MULTIPLE PRECODER CODEBOOKS FOR MIMO SYSTEMS WITH LIMITED FEEDBACK OF PRECODER AND BIT LOADING

Hung-Chun Chen and Yuan-Pei Lin

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

ABSTRACT

In this paper we propose to use multiple precoder codebooks for MIMO systems with limited feedback of precoder and bit loading. One precoder codebook is designed for each bit loading vector in the bit loading codebook as bit loading contains information on the importance of individual subchannels. For each precoder codebook, we incorporate the associated bit loading in codebook design to minimize transmission power based on sequential vector quantization. There is no need of informing the transmitter which precoder codebook has been used as the information is conveyed by the bit loading feedback. Due to the exploitation of bit loading information, the proposed multi-codebook scheme outperforms the single-codebook scheme significantly as will be demonstrated by simulations.

1. INTRODUCTION

Recently, there has been considerable interest in multi-input multi-output (MIMO) systems with limited feedback [1]-[3]. It has been demonstrated that the system performance can be improved significantly with limited amount of feedback. Commonly adopted types of feedback information are precoder, bit loading, power loading or a combination of these three. The feedback of precoder information has been the most studied [1]-[3]. Codebook designs for unitary precoders using Grassmannian subspace packing are developed in [1] for a number of criteria. A randomly generated codebook is proposed in [2]. In the multimode scheme [3], the number of substreams transmitted can vary with the channel and bits are loaded uniformly.

There has also been research on the feedback of both bit loading and precoder [4]-[5]. A number of optimal MIMO transceivers with decision feedback and bit loading are given in [4]. Bit loading is incorporated in the multimode scheme to further improve the performance in [5], and both precoder and bit loading are fed back. The feedback of precoder and power loading are considered in [6]-[7]. In [6], the codebooks of power loading are designed separately for each mode. Two efficient methods are developed in [7] for parameterizing unitary precoders. It is shown therein that the feedback of power loading provides only slight improvement. Quantization of bit loading is proposed in [8] to reduce the feedback rate.

In this paper, we consider the quantization of precoder for MIMO systems with limited feedback of both precoder and bit loading. As bit loading carries information on the importance of individual subchannels, we propose to use multiple precoder codebooks, each tailored to a bit loading codeword in the bit loading codebook. For each bit loading codeword, we design the associated precoder codebook to minimize transmission power based on sequential vector quantization (SVQ). The multi-codebook approach has an edge over one-codebook scheme as it better exploits the bit loading information. There is no need of informing the transmitter which codebook has been used as the transmitter can obtain the information from the feedback of bit loading. The use of bit-loading dependent codebooks incurs only a small cost of extra codebook storage for the precoder. We demonstrate through simulation examples that the proposed feedback scheme can achieve a very good performance.

2. SYSTEM MODEL

Consider a MIMO communication system with M_t transmit antennas and M_r receive antennas as shown in Fig. 1. The channel is modeled by an $M_r \times M_t$ matrix $\mathbf{H}$ whose entries are independent and identically distributed circularly symmetric complex Gaussian random variables with zero mean and unit variance. The $M_r \times 1$ channel noise vector $\mathbf{n}$ is additive white Gaussian with zero mean and variance N_0. The precoder $\mathbf{F}$ is an $M_t \times M$ matrix with orthonormal columns, where $M = \min(M_r, M_t)$. The input vector $\mathbf{s}$ consists of symbols $s_0,\ s_1, \ldots,\ s_{M-1}$ that are uncorrelated, and zero mean. Let the number of bits loaded on s_k be b_k, then the number of bits transmitted per channel use is $R_b = \sum_{k=0}^{M-1} b_k$. The total transmission power $E[\mathbf{x}^\dagger\mathbf{x}]$ is P_t, where $\mathbf{x} = \mathbf{Fs}$ is the transmitter output vector. The channel output $\mathbf{r}$ is given by $\mathbf{r} = \mathbf{HFs} + \mathbf{n}$. The error vector at the output of the $M \times M_r$ receive matrix $\mathbf{G}$ is $\mathbf{e} = \mathbf{Gr} - \mathbf{s}$, where $\mathbf{G}$ is zero-forcing, given by $\mathbf{G} = (\mathbf{F}^\dagger\mathbf{H}^\dagger\mathbf{HF})^{-1}\mathbf{F}^\dagger\mathbf{H}^\dagger$ [9]. The autocorrelation matrix of the error vector $E[\mathbf{ee}^\dagger]$ is $\mathbf{R}_e = N_0(\mathbf{F}^\dagger\mathbf{H}^\dagger\mathbf{HF})^{-1}$ [9]. Let the eigen decomposition of

