precoder consists of two parts, $\mathbf{V}_{n,0}$ and $\mathbf{P}_n$. The first part $\mathbf{V}_{n,0}$, formed by the singular vectors of $\mathbf{H}_n$, diagonalizes the channel while the second part $\mathbf{P}_n$ has the effect of equalizing the subchannel SNRs [4] and will be termed the SNR equalizer.

3. FEEDBACK OF DIFFERENTIAL GMD PRECODER

When the channel is time correlated, we can expect the optimal precoder of consecutive channel uses to be correlated as well. Without loss of generality, the precoder at time $n+1$ can be expressed as

$$\mathbf{F}_{n+1} = \mathbf{F}_{a,n} \mathbf{E}_n, \tag{4}$$

where $\mathbf{E}_n$, an $M_t \times M$ semi-unitary matrix satisfying $\mathbf{E}_n^\dagger \mathbf{E}_n = \mathbf{I}_M$, will be called the differential precoder. The notation $\mathbf{A}^\dagger$ denotes the transpose conjugate of a matrix $\mathbf{A}$. The $M_t \times M_t$ matrix $\mathbf{F}_{a,n}$ is given by $\mathbf{F}_{a,n} = [\mathbf{F}_n \ \mathbf{U}_n]$, where $\mathbf{U}_n$ is an $M_t \times (M_t - M)$ matrix chosen in a deterministic manner from $\mathbf{F}_n$ such that $\mathbf{F}_{a,n}$ is unitary. For example, we can choose $\mathbf{F}_{a,n} = [\mathbf{F}_n \ \mathbf{V}_{n,1}]$, where $\mathbf{V}_{n,1}$ is the matrix that contains the last $M_t - M$ columns of $\mathbf{V}_n$. With the feedback of $\mathbf{E}_n$, the transmitter can compute the precoder at $n+1$ from the current precoder. From (4), we can write $\mathbf{E}_n$ as

$$\mathbf{E}_n = \mathbf{F}_{a,n}^\dagger \mathbf{F}_{n+1}. \tag{5}$$

Consider the special case $\mathbf{H}_{n+1} = \mathbf{H}_n$. The differential precoder matrix $\mathbf{E}_n$ in (5) becomes $[\mathbf{I}_M \ \mathbf{0}]^T$. When the channel varies slowly, i.e., $\mathbf{H}_{n+1} \approx \mathbf{H}_n$, it can be expected that $\mathbf{E}_n$ is in the neighborhood of $[\mathbf{I}_M \ \mathbf{0}]^T$. As we will see such a neighborhood can be characterized. To simply the notations, the time index is omitted in the following discussion and $\mathbf{A}_{n+1}$, $\mathbf{A}_n$ are denoted by $\widetilde{\mathbf{A}}$ and $\mathbf{A}$, respectively.

Define the average distance between $\mathbf{E}$ and $[\mathbf{I}_M \ \mathbf{0}]^T$ as

$$D_\epsilon = \sqrt{E\big[\|\mathbf{E} - [\mathbf{I}_M \ \mathbf{0}]^T\|_F^2\big]}, \tag{6}$$

where $\|\mathbf{A}\|_F$ denotes the Frobenius norm of a matrix $\mathbf{A}$. The optimal precoder at time n depends on $\mathbf{V}$ and $\mathbf{P}$. Thus $\mathbf{E}$ depends on $\mathbf{V}$ and $\mathbf{P}$ as well $\widetilde{\mathbf{V}}$ and $\widetilde{\mathbf{P}}$. From [4], we know that $\mathbf{V}$ and $\mathbf{P}$ are statistically independent. It is reasonable that $\mathbf{V}$ and $\widetilde{\mathbf{P}}$ (likewise $\widetilde{\mathbf{V}}$ and $\mathbf{P}$) are also statistically independent. In the following, we derive a bound of D_ϵ that will give us some insight on how D_ϵ is related to the variation of the channel.

Lemma 1. *Assume $\mathbf{P}$ and $\widetilde{\mathbf{P}}$ are respectively statistically independent of $\widetilde{\mathbf{V}}$ and $\mathbf{V}$. The average distance D_ϵ satisfies*

$$D_\epsilon \le \sqrt{E[D_{v,0} + D_{v,1} + D_p] + 2\sqrt{E[D_{v,0}]E[D_p]}}, \tag{7}$$

$$\text{where } \ D_{v,0} = \|\widetilde{\mathbf{V}}_0^\dagger \mathbf{V}_0 - \mathbf{I}_M\|_F^2, \ D_{v,1} = \|\widetilde{\mathbf{V}}_1^\dagger \mathbf{V}_0\|_F^2, \tag{8}$$

$$\text{and } \ D_p = \|\mathbf{P}^\dagger \widetilde{\mathbf{P}} - \mathbf{I}_M\|_F^2. \tag{9}$$

