This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE Globecom 2011 proceedings.

Optimal MMSE Beamforming for Multiuser Downlink with Delayed CSI Feedback Using Codebooks

Binbin Dai, Wei Xu, and Chunming Zhao

National Mobile Communications Research Lab., Southeast University, Nanjing, 210096, P. R. China
({daibinbin, wxu, cmzhao}@seu.edu.cn)

***Abstract*—This paper investigates the beamforming design for multiuser downlink with limited feedback, where a practical channel state information (CSI) feedback model consisting both feedback delay and CSI quantization error is considered. Under this circumstance, we derive a closed-form multiuser beamforming scheme via minimizing the expected mean square error (MSE) with respect to the feedback imperfection. Compared with conventional MMSE beamforming scheme, we find that not only the regularization factor but the optimal MMSE beamforming structure changes due to the CSI imperfection, especially the feedback delay. Numerical results verify in different cases that the optimized beamforming design outperforms conventional ones in terms of both sum rate and BER performance.**

I. Introduction

Multiple-input multiple-output (MIMO) communications have been intensely investigated over the past decades. It has been well-known that, by utilizing proper pre-processing techniques, MIMO is able to offer a dramatic capacity improvement and higher reliability in comparison with single-input single-output (SISO) systems [1]. In a multiuser MIMO (MU-MIMO) downlink channel, it has been revealed in [2] and [3] that dirty paper coding (DPC) is a capacity-achieving precoding scheme. However, due to its high complexity, a family of linear pre-processing techniques [4], [5] has recently attracted more attentions for practical applications.

Among the linear pre-processing family, zero-forcing beamforming (ZFBF) [4] is one of the most efficient methods. It inverses the concatenated channel of the entire system and hence perfectly eliminates the multiuser interference. Although ZFBF has been proven to be asymptotically optimal for the multiuser system at high signal-to-noise ratios (SNRs), a noticeable performance degradation can be found at low SNRs because of the effect of noise amplification. Under this consideration, an improved beamforming approach called minimum mean square error beamforming (MMSEBF) was proposed in [5]. The MMSEBF with vector-perturbation technique can obtain good performance at all SNRs.

Although the above methods can achieve near-capacity performance with low complexity, they are based on an infeasible and impractical assumption of perfect CSI at the transmitter (CSIT). To deal with this problem, a number of practical techniques have been developed in [6]–[9] by exploiting quantized CSI feedback based on pre-determined codebooks. In [6], a finite-rate feedback strategy exploiting ZFBF and quantized CSI feedback was discussed at full length. It discovered that the feedback rate per mobile must increase linearly with the SNR in order to achieve full multiplexing gain. However, the beamforming strategy in [6] simply followed the conventional ZFBF scheme and replaced the CSI matrix with its quantized version. In order to suppress the negative effect of CSI quantization, robust beamforming designs were then proposed in [8] and [9].

However, these robust beamforming schemes were designed under the assumption of delay-free feedback links from remote users to the transmitter. As we know, in practical applications, the feedback delay is inevitable and sometimes plays an important role in performance degradation [10]. Therefore, in this paper, we aim at optimizing multiuser beamforming for the finite-rate feedback system by considering both CSI quantization error and feedback delay from remote users. Through optimization theory, we derive the optimal beamforming scheme via minimizing the expected MSE with respect to the CSI imperfection. The derived beamforming scheme achieves better performance, especially for delayed feedback links, while costs similar computational complexity as conventional ones.

Notations: In this letter, we denote the transpose, the conjugate transpose, the trace and the Frobenius norm of matrix $\mathbf{A}$ as $\mathbf{A}^{\mathrm{T}}$, $\mathbf{A}^{\dagger}$, $\mathrm{tr}(\mathbf{A})$ and $\|\mathbf{A}\|_{\mathrm{F}}$ respectively. We use $\|\mathbf{x}\|$ to denote the Euclidean norm of vector $\mathbf{x}$. $\mathbf{I}_M$ is the $M \times M$ identity matrix. $\mathrm{diag}\{\cdots\}$ represents a diagonal matrix with the given elements on the diagonal. $\mathbb{E}\{\cdot\}$ stands for the statistical expectation of a random variable and $\mathrm{Re}[\cdot]$ returns the real part of the input.

