Limited Feedback of Precoder and Bit Loading for MIMO Systems: A Joint Design

Hung-Chun Chen and Yuan-Pei Lin, *Senior Member, IEEE*

***Abstract*—This paper jointly considers limited feedback of bit loading and precoder. In the past when both precoder and bit loading are fed back to the transmitter, the feedback rate is often allocated between the two using an *ad hoc* approach and codebooks are designed separately rather than jointly. In this paper we allocate feedback rate in a systematic manner by analyzing the effect of quantization on transmission power. The analysis allows us to obtain the rate allocation that minimizes the power penalty due to limited feedback. As both precoder and bit loading are fed back to the transmitter, the information embedded in one can be exploited for the design of the other. To take advantage of bit loading feedback, which carries valuable information on the importance of individual subchannels, we employ multiple precoder codebooks, each tailored to a bit loading vector in the bit loading codebook. The multi-codebook scheme enjoys significant gain over the single-codebook case that does not take bit loading feedback into consideration. Simulations are given to demonstrate that the proposed system can achieve a very good performance due to carefully designed feedback rate allocation and joint codebook designs.**

***Index Terms*—Bit loading, joint codebook design, limited feedback, MIMO system, precoder.**

I. Introduction

RECENTLY, there has been considerable interest in multi-input multi-output (MIMO) systems with limited feedback [1]–[9]. 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 [2]–[9]. The precoder is chosen from a codebook using an appropriate selection criterion and the index is fed back to the transmitter. Codebooks designs for unitary precoders using Grassmannian subspace packing are developed in [2] for a number of criteria. An efficient approach to codebook storage and codeword search is given in [3]. A randomly generated codebook is proposed in [4] and the required feedback rate can be computed for a given target spectral efficiency. In [5], the capacity loss of MIMO systems due to precoder quantization is analyzed. In [6], the precoder is selected from the codebook to minimize bit error rate (BER) and the generalized Lloyd algorithm is used to design codebooks. In the multimode scheme [7], the number of substreams transmitted can vary with the channel and bits are loaded uniformly. A capacity maximizing codebook for the multimode scheme is designed in [8] using the generalized Lloyd algorithm. A joint design of precoder and zero-forcing decision feedback equalizer (DFE) for a number of design criteria is developed in [9].

The feedback of bit loading and power loading have been considered in the literature [10]–[13]. An efficient algorithm for per antenna power and rate control is developed in [10]. Successive quantization of power loading and bit loading is considered in [11]. In [12], the receiver feeds back the detection ordering for a fixed bit loading. This is equivalent to having a bit loading codebook that consists of all permutations of a single bit loading vector. An iterative algorithm for designing antenna selection, bit loading, and power loading to minimize the error rate is given in [13]. There has also been research on the feedback of both bit loading and precoder [14]–[16]. A number of optimal MIMO transceivers with decision feedback and bit loading are given in [14]. It is shown therein that when full channel state information is available at the transmitter, these optimal designs have similar performance. When the feedback rate is limited, the use of identity precoder combined with the feedback of only bit loading is proposed. In [15], the ideal unitary precoder is first decomposed using Givens rotation matrices and the feedback rate allocation among the Givens parameters is derived. Bit loading is incorporated in the multimode scheme to further improve the performance in [16], and both precoder and bit loading are fed back. The feedback of precoder and power loading are considered in [17], [18]. In [17], the codebooks of power loading are designed separately for each mode. Two efficient methods are developed in [18] for parameterizing unitary precoders. It is shown therein that the feedback of power loading provides only slight improvement. In [19], the information of power loading, bit loading and precoder are fed back to the transmitter to maximize the transmission rate. As the quantization of bit loading is not considered, a large feedback rate may be needed. Quantization of bit loading is proposed in [20] to reduce the feedback rate.

In this paper, we jointly consider the quantization of both precoder and bit loading for MIMO systems with limited feedback. As there are two types of information in the feedback, the first question to answer is: How to allocate feedback resource? This issue has not been formally addressed in the past. The allocation of feedback rate is often determined in an *ad hoc* manner and the codebooks are usually designed separately rather than jointly. In this paper, we allocate the feedback rate by analyzing

Manuscript received March 08, 2013; revised July 15, 2013; accepted September 03, 2013. Date of publication September 16, 2013; date of current version November 04, 2013. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Walaa Hamouda.

The authors are with the Department of Electrical Engineering, National Chiao-Tung University, Hsinchu 30050, Taiwan (e-mail: hcchen1228@gmail.com; ypl@mail.nctu.edu.tw).

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/TSP.2013.2281785

the power penalty due to limited feedback and jointly design the feedback of precoder and bit loading. We first derive the increase in transmission power when precoder is quantized, and then the additional penalty when bit loading is also quantized. Based on the analysis, the feedback rate is allocated to minimize the combined power penalty for a given transmission rate and target error rate. As both precoder and bit loading are fed back, the information embedded in one can be exploited for the feedback of the other. In particular, bit loading carries valuable information on the importance of individual subchannels. To take advantage of bit loading feedback, we propose to use multiple precoder codebooks, each tailored to a bit loading vector in the bit loading codebook. We show how to incorporate bit loading in precoder codebook design for power minimization based sequential vector quantization method [18]. The multi-codebook approach has an edge over the single-codebook scheme as it better exploits the bit loading information. Furthermore, because of precoder feedback, we can consider bit loading that is in nonincreasing order, which effectively reduces quantization error for the same feedback rate. We demonstrate through examples that the proposed feedback scheme can achieve a very good performance.

The main contributions of this paper are summarized as follows. We analyze the power penalty due to quantization of bit loading and precoder. Based on the results, we determine the feedback rate allocation between bit loading and precoder using a systematic approach. This stands in contrast to earlier works that consider feedback of precoder and bit loading, in which the rate allocation is usually determined in an *ad hoc* manner. We propose the use of multiple precoder codebooks and incorporate bit loading in codebook design to minimize transmission power. The joint design allows us to exploit the information of bit loading feedback and to achieve a better performance. This is different from earlier codebook designs, for which precoder and bit loading codebooks are designed separately, and in most cases only the design of precoder codebook or the design of bit loading codebook is considered, but not both.

The sections are organized as follows. Section II gives the system model for a precoded MIMO system. The feedback rate allocation between precoder and bit loading is derived in Section III. The design of precoder codebook is presented in Section IV. The design of bit loading codebook is discussed in Section V. The design procedure and codeword selection criteria are given in Section VI. Simulation examples are shown in Section VII and a conclusion is given in Section VIII.

Notation: 1) Boldfaced lower case letters represent vectors and boldfaced upper case letters are reserved for matrices. The notation $\mathbf{A}^\dagger$ denotes transpose-conjugate of $\mathbf{A}$. 2) The function $E[y]$ denotes the expected value of a random variable y. 3) The notation $|\mathcal{S}|$ denotes the number of elements in a set $\mathcal{S}$. 4) The notation $\|\mathbf{x}\|$ denotes the 2-norm of a vector $\mathbf{x}$.

II. System Model

Consider the MIMO communication system with M_t transmit antennas and M_r receive antennas 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 $\mathbf{x}$ is the transmitter output vector as indicated in Fig. 1. 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 $\mathbf{e} = \hat{\mathbf{s}} - \mathbf{s} = \mathbf{Gr} - \mathbf{s}$, where $\mathbf{G}$ is zero-forcing, given by $\mathbf{G} = \left(\mathbf{F}^\dagger\mathbf{H}^\dagger\mathbf{HF}\right)^{-1}\mathbf{F}^\dagger\mathbf{H}^\dagger$ [21]. The autocorrelation matrix of the error vector $\mathbf{R}_e = E[\mathbf{ee}^\dagger]$ is [21]

$$\mathbf{R}_e = N_0\left(\mathbf{F}^\dagger\mathbf{H}^\dagger\mathbf{HF}\right)^{-1}. \tag{1}$$

Let the eigen decomposition of $\mathbf{H}^\dagger\mathbf{H}$ be $\mathbf{V}\begin{bmatrix}\mathbf{\Lambda} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\mathbf{V}^\dagger$, where the $M \times M$ diagonal matrix $\mathbf{\Lambda}$ contains the eigenvalues of $\mathbf{H}^\dagger\mathbf{H}$ in nonincreasing order, i.e., $\lambda_0 \geq \lambda_1 \geq \ldots \geq \lambda_{M-1}$, and $\mathbf{V}$ is an $M_t \times M_t$ unitary matrix. For a number of design criteria, e.g., minimization of transmission power [2], [14], [18], the optimal unitary precoder has been found to be

$$\mathbf{F} = \mathbf{V}_M, \tag{2}$$

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

$$\sigma_{e_k}^2 = [\mathbf{R}_e]_{kk} = N_0\lambda_k^{-1}. \tag{3}$$

As $\{\lambda_k\}$ is in nonincreasing order, $\{\sigma_{e_k}^2\}$ is in nondecreasing order. The optimal bit loading $\{b_k\}$ that minimizes the transmission power for a given transmission rate R_b is thus in nonincreasing order [14].

In this paper, we consider the limited feedback of precoder and bit loading. At the receiver, the bit loading vector $\mathbf{b} = [b_0\ b_1\ \ldots\ b_{M-1}]$ 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 feedback rate allocation between bit loading and precoder is considered in the next section.

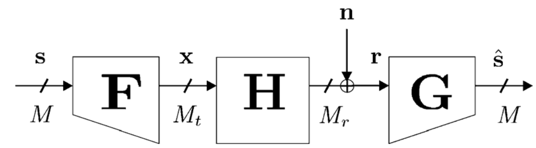

Fig. 1. The MIMO communication system.

III. Allocation of Feedback Rate

In this section, we allocate the feedback rate between precoder and bit loading by considering the increase in transmission power due to the quantization of precoder and bit loading with a high feedback rate assumption. First we analyze the power penalty when only the precoder is quantized. Then we derive the additional penalty when bit loading is also quantized. The results are used to determine feedback rate allocation between precoder and bit loading.