See Appendix A for a proof. Observe from (8) that $\sqrt{D_{v,0} + D_{v,1}}$ represents the distance from $\widetilde{\mathbf{V}}^\dagger \mathbf{V}_0$ to $[\mathbf{I}_M \ \mathbf{0}]^T$. On the other hand, $\sqrt{D_p}$ is the distance between $\mathbf{P}$ and the perturbed matrix $\widetilde{\mathbf{P}}$. Thus $E[D_p]$ corresponds to the perturbation of SNR equalizer and $E[D_{v,0} + D_{v,1}]$ corresponds to the perturbation of eigenspace. In the following two subsections, we quantify the perturbations of eigenspace and SNR equalizer for the Gauss-Markov channel in (1).

3.1. Perturbation of eigenspace

Consider the time-correlated channel model given in (1), we have

$$\widetilde{\mathbf{H}}^\dagger \widetilde{\mathbf{H}} = \mathbf{H}^\dagger \mathbf{H} + \epsilon \sqrt{1 - \epsilon^2} \mathbf{\Delta}_0 + \epsilon^2 \mathbf{\Delta}_1, \tag{10}$$

where $\mathbf{\Delta}_0 = \mathbf{H}^\dagger \widetilde{\mathbf{W}} + \widetilde{\mathbf{W}}^\dagger \mathbf{H}$ and $\mathbf{\Delta}_1 = \widetilde{\mathbf{W}}^\dagger \widetilde{\mathbf{W}} - \mathbf{H}^\dagger \mathbf{H}$. When the time-correlated channel is changing slowly, i.e., ϵ is small, $\widetilde{\mathbf{H}}^\dagger \widetilde{\mathbf{H}}$ can be viewed as a perturbation of $\mathbf{H}^\dagger \mathbf{H}$. There are many results in the literature on the perturbation of matrices [15][16]. In these studies, $D_{v,1} = \|\widetilde{\mathbf{V}}_1^\dagger \mathbf{V}_0\|_F^2$ is regarded as eigenspace variation and various bounds have been derived. However there is no discussion on $D_{v,1} + D_{v,0}$, to the best of our knowledge. In the following, we use a technique similar to that in [15] to derive a bound for $D_{v,1} + D_{v,0}$. Then $D_{v,0}$ can be bounded in a similar manner.

The column vectors of $\mathbf{V}_0$ correspond to the eigenvectors of $\mathbf{H}^\dagger \mathbf{H}$ and they are not uniquely determined. However we can always choose $\widetilde{\mathbf{V}}_0$ such that $[\widetilde{\mathbf{V}}_0^\dagger \mathbf{V}_0]_{jj}$ is a positive real number for all j. In this case, we have the following second order approximation of the perturbation of eigenspace when ϵ is small.

Lemma 2. *Assume $\tilde{\lambda}_i \neq \lambda_j$ for $0 \le i \le M_t - 1$ and $0 \le j \le M - 1$. A second order approximation of $D_{v,0} + D_{v,1}$ is given by*

$$D_{v,0} + D_{v,1} \approx \epsilon^2 \sum_{i=0}^{M_t - 1} \sum_{j=0, j \neq i}^{M-1} |[\widetilde{\mathbf{V}}^\dagger \mathbf{\Delta}_0 \mathbf{V}]_{ij}|^2 / (\tilde{\lambda}_i - \lambda_j)^2.$$

See Appendix B for a proof. Notice that an upper bound for the term $\sum_{i=0}^{M_t-1} \sum_{j=0, j\neq i}^{M-1} |[\widetilde{\mathbf{V}}^\dagger \mathbf{\Delta}_0 \mathbf{V}]_{ij}|^2$ is $\|\widetilde{\mathbf{V}}^\dagger \mathbf{\Delta}_0 \mathbf{V}_0\|_F^2$. Thus Lemma 2 implies

$$D_{v,0} + D_{v,1} \lesssim \epsilon^2 \max_{i,j \in \mathcal{S}_1} \|\mathbf{\Delta}_0 \mathbf{V}_0\|_F^2 / (\tilde{\lambda}_i - \lambda_j)^2,$$