This work was supported in part by the National Science Council, Taiwan, under Grant NSC101-2221-E-009-091-MY3 and in part by the International Research Center of Excellence on Advanced Bioengineering under Grant NSC99-2911-I-009-101.

978-1-4673-5577-3/13/$31.00 ©2013 IEEE

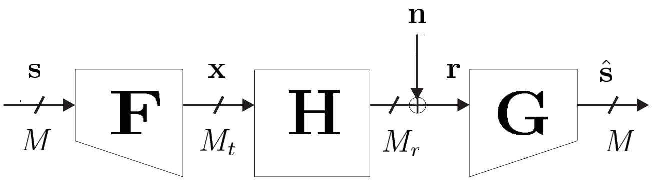

Fig. 1. A MIMO system with M_t transmit antennas and M_r receive antennas.

$\mathbf{H}^\dagger\mathbf{H}$ be $\mathbf{V}\mathbf{\Lambda}\mathbf{V}^\dagger$ where $\mathbf{V}$ is an $M_t \times M_t$ unitary matrix and $\mathbf{\Lambda}$ is a diagonal matrix containing the eigenvalues of $\mathbf{H}^\dagger\mathbf{H}$ in non-increasing order, i.e., $\lambda_0 \geq \lambda_1 \geq \ldots \geq \lambda_{M-1}$. For a number of design criteria, e.g. minimization of transmission power [1][4][7], the optimal unitary precoder has been found to be $\mathbf{F} = \mathbf{V}_M$, where $\mathbf{V}_M$ is the $M_t \times M$ matrix obtained by keeping the first M columns of $\mathbf{V}$. With the above precoder, the kth error variance is given by

$$\sigma^2_{e_k} = [\mathbf{R}_e]_{kk} = N_0\lambda_k^{-1}.$$

As $\{\lambda_k\}$ is in non-increasing order, $\{\sigma^2_{e_k}\}$ is in nondecreasing order. Thus the optimal bit loading vector $\mathbf{b} = \begin{bmatrix} b_0 & b_1 & \cdots & b_{M-1}\end{bmatrix}$ that minimizes the transmission power is in non-increasing order [4]. When only i substreams are transmitted, only the first i entries of $\mathbf{b}$ are nonzero. At the receiver, the bit loading vector $\mathbf{b}$ and precoder matrix $\mathbf{F}$ are chosen from their respective codebooks and the indexes are sent back to the transmitter. Suppose B_b and B_f bits are used to represent $\mathbf{b}$ and $\mathbf{F}$, respectively. The total feedback rate is $B = B_b + B_f$. The bit loading codebook can be designed using the generalized Lloyd algorithm [10] as in [8]. In the next section, we show how bit loading can be incorporated in the design of precoder codebooks.

3. DESIGN OF MULTIPLE PRECODER CODEBOOKS

When both precoder and bit loading are fed back to the transmitter, we can take advantage of the bit loading information to design the feedback of precoder. To do this, we use 2^{B_b} precoder codebooks, one codebook tailored to one bit loading vector. There is no need of informing the transmitter which codebook has been used due to the feedback of bit loading and each precoder codebook has 2^{B_f} codewords. For the incorporation of bit loading in precoder codebook design, we find sequential vector quantization [7] particularly useful. For a given bit loading vector, we show in the following how to design the associated precoder codebook to minimize the average transmission power assuming a large feedback rate.