A. Performance Loss Due to Precoder Quantization

For a given channel, the kth subchannel error variance $\sigma_{e_k}^2$ can be computed from (1). The total transmission power for a given bit loading and symbol error rate (SER) can be expressed as [14]

$$P_t = \Gamma \sum_{k=0}^{M-1} (2^{b_k} - 1)\sigma_{e_k}^2, \tag{4}$$

where $\Gamma = \frac{1}{3}Q^{-1}\left(\frac{SER}{4}\right)^2$ and $Q(y) = \frac{1}{\sqrt{2\pi}}\int_y^\infty e^{-t^2/2}dt$, $y \geq 0$. When the transmission rate is large, $2^{b_k} - 1 \approx 2^{b_k}$, the total transmission power can be approximated as $P_t \approx \Gamma\sum_{k=0}^{M-1} 2^{b_k}\sigma_{e_k}^2$. In this case, it is shown in [14] that for a given transmission rate R_b the minimized transmission power with the optimal bit loading is given by

$$P_t^* \approx M\Gamma 2^{R_b/M}\prod_{k=0}^{M-1}\sigma_{e_k}^{2/M}. \tag{5}$$

For the case the precoder is not quantized (i.e., $\mathbf{F} = \mathbf{V}_M$), we can use $\sigma_{e_k}^2 = N_0\lambda_k^{-1}$ in (3) to obtain $P_t^* \approx M\Gamma 2^{R_b/M}\prod_{k=0}^{M-1}\left(N_0\lambda_k^{-1}\right)^{1/M}$. Let $\hat{\sigma}_{e_k}^2$ be the kth subchannel variance when the precoder is quantized. In this case, the minimized transmission power becomes $\widehat{P}_t^* \approx M\Gamma 2^{R_b/M}\prod_{k=0}^{M-1}\hat{\sigma}_{e_k}^{2/M}$. A useful approximation of $\hat{\sigma}_{e_k}^2$ is given in the following Lemma.

Lemma 1: Consider the case the precoder is a quantized version of $\mathbf{V}_M$ in (2), $\mathbf{F} = \widehat{\mathbf{V}}_M$. When the feedback rate B_f is sufficiently large and the channel $\mathbf{H}$ has full rank, the kth subchannel error variance can be approximated as

$$\hat{\sigma}_{e_k}^2 \approx N_0\lambda_k^{-1}\left|\mathbf{v}_k^\dagger\hat{\mathbf{v}}_k\right|^{-2}, \tag{6}$$

where $\mathbf{v}_k$ and $\hat{\mathbf{v}}_k$ are respectively the kth column of $\mathbf{V}_M$ and $\widehat{\mathbf{V}}_M$.

See Appendix A for a proof. Using (6), we have the approximation

$$\begin{aligned}\widehat{P}_t^* &\approx M\Gamma 2^{R_b/M}\prod_{k=0}^{M-1}\left(N_0\lambda_k^{-1}\left|\mathbf{v}_k^\dagger\hat{\mathbf{v}}_k\right|^{-2}\right)^{1/M}\\ &\approx P_t^*\prod_{k=0}^{M-1}\left|\mathbf{v}_k^\dagger\hat{\mathbf{v}}_k\right|^{-2/M}.\end{aligned}$$

Therefore the transmission power is increased by $10\log_{10}\frac{\widehat{P}_t^*}{P_t^*} \approx 10\log_{10}(\prod_{k=0}^{M-1}|\mathbf{v}_k^\dagger\hat{\mathbf{v}}_k|^{-2/M})$ dB. We define the power penalty due to precoder quantization as

$$D_f = E\left[\frac{-10}{M}\sum_{k=0}^{M-1}\log_{10}\left(\left|\mathbf{v}_k^\dagger\hat{\mathbf{v}}_k\right|^2\right)\right]. \tag{7}$$

Lemma 2: Let the entries of the channel $\mathbf{H}$ be independent Gaussian random variables with zero mean and unit variance. When $\mathbf{v}_k$ is quantized to $\hat{\mathbf{v}}_k$ using B_{v_k} bits, the power penalty of precoder quantization is given by

$$D_f \approx \frac{10(M_t-1)}{\ln 10}\sum_{k=0}^{M-1}\left(\frac{2^{B_{v_k}}}{M}\sum_{j=0}^{M_t-2}\left[(-1)^j\frac{C(M_t-2,j)}{j+1}\left(\left(1-2^{\frac{-B_{v,k}}{M_t-1}}\right)^{j+1}\ln\left(1-2^{\frac{-B_{v,k}}{M_t-1}}\right)+\frac{1-\left(1-2^{\frac{-B_{v,k}}{M_t-1}}\right)^{j+1}}{j+1}\right)\right]\right), \tag{8}$$

where $C(\cdot,\cdot)$ denotes the choose function.

Proof: See Appendix B.

B. Performance Loss Due to Bit Loading Quantization

For a given quantized precoder, we can compute the subchannel error variances $\hat{\sigma}_{e_k}^2$, the optimal bit loading $\hat{b}_k^*$ corresponding to $\hat{\sigma}_{e_k}^2$, and the minimum transmission power $\widehat{P}_t^*$. From [14], we know the optimal bit loading $\{\hat{b}_k^*\}$ that minimizes the transmission power satisfies $2^{\hat{b}_k^*} = \widehat{P}_t^*/(M\Gamma\hat{\sigma}_{e_k}^2)$. Suppose now we quantize $\hat{b}_k^*$ to $\hat{b}_k$ (quantization of $\hat{b}_k$ to be discussed later), the required transmission power using the quantized bit loading $\widehat{P}_t \approx \Gamma\sum_{k=0}^{M-1}2^{\hat{b}_k}\hat{\sigma}_{e_k}^2$ can be rewritten as $\widehat{P}_t \approx \Gamma\sum_{k=0}^{M-1}2^{\hat{b}_k^*}\hat{\sigma}_{e_k}^2 2^{(\hat{b}_k-\hat{b}_k^*)} = \frac{\widehat{P}_t^*}{M}\sum_{k=0}^{M-1}2^{(\hat{b}_k-\hat{b}_k^*)}$. Hence the transmission power is increased by $\frac{\widehat{P}_t}{\widehat{P}_t^*} \approx \frac{1}{M}\sum_{k=0}^{M-1}2^{(\hat{b}_k-\hat{b}_k^*)}$. Note that $\widehat{P}_t/\widehat{P}_t^*$ is larger than one since $\widehat{P}_t^*$ is the minimum transmission power when the quantized precoder is given. We define the power penalty due to the quantization of bit loading as

$$D_b = E\left[10\log_{10}\left(\frac{1}{M}\sum_{k=0}^{M-1}2^{\left(\hat{b}_k-\hat{b}_k^*\right)}\right)\right]. \tag{9}$$

When the precoder is not quantized, the optimal bit loading $\{b_k^*\}$ is in nonincreasing order [14]. If B_f is large and the quantization error of the precoder is small, we can assume that the optimal bit loading $\{\hat{b}_k^*\}$ is also in nonincreasing order. We can verify that the nonincreasing property of $\{\hat{b}_k^*\}$ and the fact $\sum_{k=0}^{M-1}\hat{b}_k^* = R_b$ imply that $\hat{b}_k^*$ are bounded as follows: $b_{k,min} \leq \hat{b}_k^* \leq b_{k,max}$, where $b_{k,max} = R_b/(k+1), 0 \leq k \leq M-1, b_{0,min} = R_b/M$, $b_{k,min} = 0, 1 \leq k \leq M-1$.

Suppose $B_{b,k}$ bits are used for scalar quantization of $\hat{b}_k^*$ for $1 \leq k \leq M-1$ and $\hat{b}_0$ is chosen as $R_b - \sum_{k=1}^{M-1}\hat{b}_k$ to satisfy the transmission rate constraint. Define the quantization error $\delta_k = \hat{b}_k - \hat{b}_k^*$. It is known that [22], the quantization error δ_k has a uniform distribution over $(-\Delta_k/2, \Delta_k/2]$ for $1 \leq k \leq M-1$ when $B_{b,k}$ is reasonably large, where $\Delta_k = (b_{k,max} - b_{k,min})2^{-B_{b,k}}$ is the quantization step size. In this case, we can obtain an approximation of D_b, as given in the following lemma.

Lemma 3: Suppose the quantization error $\delta_k = \hat{b}_k - \hat{b}_k^*$ for $1 \leq k \leq M-1$ are independent and uniformly distributed over

$(-\Delta_k/2, \Delta_k/2]$ and $\delta_0 = -\sum_{k=1}^{M-1} \delta_k$. The power penalty of bit loading quantization D_b can be approximated by (10) at the bottom of this page.

Proof: See Appendix C.

Rate Allocation: Starting from the optimal precoder $\mathbf{F} = \mathbf{V}_M$ and the optimal bit loading, the performance is degraded by D_f (dB) when the precoder is quantized. When we further quantize the bit loading, there is an additional degradation of D_b (dB). Therefore we can minimize the power penalty by allocating the rate such that the combined penalty $D_f + D_b$ is minimized. From (7), we see that the quantization of each $\mathbf{v}_k$ contribute to D_f in the same manner, so we choose $B_{v_k} = B_f/M$ for $0 \leq k \leq M-1$. For $0 \leq B_b \leq B$, we evaluate $D_f + D_b$ for all possible integer $\{B_{b,k}\}$ that satisfy $\sum_{k=1}^{M-1} B_{b,k} = B_b$ and choose the one that has the smallest combined power penalty. For each B_b, the number of iterations is $C(B_b+M-2, B_b)$. The number of iterations for finding the optimal rate allocation is $\sum_{\ell=0}^{B} C(M-2+\ell, \ell) = C(M-1+B, B)$, where we have used the Pascal's triangle $C(n,k) = C(n-1,k-1)+C(n-1,k)$ for any nonnegative integer n such that $n > k > 0$. The complexity is not high for practical cases of M and B. For instance, when $M = 4$, $B = 8$, $C(M-1+B, B)$ is 165. Having determined B_b, we design the bit loading codebook with 2^{B_b} codewords in Section V.

The expressions of power penalty in (8) and (10) are obtained with the assumption that B is sufficiently large. However, simulations show that (8) and (10) are good approximations of the actual power penalties even for a moderate B. In Fig. 2 we plot the differences between the approximated power penalty computed using (8) and (10) and the simulated penalty for $M_r = M_t = 3$ and $R_b = 16$. For a given feedback rate B, we find the optimal B_b that minimizes the sum of (8) and (10). The optimal B_b is equal to 3, 4, 5, and 6 respectively for the following 4 ranges of B: (1) $B = 8$, (2) $9 \leq B \leq 13$, (3) $14 \leq B \leq 19$, and (4) $B = 20$. For example, for $B = 10$ we have $B_b = 4$; the precoder and bit loading are quantized using respectively 6 and 4 bits. The simulated penalty is computed by actually quantizing the precoder and the bit loading using the above choice of B_b and B_f, and the penalty is averaged over 10^5 channel realizations. The difference between the simulated and approximated power penalty "D(simulated)$-D((8)+(10))$" is less than 0.3 dB for $B \geq 12$.

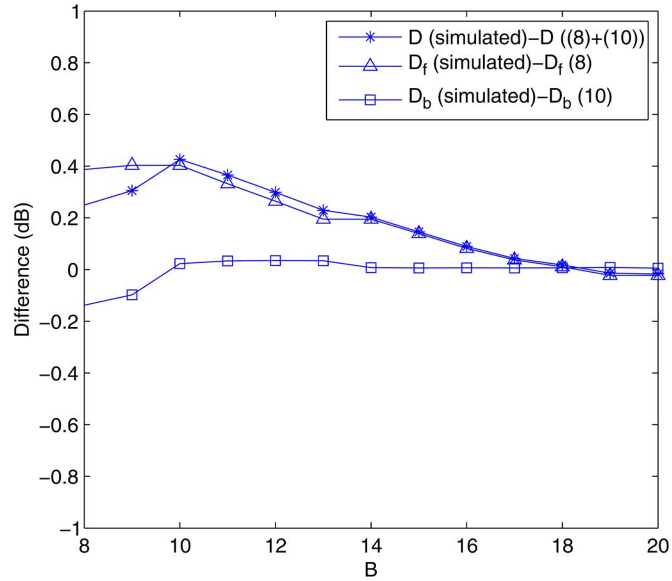

Fig. 2. The differences between the approximated and simulated power penalty for $M_r = M_t = 3$ and $R_b = 16$.