where $\mathcal{S}_1 = \{j \neq i, \ i, j \in \mathbb{N} \mid 0 \le i \le M_t - 1, \ 0 \le j \le M - 1\}$. The above expression depends the distance between $\tilde{\lambda}_i$ and λ_j. Using perturbation theory for matrix eigenvalues [17], the distance satisfies $|\tilde{\lambda}_i - \lambda_i| \lesssim \epsilon \zeta_1$, where $\zeta_1 = \|(\mathbf{H}^\dagger \widetilde{\mathbf{W}} + \widetilde{\mathbf{W}}^\dagger \mathbf{H})\|_2$ and $\|\mathbf{A}\|_2$ denotes the two norm of a matrix $\mathbf{A}$. Using this result and ignoring the third or higher order terms of ϵ, we obtain an upper bound that depends on the singular values of the current channel but not those of the previous channel, given by

$$D_{v,0} + D_{v,1} \lesssim \epsilon^2 \max_{i,j \in \mathcal{S}_1} \frac{1}{(\lambda_i - \lambda_j)^2} \|\mathbf{\Delta}_0 \mathbf{V}_0\|_F^2. \tag{11}$$

Similarly, we can show $D_{v,0} \lesssim \epsilon^2 \max_{i,j \in \mathcal{S}_0} \|\mathbf{\Delta}_0 \mathbf{V}_0\|_F^2 / (\lambda_i - \lambda_j)^2$, where $\mathcal{S}_0 = \{j \neq i, \ i, j \in \mathbb{N} \mid 0 \le i \le M - 1, \ 0 \le j \le M - 1\}$. These results lead to the following result on the average perturbation of eigenspace.

Theorem 1. *Consider the Gauss-Markov channel in (1). For a small ϵ, we have $E[D_{v,0} + D_{v,1}] \lesssim \epsilon^2 \rho_{v,1}$ and $E[D_{v,0}] \lesssim \epsilon^2 \rho_{v,0}$, where*

$$\rho_{v,k} = E_{\mathbf{H}}\big[\max_{i,j \in \mathcal{S}_k} \frac{\eta_0}{(\lambda_i - \lambda_j)^2}\big], \tag{12}$$

and $\eta_0 = M_t \sum_{\ell=0}^{M-1} \lambda_\ell + M \sum_{\ell=0}^{M_t-1} \lambda_\ell$.

Proof. (12) can be obtained by taking expectation of the bound in (11). A sketch of proof is given below. The computation requires averaging over the random matrices $\mathbf{H}$ and $\widetilde{\mathbf{W}}$. To do this, we observe that $\widetilde{\mathbf{W}}$ is independent of the channel $\mathbf{H}$, so the expectation $E[D_{v,0} + D_{v,1}]$ can be obtained using $E_{\mathbf{H}}[E_{\widetilde{\mathbf{W}}}[D_{v,0} + D_{v,1}|\mathbf{H}]]$. Also observe that the elements of $\widetilde{\mathbf{W}}$ are i.i.d. Gaussian with zero mean and unit variance, so $E_{\widetilde{\mathbf{W}}}[\widetilde{\mathbf{W}}^\dagger \mathbf{K} \widetilde{\mathbf{W}}] = tr(\mathbf{K})\mathbf{I}_{M_t}$ and $E_{\widetilde{\mathbf{W}}}[\widetilde{\mathbf{W}}\mathbf{G}\widetilde{\mathbf{W}}] = \mathbf{0}$ for a deterministic $M_r \times M_r$ matrix $\mathbf{K}$ and $M_t \times M_r$ matrix $\mathbf{G}$. The theorem can be proved using these two observations. □

The bound in Theorem 1 shows that the perturbation of eigenspace is proportional to ϵ^2. The constants $\rho_{v,k}$ in (12) depend on $\{\lambda_\ell\}$, the eigenvalues of $\mathbf{H}^\dagger\mathbf{H}$. As the elements of $\mathbf{H}$ are i.i.d complex Gaussian random variables with zero mean and unit variance, the matrix $\mathbf{H}^\dagger\mathbf{H}$ has a Wishart distribution. Thus the joint probability density function for the ordered eigenvalues of a Wishart matrix in [18] can be used to compute the expectation $\rho_{v,k}$ in (12).

3.2. Perturbation of SNR equalizer

We first review a closed form solution of $\mathbf{P}$ that helps to establish a connection between D_p and the perturbation of eigenvalues. To ease the derivation of the perturbation of SNR equalizer, we consider $M = 3$. The case $M = 2$ follows directly. Derivation for a general M can be found in [19].