II. System Model

Consider a multiuser downlink with an M-antenna base station (BS) and K single-antenna users, as shown in Fig. 1. For simplicity, we only focus on the case with $K = M$ in this paper. As for $K > M$, the system can implement user selection at first. Denote the channel vector from the BS to user k by $\mathbf{h}_k \in \mathbb{C}^{M\times 1}$ whose entries are independent and identically distributed (i.i.d) complex Gaussian variables

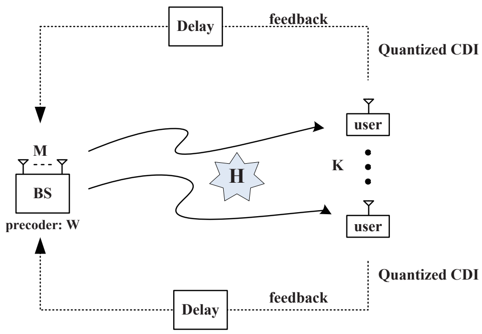

Fig. 1. System Model.

with zero mean and unit variance. Let $\mathbf{W} \in \mathbb{C}^{M\times M}$ and $\mathbf{x} \in \mathbb{C}^{M\times 1}$ be the precoder matrix and transmitted vector signal respectively. Then, the received symbol at the kth user can be expressed as

$$y_k = \mathbf{h}_k^\dagger \mathbf{W}\mathbf{x} + n_k, k = 1, \cdots, M. \tag{1}$$

where n_k is the zero-mean and unit variance complex Gaussian noise at user k. By stacking y_k and n_k, we get

$$\mathbf{y} = \mathbf{H}\mathbf{W}\mathbf{x} + \mathbf{n} \tag{2}$$

where $\mathbf{y} = [y_1 \cdots y_M]^{\mathrm{T}}$, $\mathbf{H} = [\mathbf{h}_1 \cdots \mathbf{h}_M]^\dagger$ and $\mathbf{n} = [n_1 \cdots n_M]^{\mathrm{T}}$. Note that the energy of the transmitted symbol and the noise are both normalized, $\mathbb{E}\{\mathbf{x}\mathbf{x}^\dagger\} = \mathbb{E}\{\mathbf{n}\mathbf{n}^\dagger\} = \mathbf{I}_M$. The transmit power at the BS is then limited by $\mathrm{tr}\left(\mathbf{W}\mathbf{W}^\dagger\right) = P$.

Assume that each user k has a full knowledge of $\mathbf{h}_k$ obtained by channel estimation as in [6]. Each user quantizes its channel direction information (CDI), i.e., normalized CSI $\tilde{\mathbf{h}}_k = \mathbf{h}_k / \|\mathbf{h}_k\|$, to a codeword in a given codebook $\mathbf{C}_k \in \mathbb{C}^{M\times 2^B}$, where B represents the number of feedback bits. Then, the user sends the index of the codeword to the BS. Since the codebook is predetermined and known at both two sides, the BS can thus recover the quantized CSI, denoted by $\hat{\mathbf{h}}_k$, from each user k. The quantized CSI is then utilized by the BS for calculating the beamforming matrix. Besides, since there is always a feedback delay due to the process of channel estimation and feedback transmission, the quantized CSI available at the BS is actually a quantized version of the true CSI at the previous time-slot.

III. Channel Model

In order to formulate the beamforming optimization problem, in this section, we first introduce a synthesized channel model considering the effect of both quantization error and feedback delay. The proposed robust beamforming in the next section is based on the synthesized channel model.

A. CSI Quantization Model

Using the finite-rate feedback strategy in [6], the relationship between CDI, i.e. the normalized CSI $\tilde{\mathbf{h}}_k$ at user k, and its quantization version, i.e. $\hat{\mathbf{h}}_k$, can be expressed by:

$$\tilde{\mathbf{h}}_k = \sqrt{1 - z_k}\hat{\mathbf{h}}_k + \sqrt{z_k}\mathbf{s}_k \tag{3}$$

where $\mathbf{s}_k$ is the unit-norm quantization error vector isotropically distributed in the null space of $\hat{\mathbf{h}}_k$ with the same dimension of $\hat{\mathbf{h}}_k$ as well as $\tilde{\mathbf{h}}_k$ and is independent of $z_k = 1 - \|\tilde{\mathbf{h}}_k^\dagger \hat{\mathbf{h}}_k\|^2$. The expectation of the quantization error z_k is characterized in [6] by

$$\mathbb{E}\{z_k\} = \frac{M-1}{M}\delta \tag{4}$$

where $\delta = 2^{-B/(M-1)}$.

B. Delayed CSI Feedback

We consider a stationary ergodic Gauss-Markov block fading process where the channel stays constant during one symbol duration but changes independently between consecutive transmissions. A classic model of this kind of channel can be characterized by [10]

$$\mathbf{h}_k[n] = \rho_k \mathbf{h}_k[n-1] + \mathbf{e}_k[n] \tag{5}$$

where $\mathbf{h}_k[n]$ and $\mathbf{h}_k[n-1]$ represent the real channel information of user k at current time n and the previous time-slot $(n-1)$ respectively. $\mathbf{e}_k[n]$ is independent of $\mathbf{h}_k[n-1]$ with its entries $e_{ki}[n] \sim \mathcal{CN}\left(0, \epsilon_k^2\right)$ i.i.d. According to Jake's model, we can define $\rho_k = J_0\left(2\pi f_{dk} T_s\right)$ where $J_0(\cdot)$ is the zero-th order Bessel function of the first kind, f_{dk} and T_s are denoted as the Doppler spread of user k and the length of each symbol respectively. The variance of $\mathbf{e}_k[n]$ is constrained by $\epsilon_k^2 = 1 - \rho_k^2$.