IV. Design of Precoder Codebooks

As both precoder and bit loading are fed back to the transmitter, we can take advantage of bit loading in designing the precoder codebook. We propose to use multiple precoder codebooks, one codebook tailored to one bit loading codeword. There is no need to inform the transmitter which codebook has been used due to the feedback of bit loading and each codebook has 2^{B_f} codewords. Given a bit loading vector, the corresponding precoder codebook is chosen to quantize the precoder. The codebooks can be obtained using codebooks designs for unitary precoders, for example, Grassmannian method [2], random vector quantization (RVQ) [4], and sequential vector quantization (SVQ) [18]. Efficient implementation is possible with SVQ because it decomposes the ideal precoder into some M unit-norm vectors and these vectors are quantized using M smaller subcodebooks. (The total number of codewords in the M subcodebooks is 2^{B_f}.) However there has been no systematic method for rate allocation among the subcodebooks. In the following, we show how to take bit loading into consideration and allocate rate assuming a large feedback rate.

$$D_b \approx 10\log_{10}\frac{1}{M} + \frac{10}{\ln 10}\Bigg(\ln(v) - \sum_{k=0}^{M-1}\left[\frac{1}{2v^2}\left(\mu_{2,k} - \mu_{1,k}^2\right) + \frac{1}{3v^3}\left(\mu_{3,k} - 3\mu_{2,k}\mu_{1,k} + 2\mu_{1,k}^3\right)\right] + \frac{1}{v^3}\sum_{k=1}^{M-1}\Big[\mu_{2,0}\mu_{1,k}\left(\mu_{2,k}^{-1} - 1\right)$$
$$+ \left(v\mu_{1,0} + 2\mu_{1,0}^2\right)\left(\mu_{1,k} - \mu_{1,k}^{-1}\right) + \mu_{1,0}\left(2\mu_{1,k}^2 - \mu_{2,k} - 1\right)\Big] + 2\frac{\mu_{1,0}}{v^3}\sum_{k=1}^{M-2}\sum_{\ell=k+1}^{M-1}\left[\mu_{1,k}^{-1}\mu_{1,\ell}^{-1} - \mu_{1,k}^{-1}\mu_{1,\ell} - \mu_{1,k}\mu_{1,\ell}^{-1} + \mu_{1,k}\mu_{1,\ell}\right]\Bigg), \quad (10)$$

where $\mu_{n,k} = E[2^{n\delta_k}] = \begin{cases} \prod_{\ell=1}^{M-1} \frac{2^{n\Delta_\ell/2} - 2^{-n\Delta_\ell/2}}{n\Delta_\ell \ln 2}, & k = 0, \\ \frac{2^{n\Delta_k/2} - 2^{-n\Delta_k/2}}{n\Delta_k \ln 2}, & 1 \leq k \leq M-1, \end{cases}$ and $v = \sum_{k=0}^{M-1} \mu_{1,k}$.

.

Suppose we are to design a precoder codebook corresponding to a bit loading codeword $\mathbf{b}$ in the bit loading codebook. Using (6) and (4), the required total transmission power for a quantized precoder can be approximated as $P_t \approx \Gamma N_0 \sum_{j=0}^{M-1}(2^{b_j}-1)\lambda_j^{-1}|\mathbf{v}_j^\dagger \hat{\mathbf{v}}_j|^{-2}$. The transmission power averaged over the random channel is given by

$$\overline{P_t} = E[P_t] \approx \Gamma N_0 \sum_{j=0}^{M-1}(2^{b_j}-1)E\left[\lambda_j^{-1}\right] E\left[\left|\mathbf{v}_j^\dagger \hat{\mathbf{v}}_j\right|^{-2}\right], \quad (11)$$

where 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 [23] and $E[\lambda_j^{-1}|\mathbf{v}_j^\dagger\hat{\mathbf{v}}_j|^{-2}] = E[\lambda_j^{-1}]E[|\mathbf{v}_j^\dagger\hat{\mathbf{v}}_j|^{-2}]$. Note that $\lambda_j > 0$ with probability one for $0 \le j \le M-1$ [23], so $E[\lambda_j^{-1}]$ exists. In SVQ, 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 [18],

$$\mathbf{v}_0 = \mathbf{q}_0, \mathbf{W}^\dagger(\mathbf{q}_{j-1})\cdots\mathbf{W}^\dagger(\mathbf{q}_0)\mathbf{v}_j = \begin{bmatrix}\mathbf{0}\\ \mathbf{q}_j\end{bmatrix},\ 1 \le j \le M-1, \quad (12)$$

where $\mathbf{W}(\mathbf{q}_j) = \begin{bmatrix}\mathbf{I}_j & \mathbf{0}\\ \mathbf{0} & \widetilde{\mathbf{W}}(\mathbf{q}_j)\end{bmatrix}$ and $\widetilde{\mathbf{W}}(\mathbf{q}_j)$ is an $(M_t - j)\times(M_t - j)$ unitary matrix such that $\mathbf{q}_j^\dagger\widetilde{\mathbf{W}}(\mathbf{q}_j) = [1\ 0\ \ldots\ 0]$. The columns of $\widetilde{\mathbf{W}}(\mathbf{q}_j)$ can be obtained by extending $\mathbf{q}_j$ to an orthonormal basis for $\mathbb{C}^{M_t-j}$, where $\mathbb{C}^{M_t-j}$ is the set of all complex vectors with $M_t - j$ elements. They can be computed in a deterministic approach, e.g., Gram-Schmidt process. We see that $\{\mathbf{v}_j\}$ is an orthonormal set while the unit-norm vectors $\mathbf{q}_0, \mathbf{q}_1, \cdots, \mathbf{q}_{M-1}$, of decreasing sizes, $M_t, M_t-1, \cdots, M_t - M + 1$, are not constrained like $\{\mathbf{v}_j\}$ [18]. Let $\hat{\mathbf{q}}_j$ be the quantized version of $\mathbf{q}_j$. In Appendix D, we show that

$$\mathbf{v}_0^\dagger\hat{\mathbf{v}}_0 = \mathbf{q}_0^\dagger\hat{\mathbf{q}}_0, \quad \mathbf{v}_j^\dagger\hat{\mathbf{v}}_j \approx \mathbf{q}_j^\dagger\hat{\mathbf{q}}_j, \text{ for } 1 \le j \le M-1 \quad (13)$$

when the feedback rate B_f is sufficiently large. Thus $\overline{P_t}$ becomes

$$\overline{P_t} \approx \Gamma N_0 \sum_{j=0}^{M-1}(2^{b_j}-1)E\left[\lambda_j^{-1}\right] E\left[\left|\mathbf{q}_j^\dagger\hat{\mathbf{q}}_j\right|^{-2}\right], \quad (14)$$

where $E[\lambda_j^{-1}]$ can be computed numerically using the probability density function of λ_j [24] or using the sample mean estimator [25]. By properly allocate the rate among the subcodebooks, we can minimize the average transmission power $\overline{P_t}$. Suppose $B_f(j)$ bits are used for quantizing $\mathbf{q}_j$. Using derivations similar to those in Lemma 2, we can obtain

$$\overline{P_t} \approx \Gamma N_0 \sum_{j=0}^{M-1}(2^{b_j}-1)E\left[\lambda_j^{-1}\right]\Bigg(2^{B_f(j)}(M_t-j-1) \left[\sum_{\ell=2}^{M_t-j-1}\left(\frac{2^{\frac{-(M_t-j-\ell)}{M_t-j-1}B_f(j)}}{\ell - M_t + j}\right) - \ln\left(1 - 2^{\frac{-B_f(j)}{M_t-j-1}}\right)\right]\Bigg). \quad (15)$$

The optimal rate allocation $B_f(0), B_f(1), \ldots, B_f(M-1)$ that minimizes $\overline{P_t}$ can be obtained by using an exhaustive search of all possible nonnegative integers $B_f(j)$ such that $\sum_{j=0}^{M-1} B_f(j) = B_f$. Using an approach similar to that in Section III, we find the number of iterations to be $C(M + B_f - 1, B_f)$, which is a small number for practical cases of M and B_f. Having decided the rate allocation among $\mathbf{q}_j$, the subcodebooks for quantizing $\mathbf{q}_j$ can be designed using the generalized Lloyd algorithm as in [26]. For each of the 2^{B_b} bit loading codewords, we can design the corresponding precoder codebook using the above method.

V. Design of Bit Loading Codebook

For a given bit loading codebook size, we can generate the 2^{B_b} codewords using the generalized Lloyd algorithm, e.g., [20]. However, applying the general Lloyd algorithms directly to all the training data does not take mode, i.e., the number of substreams transmitted into consideration. When we perform the second step of the algorithm—computation of centroid, bit loading vectors of different modes are averaged, which often results in slow convergence or no convergence at all. Such a problem can be avoided using classified vector quantization (VQ) [22] in the VQ literature. In this case, the bit loading codebook $\mathcal{B}$ is a union of smaller subcodebooks, $\mathcal{B} = \bigcup_{i=1}^{M}\mathcal{B}_i$, where $\mathcal{B}_i$ is the subcodebook for the ith mode, i.e., exactly i substreams used for transmission, $1 \le i \le M$. Then the generalized Lloyd algorithm is applied to obtain $\mathcal{B}_i$. As there are subcodebooks, we need to address the issue of rate allocation among the subcodebooks. In the following we show how to allocate the rate to minimize quantization error.