It is known that the SNR equalizer $\mathbf{P}$ can be expressed as a product of permutation matrices and Givens rotations [14]. Define $\alpha_k = \sqrt{(\lambda_k - \lambda)/(\lambda_k - d_k)}$, $\beta_k = \sqrt{(\lambda - d_k)/(\lambda_k - d_k)}$ for $k = 0,\ 1$, where $\lambda = (\lambda_0\lambda_1\lambda_2)^{1/3}$, $d_0 = \lambda_2$ and $d_1 = \lambda_0\lambda_2/\lambda$. Then $\mathbf{P} = \mathbf{A}_0\mathbf{A}_1$ [14], where

$$\mathbf{A}_0 = \begin{bmatrix} \alpha_0 & -\beta_0 & 0 \\ 0 & 0 & 1 \\ \beta_0 & \alpha_0 & 0 \end{bmatrix},\ \mathbf{A}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \alpha_1 & -\beta_1 \\ 0 & \beta_1 & \alpha_1 \end{bmatrix}. \tag{13}$$

Similarly, $\widetilde{\mathbf{P}}$ can be given in terms of $\tilde{\alpha}_k$ and $\tilde{\beta}_k$. We also define $\theta_k = \sin^{-1}\alpha_k$, $\tilde{\theta}_k = \sin^{-1}\tilde{\alpha}_k$. Then $\delta_k = \tilde{\theta}_k - \theta_k$ represents the perturbation in rotation angles. We show in Appendix C that D_p can be approximated nicely as

$$D_p \approx 2(\delta_0^2 + \delta_1^2). \tag{14}$$

Although the statistical properties of δ_k^2 are not readily available, they can be bounded from above using eigenvalues λ_i that are statistically more tractable, as we will see next. Using $\delta_k \approx \sin\delta_k$ and $\sin\delta_k = \sin(\tilde{\theta}_k - \theta_k)$, we have $\delta_k \approx \alpha_k\beta_k(\tilde{\alpha}_k\alpha_k^{-1} - \tilde{\beta}_k\beta_k^{-1})$ for a small ϵ. Notice that $\tilde{\alpha}_k\alpha_k^{-1}$ and $\tilde{\beta}_k\beta_k^{-1}$ are both close to one. For $x \approx 1$, Taylor approximation yields $\sqrt{x} \approx (x+1)/2$. This means $\tilde{\alpha}_k\alpha_k^{-1} \approx \tilde{\alpha}_k^2\alpha_k^{-2}/2 + 1/2$ and $\tilde{\beta}_k\beta_k^{-1} \approx \tilde{\beta}_k^2\beta_k^{-2}/2 + 1/2$. Thus we have $\delta_k \approx (1 + \tilde{\alpha}_k\alpha_k^{-1})(\tilde{\alpha}_k - \alpha_k)/(2\beta_k)$. Substituting the definitions of α_k and β_k to the approximation and ignoring the second and higher order terms of ϵ, we obtain

$$\delta_k \approx \nu_k \mathbf{c}_k^T \mathbf{y}, \tag{15}$$

where $\mathbf{y} = [\tilde{\lambda}_0 - \lambda_0\ \tilde{\lambda}_2 - \lambda_2\ \tilde{\lambda}_1 - \lambda_1]^T$,

$$\mathbf{c}_0 = \begin{bmatrix} (\frac{3d_0}{\lambda_2 - d_0} - \frac{\lambda}{\lambda_2 - \lambda})\frac{1}{\lambda_0} \\ (-\frac{3\lambda_2}{\lambda_2 - d_0} + \frac{3\lambda_2 - \lambda}{\lambda_2 - \lambda})\frac{1}{\lambda_2} \\ \frac{-\lambda}{(\lambda_2 - \lambda)\lambda_1} \end{bmatrix}, \mathbf{c}_1 = \begin{bmatrix} (\frac{2d_1}{\lambda_1 - d_1} - \frac{\lambda}{\lambda_1 - \lambda})\frac{1}{\lambda_0} \\ (\frac{2d_1}{\lambda_1 - d_1} - \frac{\lambda}{\lambda_1 - \lambda})\frac{1}{\lambda_2} \\ (-\frac{3\lambda_1 + d_1}{\lambda_1 - d_1} + \frac{3\lambda_1 - \lambda}{\lambda_1 - \lambda})\frac{1}{\lambda_1} \end{bmatrix},$$

$\nu_0 = \sqrt{(\lambda_2 - \lambda)/36(\lambda - d_0)}$ and $\nu_1 = \sqrt{(\lambda_1 - \lambda)/36(\lambda - d_1)}$. Applying the Cauchy-Schwarz inequality, we have $\delta_k^2 \lesssim \nu_k^2\|\mathbf{c}_k\|^2 \sum_{j=0}^2 (\tilde{\lambda}_j - \lambda_j)^2$. This leads to

$$D_p \lesssim 2(\nu_0^2\|\mathbf{c}_0\|^2 + \nu_1^2\|\mathbf{c}_1\|^2) \sum_{j=0}^{2} (\tilde{\lambda}_j - \lambda_j)^2. \tag{16}$$