In SVQ [7], the ideal precoder $\mathbf{V}_M$ is decomposed to a set of unit-norm vectors $\mathbf{q}_0, \ldots, \mathbf{q}_{M-1}$, where $\mathbf{q}_j$ is $(M_t - j) \times 1$. Let $\mathbf{v}_j$ be the jth column of $\mathbf{V}_M$. The vectors $\{\mathbf{q}_j\}$ and $\{\mathbf{v}_j\}$ are related in an iterative manner: $\mathbf{v}_0 = \mathbf{q}_0$, and

$$\mathbf{W}^\dagger(\mathbf{q}_{j-1}) \cdots \mathbf{W}^\dagger(\mathbf{q}_0)\mathbf{v}_j = \begin{bmatrix}\mathbf{0} \\ \mathbf{q}_j\end{bmatrix}, \; 1 \leq j \leq M-1. \quad (1)$$

where $\mathbf{W}(\mathbf{q}_j) = \begin{bmatrix}\mathbf{I}_j & \mathbf{0} \\ \mathbf{0} & \mathbf{C}(\mathbf{q}_j)\end{bmatrix}$ and $\mathbf{C}(\mathbf{q}_j)$ is an $(M_t - j) \times (M_t - j)$ unitary matrix such that $\mathbf{q}_j^\dagger\mathbf{C}(\mathbf{q}_j) = \begin{bmatrix}1 & 0 & \ldots & 0\end{bmatrix}$. The matrix $\mathbf{C}(\mathbf{q}_j)$ can be obtained in a deterministic approach, e.g., Gram-Schmidt process. Suppose $\mathbf{q}_j$ is quantized to $\hat{\mathbf{q}}_j$. Let $\hat{\mathbf{v}}_j$ be the quantized version of $\mathbf{v}_j$ obtained from $\hat{\mathbf{q}}_0, \hat{\mathbf{q}}_1, \cdots, \hat{\mathbf{q}}_j$ using (1).

To measure the average transmission power, note that the total transmission power for a given symbol error rate (SER) can be expressed as

$$P_t = \Gamma \sum_{k=0}^{M-1}(2^{b_k} - 1)\sigma^2_{e_k},$$

where $\Gamma = \frac{1}{3}Q^{-1}(\frac{SER}{4})^2$ and $Q(y) = \frac{1}{\sqrt{2\pi}}\int_y^\infty e^{-t^2/2}dt$, $y \geq 0$ [4]. It is shown in [11] that $\hat{\sigma}^2_{e_j} \approx N_0\lambda_j^{-1}|\mathbf{v}_j^\dagger\hat{\mathbf{v}}_j|^{-2}$ when B_f is large. Using this approximation, the required total transmission power for a quantized precoder is given by

$$P_t \approx \Gamma N_0 \sum_{j=0}^{M-1}(2^{b_j} - 1)\lambda_j^{-1}|\mathbf{v}_j^\dagger\hat{\mathbf{v}}_j|^{-2}.$$

Thus the transmission power averaged over the random channel is

$$\overline{P_t} = E[P_t] \approx \Gamma N_0 \sum_{j=0}^{M-1}(2^{b_j} - 1)E[\lambda_j^{-1}]E[|\mathbf{v}_j^\dagger\hat{\mathbf{v}}_j|^{-2}]. \quad (2)$$

In (2) we have used the property that the singular values and singular vectors of a matrix with independent and identically distributed Gaussian random variables are independent [12]. We show in the appendix that

$$\mathbf{v}_j^\dagger\hat{\mathbf{v}}_j \approx \mathbf{q}_j^\dagger\hat{\mathbf{q}}_j, \text{ for } 0 \leq j \leq M-1 \quad (3)$$

when the feedback rate B_f is sufficiently large. With this approximation, $\overline{P_t}$ becomes

$$\overline{P_t} \approx \Gamma N_0 \sum_{j=0}^{M-1}(2^{b_j} - 1)E[\lambda_j^{-1}]E[|\mathbf{q}_j^\dagger\hat{\mathbf{q}}_j|^{-2}]. \quad (4)$$