Let us first analyze the quantization error from each mode. For the ith mode, only the first i entries are nonzero and we only need to consider the quantization of these entries. In this case, the bit loading vector $\mathbf{b}$ is of the form $\mathbf{b} = [b_0 \ldots b_{i-1}\, 0 \ldots 0]$. Let $\hat{\mathbf{b}}$ be the quantized version of $\mathbf{b}$. As $\sum_{k=0}^{i-1} b_k = R_b$, we only need to quantize b_k to $\hat{b}_k$ for $1 \le k \le i-1$ and choose $\hat{b}_0 = R_b - \sum_{k=1}^{i-1}\hat{b}_k$. Let m_i be the number of quantization levels for scalar quantization of b_k for $1 \le k \le i-1$. Then $\mathcal{B}_i$ contains m_i^{i-1} codewords and thus $\sum_{i=1}^{M} m_i^{i-1} = 2^{B_b}$. We can find the upper and lower bounds of b_k to obtain the quantization dynamic ranges. In particular, in the ith mode, $b_{k,min}^{(i)} \le b_k \le b_{k,max}^{(i)}$ for $i = 1$ and $b_{k,min}^{(i)} < b_k < b_{k,max}^{(i)}$ for $2 \le i \le M$ where $b_{0,min}^{(i)} = R_b/i$, $b_{k,min}^{(i)} = 0$ for $1 \le k \le i-1$, and $b_{k,max}^{(i)} = R_b/(k+1)$ for $0 \le k \le i-1$. The error $\delta_k = \hat{b}_k - b_k$ for $1 \le k \le i-1$ satisfies $|\delta_k| \le \triangle_k^{(i)}/2$, where $\triangle_k^{(i)} = (b_{k,max}^{(i)} - b_{k,min}^{(i)})m_i^{-1}$. When the number of quantization levels is sufficiently large, it is reasonable to assume that δ_k are independent and uniformly distributed over $(-\triangle_k^{(i)}/2, \triangle_k^{(i)}/2]$, for $1 \le k \le i-1$, with variance [22]

$$\sigma_{\delta_k}^2 = \frac{1}{12}\left(\triangle_k^{(i)}\right)^2 = \frac{1}{12}\left(b_{k,max}^{(i)} - b_{k,min}^{(i)}\right)^2 m_i^{-2}, \text{ for } 1 \le k \le i-1. \quad (16)$$

As $\hat{b}_0 = R_b - \sum_{k=1}^{i-1}\hat{b}_k$, we have $\delta_0 = -\sum_{k=1}^{i-1}\delta_k$; therefore $\sigma_{\delta_0}^2 = \sum_{k=1}^{i-1}\sigma_{\delta_k}^2$ when δ_k are independent. Define the average quantization error for the ith mode as $\mathcal{E}_{b,i} = \frac{1}{i}\sum_{k=0}^{i-1}\sigma_{\delta_k}^2$. To compute the overall average quantization error of all modes, we need to take into account that the modes are not used with equal probability. Suppose the ith mode is used with probability P_i, where P_i can be computed using training channels. The overall average quantization error is

$$\mathcal{E} = \sum_{i=1}^{M} P_i \mathcal{E}_{b,i} = \sum_{i=1}^{M} P_i c_i m_i^{-2},$$

where $c_i = \frac{1}{6i}\sum_{k=1}^{i-1}\left(b_{k,max}^{(i)} - b_{k,min}^{(i)}\right)^2$. The optimal rate allocation among the subcodebooks can be obtained by solving

$$\begin{aligned} &\text{minimize } \sum_{i=1}^{M} P_i c_i m_i^{-2} \\ &\text{subject to } \sum_{i=1}^{M} m_i^{i-1} \le 2^{B_b}, \text{and } 0 \le m_i^{i-1} \le A_i, \end{aligned} \quad (17)$$

where A_i is the maximum number of possible codewords in $\mathcal{B}_i$, i.e., the number of all nonincreasing integer bit loading vectors that have exactly i nonzero entries and the sum of the entries is equal to R_b. When $2^{B_b} \ge \sum_{i=1}^{M} A_i$, we can choose $m_i^{i-1} = A_i$, for $i = 1, \ldots, M$. Hence we only need to consider the case when $2^{B_b} < \sum_{i=1}^{M} A_i$. We can use the Karush-Kuhn-Tucker (KKT) condition [27] to solve such a problem. Let $\mathcal{S}_1 = \{i : m_i^{i-1} = A_i\}$ and $\mathcal{S}_2 = \{i : m_i^{i-1} < A_i\}$. The optimal m_i^{i-1} is given by (proof given in Appendix E)

$$m_i^{i-1} = \max\left(\min\left(\left(\frac{2P_i c_i}{\nu(i-1)}\right)^{\frac{i-1}{i+1}}, A_i\right), 0\right), \quad (18)$$

where ν is the unique positive real root of $\sum_{i\in\mathcal{S}_2}\left(\frac{2P_i c_i}{(i-1)\nu}\right)^{\frac{i-1}{i+1}} + \sum_{i\in\mathcal{S}_1} A_i - 2^{B_b} = 0$ for the given $\mathcal{S}_1$ and $\mathcal{S}_2$. The set $\mathcal{S}_1$ and $\mathcal{S}_2$ can be iteratively solved as in [27]. The solution of rate allocation given by (18) is not an integer in general. We can quantize it to an integer using the method in [28].

For the design of $\mathcal{B}_i$, we first generate a set of training channels, which are classified into different modes as follows. For each training channel, we compute the quantized precoder (as in Section IV) for every possible mode and compute the corresponding optimal bit loading vector as in [14], [20], [29]. We then choose the pair of precoder and bit loading that minimizes the chosen criterion, e.g., transmission power, or bit error rate. Collect all the bit loading vectors associated with the ith mode together in one set T_i. The codewords in $\mathcal{B}_i$ are then designed by applying the generalized Lloyd algorithm on T_i. The generalized Lloyd algorithm iterates the following two steps: (1) Given a codebook $\mathcal{B}_i = \{\mathbf{b}_k^{(i)}, k = 1, 2, \ldots, |\mathcal{B}_i|\}$ for the ith mode, where $|\mathcal{B}_i|$ denotes the number of vectors in $\mathcal{B}_i$. Determine the partition cell $\mathcal{R}_k^{(i)}$ corresponding to the kth codeword $\mathbf{b}_k^{(i)}$ in $\mathcal{B}_i$ by $\mathcal{R}_k^{(i)} = \{\mathbf{w} \in T_i : \|\mathbf{w} - \mathbf{b}_k^{(i)}\| \le \|\mathbf{w} - \mathbf{b}_j^{(i)}\|, \forall j \ne k\}$, for $k = 1, 2, \ldots, |\mathcal{B}_i|$. (2) Given a partition $\{\mathcal{R}_k^{(i)}, k = 1, 2, \ldots, |\mathcal{B}_i|\}$, the new codeword $\mathbf{b}_k^{(i)}$ is chosen as the arithmetic mean of the vectors in $\mathcal{R}_k^{(i)}$. The above two steps are iterated until the average mean square error is below a given threshold or when there is little improvement. The codewords generated in this way are real-valued; like rate allocation obtained in Section III they can be quantized to integer codewords using the method in [28].

VI. Design Procedure and Codeword Selection

The procedure of designing codebooks for the precoder and bit loading is summarized as follows:

1) Compute B_b and B_f using the method in Section III.
2) Generate a set of training channels, and compute P_i, i.e., the probability that the ith mode is used.
3) Determine the rate allocation among the bit loading subcodebooks using (18) and design the subcodebooks using the generalized Lloyd algorithm given in Section V.
4) For each vector in the bit loading codebook, compute the rate allocation for the precoder subcodebooks using (15) and design the subcodebooks using the generalized Lloyd algorithm in [26].

During the course of transmission, a precoder and bit loading are chosen from their respective codebooks for the given channel, and their indexes are fed back to the transmitter. Let the codewords in the bit loading codebook $\mathcal{B}$ be $\mathbf{b}_k$, $k = 0, 1, \ldots, 2^{B_b} - 1$ and $\mathcal{G}(\mathbf{b}, \mathbf{F})$ be a given objective function, e.g., total transmission power, bit error rate. We consider two codeword selection criteria.

- *Selection criterion 1*: We first compute the eigen decomposition of $\mathbf{H}^\dagger\mathbf{H}$ to obtain the unitary matrix $\mathbf{V}_M$ in (2) and the corresponding unit vectors $\mathbf{q}_0, \mathbf{q}_1, \ldots, \mathbf{q}_{M-1}$. For each bit loading vector $\mathbf{b}_k$ in the bit loading codebook, we quantize $\{\mathbf{q}_j\}$[1] using the associated precoder codebook to obtain the quantized precoder $\mathbf{F}_k$. Compute $\mathcal{G}(\mathbf{b}_k, \mathbf{F}_k)$, for $k = 0, 1, \ldots, 2^{B_b} - 1$ and the pair with the smallest objective value is chosen.
- *Selection criterion 2*: For each $\mathbf{b}_k$ in $\mathcal{B}$, we use the associated subcodebooks for $\mathbf{q}_j$ to construct all 2^{B_f} possible quantized precoders, and call the collection $\mathcal{F}_k$. Then we find the best pair of bit loading and precoder for $\min_{0\le k\le 2^{B_b}-1} \min_{\mathbf{F}\in\mathcal{F}_k} \mathcal{G}(\mathbf{b}_k, \mathbf{F})$. For a given bit loading vector, the unit vectors $\{\mathbf{q}_j\}$ are quantized directly with the first criterion. With the second criterion we search among all 2^{B_f} quantized precoders to find the one that minimizes the objective function; the objective function is evaluated 2^B times and the complexity is similar to that in [6], [9], [20]. With the first criterion, the objective function is computed only 2^{B_b} times and the complexity is roughly $1/2^{B_f}$ that of criterion two. The first one has a lower complexity but the second one enjoys a better performance.

Remark: In the derivation of feedback rate allocation, the receiver is assumed to be linear. After the rate is allocated and codebooks designed, we can still replace the receiver by a decision feedback receiver. With the aid of decision feedback, the performance of the proposed limited feedback system can be improved significantly (to be demonstrated in the next section) although the rate allocation and codebooks have been designed for a linear receiver.

[1]Three types of encoding schemes for $\{\mathbf{q}_j\}$ were proposed in [18]. We use encoding scheme B in [18] here.

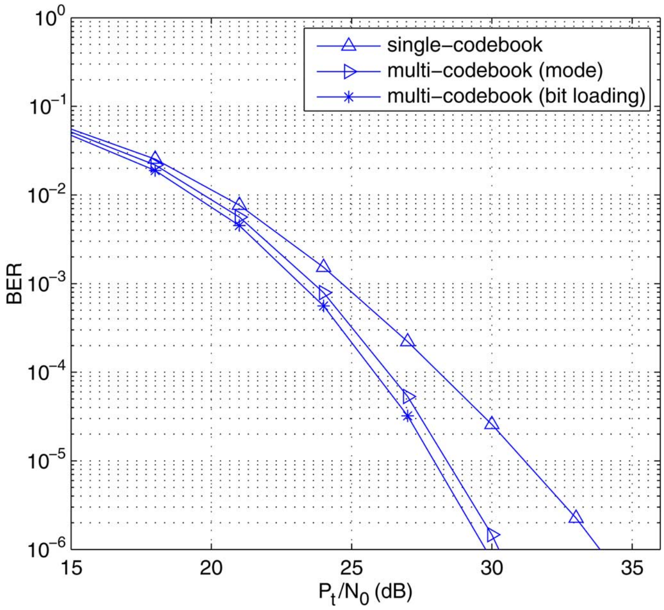

Fig. 3. Example 1. Bit error rate performance of multiple precoder codebooks and a single codebook.