The right hand side of the above equation depends only on $\{\lambda_j\}$ and perturbation of eigenvalues. Using perturbation result for eigenvalues [17], we know $\sum_{j=0}^{M_t-1}(\tilde{\lambda}_j - \lambda_j)^2 \lesssim \epsilon^2\zeta_0$, where $\zeta_0 = \|(\mathbf{H}^\dagger\widetilde{\mathbf{W}} + \widetilde{\mathbf{W}}^\dagger\mathbf{H})\|_F^2$. Thus we can obtain $D_p \lesssim 2\epsilon^2\zeta_0 \sum_{k=0}^{1} \nu_k^2\|\mathbf{c}_k\|^2$. Taking expectation of both sides by using the technique in the proof of Theorem 1, we arrive at the following theorem.

Theorem 2. *Consider the Gauss-Markov channel in (1). When* ϵ *is small,* $E[D_p]$ *can be bounded by* $E[D_p] \lesssim \epsilon^2\rho_p$*, where*

$$\rho_p = 4M_t E_{\mathbf{H}}\Big[\sum_{j=0}^{M_t-1} \lambda_j \sum_{k=0}^{M-2} \nu_k^2\|\mathbf{c}_k\|^2\Big]. \tag{17}$$

Combining Theorem 1 and 2, we obtain

$$D_\epsilon \lesssim \epsilon\gamma, \tag{18}$$

where $\gamma = \sqrt{\rho_{v,1} + \rho_p + 2\sqrt{\rho_{v,0}\rho_p}}$. The constant γ can be computed numerically using the joint probability density function for the ordered eigenvalues of a Wishart matrix [18]. The result means that the differential precoder $\mathbf{E}$ lies in a neighborhood of $[\mathbf{I}_M\ \mathbf{0}]^T$ with radius $\epsilon\gamma$. The radius is proportional to ϵ as the constant γ does not depend on ϵ. We can use this result to design codebook for the differential precoder $\mathbf{E}$. One possible approach is to perturb $[\mathbf{I}_M\ \mathbf{0}]^T$ and apply Gram Schmidt orthogonalization to the perturbed matrix as described in [19]. This codebook design method is adopted in the simulation examples.

4. SIMULATIONS

In the examples, we consider $M_r = 4$, $M_t = 4$ and $M = 3$. The time-correlated channel is generated using the first-order Gauss-Markov process in (1) and the optimal decision feedback receiver [4] is adopted in all the systems compared. Let B_f be the feedback bits within a time instant. The transmission rate is 12 bits per channel, the transmission power is equally divided among all the substreams, and the feedback rate is $B_f = 2$. We have used for $f_c = 2.5{*}10^9$Hz and $T_s = 2$ms as in [20]. In this case, a terminal speed of 3 km/hr, a speed of interest in an indoor or microcellular environment [20], corresponds to $\epsilon = 0.06$.

Example 1. Fig 2 shows the performance of the proposed differential GMD system designed according to the upper bound of D_ϵ in (6) for $\epsilon = 0.02,\ 0.06$, and 0.1, where 0.02 and 0.1 correspond to terminal speed 1 km/hr and 5 km/hr, respectively. We have also shown the case when the system is designed using D_ϵ in (18) that is computed by averaging over 10^5 random channels. The performance of the two are very close for all three cases of ϵ. The upper bound in (18), although an overestimate of D_ϵ, provides a useful estimate of D_ϵ. We have also compared with "unquantized", for which the transmitter has perfect channel knowledge and the precoder is not quantized. For $\epsilon = 0.02$ and $\epsilon = 0.06$, the quantized differential GMD with $B_f = 2$ is within 0.5 dB of "unquantized" when BER$= 10^{-4}$.

Example 2. In this example we show the BER of the proposed method and other feedback systems for the same feedback rate $B_f = 2$. The differential precoder system based on the rotation matrix in

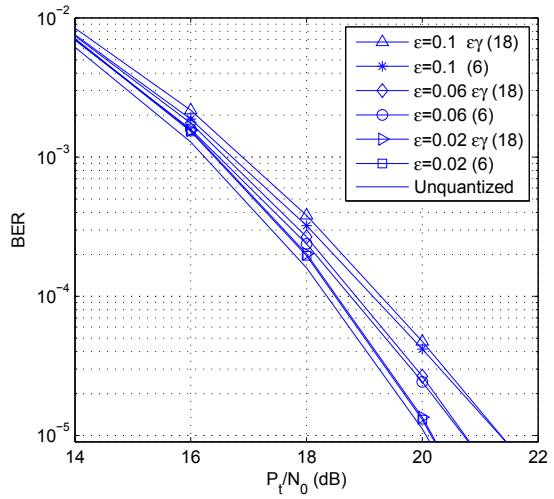

Fig. 2. Example 1. BER performance of the differential GMD system designed according to two radiuses for the differential precoder.