We can compute $E[\lambda_j^{-1}]$ numerically using the probability density function of λ_j given in [13]. Given a bit loading vector $\mathbf{b}$ in the bit loading codebook, we allocate B_f bits among $\{\mathbf{q}_j\}$ to minimize $\overline{P_t}$. Suppose $B_f(j)$ bits are used for quantizing $\mathbf{q}_j$, then $\sum_{j=0}^{M-1} B_f(j) = B_f$. Thus we have 2^{B_b} precoder codebooks and each is a concatenation of subcodebooks. (For clarity we call the codebooks for quantizing $\mathbf{q}_j$ *subcodebooks*.) It has been shown in [14] that the probability density function of $|\mathbf{q}_j^\dagger \hat{\mathbf{q}}_j|$ can be approximated as $f_{|\mathbf{q}_j^\dagger \hat{\mathbf{q}}_j|^2}(x) \approx 2^{B_f(j)}(M_t-j-1)(1-x)^{M_t-j-2}1_{[1-\epsilon_j,1)}(x)$, where $0 < x < 1$, $\epsilon_j = 2^{-B_f(j)/(M_t-j-1)}$ and $1_{\mathcal{I}}(x)$ is the indicator function, which is equal to 1 if x in the interval $\mathcal{I}$ and zero otherwise. With the pdf approximation, we can verify that the average transmission power $\overline{P_t}$ is

$$\overline{P_t} \approx \Gamma N_0 \sum_{j=0}^{M-1} (2^{b_j} - 1) E[\lambda_j^{-1}] \Bigg(2^{B_f(j)} (M_t - j - 1) \left[\sum_{\ell=2}^{M_t-j-1} \left(\frac{2^{\frac{-(M_t-j-\ell)}{M_t-j-1} B_f(j)}}{\ell - M_t + j} \right) - \ln \left(1 - 2^{\frac{-B_f(j)}{M_t-j-1}} \right) \right] \Bigg). \quad (5)$$

The optimal rate allocation $B_f(0)$, $B_f(1)$,..., $B_f(M-1)$ that minimizes $\overline{P_t}$ can be obtained by using an exhaustive search of all possible nonnegative integers $B_f(j)$ that add up to B_f. The number of searches required is $(M + B_f - 1)!/((M-1)!B_f!)$, which is a small number for practical values of B_f and M. Having decided the rate allocation among $\mathbf{q}_j$, the subcodebook for quantizing $\mathbf{q}_j$ can be designed using the generalized Lloyd algorithm as in [14]. For every bit loading codeword in the bit loading codebook, we can design the corresponding precoder codebook using the above method. Although we have designed the rate allocation assuming a large feedback rate, the result is useful for practical values of B_f as we see in simulation examples.

4. SIMULATIONS

In the following examples, the elements of the channel matrix $\mathbf{H}$ are independent complex Gaussian random variables with zero mean and unit variance, $M_r = M_t = 4$, $B_f = 5$, $B_b = 3$ and the the number of feedback bits $B = B_f + B_b$ is 8. We have used 10^5 training channels in the design of codebooks for bit loading and precoder. The bit loading codebook is designed using the generalized Lloyd algorithm as in [8] and 2^{B_b} bit-loading dependent precoder codebooks are designed using the method in Sec. 3. For BER simulations, 10^5 channel realizations are used. The power is equally divided among all symbols carrying nonzero bits. The receiver is linear and zero-forcing.

Example 1. In Fig. 2, we compare the multi-codebook and one-codebook schemes of precoder codebook designs. We have shown the results for two different transmission rate, $R_b = 12$ and $R_b = 18$. In the one-codebook case, we compute a fixed rate allocation for $\{\mathbf{q}_j\}$ using (5), assuming uniform bit loading. For the multi-codebook scheme, the receiver does not need to inform the transmitter the codebook used as the transmitter can obtain the information from bit loading. The gain of using multiple precoder codebooks at BER= 10^{-4} is around 2.5 dB for $R_b = 12$ and 1.9 dB for $R_b = 18$. The one-codebook scheme does not perform as well as the mult-codebook case because bits are allocated to $\{\mathbf{q}_j\}$ in a fixed manner even if some symbols are not loaded with bits and the codebook is designed independent of the bit loading used.

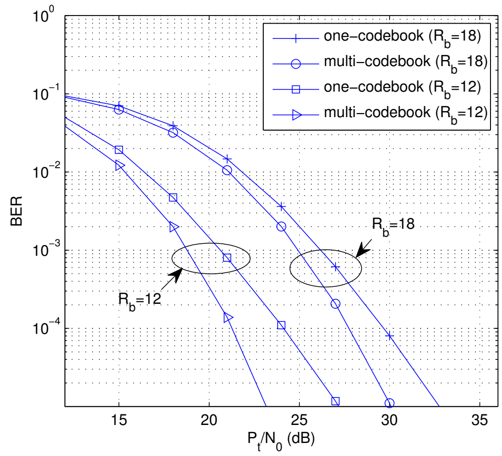

Fig. 2. Example 1. Multiple v.s. one precoder codebooks.