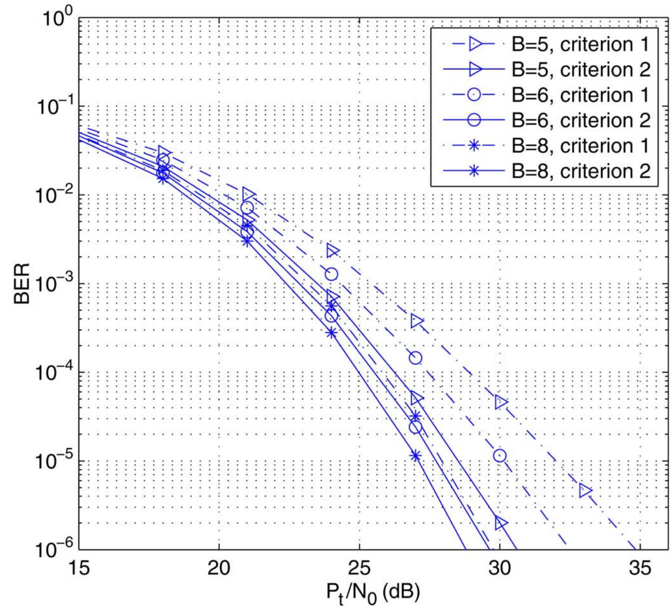

Fig. 4. Example 2. Comparison of the two selection criteria for different feedback rate B.

VII. Simulations

In the following examples, the elements of channel matrix $\mathbf{H}$ are independent complex Gaussian random variables with zero mean and unit variance. The precoder and bit loading are chosen to minimize bit error rate. We have used 10^5 training channels for designing codebooks of bit loading and precoder, and 10^5 channel realizations for BER simulations. The power is equally divided among all symbols carrying nonzero bits. The receiver is linear and zero-forcing in Examples 1-4, and a decision feedback receiver is used in Example 5. For a given rate allocation, the precoder codebook and bit loading codebook are designed as described in Section IV and Section V, respectively.

Example 1. Multiple Precoder Codebooks: In Fig. 3, we compare the multi-codebook and single-codebook schemes for precoder codebook designs with $M_r = M_t = 4$ and $R_b = 16$. For $B = 8$, the optimal feedback rate allocation obtained using the method in Section III is $(B_f, B_b) = (4, 4)$. Selection criterion 1 in Section VI is used. We show the results of two types of multi-codebook design for the precoder. In the first one (labeled as "multi-codebook (bit loading)" in Fig. 3), one precoder codebook is designed for each bit loading as discussed in Section IV; there are a total of 2^{B_b} codebooks for the precoder. In the second multi-codebook scheme (labeled as "multi-codebook (mode)" in Fig. 3), we have only one codebook for each mode. For the ith mode, we solve (15) to obtain rate allocation among $\{\mathbf{q}_j\}$, assuming data bits are uniformly loaded on all i substreams. For both multi-codebook schemes, the receiver does not need to inform the transmitter the codebook used as the transmitter can obtain the information from bit loading. In the single-codebook case, a fixed rate allocation is used for $\{\mathbf{q}_j\}$, independent of mode and bit loading. For this case, we solve (15) with the assumption that data bits are uniformly loaded on all M substreams. The same bit loading codebook is used for all three cases in Fig. 3. The single-codebook scheme is the worst because the feedback bits are allocated in a fixed manner even if some substreams are not loaded with bits. The two multi-codebook schemes enjoy significant gain over the single-codebook scheme. The gain of bit-loading-dependent codebook scheme is around 3.2 dB at $\mathrm{BER} = 10^{-5}$.

Example 2. Selection Criteria: Fig. 4 compares the BER of the two selection criteria introduced in Section VI for $M_r = M_t = 4$ and $R_b = 16$. With the first criterion, vector quantization is applied directly on $\{\mathbf{q}_i\}$ for each bit loading vector while an exhaustive search among all 2^{B_f} precoders is performed in the second case. For the same feedback rate B, we use the same precoder and bit loading codebooks in Fig. 4; only the selection criteria are different. When the feedback rate increases, the degradation of using the low-cost criterion 1 becomes smaller. For example, when $B = 5$, the gap between the two criteria is around 3.3 dB at $\mathrm{BER} = 10^{-4}$ and it narrows to 0.8 dB when $B = 8$. For a large B, criterion 1 can be used to reduce complexity at the cost of a small performance loss. For a small B, it is worthwhile to use criterion 2, for which the number of searches 2^B is small.

Example 3. Feedback Rate Allocation: We demonstrate the importance of proper feedback rate allocation between precoder and bit loading in this example for $M_r = M_t = 3$ and $R_b = 15$. Selection criterion 1 in Section VI is used. For $B = 5$, the optimal rate allocation is $(B_f, B_b) = (3, 2)$ using the method in Section III. In Fig. 5, we show the BER for all possible (B_f, B_b) such that $B_f + B_b = B$. We see that the rate allocation $(B_f, B_b) = (3, 2)$ gives the best performance. For example, at $\mathrm{BER} = 10^{-5}$, $(B_f, B_b) = (3, 2)$ is better than $(B_f, B_b) = (4, 1)$ by around 2.1 dB. The performance is sensitive to rate allocation; by moving one bit from B_f to B_b the performance can differ by 2.1 dB. In the case $(B_f, B_b) = (5, 0)$, all feedback bits are used for precoder feedback and the bit loading is a fixed vector. Two cases of fixed bit loading are shown, a nonuniform one [8 7 0] and a uniform one [5 5 5].

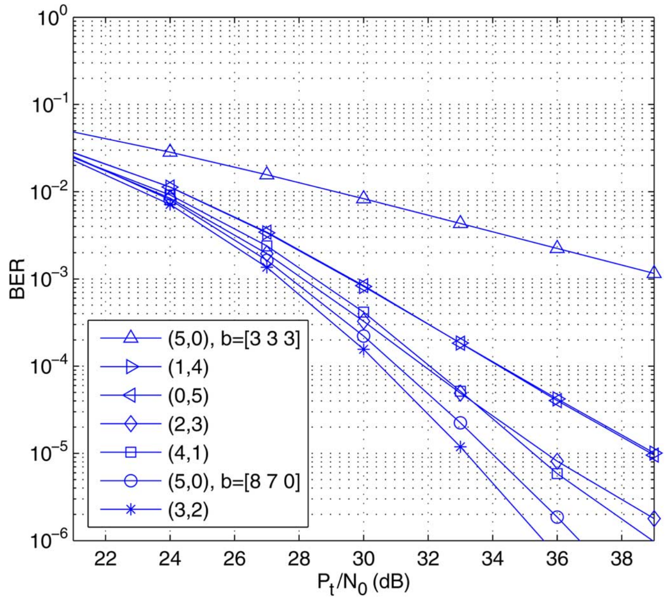

Fig. 5. Example 3. Bit error rate performance for all different feedback rate allocations when $B = 5$.

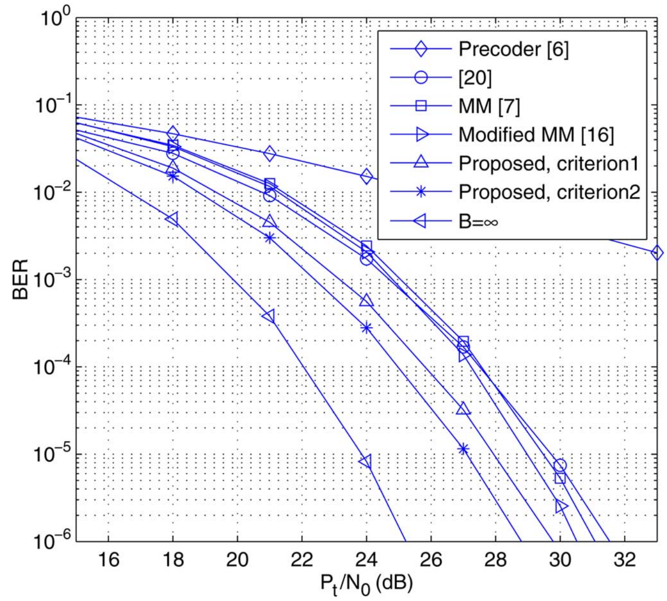

Fig. 6. Example 4. Comparisons of BER for systems with linear receivers for $B = 8$.

The fixed nonuniform bit loading is obtained by using the generalized Lloyd algorithm in Section V with only one codeword; the performance is considerably better than that of uniform bit loading. Therefore the design of bit loading is particularly important when $B_b = 0$.

Example 4. BER Comparisons for Linear Receivers: In this example we show the BER of the proposed method and other limited feedback systems with a linear receiver for $M_r = M_t = 4$, $R_b = 16$, and $B = 8$. The feedback rate allocation computed using the method in Section III is $(B_f, B_b) = (4, 4)$. The precoder system [6] feeds back the index of the precoder in the codebook and data bits are uniformly loaded on all M substreams. In the multimode (MM) precoding system [7], the constellation on all substreams are the same, but the number of substreams transmitted can vary with the channel. The modified multimode precoding in [16] improves the performance of MM in [7] by introducing additional feedback of nonuniform bit loading. The feedback of bit loading only is proposed in [20]; the precoder is allocated zero feedback bit. The results are shown in Fig. 6. The systems that allow the number of substreams to vary enjoy a better performance. At $\mathrm{BER} = 10^{-4}$, the gap between the proposed system and other systems is around 1.5 dB when selection criterion 1 is used and around 2.3 dB when selection criterion 2 is used. By judicious allocation of feedback rates and joint consideration of precoder and bit loading feedback, the proposed system can achieve a better performance. As a benchmark, the performance of the case $B = \infty$ is also shown, in which the precoder $\mathbf{F} = \mathbf{V}_M$, and the optimal positive bit loading is used. With 8 bits of feedback, the performance of the proposed system with criterion 2 is around 3 dB away from the curve "$B = \infty$" at $\mathrm{BER} = 10^{-4}$.

Example 5. BER Comparisons for Decision Feedback Receivers: We show the BER of the proposed system with a decision feedback receiver in Fig. 7 for $M_r = M_t = 4$, $R_b = 16$, and $B = 8$. We use the feedback rate allocation and codebooks designed for a linear receiver. With $B = 8$, the proposed method with criterion 2 is very close to the curve "$B = \infty$" that we have shown for a linear receiver in Fig. 6. For comparison, we have also shown the BER of the decision feedback systems in [9], [12], and [14]. In [12], the detection ordering is fed back to the transmitter. The required feedback rate is a fixed number $\log_2(M_t!)$ bits; which is around 7 bits in this case. The QR-based system in [14] feeds back the index of bit loading. We constrain b_k in a manner similar to that in [14] to satisfy the given feedback rate: b_k are integers such that $b_0 \geq 2$, $b_1 \geq 2$, $b_2 \geq 2$, $b_3 \geq 0$, and $b_0 + b_1 + b_2 + b_3 = R_b$. The number of bit loading vectors is 286, which requires $\log_2 286 \approx 8$ bits. The system proposed in [9] feeds back the index of the optimal precoder in the Grassmannian codebook of 256 codewords for uniform bit loading. Due to the restriction of parameters $(M_t > M)$ in [9], we have increased the number of transmit antennas M_t to 5 in the simulation for [9] and the other parameters remain the same. We can see that the proposed system is able to achieve a smaller BER than those that do not consider the feedback of precoder and bit loading together.