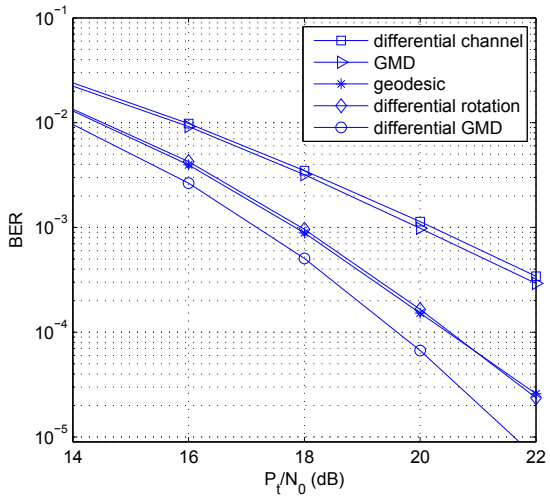

Fig. 3. Example 2. BER performance for $\epsilon = 0.06$.

[9] and the geodesic curves in [8] are labeled as 'differential rotation' and 'geodesic', respectively. For 'differential channel' [10], the feedback information is the difference of consecutive channels. We also compare with the GMD precoder [4] that does not take the correlation into consideration. The results are shown in Fig. 3 for $\epsilon = 0.06$. We can see that the proposed feedback scheme has a good BER performance. This is because GMD with full channel information achieves the minimum BER and the performance of our proposed method is very close to the unquantized GMD.

5. CONCLUSION

In this paper, we consider differential quantization of GMD precoder for a time-correlated channel. Modeling the time-correlated channel as a first-order Gauss-Markov process, we show the differential precoder is in the neighborhood of the identity matrix for a small ϵ. Moreover we derive an upper bound of the radius of the neighborhood and show it is proportional to ϵ. The result is very useful towards the codebook design for the differential precoder, as demonstrated by simulations. With a small feedback rate, the performance of the proposed differential GMD comes close to that of GMD with perfect channel state information at the transmitter.

6. APPENDIX

6.1. Appendix A: Proof of Lemma 1

From (3) and (5), $\mathbf{E}$ can be expressed as $\mathbf{E} = \begin{bmatrix}\mathbf{V}_0\mathbf{P} & \mathbf{V}_1\end{bmatrix}^\dagger \widetilde{\mathbf{V}}_0\widetilde{\mathbf{P}}$, so we have $\|\mathbf{E}-[\mathbf{I}_M\ \mathbf{0}]^T\|_F^2 = \|\mathbf{P}^\dagger\mathbf{V}_0^\dagger\widetilde{\mathbf{V}}_0\widetilde{\mathbf{P}}-\mathbf{I}_M\|_F^2+\|\mathbf{V}_1^\dagger\widetilde{\mathbf{V}}_0\widetilde{\mathbf{P}}\|_F^2$. Using $\mathbf{P}^\dagger\mathbf{V}_0^\dagger\widetilde{\mathbf{V}}_0\widetilde{\mathbf{P}} - \mathbf{I}_M = \mathbf{P}^\dagger(\mathbf{V}_0^\dagger\widetilde{\mathbf{V}}_0 - \mathbf{I}_M)\widetilde{\mathbf{P}} + \mathbf{P}^\dagger\widetilde{\mathbf{P}} - \mathbf{I}_M$ and $\|\mathbf{A}+\mathbf{B}\|_F \le \|\mathbf{A}\|_F + \|\mathbf{B}\|_F$, we obtain $\|\mathbf{P}^\dagger\mathbf{V}_0^\dagger\widetilde{\mathbf{V}}_0\widetilde{\mathbf{P}} - \mathbf{I}_M\|_F^2 \le \|\mathbf{P}^\dagger(\mathbf{V}_0^\dagger\widetilde{\mathbf{V}}_0-\mathbf{I}_M)\widetilde{\mathbf{P}}\|_F^2+2\|\mathbf{P}^\dagger(\mathbf{V}_0^\dagger\widetilde{\mathbf{V}}_0-\mathbf{I}_M)\widetilde{\mathbf{P}}\|_F\|\mathbf{P}^\dagger\widetilde{\mathbf{P}}-\mathbf{I}_M\|_F+\|\mathbf{P}^\dagger\widetilde{\mathbf{P}} - \mathbf{I}_M\|_F^2$. We know that the Frobenius norm is an unitary invariance norm and $\|\mathbf{V}_1^\dagger\widetilde{\mathbf{V}}_0\|_F^2 = \|\widetilde{\mathbf{V}}_1^\dagger\mathbf{V}_0\|_F^2$. Thus we have the upper bound for D_ϵ, $\sqrt{E[D_{v,0}+D_{v,1}+D_p]+2E[\sqrt{D_{v,0}}\sqrt{D_p}]}$. When $\mathbf{P}$ and $\widetilde{\mathbf{P}}$ are respectively statistically independent of $\widetilde{\mathbf{V}}$ and $\mathbf{V}$, we have $E[\sqrt{D_{v,0}}\sqrt{D_p}] = E[\sqrt{D_{v,0}}]E[\sqrt{D_p}]$. Using the Jensen's inequality $E[\sqrt{x^2}] \le \sqrt{E[x^2]}$, the result follows.