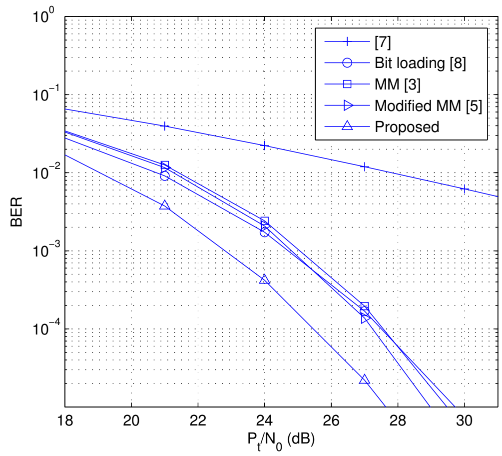

Fig. 3. Example 2. BER comparison for $B = 8$ and $R_b = 16$.

Example 2. In this example we show the BER of the proposed method and other limited feedback systems for $B = 8$ and $R_b = 16$. In [7], the precoder is quantized using SVQ and uniform bit loading is used. In the multimode (MM) precoding system [3], the constellation on all substreams are the same, but the number of substreams transmitted can vary with the channel. The modified multimode precoding in [5] improves the performance of MM in [3] by introducing additional feedback of nonuniform bit loading. In [8] the receiver feeds back only bit loading and the precoder is allocated zero feedback bit. The results are shown in Fig. 3. At BER=10^{-4}, the gap between the proposed system and other systems is

978-1-4673-5577-3/13/$31.00 ©2013 IEEE

1.8 dB. Due to the incorporation of bit loading in precoder codebook design, the proposed system can achieve a better performance.

5. CONCLUSION

In this paper, we considered the use of multiple precoder codebooks for MIMO systems with limited feedback of precoder and bit loading. We proposed to design the precoder codebooks according to the bit loading used because bit loading carries information on the importance of individual subchannels. The use of multi-codebook design yields significant gain over the one-codebook design. There is no need of informing the transmitter which codebook has been used because of bit loading feedback. The use of multiple bit-loading dependent precoder codebooks leads to a very good performance compared to systems that design the feedback of bit loading and precoder separately.

Appendix: Proof of (3)

Using (1), the jth column of the quantized precoder $\widehat{\mathbf{V}}_M$ can be obtained from quantized $\hat{\mathbf{q}}_0, \ldots, \hat{\mathbf{q}}_{M-1}$ as

$$\hat{\mathbf{v}}_0 = \hat{\mathbf{q}}_0,$$
$$\hat{\mathbf{v}}_j = \mathbf{W}(\hat{\mathbf{q}}_0) \cdots \mathbf{W}(\hat{\mathbf{q}}_{j-1}) \begin{bmatrix} \mathbf{0} \\ \hat{\mathbf{q}}_j \end{bmatrix}, \; 1 \leq j \leq M-1. \quad (6)$$

Using (1) and (6), we have $\mathbf{v}_0^\dagger \hat{\mathbf{v}}_0 = \mathbf{q}_0^\dagger \hat{\mathbf{q}}_0$ and

$$\mathbf{v}_j^\dagger \hat{\mathbf{v}}_j = \begin{bmatrix} \mathbf{0} & \mathbf{q}_j^\dagger \end{bmatrix} \mathbf{U}_{j-1}^\dagger \widehat{\mathbf{U}}_{j-1} \begin{bmatrix} \mathbf{0} \\ \hat{\mathbf{q}}_j \end{bmatrix}, \; 1 \leq j \leq M-1, \quad (7)$$

where

$$\mathbf{U}_{j-1} = \mathbf{W}(\mathbf{q}_0) \cdots \mathbf{W}(\mathbf{q}_{j-1})$$

and

$$\widehat{\mathbf{U}}_{j-1} = \mathbf{W}(\hat{\mathbf{q}}_0) \cdots \mathbf{W}(\hat{\mathbf{q}}_{j-1}).$$

In the following, we show that $\mathbf{v}_j^\dagger \hat{\mathbf{v}}_j \approx \mathbf{q}_j^\dagger \hat{\mathbf{q}}_j$ when the feedback rate B_f is large.