C. Synthesized Channel Model

By utilizing the CSI quantization and feedback delay model, the synthesized channel model in this study can be directly formulated according to the finite-rate feedback procedure. Without loss of generality, the kth user gets $\mathbf{h}_k[n-1]$ at time-slot $(n-1)$ and then quantizes its CDI $\tilde{\mathbf{h}}_k[n-1]$ into $\hat{\mathbf{h}}_k[n-1]$. Via a finite-rate feedback link, the BS obtains $\hat{\mathbf{h}}_k[n-1]$ to calculate the precoder $\mathbf{W}$. Due to the feedback delay, the BS will transmit data by using $\mathbf{W}$ at the following time-slot n. Accordingly, from (3) and (5), the relationship between current CSI, $\mathbf{h}_k[n]$, and the quantized CSI available at the BS, $\hat{\mathbf{h}}_k[n-1]$, can be formulated by

$$\begin{aligned}\mathbf{h}_k[n] =& \rho_k \|\mathbf{h}_k[n-1]\| \Bigg(\sqrt{1 - z_k[n-1]}\hat{\mathbf{h}}_k[n-1] \\ &+ \sqrt{z_k[n-1]}\mathbf{s}_k[n-1] \Bigg) + \mathbf{e}_k[n]. \end{aligned} \tag{6}$$

Further by stacking user channels as a matrix $\mathbf{H} = [\mathbf{h}_1[n] \cdots \mathbf{h}_M[n]]^\dagger$, the synthesized channel can be re-written concisely as

$$\begin{aligned}\mathbf{H} &= \mathbf{D}\mathbf{G}\left((\mathbf{I}_M - \mathbf{Z})^{1/2}\hat{\mathbf{H}} + \mathbf{Z}^{1/2}\mathbf{S} \right) + \mathbf{F} \\ &\triangleq \mathbf{A}\hat{\mathbf{H}} + \mathbf{B}\mathbf{S} + \mathbf{F} \end{aligned} \tag{7}$$

where $\mathbf{D} = \mathrm{diag}\{\rho_1, \cdots, \rho_M\}$,
$\mathbf{G} = \mathrm{diag}\{\|\mathbf{h}_1[n-1]\|, \cdots, \|\mathbf{h}_M[n-1]\|\}$,
$\mathbf{Z} = \mathrm{diag}\{z_1[n-1], \cdots, z_M[n-1]\}$,
$\hat{\mathbf{H}} = \left[\hat{\mathbf{h}}_1[n-1] \cdots \hat{\mathbf{h}}_M[n-1]\right]^\dagger$,

$\mathbf{S} = [\mathbf{s}_1[n-1]\cdots\mathbf{s}_M[n-1]]^\dagger$, $\mathbf{F} = [\mathbf{e}_1[n]\cdots\mathbf{e}_M[n]]^\dagger$, $\mathbf{A} = \mathbf{DG}(\mathbf{I}_M - \mathbf{Z})^{1/2}$ and $\mathbf{B} = \mathbf{DGZ}^{1/2}$.

IV. Robust Beamforming Design

As we stated in above sections, conventional beamforming methods for limited feedback system either simply use the quantized CSI to replace the true one for beamforming calculation [6] or ignore the delay effect of the feedback process [8], [9]. In this section, we intend to develop an improved beamforming regarding the impact of both imperfection factors in order to better suit practical applications.

We optimize the beamforming design via minimizing the MSE of the received symbols. According to the definition of MSE in [11], the MSE is calculated by

$$\begin{aligned}\varepsilon(\mathbf{W},\beta) &= \mathbb{E}_{\mathbf{x},\mathbf{n}}\left\{\|\mathbf{x}-\beta^{-1}\mathbf{y}\|^2\right\}\\ &= \|\beta^{-1}\mathbf{HW}-\mathbf{I}_M\|_{\mathrm{F}}^2 + M\beta^{-2}\end{aligned} \tag{8}$$

where β is a scaling factor at the receiver. Although equation (8) is calculated for distributed users, the MMSE precoding design is concentrated by optimizing $\mathbf{W}$ at the BS with all channel informations $\mathbf{H}$ feedback. In our system, since only outdated CSI quantization version is available at the BS, the MSE cost function should be modified as a conditional expectation over the CSI imperfection. Mathematically, we can re-formulate the beamforming optimization problem as

$$\begin{aligned}&\text{minimize} && \mathbb{E}_{\mathbf{H}|\hat{\mathbf{H}}}\{\varepsilon(\mathbf{W},\beta)\}\\ &\text{subject to} && \mathrm{tr}\left(\mathbf{WW}^\dagger\right) = P.\end{aligned} \tag{9}$$