VIII. Conclusion

In this paper, we have jointly considered the feedback of both precoder and bit loading for MIMO systems. We have developed a systematic approach to designing feedback rate allocation between precoder and bit loading by analyzing the power penalty due to quantization. As bit loading carries information on the importance of individual subchannels, we proposed to use multiple bit-loading-dependent codebooks for the precoder. The use of multi-codebook design yields significant gain over the single-codebook design. There is no need of informing the transmitter which codebook has been used because of bit loading feedback. The code rate of each precoder codebook is equal to

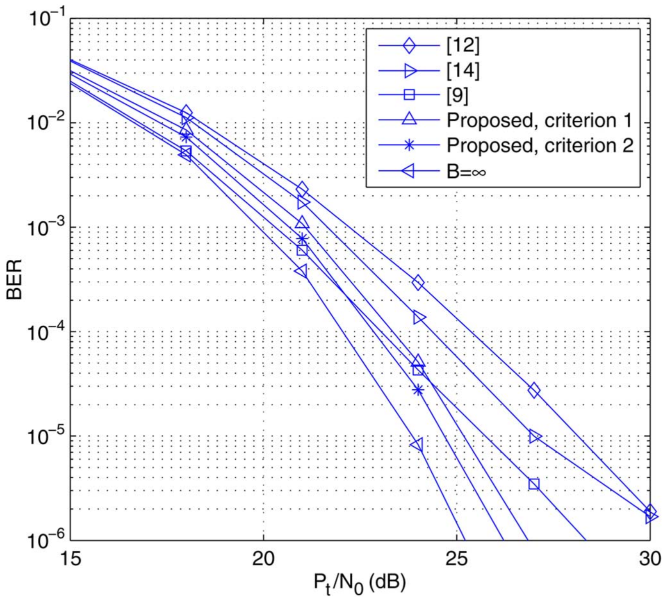

Fig. 7. Example 5. Comparisons of BER for systems with DFE receivers for $B = 8$.

the full feedback rate allocated for the precoder. The joint consideration of feedback of precoder and bit loading leads to a very good performance compared to systems that design the feedback of bit loading and precoder separately.

Appendix A
Proof of Lemma 1

When the precoder is $\mathbf{F} = \widehat{\mathbf{V}}_M$, the error autocorrelation matrix in (1) is $\widehat{\mathbf{R}}_e = N_0(\mathbf{V}_M^\dagger \widehat{\mathbf{V}}_M)^{-1} \mathbf{\Lambda}^{-1} (\widehat{\mathbf{V}}_M^\dagger \mathbf{V}_M)^{-1}$ if $\mathbf{H}$ has full rank. Express $\mathbf{V}_M^\dagger \widehat{\mathbf{V}}_M$ as $\mathbf{V}_M^\dagger \widehat{\mathbf{V}}_M = \mathbf{D}(\mathbf{I}_M + \mathbf{E})$, where $\mathbf{D}$ is a diagonal matrix with $[\mathbf{D}]_{kk} = \mathbf{v}_k^\dagger \hat{\mathbf{v}}_k$, $k = 0, 1, \ldots, M-1$, $\mathbf{I}_M$ is the $M \times M$ identity matrix, and the matrix $\mathbf{E}$ is given by $[\mathbf{E}]_{kj} = \mathbf{v}_k^\dagger \hat{\mathbf{v}}_j / \mathbf{v}_k^\dagger \hat{\mathbf{v}}_k$ when $k \neq j$ and $[\mathbf{E}]_{kk} = 0$. When B_f is large, the quantization error is small, i.e., $\mathbf{v}_k^\dagger \hat{\mathbf{v}}_j \approx 0$, $k \neq j$. Thus $[\mathbf{E}]_{kj} \approx 0, \forall k, j$. It is known that [30] we can write $(\mathbf{I}_M + \mathbf{E})^{-1}$ as a power series in $\mathbf{E}$, i.e., $(\mathbf{I}_M + \mathbf{E})^{-1} = \sum_{j=0}^{\infty} (-1)^j \mathbf{E}^j$ when $\|\mathbf{E}\|_F < 1$, where $\|\mathbf{E}\|_F$ denotes the Frobenius norm of $\mathbf{E}$. As the elements of $\mathbf{E}$ are small, we have the approximation $(\mathbf{I}_M + \mathbf{E})^{-1} \approx \mathbf{I}_M - \mathbf{E}$. It follows that $(\mathbf{V}_M^\dagger \widehat{\mathbf{V}}_M)^{-1} = (\mathbf{I}_M + \mathbf{E})^{-1} \mathbf{D}^{-1} \approx (\mathbf{I}_M - \mathbf{E}) \mathbf{D}^{-1}$. Thus we have $\widehat{\mathbf{R}}_e \approx N_0 (\mathbf{I}_M - \mathbf{E}) \mathbf{D}^{-1} \mathbf{\Lambda}^{-1} \mathbf{D}^{-\dagger} (\mathbf{I}_M - \mathbf{E}^\dagger)$. Notice that $\mathbf{D}^{-1} \mathbf{\Lambda}^{-1} \mathbf{D}^{-\dagger}$ is a diagonal matrix, and the diagonal elements of $\mathbf{E}$ are equal to zero, so $[\mathbf{E}\mathbf{D}^{-1}\mathbf{\Lambda}^{-1}\mathbf{D}^{-\dagger}]_{kk} = [\mathbf{D}^{-1}\mathbf{\Lambda}^{-1}\mathbf{D}^{-\dagger}\mathbf{E}^\dagger]_{kk} = 0$. Therefore the kth subchannel error variance $\hat{\sigma}_{e_k}^2 = [\widehat{\mathbf{R}}_e]_{kk}$ can be written as

$$\hat{\sigma}_{e_k}^2 = [\widehat{\mathbf{R}}_e]_{kk} \approx N_0 \left(|[\mathbf{D}]_{kk}|^{-2} \lambda_k^{-1} + \left\| \mathbf{1}_k^\dagger \mathbf{E} \mathbf{D}^{-1} \mathbf{\Lambda}^{-1/2} \right\|^2 \right), \tag{19}$$

where $\mathbf{1}_k$ is the kth standard vector with $[\mathbf{1}_k]_k = 1$ and $[\mathbf{1}_k]_j = 0$ when $j \neq k$. The jth element of $\mathbf{1}_k^\dagger \mathbf{E}\mathbf{D}^{-1}\mathbf{\Lambda}^{-1/2}$ is equal to zero when $j = k$, and equal to $\lambda_j^{-1/2} \mathbf{v}_k^\dagger \hat{\mathbf{v}}_j / (\mathbf{v}_k^\dagger \hat{\mathbf{v}}_k \mathbf{v}_j^\dagger \hat{\mathbf{v}}_j)$ when $j \neq k$, so $\|\mathbf{1}_k^\dagger \mathbf{E}\mathbf{D}^{-1}\mathbf{\Lambda}^{-1/2}\|^2$ can be expressed as

$$\left\| \mathbf{1}_k^\dagger \mathbf{E} \mathbf{D}^{-1} \mathbf{\Lambda}^{-1/2} \right\|^2 = \sum_{j=0, j \neq k}^{M-1} \left| \frac{\mathbf{v}_k^\dagger \hat{\mathbf{v}}_j}{\mathbf{v}_k^\dagger \hat{\mathbf{v}}_k} \right|^2 \frac{\lambda_j^{-1}}{\left| \mathbf{v}_j^\dagger \hat{\mathbf{v}}_j \right|^2}. \tag{20}$$

Substituting (20) to (19), and using $|\mathbf{v}_k^\dagger \hat{\mathbf{v}}_j| \approx 0$, we have the approximation in (6)

Appendix B
Proof of Lemma 2

The power penalty due to precoder quantization in (7) can be rearranged as

$$D_f = -\frac{10}{M} \sum_{k=0}^{M-1} E\left[\log_{10} \left| \mathbf{v}_k^\dagger \hat{\mathbf{v}}_k \right|^2 \right]. \tag{21}$$

It is known that each $\mathbf{v}_k$ is uniformly distributed over the M_t-dimensional space $\{\mathbf{u} \in \mathbb{C}^{M_t} : \|\mathbf{u}\| = 1\}$, where $\mathbb{C}^{M_t}$ is the set of all complex vectors with M_t elements [23]. When B_{v_k} bits is used to quantize $\mathbf{v}_k$ to $\hat{\mathbf{v}}_k$, the probability density function of $|\mathbf{v}_k^\dagger \hat{\mathbf{v}}_k|^2$ can be approximated as [26]

$$f_{|\mathbf{v}_k^\dagger \hat{\mathbf{v}}_k|^2}(x) \approx 2^{B_{v_k}} (M_t - 1)(1-x)^{M_t - 2} 1_{[1-\epsilon_k, 1)}(x),$$

where $0 < x < 1$, $\epsilon_k = 2^{-B_{v_k}/(M_t - 1)}$ 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 above pdf approximation, it can be verified that $E[\log_{10} |\mathbf{v}_k^\dagger \hat{\mathbf{v}}_k|^2] \approx 2^{B_{v_k}} (M_t - 1) \int_{1-\epsilon_k}^{1} \log_{10} x (1-x)^{M_t - 2} dx$. Using binomial expansion, we have $(1-x)^{M_t-2} = \sum_{j=0}^{M_t - 2} C(M_t - 2, j)(-x)^j$. Using integration by parts repeatedly, we obtain

$$E\left[\log_{10} \left| \mathbf{v}_k^\dagger \hat{\mathbf{v}}_k \right|^2 \right] \approx \frac{2^{B_{v_k}} (M_t - 1)}{-\ln 10} \sum_{j=0}^{M_t - 2} \left[(-1)^j C(M_t - 2, j) \left(\frac{(1-\epsilon_k)^{j+1} \ln(1-\epsilon_k)}{j+1} + \frac{1 - (1-\epsilon_k)^{j+1}}{(j+1)^2} \right) \right]. \tag{22}$$

Substituting (22) to (21) leads to (8).