6.2. Appendix B: Proof of Lemma 2

Pre- and post-multiplying (10) by $\widetilde{\mathbf{V}}^\dagger$ and $\mathbf{V}$, respectively and considering the ijth element of the left and right hand side, we obtain $[\widetilde{\mathbf{V}}^\dagger\mathbf{V}]_{ij}(\tilde{\lambda}_i-\lambda_j) = \epsilon\sqrt{1-\epsilon^2}[\widetilde{\mathbf{V}}^\dagger\mathbf{\Delta}_0\mathbf{V}]_{ij}+\epsilon^2[\widetilde{\mathbf{V}}^\dagger\mathbf{\Delta}_1\mathbf{V}]_{ij}$. When $\tilde{\lambda}_i \ne \lambda_j$ for $0 \le i \le M_t-1$ and $0 \le j \le M-1$, the ijth element of $\widetilde{\mathbf{V}}^\dagger\mathbf{V}$ can be written as

$$[\widetilde{\mathbf{V}}^\dagger\mathbf{V}]_{ij} = \left(\epsilon\sqrt{1-\epsilon^2}[\widetilde{\mathbf{V}}^\dagger\mathbf{\Delta}_0\mathbf{V}]_{ij} + \epsilon^2[\widetilde{\mathbf{V}}^\dagger\mathbf{\Delta}_1\mathbf{V}]_{ij}\right)/(\tilde{\lambda}_i - \lambda_j). \tag{19}$$

When $[\widetilde{\mathbf{V}}_0^\dagger\mathbf{V}_0]_{jj}$ is a positive real number for all j, $D_{v,0}+D_{v,1}$ in (8) becomes

$$D_{v,0}+D_{v,1} = \sum_{j=0}^{M-1}(1-[\widetilde{\mathbf{V}}_0^\dagger\mathbf{V}_0]_{jj})^2 + \sum_{i=0}^{M_t-1}\sum_{j=0,j\ne i}^{M-1}|[\widetilde{\mathbf{V}}^\dagger\mathbf{V}_0]_{ij}|^2. \tag{20}$$

Using (19) and $[\widetilde{\mathbf{V}}_0^\dagger\mathbf{V}_0]_{jj}^2 = 1-\sum_{i=0,i\ne j}^{M_t-1}|[\widetilde{\mathbf{V}}^\dagger\mathbf{V}_0]_{ij}|^2$, it can be shown that $(1-[\widetilde{\mathbf{V}}_0^\dagger\mathbf{V}_0]_{jj})^2$ is in the order of ϵ^4. The proof can be found in [19]. Thus (20) can be approximated as $D_{v,0}+D_{v,1} \approx \sum_{i=0}^{M_t-1}\sum_{j=0,j\ne i}^{M-1}|[\widetilde{\mathbf{V}}^\dagger\mathbf{V}_0]_{ij}|^2$. Combining the approximation and (19), we obtain the result.