Then, by substituting (7) into (8), the cost function in (9) is equivalent to

$$\begin{aligned}&\mathbb{E}_{\mathbf{H}|\hat{\mathbf{H}}}\{\varepsilon(\mathbf{W},\beta)\}\\ =&\mathbb{E}_{\mathbf{A},\mathbf{B},\mathbf{S},\mathbf{F}}\left\{\|\beta^{-1}\left(\mathbf{A}\hat{\mathbf{H}}+\mathbf{BS}+\mathbf{F}\right)\mathbf{W}-\mathbf{I}_M\|_{\mathrm{F}}^2\right\}+M\beta^{-2}\\ =&\mathbb{E}\{\mathrm{tr}(\mathbf{I}_M)\}+M\beta^{-2}+\mathbb{E}_{\mathbf{A}}\left\{\mathrm{tr}\left(\beta^{-2}\mathbf{W}^\dagger\hat{\mathbf{H}}^\dagger\mathbf{A}^\dagger\mathbf{A}\hat{\mathbf{H}}\mathbf{W}\right)\right\}\\ &+\mathbb{E}_{\mathbf{B},\mathbf{S}}\left\{\mathrm{tr}\left(\beta^{-2}\mathbf{W}^\dagger\mathbf{S}^\dagger\mathbf{B}^\dagger\mathbf{BSW}\right)\right\}\\ &+\mathbb{E}_{\mathbf{F}}\left\{\mathrm{tr}\left(\beta^{-2}\mathbf{W}^\dagger\mathbf{F}^\dagger\mathbf{FW}\right)\right\}-2\mathbb{E}_{\mathbf{A}}\left\{\mathrm{Re}\left[\mathrm{tr}\left(\beta^{-1}\mathbf{A}\hat{\mathbf{H}}\mathbf{W}\right)\right]\right\}\\ &-2\mathbb{E}_{\mathbf{F}}\left\{\mathrm{Re}\left[\mathrm{tr}\left(\beta^{-1}\mathbf{FW}\right)\right]\right\}-2\mathbb{E}_{\mathbf{B},\mathbf{S}}\left\{\mathrm{Re}\left[\mathrm{tr}\left(\beta^{-1}\mathbf{BSW}\right)\right]\right\}\\ &+2\mathbb{E}_{\mathbf{A},\mathbf{F}}\left\{\mathrm{Re}\left[\mathrm{tr}\left(\beta^{-2}\mathbf{W}^\dagger\mathbf{F}^\dagger\mathbf{A}\hat{\mathbf{H}}\mathbf{W}\right)\right]\right\}\\ &+2\mathbb{E}_{\mathbf{A},\mathbf{B},\mathbf{S}}\left\{\mathrm{Re}\left[\mathrm{tr}\left(\beta^{-2}\mathbf{W}^\dagger\hat{\mathbf{H}}^\dagger\mathbf{A}^\dagger\mathbf{BSW}\right)\right]\right\}\\ &+2\mathbb{E}_{\mathbf{B},\mathbf{S},\mathbf{F}}\left\{\mathrm{Re}\left[\mathrm{tr}\left(\beta^{-2}\mathbf{W}^\dagger\mathbf{F}^\dagger\mathbf{BSW}\right)\right]\right\}\end{aligned} \tag{10}$$

Note that $\mathbf{S}$ and $\mathbf{F}$ are independent of each other as well as $\mathbf{A}$, $\mathbf{B}$. Thus, by considering the results $\mathbb{E}\{\mathbf{S}\} = \mathbf{0}$ and $\mathbb{E}\{\mathbf{F}\} = \mathbf{0}$, it is not hard to verify that the last five items in (10) are all zeros. Hence, (10) is further simplified to

$$\begin{aligned}&\mathbb{E}_{\mathbf{H}|\hat{\mathbf{H}}}\{\varepsilon(\mathbf{W},\beta)\}\\ =&M\left(1+\beta^{-2}\right)+\mathbb{E}_{\mathbf{A}}\left\{\mathrm{tr}\left(\beta^{-2}\mathbf{W}^\dagger\hat{\mathbf{H}}^\dagger\mathbf{A}^\dagger\mathbf{A}\hat{\mathbf{H}}\mathbf{W}\right)\right\}\\ &+\mathbb{E}_{\mathbf{B},\mathbf{S}}\left\{\mathrm{tr}\left(\beta^{-2}\mathbf{W}^\dagger\mathbf{S}^\dagger\mathbf{B}^\dagger\mathbf{BSW}\right)\right\}\\ &+\mathbb{E}_{\mathbf{F}}\left\{\mathrm{tr}\left(\beta^{-2}\mathbf{W}^\dagger\mathbf{F}^\dagger\mathbf{FW}\right)\right\}-2\mathbb{E}_{\mathbf{A}}\left\{\mathrm{Re}\left[\mathrm{tr}\left(\beta^{-1}\mathbf{A}\hat{\mathbf{H}}\mathbf{W}\right)\right]\right\}\end{aligned} \tag{11}$$