Appendix C
Proof of Lemma 3

The power penalty due to bit loading quantization in (9) can be written as $D_b = 10 \log_{10} M^{-1} + \frac{10}{\ln 10} E[\ln(g(\mathbf{z})]$, where $g(\mathbf{z}) = \ln(z_0 + z_1 + \ldots + z_{M-1})$, $z_k = 2^{\delta_k}$ and $\mathbf{z} = [z_0 \ z_1 \ \ldots \ z_{M-1}]$. Consider the Taylor series of $g(z_0, z_1, \ldots, z_{M-1})$ about the mean $(\mu_{1,0}, \mu_{1,1}, \ldots, \mu_{1,M-1})$, where $\mu_{1,k} = E[2^{\delta_k}]$ as defined in Lemma 3. Using the 3rd order Taylor approximation [31], we have (23), shown at the bottom of the next page, where $v = \sum_{k=0}^{M-1} \mu_{1,k}$, α_k are nonnegative integer and $\alpha_k!$ denotes the factorial of α_k. It can be verified that $\left.\frac{\partial^{\alpha_0 + \alpha_1 + \ldots + \alpha_{M-1}} g(\mathbf{z})}{\partial z_0^{\alpha_0} \partial z_1^{\alpha_1} \ldots \partial z_{M-1}^{\alpha_{M-1}}}\right|_{\mathbf{z}=\mathbf{a}} = (-1)^{(-1 + \sum_{k=0}^{M-1} \alpha_k)} (-1 +$

$\sum_{k=0}^{M-1}\alpha_k)!v^{-\sum_{k=0}^{M-1}\alpha_k}$. The expectation of (23) is (24), shown at the bottom of the page. Let us examine the second and third terms on the right hand side (r.h.s.) of (24). As z_j and z_ℓ are independent for $j\neq\ell$ and $j,\ell\neq 0$, we have $E[(z_j-\mu_{1,j})^{\alpha_j}(z_\ell-\mu_{1,\ell})^{\alpha_\ell}] = E[(z_j-\mu_{1,j})^{\alpha_j}]E[(z_\ell-\mu_{1,\ell})^{\alpha_\ell}]=0$ when $\alpha_j=1$ or $\alpha_\ell=1$. So the second term on the r.h.s. of (24) can be simplified as

$$\frac{-1}{2v^2}\sum_{k=0}^{M-1}E\left[(z_k-\mu_{1,k})^2\right]+\frac{-1}{v^2}\sum_{k=1}^{M-1}E\left[(z_0-\mu_{1,0})(z_k-\mu_{1,k})\right]$$
$$=\frac{-1}{2v^2}\sum_{k=0}^{M-1}\left(\mu_{2,k}-\mu_{1,k}^2\right)+\frac{\mu_{1,0}}{v^2}\sum_{k=1}^{M-1}\left(\mu_{1,k}-\mu_{1,k}^{-1}\right),\quad (25)$$

where we have used the facts that $E[z_k^{-1}]=\mu_{1,k}$, $\mu_{1,0}=\prod_{j=1}^{M-1}\mu_{1,j}$ and thus $E[z_0z_k]=\mu_{1,0}/\mu_{1,k}$ for $k\neq 0$. Similarly, for $i\neq j, j\neq\ell, i,j,\ell\neq 0$, we have $E[(z_i-\mu_{1,i})^{\alpha_i}(z_j-\mu_{1,j})^{\alpha_j}(z_\ell-\mu_{1,\ell})^{\alpha_\ell}]=0$ when $\alpha_i=1$, $\alpha_j=1$, or $\alpha_\ell=1$. Thus the third term on the r.h.s. of (24) becomes

$$\frac{1}{3v^3}\sum_{k=0}^{M-1}E\left[(z_k-\mu_{1,k})^3\right]+\frac{1}{v^3}\sum_{k=1}^{M-1}\Bigg(E[(z_0-\mu_{1,0})^2 (z_k-\mu_{1,k})]+E\left[(z_0-\mu_{1,0})(z_k-\mu_{1,k})^2\right]\Bigg)$$
$$+\frac{2}{v^3}\sum_{k=1}^{M-2}\sum_{\ell=k+1}^{M-1}E\left[(z_0-\mu_{1,0})(z_k-\mu_{1,k})(z_\ell-\mu_{1,\ell})\right].\quad (26)$$

Using $E[z_0^2z_k]=\mu_{2,0}\mu_{1,k}/\mu_{2,k}$, $E[z_0z_k^2]=\mu_{1,0}$, and $E[z_0z_kz_\ell]=\mu_{1,0}/(\mu_{1,k}\mu_{1,\ell})$ for $k\neq\ell$, $k,\ell\neq 0$, (26) can be expressed as

$$\frac{1}{3v^3}\sum_{k=0}^{M-1}(\mu_{3,k}-3\mu_{2,k}\mu_{1,k}+2\mu_{1,k}^3)+\frac{\mu_{2,0}}{v^3}\sum_{k=1}^{M-1}\left(\frac{\mu_{1,k}}{\mu_{2,k}}-\mu_{1,k}\right)$$
$$+\frac{2\mu_{1,0}^2}{v^3}\sum_{k=1}^{M-1}\left(\mu_{1,k}-\mu_{1,k}^{-1}\right)+\frac{\mu_{1,0}}{v^3}\sum_{k=1}^{M-1}(2\mu_{1,k}^2-\mu_{2,k}-1)$$
$$+\frac{2\mu_{1,0}}{v^3}\sum_{k=1}^{M-2}\sum_{\ell=k+1}^{M-1}\left(\mu_{1,k}^{-1}\mu_{1,\ell}^{-1}-\mu_{1,k}^{-1}\mu_{1,\ell}-\mu_{1,k}\mu_{1,\ell}^{-1}+\mu_{1,k}\mu_{1,\ell}\right).\quad (27)$$

Substituting (25) and (27) to (24), we obtain D_b in (10). For any nonnegative integer n, it can be verified that $\mu_{n,k}=E[z_k^n]=\frac{1}{n\Delta_k\ln 2}[2^{n\Delta_k/2}-2^{-n\Delta_k/2}]$ for $1\le k\le M-1$. Thus $\mu_{n,0}=\prod_{\ell=1}^{M-1}\mu_{n,\ell}=\prod_{\ell=1}^{M-1}\frac{1}{n\Delta_\ell\ln 2}[2^{n\Delta_\ell/2}-2^{-n\Delta_\ell/2}]$.

Appendix D
Proof of (13)

Using (12), 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

$$\hat{\mathbf{v}}_0=\hat{\mathbf{q}}_0,$$
$$\hat{\mathbf{v}}_j=\mathbf{W}(\hat{\mathbf{q}}_0)\cdots\mathbf{W}(\hat{\mathbf{q}}_{j-1})\begin{bmatrix}\mathbf{0}\\ \hat{\mathbf{q}}_j\end{bmatrix},\ 1\le j\le M-1.\quad (28)$$

Using (12) and (28), we have $\mathbf{v}_0^\dagger\hat{\mathbf{v}}_0=\mathbf{q}_0^\dagger\hat{\mathbf{q}}_0$ and

$$\mathbf{v}_j^\dagger\hat{\mathbf{v}}_j=\begin{bmatrix}\mathbf{0} & \mathbf{q}_j^\dagger\end{bmatrix}\mathbf{U}_{j-1}^\dagger\widehat{\mathbf{U}}_{j-1}\begin{bmatrix}\mathbf{0}\\ \hat{\mathbf{q}}_j\end{bmatrix},\ 1\le j\le M-1,\quad (29)$$

where $\mathbf{U}_{j-1}=\mathbf{W}(\mathbf{q}_0)\cdots\mathbf{W}(\mathbf{q}_{j-1})$ and $\widehat{\mathbf{U}}_{j-1}=\mathbf{W}(\hat{\mathbf{q}}_0)\cdots\mathbf{W}(\hat{\mathbf{q}}_{j-1})$. 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 $\mathbf{A}(\mathbf{q}_j)=[\mathbf{q}_j\ \mathbf{1}_1\ \ldots\ \mathbf{1}_{M_t-j-1}]$ 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 [30]. 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 $1\le k\le M_t-j-1$. 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., $\hat{\mathbf{w}}_0^{(j)}\approx\mathbf{w}_0^{(j)}$, which implies $\mathbf{t}_1^{(j)}\approx\hat{\mathbf{t}}_1^{(j)}$, $\|\mathbf{t}_1^{(j)}\|\approx\|\hat{\mathbf{t}}_1^{(j)}\|$ and thus $\hat{\mathbf{w}}_1^{(j)}\approx\mathbf{w}_1^{(j)}$. Using a similar approach, we get $\hat{\mathbf{w}}_k^{(j)}\approx\mathbf{w}_k^{(j)}$ for $2\le k\le M_t-j-1$. Defining $\mathbf{\Phi}_j=\mathbf{W}(\hat{\mathbf{q}}_j)-\mathbf{W}(\mathbf{q}_j)$, we can write

$$\mathbf{W}(\hat{\mathbf{q}}_j)=\mathbf{W}(\mathbf{q}_j)+\mathbf{\Phi}_j,\ j=0,\,1,\ldots,\,M-1,\quad (30)$$

where the entries of $\mathbf{\Phi}_j$ are small. Using (30), the matrix $\mathbf{U}_{j-1}^\dagger\widehat{\mathbf{U}}_{j-1}$ can be expressed as

$$\mathbf{U}_{j-1}^\dagger\widehat{\mathbf{U}}_{j-1}=\mathbf{I}_{M_t}+\mathbf{\Omega}_{j-1}^{(j-1)}+\mathbf{\Omega}_{j-2}^{(j-1)}+\cdots+\mathbf{\Omega}_0^{(j-1)},\quad (31)$$

where $\mathbf{\Omega}_k^{(j-1)}=\mathbf{W}^\dagger(\mathbf{q}_{j-1})\cdots\mathbf{W}^\dagger(\mathbf{q}_k)\mathbf{\Phi}_k\mathbf{W}(\hat{\mathbf{q}}_{k+1})\mathbf{W}(\hat{\mathbf{q}}_{k+2})\cdots\mathbf{W}(\hat{\mathbf{q}}_{j-1})$,

$$g(\mathbf{z})\approx\ln(v)+\sum_{\alpha_0+\alpha_1+\ldots+\alpha_{M-1}=1,2,3}\left[\left(\prod_{k=0}^{M-1}\frac{(z_k-\mu_{1,k})^{\alpha_k}}{\alpha_k!}\right)\frac{\partial^{\alpha_0+\alpha_1+\ldots+\alpha_{M-1}}g(\mathbf{z})}{\partial z_0^{\alpha_0}\partial z_1^{\alpha_1}\ldots\partial z_{M-1}^{\alpha_{M-1}}}\bigg|_{\mathbf{z}=\mathbf{a}}\right],\quad (23)$$

$$E\left[g(\mathbf{z})\right]\approx\ln(v)+\frac{-1}{v^2}E\left[\sum_{\alpha_0+\ldots+\alpha_{M-1}=2}\prod_{k=0}^{M-1}\frac{(z_k-\mu_{1,k})^{\alpha_k}}{\alpha_k!}\right]+\frac{2}{v^3}E\left[\sum_{\alpha_0+\ldots+\alpha_{M-1}=3}\prod_{k=0}^{M-1}\frac{(z_k-\mu_{1,k})^{\alpha_k}}{\alpha_k!}\right].\quad (24)$$

for $0 \leq k \leq j-2$ and $\mathbf{\Omega}_{j-1}^{(j-1)} = \mathbf{W}^\dagger(\mathbf{q}_{j-1})\mathbf{\Phi}_{j-1}$. Substituting (31) to (29), we get

$$\mathbf{v}_j^\dagger \hat{\mathbf{v}}_j = \mathbf{q}_j^\dagger \hat{\mathbf{q}}_j + \sum_{k=0}^{j-1} \begin{bmatrix} \mathbf{0} & \mathbf{q}_j^\dagger \end{bmatrix} \mathbf{\Omega}_k^{(j-1)} \begin{bmatrix} \mathbf{0} \\ \hat{\mathbf{q}}_j \end{bmatrix}. \tag{32}$$

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 $1 \leq j \leq M-1$.