Let us use Gram-Schmidt process to obtain $\widetilde{\mathbf{W}}(\mathbf{q}_j)$ from $\mathbf{q}_j$ and $\widetilde{\mathbf{W}}(\hat{\mathbf{q}}_j)$ from $\hat{\mathbf{q}}_j$. Let $\mathbf{A}(\mathbf{q}_j) = \begin{bmatrix} \mathbf{q}_j & \mathbf{1}_1 & \ldots & \mathbf{1}_{M_t-j-1} \end{bmatrix}$ be an $(M_t - j) \times (M_t - j)$ matrix. We can obtain $\widetilde{\mathbf{W}}(\mathbf{q}_j)$ and $\widetilde{\mathbf{W}}(\hat{\mathbf{q}}_j)$ by applying the Gram-Schmidt process to $\mathbf{A}(\mathbf{q}_j)$ and $\mathbf{A}(\hat{\mathbf{q}}_j)$ respectively [15]. Let the kth column of $\widetilde{\mathbf{W}}(\mathbf{q}_j)$ and $\widetilde{\mathbf{W}}(\hat{\mathbf{q}}_j)$ be $\mathbf{w}_k^{(j)}$ and $\hat{\mathbf{w}}_k^{(j)}$ respectively. We have $\mathbf{w}_0^{(j)} = \mathbf{q}_j$. Let $\mathbf{t}_k^{(j)} = \mathbf{1}_k - \sum_{\ell=0}^{k-1} (\mathbf{1}_k^\dagger \mathbf{w}_\ell^{(j)}) \mathbf{w}_\ell^{(j)}$. Then $\mathbf{w}_k^{(j)} = \mathbf{t}_k^{(j)} / \|\mathbf{t}_k^{(j)}\|$ for $1 \leq k \leq M_t - j - 1$. In a similar way, we can obtain $\hat{\mathbf{w}}_k^{(j)}$ from $\hat{\mathbf{q}}_k^{(j)}$. When B_f is sufficiently large, we have $\hat{\mathbf{q}}_j \approx \mathbf{q}_j$, i.e. $\hat{\mathbf{w}}_0^{(j)} \approx \mathbf{w}_0^{(j)}$. Then $\mathbf{t}_1^{(j)} = \mathbf{1}_1 - (\mathbf{1}_1^\dagger \mathbf{w}_0^{(j)}) \mathbf{w}_0^{(j)}$ and $\hat{\mathbf{t}}_1^{(j)} = \mathbf{1}_1 - (\mathbf{1}_1^\dagger \hat{\mathbf{w}}_0^{(j)}) \hat{\mathbf{w}}_0^{(j)}$. Using $\hat{\mathbf{w}}_0^{(j)} \approx \mathbf{w}_0^{(j)}$, we have $\mathbf{t}_1^{(j)} \approx \hat{\mathbf{t}}_1^{(j)}$ and $\|\mathbf{t}_1^{(j)}\| \approx \|\hat{\mathbf{t}}_1^{(j)}\|$. It follows that $\hat{\mathbf{w}}_1^{(j)} \approx \mathbf{w}_1^{(j)}$. With a similar approach, we get $\hat{\mathbf{w}}_k^{(j)} \approx \mathbf{w}_k^{(j)}$ for $2 \leq k \leq M_t - j - 1$. Hence we have $\widetilde{\mathbf{W}}(\mathbf{q}_j) \approx \widetilde{\mathbf{W}}(\hat{\mathbf{q}}_j)$ and $\mathbf{W}(\mathbf{q}_j) \approx \mathbf{W}(\hat{\mathbf{q}}_j)$. Defining $\mathbf{\Phi}_j = \mathbf{W}(\hat{\mathbf{q}}_j) - \mathbf{W}(\mathbf{q}_j)$, we can write

$$\mathbf{W}(\hat{\mathbf{q}}_j) = \mathbf{W}(\mathbf{q}_j) + \mathbf{\Phi}_j, \; j = 0, 1, \ldots, M-1, \quad (8)$$