As channel magnitude information $\|\mathbf{h}_k\|$ is independent of channel direction information $\tilde{\mathbf{h}}_k$, $\|\mathbf{h}_k\|$ is therefore also independent of z_k so that the following equations can be derived.

$$\mathbb{E}\{\mathbf{A}\} = \mathbf{D}\mathbb{E}\{\|\mathbf{h}_1\|\}\mathbb{E}\left\{\sqrt{1-z_1}\right\} \triangleq \alpha\mathbf{D} \tag{12}$$

$$\mathbb{E}\left\{\mathbf{A}^\dagger\mathbf{A}\right\} = [M-(M-1)\delta]\,\mathbf{D}^2 \tag{13}$$

By substituting (12), (13), and (19), (20) (see Appendix A) into (11), the objective function can be expressed by

$$\begin{aligned}&\mathbb{E}_{\mathbf{H}|\hat{\mathbf{H}}}\{\varepsilon(\mathbf{W},\beta)\}\\ =&[M-(M-1)\delta]\,\beta^{-2}\|\mathbf{D}\hat{\mathbf{H}}\mathbf{W}\|_{\mathrm{F}}^2\\ &+\delta\beta^{-2}\left(P\|\mathbf{D}\|_{\mathrm{F}}^2-\|\mathbf{D}\hat{\mathbf{H}}\mathbf{W}\|_{\mathrm{F}}^2\right)+P\beta^{-2}\left(\sum_{k=1}^{M}\epsilon_k^2\right)\\ &-2\alpha\beta^{-1}\mathrm{Re}\left\{\mathrm{tr}\left(\mathbf{D}\hat{\mathbf{H}}\mathbf{W}\right)\right\}+M\left(1+\beta^{-2}\right)\\ =&M(1-\delta)\,\beta^{-2}\|\mathbf{D}\hat{\mathbf{H}}\mathbf{W}\|_{\mathrm{F}}^2-2\alpha\beta^{-1}\mathrm{Re}\left\{\mathrm{tr}\left(\mathbf{D}\hat{\mathbf{H}}\mathbf{W}\right)\right\}\\ &+\beta^{-2}\left[M(1+P)-P(1-\delta)\|\mathbf{D}\|_{\mathrm{F}}^2\right]+M\\ \triangleq&\mu(\mathbf{W},\beta)\end{aligned} \tag{14}$$

Hence, the primal optimization problem in (9) is equivalent to

$$\begin{aligned}&\text{minimize} && \mu(\mathbf{W},\beta)\\ &\text{subject to} && \mathrm{tr}\left(\mathbf{WW}^\dagger\right) = P.\end{aligned} \tag{15}$$

At this point, we can use the *Lagrange multiplier method* to solve problem (15). Following the detailed derivations in Appendix B, we fortunately obtain a closed-form solution to $\mathbf{W}$ as:

$$\mathbf{W} = \eta\left[\left(\mathbf{D}\hat{\mathbf{H}}\right)^\dagger\mathbf{D}\hat{\mathbf{H}}+\xi\mathbf{I}_M\right]^{-1}\left(\mathbf{D}\hat{\mathbf{H}}\right)^\dagger \tag{16}$$

where $\xi = \frac{M(1+P)-P(1-\delta)\|\mathbf{D}\|_{\mathrm{F}}^2}{PM(1-\delta)}$ and η is the transmit power constraint factor given by

$$\eta = \sqrt{\frac{P}{\left\|\left[\left(\mathbf{D}\hat{\mathbf{H}}\right)^\dagger\mathbf{D}\hat{\mathbf{H}}+\xi\mathbf{I}_M\right]^{-1}\left(\mathbf{D}\hat{\mathbf{H}}\right)^\dagger\right\|_{\mathrm{F}}^2}}. \tag{17}$$

This result presents the optimal precoder under both CSI quantization error and feedback delay. In [9], it has been revealed that when only the quantization error is considered for beamforming optimization, the only difference between the robust MMSEBF scheme in [9] and the conventional MMSE precoder in [6] lies in the regularization factor. However, by comparing (16) with [6], we find that *not only the regularization factor but also the quantized CDI matrix* $\hat{\mathbf{H}}$ *in the optimized MMSE precoder calculation expression is changed*. Moreover, from a generalized viewpoint, the proposed beamforming technique (16) can be treated as a general expression of optimal MMSE precoder for finite-rate feedback system. When there is no delay with $\mathbf{D} = \mathbf{I}_M$, the expression (16) reduces to the previous one given in [9]. Besides these, since all the mentioned schemes share a similar MMSE precoder

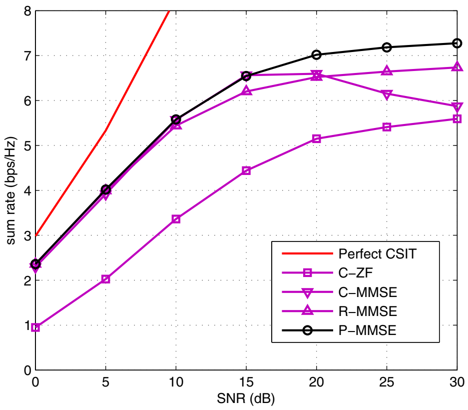

Fig. 2. sum rate v.s. SNR under $M = K = 4$, $B = 16$ bits and $f_d T = 0.1$.