6.3. Appendix C: Proof of (14)

Using the fact that $\mathbf{P}$ and $\widetilde{\mathbf{P}}$ are real and unitary and $\|\mathbf{A}\|_F^2 = tr(\mathbf{A}^\dagger\mathbf{A})$, we can rewrite D_p in (9) as $D_p = 2\sum_{k=0}^{M-1}(1-\widetilde{\mathbf{p}}_k^\dagger\mathbf{p}_k)$. With (13), it can be verified that $\widetilde{\mathbf{p}}_0^\dagger\mathbf{p}_0 = \alpha_0\tilde{\alpha}_0+\beta_0\tilde{\beta}_0$, $\widetilde{\mathbf{p}}_1^\dagger\mathbf{p}_1 = \alpha_1\tilde{\alpha}_1(\alpha_0\tilde{\alpha}_0+\beta_0\tilde{\beta}_0)+\beta_1\tilde{\beta}_1$ and $\widetilde{\mathbf{p}}_2^\dagger\mathbf{p}_2 = \beta_1\tilde{\beta}_1(\alpha_0\tilde{\alpha}_0+\beta_0\tilde{\beta}_0)+\alpha_1\tilde{\alpha}_1$. On the other hand, $\cos\delta_k = \cos(\theta_k-\theta_k)$, i.e., $\cos\delta_k = \alpha_k\tilde{\alpha}_k+\beta_k\tilde{\beta}_k$. When δ_k is small, we obtain $\cos\delta_k \approx 1-\delta_k^2/2$. Using these results, we arrive at (14).

7. REFERENCES

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

[2] D. J. Love and R. W. Heath, "Limited feedback unitary precoding for spatial multiplexing systems," *IEEE Transactions on Information Theory*, vol. 51, pp. 2967-2976, Aug. 2005.

[3] J. C. Roh and B. D. Rao, "Design and analysis of MIMO spatial multiplexing systems with quantized feedback," *IEEE Transactions on Signal Processing*, vol. 54, no. 8, pp. 2874-2886, Aug. 2006.

[4] M. B. Shenouda and T. N. Davidson, "A design framework for limited feedback MIMO systems with zero-forcing DFE," *IEEE Journal on Selected Area in Communications*, vol. 26, no.8, pp. 1578-1587, Oct. 2008.

[5] C. C. Weng, C. Y. Chen, and P. P. Vaidyanathan, "MIMO transceivers with decision feedback and bit loading: theory and optimization," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1334-1346, March 2010.

[6] Y.-P. Lin, H.-C. Chen and P. Jeng, "Bit allocation and statistical precoding for correlated MIMO channels with limited feedback," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 2, pp. 597-606, Feb. 2012.

[7] J. C. Roh and B. D. Rao, "Efficient feedback methods for MIMO channels based on parameterization," *IEEE Transactions on Wireless Communications*, vol. 6, no. 1, pp. 282-292, Jan. 2007.

[8] D. Sacristán-Murga and A. Pascual-Iserte, "Differential feedback of MIMO channel Gram matrices based on geodesic curves," *IEEE Transactions on Wireless Communications*, vol. 9, no. 12, pp. 3714-3727, Dec. 2010.

[9] T. Kim, D. J. Love, and B. Clerckx, "MIMO systems with limited rate differential feedback in slowly varying channels," *IEEE Transactions on Communications*, vol. 59, no. 4, pp. 1175-1189, April 2011.

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

[11] C.-C. Li and Y.-P. Lin, "Predictive coding of bit loading for time-correlated MIMO channels with a decision feedback receiver,"*IEEE Trans. Signal Process.*, vol. 63, no. 13, pp. 3376-3386, July 2015.

[12] R. Etkin and D .N. C. Tse, Degrees of freedom in some underspread MIMO fading channels, IEEE Transaction on Information Theory, vol. 52, no. 4, pp. 1576-1608, April 2006.

[13] J. G. Proakis, *Digital Communications.* New York: McGraw-Hill,1995.

[14] Y. Jiang, W. Wang and J. Li, "The geometric mean decomposition," Linear Algebra and its Applications, vol. 396, pp. 373-384, Feb. 2005.

[15] R.-C. Li, Relative perturbation theory: II. eigenspace and singular subspace variations, SIAM J. Matrix Anal. Appl., vol. 20, pp. 471-492, 1998.

[16] X. Chen, W. Li, "A note on perturbation bounds of eigenspaces for Hermitian matrices," J. Comput. Appl. Math., vol. 196, pp. 338-346, 2006.

[17] R.-C. Li, Relative perturbation theory: I. eigenvalue and singular value variations, SIAM J. Matrix Anal. Appl., vol. 19, pp. 956-982, 1998.

[18] L. G. Ordóñez, D. P. Palomar, and J. R. Fonollosa, "Ordered eigenvalues of a general class of Hermitian random matrices with application to the performance analysis of MIMO systems," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 672-689, Feb. 2009.

[19] H.-C. Chen and Y.-P. Lin, "Differential feedback of geometrical mean decomposition precoder for time-correlated MIMO channels," to be submitted to *IEEE Transactions on Signal Processing*.

[20] ITU-R, Guidelines for Evaluation of Radio Interface Technologies for IMT-Advanced, Report ITU-R M.2135-1, Dec. 2009.