Appendix E
Proof of (18)

We can use KKT condition to solve the problem in (17). Let m_i^* be a local minimum. There exist constants ν, η_ℓ, $\kappa_\ell \geq 0$ such that 1) $\frac{\partial}{\partial m_i}[\sum_{\ell=1}^M P_\ell c_\ell m_\ell^{-2} + \nu(\sum_{\ell=1}^M m_\ell^{\ell-1} - 2^{B_b}) + \sum_{\ell=1}^M \eta_\ell(m_\ell^{\ell-1} - A_\ell) + \sum_{\ell=1}^M -\kappa_\ell m_\ell^{\ell-1})]|_{m_i=m_i^*} = 0$. 2) $\nu(\sum_{\ell=1}^M m_\ell^{\ell-1} - 2^{B_b})|_{m_i=m_i^*} = 0$. 3) $\eta_i(m_i^{i-1} - A_i)|_{m_i=m_i^*} = 0$. 4) $-\kappa_i m_i^{i-1}|_{m_i=m_i^*} = 0$. Solving condition 1, we have

$$-2P_i c_i m_i^{-3} + \nu(i-1)m_i^{i-2} + \eta_i(i-1)m_i^{i-2} - \kappa_i(i-1)m_i^{i-2} = 0. \tag{33}$$

Suppose $\nu = 0$ in the above equation, we obtain $\eta_i = (2P_i c_i m_i^{-3} + \kappa_i(i-1)m_i^{i-2})/((i-1)m_i^{i-2}) > 0$, which means $m_i^{i-1} = A_i$ for all i. This contradicts the condition $2^{B_b} < \sum_{\ell=1}^M A_\ell$, hence $\nu > 0$ and condition 2 becomes

$$\sum_{\ell=1}^M m_\ell^{\ell-1} = 2^{B_b}. \tag{34}$$

From condition 3, we have $m_i^{i-1} = A_i$ if $\eta_i > 0$. Thus condition 4 implies $\kappa_i = 0$. If $\kappa_i > 0$ in condition 4, we have $m_i^{i-1} = 0$ and hence $\eta_i = 0$. Note that when $\eta_i = 0$, and $\kappa_i = 0$, (33) implies $m_i^{i-1} = \left(\frac{2P_i c_i}{(i-1)\nu}\right)^{\frac{i-1}{i+1}}$. Combining conditions 3 and 4, we obtain the optimal m_i^{i-1} given in (18). To obtain the constant ν, we substitute the optimal m_i^{i-1} to (34),

$$\sum_{i\in\mathcal{S}_2} \left(\frac{2P_i c_i}{(i-1)\nu}\right)^{\frac{i-1}{i+1}} + \sum_{i\in\mathcal{S}_1} A_i = 2^{B_b}. \tag{35}$$

We only need to solve ν when $|\mathcal{S}_2| \neq 0$. Let p be the lower common multiple of the set $\{i+1 : i \in \mathcal{S}_2\}$. Defining $y = (1/\nu)^{1/p}$, then (35) becomes a polynomial of y

$$\sum_{i\in\mathcal{S}_2} \left(\frac{2P_i c_i}{i-1}\right)^{\frac{i-1}{i+1}} y^{\frac{(i-1)p}{i+1}} + \sum_{i\in\mathcal{S}_1} A_i - 2^{B_b} = 0. \tag{36}$$

Such a polynomial has only one positive real root and the solution of ν is unique. This can be shown using sign variations [32], the definition of which is given in the following for completeness. Let $a_0, a_1, \ldots, a_n$ be a finite sequence of real numbers. Suppose there are k nonzero numbers in the sequence, $a_{j_0}, a_{j_1}, \ldots, a_{j_{k-1}}$. The number of sign variations in the sequence, denoted as $var(a_0, a_1, \ldots, a_n)$, is the number of pairs $(a_{j_i}, a_{j_{i+1}})$ such that $a_{j_i} a_{j_{i+1}} < 0$ for $i = 0, \ldots, k-2$ [32].

Lemma 4: [32] Let $h(x) = a_0 + a_1 x + \ldots + a_n x^n$ be a polynomial with real coefficients and $a_n > 0$. If $var(a_0, a_1, \ldots, a_n) = 1$, then $h(x)$ has exactly one positive real root.

Observe that the constant term $\sum_{i\in\mathcal{S}_1} A_i - 2^{B_b}$ in (36) is negative because the number of codewords is equal to 2^{B_b} and $|\mathcal{S}_2| \neq 0$. Therefore, (36) is a polynomial equation with positive coefficients except for the constant term. The number of sign variations in the sequence formed by the coefficients of the polynomial in (36) is equal to one. Thus given $\mathcal{S}_1$ and $\mathcal{S}_2$, the polynomial has exactly one positive real root and the solution of ν is unique.

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 wireless communication systems," *IEEE Sel. Areas Commun.*, 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 Trans. Inf. Theory*, vol. 51, no. 8, pp. 2967–2976, Aug. 2005.

[3] T. Inoue and R. W. Heath, "Kerdock codes for limited feedback precoded MIMO systems," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3711–3716, Sep. 2009.

[4] 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.

[5] J. C. Roh and B. D. Rao, "Design and analysis of MIMO spatial multiplexing systems with quantized feedback," *IEEE Trans. Signal Process.*, vol. 54, no. 8, pp. 2874–2886, Aug. 2006.

[6] S. Zhou and B. Li, "BER criterion and codebook construction for finite-rate precoded spatial multiplexing with linear receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 5, pp. 1653–1665, May 2006.

[7] 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.

[8] X. Song and H. N. Lee, "Multimode precoding for MIMO systems: Performance bounds and limited feedback codebook design," *IEEE Trans. Signal Process.*, vol. 56, pp. 5296–5301, Oct. 2008, v10.

[9] M. B. Shenouda and T. N. Davidson, "A design framework for limited feedback MIMO systems with zero-forcing DFE," *IEEE J. Sel. Area Commun.*, vol. 26, no. 8, pp. 1578–1587, Oct. 2008.

[10] H. Zhuang, L. Dai, S. Zhou, and Y. Yao, "Low complexity per-antenna rate and power control approach for closed-loop V-BLAST," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1783–1787, Nov. 2003.

[11] S. T. Chung, A. Lozano, H. C. Huang, A. Sutivong, and J. M. Cioffi, "Approaching the MIMO capacity with a low-rate feedback channel in V-BLAST," *EURASIP J. Appl. Signal Process.*, no. 5, pp. 762–771, 2004.

[12] Y. Jiang and M. K. Varanasi, "Spatial multiplexing architectures with jointly designed ratetailoring and ordered BLAST decoding part-II: A practical method for rate and power allocation," *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 3626–3271, Aug. 2008.

[13] Q. Zhou and H. Dai, "Joint antenna selection and link adaptation for MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 55, no. 1, pp. 243–255, Jan. 2006.

[14] 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, pp. 1334–1346, Mar. 2010.

[15] M. A. Sadrabadi, A. K. Khandani, and F. Lahouti, "Channel feedback quantization for high data rate MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3335–3338, Dec. 2006.

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

[17] R. Yellapantula, Y. Yao, and R. Ansari, "Unitary precoding and power control in MIMO systems with limited feedback," presented at the IEEE Wireless Commun. Netw. Conf., 2006.

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

[19] J. Liu, D. Xu, R. Zhoa, L. Yang, and X. You, "Adaptive transmission for finite-rate feedback MIMO systems," presented at the IEEE Wireless Commun. Netw. Conf., 2008.

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

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

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

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

[24] 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.

[25] R. Stark and J. W. Woods, *Probability and Random Processes With Application to Signal Processing*, 3rd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 2002.

[26] 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.

[27] E. K. P. Chong and S. H. Zak, *An Introduction to Optimization*. New York, NY, USA: John Wiley & Sons, 1996.

[28] B. Farber and K. Zeger, "Quantization of multiple sources using nonnegative integer bit allocation," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 4945–4964, Nov. 2006.

[29] B. Fox, "Discrete optimization via marginal analysis," *Manag. Sci.*, vol. 13, pp. 210–216, Nov. 1966.

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

[31] S. G. Krantz and H. R. Parks, *A primer of Real Analytic Functions*. Berlin, Germany: Birkhauser Verlag, Basel, 1992.

[32] W. Krandick and K. Mehlhorn, "New bounds for the descartes method," *J. Symbolic Comput.*, vol. 41, pp. 49–66, 2006.

Hung-Chun Chen was born in Pingtung, Taiwan, in 1985. She received the B.S. degree in electrical and control engineering in 2008 from National ChiaoTung University, Hsinchu, Taiwan, where she is currently working towards the Ph.D. degree with the Department of Electrical Engineering. Her research interests include signal processing for digital communications and wireless communications.

Yuan-Pei Lin (S'93–M'97–SM'03) was born in Taipei, Taiwan, 1970. She received the B.S. degree in control engineering from the National Chiao-Tung University, Taiwan, in 1992, and the M.S. degree and the Ph.D. degree, both in electrical engineering from California Institute of Technology, in 1993 and 1997, respectively. She joined the Department of Electrical and Control Engineering of National Chiao-Tung University, Taiwan, in 1997. Her research interests include digital signal processing, multirate filter banks, and signal processing for digital communications.

She was a recipient of Ta-You Wu Memorial Award in 2004. She served as an associate editor for IEEE TRANSACTION ON SIGNAL PROCESSING, IEEE TRANSACTION ON CIRCUITS AND SYSTEMS II, IEEE SIGNAL PROCESSING LETTERS, IEEE TRANSACTION ON CIRCUITS AND SYSTEMS I, EURASIP Journal on Applied Signal Processing, and Multidimensional Systems and Signal Processing, Academic Press. She was a Distinquished Lecturer of the IEEE Circuits and Systems Society for 2006–2007. She has also coauthored two books, *Signal Processing and Optimization for Transceiver Systems, and Filter Bank Transceivers for OFDM and DMT Systems*, both by Cambridge University Press, 2010.