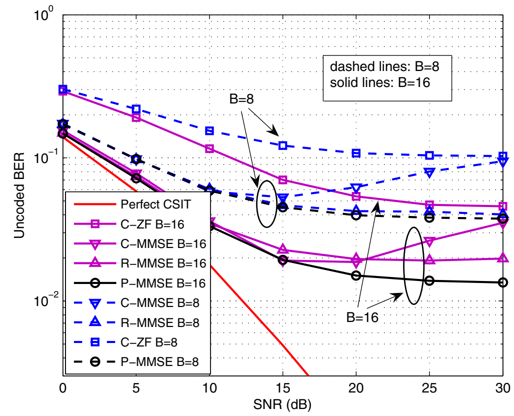

Fig. 3. Uncoded BER v.s. SNR under $M = K = 4$ and $f_d T = 0.05$.

structure with closed-form expressions, it is also necessary to stress that our proposed scheme has a comparable complexity as conventional ones.

V. Numerical Results

In this section, we compare our proposed beamforming with conventional schemes by computer simulation. We tested $M = K = 4$ and used randomly generated codebooks as in [6] for vector channel quantization. For simplicity, we set a same Doppler spread f_d and symbol duration T for all users, which indicates each user has the same delay correlation efficiency determined by the normalized Doppler frequency $f_d T$. In our simulation, we tested the following scheme for comparison:

- the proposed MMSE beamforming scheme based on delayed and quantized CDI feedback (P-MMSE),
- conventional ZFBF scheme in [6] (C-ZF),
- conventional MMSE beamforming in [6] (C-MMSE),
- robust MMSE beamforming proposed in [9] (R-MMSE).

Moreover, we also provided the numerical results obtained by using MMSE beamforming with perfect CSIT for reference.

Fig. 2 presents the sum rate performance of different schemes under $B = 16$ bits and $f_d T = 0.1$. It verifies that P-MMSE achieves a higher sum rate than other schemes only except for the perfect CSIT case. It also can be found that P-MMSE overcomes the problem that the C-MMSE scheme [6] has met a sum rate degradation at high SNRs. Fig. 3 compares these schemes in terms of BER performance, where BPSK modulation and MRC are employed. P-MMSE shows a noticeable BER performance gain over other finite-rate feedback beamforming schemes, especially when the system SNR grows. Moreover, from Fig. 3, we find that the achieved performance gain becomes more pronounced as B increases.

In Fig. 4, we compare the performance of different schemes under different number of feedback bits and an interesting phenomenon is found. Since no delay effect is taken into account in [6] and [9], both C-MMSE and R-MMSE suffer an obvious performance degradation in terms of sum rate. From Fig. 4, we find that when B increases, the achievable rate of

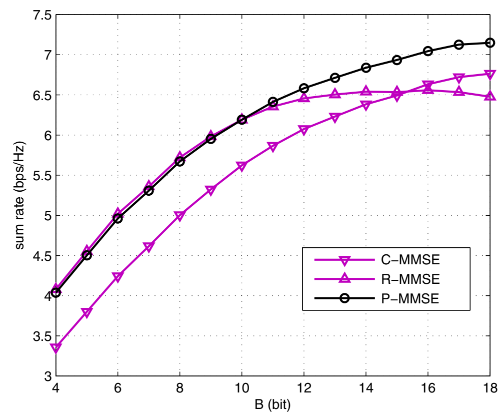

Fig. 4. sum rate v.s. B under $M = K = 4$, SNR=20dB and $f_d T = 0.1$.

R-MMSE, which only considers quantization error in beamforming design, converges to and even later becomes worse than that of the C-MMSE which considers neither quantization error nor delay. However, P-MMSE always outperforms both these two schemes when B grows large. Moreover, it is also necessary to note that P-MMSE shows a slight weakness over the R-MMSE scheme at a small B in Fig. 4. This is because the MMSE criteria is not equivalent to sum capacity maximization, especially when B decreases and quantization error dominates the entire system performance. Due to this incoincidence, it is reasonable that P-MMSE is not optimal in terms of sum rate in every single case and hence showing a slight sum rate degradation at a small B when compared with [9]. Even though, both P-MMSE and R-MMSE obtain a significant capacity improvement over the C-MMSE under low levels of CSI feedback. Consequently, from above observations, we can conclude that our proposed beamforming technique is a generally optimized scheme under various cases.

VI. Conclusion

We propose an improved MMSE beamforming scheme for MU-MISO downlink channel with delayed and quantized CDI feedback. The closed-form of the proposed precoder shows that quantization error affects the regularization factor while delay has an impact on the quantized CDI matrix as well as the regularization factor when compared to the conventional MMSE precoder formula. Numerical results verify the effectiveness of our proposed scheme over existing methods.