where the entries of $\mathbf{\Phi}_j$ are small. Using (8), the matrix $\mathbf{U}_{j-1}^\dagger \widehat{\mathbf{U}}_{j-1}$ can be expressed as

$$\mathbf{U}_{j-1}^\dagger \widehat{\mathbf{U}}_{j-1} = \mathbf{I}_{M_t} + \mathbf{\Omega}_{j-1}^{(j-1)} + \mathbf{\Omega}_{j-2}^{(j-1)} + \cdots + \mathbf{\Omega}_0^{(j-1)}, \quad (9)$$

where $\mathbf{\Omega}_{j-1}^{(j-1)} = \mathbf{W}^\dagger(\mathbf{q}_{j-1}) \mathbf{\Phi}_{j-1}$ and

$$\mathbf{\Omega}_k^{(j-1)} = \mathbf{W}^\dagger(\mathbf{q}_{j-1}) \cdots \mathbf{W}^\dagger(\mathbf{q}_k) \mathbf{\Phi}_k \mathbf{W}(\hat{\mathbf{q}}_{k+1}) \cdots \mathbf{W}(\hat{\mathbf{q}}_{j-1}),$$

for $0 \leq k \leq j-2$. Substituting (9) to (7), we get

$$\mathbf{v}_j^\dagger \hat{\mathbf{v}}_j = \mathbf{q}_j^\dagger \hat{\mathbf{q}}_j + \sum_{k=0}^{j-1} \begin{bmatrix} \mathbf{0} & \mathbf{q}_j^\dagger \end{bmatrix} \mathbf{\Omega}_k^{(j-1)} \begin{bmatrix} \mathbf{0} \\ \hat{\mathbf{q}}_j \end{bmatrix}. \quad (10)$$

When the entries of $\mathbf{\Phi}_k$ are small, so are those of $\mathbf{\Omega}_k^{(j-1)}$. Thus we have $\mathbf{v}_j^\dagger \hat{\mathbf{v}}_j \approx \mathbf{q}_j^\dagger \hat{\mathbf{q}}_j$ for $1 \leq j \leq M-1$.

6. REFERENCES

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

[2] W. Santipach and M. L. Honig, "Capacity of a multiple-antenna fading channel with a quantized precoding matrix," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 1218-1234, Mar. 2009.

[3] D. J. Love and R. W. Heath, "Multimode precoding for MIMO wireless systems," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3674-3687, Oct. 2005.

[4] C. C. Weng, C. Y. Chen, and P. P. Vaidyanathan, "MIMO transceivers with decision feedback and bit loading: theory and optimization," *IEEE Trans. Signal Process.*, vol. 58, no. 3, March 2010.

[5] M. Shin, S. Kim, J. Kang, and C. Lee, "An efficient multimode quantized precoding technique for MIMO wireless systems," *IEEE Trans. Vehicular Technology*, vol. 58, no. 2, pp. 733-743, Feb. 2009.

[6] R. Yellapantula, Y. Yao and R. Ansari, "Unitary precoding and power control in MIMO systems with limited feedback," *Proc. IEEE Wireless Communications and Networking Conference*, 2006.

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

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

[9] P. P. Vaidyanathan, S. M. Phoong, and Y.-P. Lin, *Signal Processing and Optimization for Transceiver Systems.* Cambridge Univ. Press, 2010.

[10] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression* . Norwell, MA: Kluwer Academic, 1992.

[11] H.-C. Chen and Y.-P. Lin, "Joint design of limited feedback of precoder and bit loading for MIMO systems," to be submitted.

[12] A. Gupta and D. Nagar, *Matrix Variate Distributions.* London, U.K.:Chapman & Hall/CRC, 2000.

[13] A. Zanella, M. Chiani, and M. Z. Win, "On the marginal distribution of the eigenvalues of Wishart Matrices," *IEEE Trans. Commun.*, vol. 57, no. 4, pp. 1050–1060, Apr. 2009.

[14] J. C. Roh and B. D. Rao, "Transmit beamforming in multiple-antenna systems with finite rate feedback: A VQ-based approach," *IEEE Trans. inf. Theory*, vol. 52, no. 3, pp. 1101-1112, Mar. 2006.

[15] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.