VII. Acknowledgement

This work was supported in part by the international cooperation project in China under grant 2010DFB10410, the National Science & Technology Projects of China under grant 2010ZX03002-012, and the Research Fund of the National Mobile Communications Research Laboratory, Southeast University (No. 2011B01). This work was also supported in part by Ericsson.

Appendix A
Some Preliminary Results

Because the entries of $\mathbf{e}_k$ are i.i.d. complex Gaussian variables with zero mean and variance ϵ_k^2, we have

$$\mathbb{E}\left\{\mathbf{e}_k\mathbf{e}_k^\dagger\right\} = \epsilon_k^2\mathbf{I}_M \tag{18}$$

which yields

$$\mathbb{E}\left\{\mathbf{F}^\dagger\mathbf{F}\right\} = \mathbb{E}\left\{\sum_{k=1}^{M}\mathbf{e}_k\mathbf{e}_k^\dagger\right\} = \left(\sum_{k=1}^{M}\epsilon_k^2\right)\mathbf{I}_M \tag{19}$$

Then, we consider another term as follows:

$$\begin{aligned}
&\mathbb{E}_{\mathbf{B},\mathbf{S}}\left\{\mathrm{tr}\left[\mathbf{W}^\dagger\mathbf{S}^\dagger\mathbf{B}^\dagger\mathbf{B}\mathbf{S}\mathbf{W}\right]\right\}\\
=&\delta\left(M-1\right)\mathbb{E}_{\mathbf{S}}\left\{\mathrm{tr}\left[\mathbf{W}^\dagger\mathbf{S}^\dagger\mathbf{D}^2\mathbf{S}\mathbf{W}\right]\right\}\\
=&\delta\left(M-1\right)\mathbb{E}\left\{\mathrm{tr}\left[\mathbf{W}^\dagger\left(\sum_{k=1}^{M}\rho_k^2\mathbf{s}_k\left[n-1\right]\mathbf{s}_k^\dagger\left[n-1\right]\right)\mathbf{W}\right]\right\}\\
=&\delta\mathbb{E}\left\{\mathrm{tr}\left[\mathbf{W}^\dagger\left(\sum_{k=1}^{M}\rho_k^2\left(\mathbf{I}_M-\hat{\mathbf{h}}_k\left[n-1\right]\hat{\mathbf{h}}_k^\dagger\left[n-1\right]\right)\right)\mathbf{W}\right]\right\}\\
=&\delta\left(\|\mathbf{D}\|_{\mathrm{F}}^2P-\|\mathbf{D}\hat{\mathbf{H}}\mathbf{W}\|_{\mathrm{F}}^2\right)
\end{aligned}\tag{20}$$

where we used previous results $\mathbb{E}\left\{\mathbf{B}^\dagger\mathbf{B}\right\} = \left(M-1\right)\delta\mathbf{D}^2$ and $\mathbb{E}\left\{\mathbf{s}_k\mathbf{s}_k^\dagger\right\} = \frac{1}{M-1}\left(\mathbf{I}_M-\hat{\mathbf{h}}_k\hat{\mathbf{h}}_k^\dagger\right)$ in [9].

Appendix B
Proof of equation (16) and (17)

By using the *Lagrange Multiplier method*, we first write the Lagrangian cost function of problem (15) as

$$\mathcal{L}\left(\mathbf{W},\beta,\lambda\right) = \mu\left(\mathbf{W},\beta\right)+\lambda\left(\mathrm{tr}\left(\mathbf{W}\mathbf{W}^\dagger\right)-P\right) \tag{21}$$

where $\lambda\in\mathbb{R}$ is the Lagrange multiplier. Then, taking partial derivatives of $\mathcal{L}\left(\mathbf{W},\beta,\lambda\right)$ with respect to $\mathbf{W}^*$ and β respectively and setting them to zero, we have

$$\frac{\partial\mathcal{L}}{\partial\mathbf{W}^*} = M\left(1-\delta\right)\beta^{-2}\hat{\mathbf{H}}^\dagger\mathbf{D}^2\hat{\mathbf{H}}\mathbf{W}-\alpha\beta^{-1}\hat{\mathbf{H}}^\dagger\mathbf{D}+\lambda\mathbf{W}=0 \tag{22}$$

$$\begin{aligned}
\frac{\partial\mathcal{L}}{\partial\beta} =& -2\beta^{-3}\Big(M\left(1-\delta\right)\|\mathbf{D}\hat{\mathbf{H}}\mathbf{W}\|_{\mathrm{F}}^2+M\left(1+P\right)\\
&-P\left(1-\delta\right)\|\mathbf{D}\|_{\mathrm{F}}^2\Big)+2\alpha\beta^{-2}\mathrm{Re}\left[\mathrm{tr}\left(\mathbf{D}\hat{\mathbf{H}}\mathbf{W}\right)\right]=0
\end{aligned}\tag{23}$$

After some basic manipulations from (22) and (23), it yields

$$\begin{aligned}
\mathbf{W} &= \alpha\beta\left(M\left(1-\delta\right)\hat{\mathbf{H}}^\dagger\mathbf{D}^2\hat{\mathbf{H}}+\lambda\beta^2\mathbf{I}_M\right)^{-1}\hat{\mathbf{H}}^\dagger\mathbf{D}\\
&\triangleq \alpha\beta\mathbf{Q}\hat{\mathbf{H}}^\dagger\mathbf{D}
\end{aligned}\tag{24}$$

and

$$\begin{aligned}
&\alpha\beta\mathrm{Re}\left[\mathrm{tr}\left(\mathbf{D}\hat{\mathbf{H}}\mathbf{W}\right)\right]-M\left(1-\delta\right)\|\mathbf{D}\hat{\mathbf{H}}\mathbf{W}\|_{\mathrm{F}}^2\\
&=\alpha^2\beta^2\mathrm{tr}\left\{\hat{\mathbf{H}}^\dagger\mathbf{D}^2\hat{\mathbf{H}}\mathbf{Q}\left(\mathbf{I}_M-M\left(1-\delta\right)\hat{\mathbf{H}}^\dagger\mathbf{D}^2\hat{\mathbf{H}}\mathbf{Q}\right)\right\}\\
&\overset{(a)}{=}\alpha^2\beta^2\mathrm{tr}\left\{\hat{\mathbf{H}}^\dagger\mathbf{D}^2\hat{\mathbf{H}}\mathbf{Q}\lambda\beta^2\mathbf{Q}\right\}\\
&=\lambda\beta^2P\\
&=M\left(1+P\right)-P\left(1-\delta\right)\|\mathbf{D}\|_{\mathrm{F}}^2
\end{aligned}\tag{25}$$

where (a) holds because $\mathbf{I}_M-M\left(1-\delta\right)\hat{\mathbf{H}}^\dagger\mathbf{D}^2\hat{\mathbf{H}}\mathbf{Q}=\lambda\beta^2\mathbf{Q}$. From (25), we can get

$$\lambda\beta^2=\left(M\left(1+P\right)-P\left(1-\delta\right)\|\mathbf{D}\|_{\mathrm{F}}^2\right)/P \tag{26}$$

Finally by combining (26), (24) and transmit power constraint $\mathrm{tr}\left(\mathbf{W}\mathbf{W}^\dagger\right)=P$, we have the optimal solution to $\mathbf{W}$ in (16) and (17).

References

[1] J. Duplicy, D. P. Palomar, and L. Vandendorpe, “Adaptive orthogonal beamforming for the MIMO broadcast channel,” in *Proc. IEEE CAMSAP*, Virgin Islands, USA, Dec. 2007, pp. 77-80.

[2] M. H. M. Costa, “Writing on dirty paper,” *IEEE Trans. Inform. Theory*, vol, 29, no.3, pp. 439-441, May 1983.

[3] H. Weingarten, Y. Steinberg, and S. Shamai, “The capacity region of the Gaussian MIMO broadcast channel,” in *Proc. IEEE Int. Symp. Inform. Theory(ISIT)*, vol. 48, pp. 1277-94, June 2002.

[4] G. Caire and S. Shamai, “On the achievable througput of a multiantenna Gaussian broadcast channel,” *IEEE Trans. Inform. Theory*, vol. 49, no. 7, pp. 1691-1706, Jul. 2003.

[5] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, “A vector-perturbation technique for near-capacity multiantenna multiuser communication—part I: channel inversion and regularization,” *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 195-202, Jan. 2005.

[6] N. Jindal, “MIMO broadcast channels with finite-rate feedback,” *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.

[7] G. Caire, N. Jindal, M. Kobayashi and N. Ravindran, “Multiuser MIMO achievable rates with downlink training and channel state feedback,” *IEEE Trans. Inform. Theory*, vol. 56, no. 6, pp. 2845-2866, June 2010.

[8] A. D. Dabbagh and D. J. Love, “Multiple antenna MMSE based downlink precoding with quantized feedback or channel mismatch,” *IEEE Trans. Commun.*, vol. 56, no. 11, pp. 1859–1868, Nov. 2008.

[9] C. Zhang, W. Xu, and M. Chen, “Robust MMSE beamforming for multiuser MISO systems with limited feedback,” *IEEE Sig. Proc. Lett.*, vol. 16, no. 7, pp. 588–591, July 2009.

[10] J. Zhang, R. W. Heath, Jr., M. Kountouris, and J. G. Andrews, “Mode switching for the MIMO broadcast channel based on delay and channel quantization,” *EURASIP Journal on Advances in Signal Processing, special issue on Multiuser Limited Feedback*, vol. 2009, 2009.

[11] M. Joham, W. Utschick, and J. A. Nossek, “Linear transmit processing in MIMO communications systems,” *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2700–2712, Aug. 2005.