

# Progressive Linear Precoder Optimization for MIMO Packet Retransmissions

Haitong Sun, *Student Member, IEEE*, Jonathan H. Manton, *Senior Member, IEEE*, and Zhi Ding, *Fellow, IEEE*

**Abstract**—This paper investigates the optimal linear precoder design for packet retransmissions in multi-input–multi-output (MIMO) systems. To fully utilize the time diversity provided by automatic repeat request (ARQ), we derive a sequence of successive optimal linear ARQ precoders for flat fading MIMO channels, which minimize the mean-square error between the transmitted data and the joint receiver output. The optimization is subject to an overall transmit power constraint. This progressive linear ARQ precoder combines the appropriate power loading and the optimal pairing of channel matrix singular values in the current retransmission with previous transmissions. This optimal pairing is a special feature unique to our sequential ARQ precoding approach. Simulation results demonstrate the effectiveness of this optimized ARQ precoding in reducing symbol MSE and detection bit-error rate.

**Index Terms**—Automatic repeat request (ARQ), multi-input–multi-output (MIMO), minimum mean-square error (MMSE), packet retransmission, precoding.

## I. INTRODUCTION

IT HAS BEEN widely recognized that by utilizing multiple transmit and/or receive antennas, substantial improvement of channel capacity can be achieved by multi-input–multi-output (MIMO) systems. However, to exploit the advantages provided by MIMO, well designed precoding is critical when channel state information (CSI) is available. There has been extensive research on the selection of optimal precoders under various criteria. The optimal precoder design was first considered jointly with the optimal decoder design in MIMO systems. With the objective of minimizing the mean-square error (MSE) under total transmit power constraint, optimal joint linear precoder, and decoder have appeared in [1]–[4]. A more comprehensive linear precoder and decoder design with different criteria of optimality and constraints can be found in [5] and [6]. Recently, precoder design from the perspective of maximizing mutual information (channel capacity) for multiple-antenna systems has also received considerable attention. This subject has been well studied by Telatar [7] for the case of perfect CSI at the transmitter and receiver. Optimal precoder designs with partial (statistical) CSI feedback have also been

investigated for both the multiple-input–single-output (MISO) system [8] and the MIMO system [9], [10].

The design of optimal linear precoders for single-input–single-output (SISO) convolutive systems was considered in [11] and [12]. There, it was explained that a linear precoder can reduce the effects of fading but cannot effectively combat additive noise. Therefore, it was argued that for performance improvement, the linear precoder and associated equalizer should be used as an inner code, with an forward error correction (FEC) encoder and associated decoder added to the system to form an outer code. In particular, this means that appropriate linear precoder/equalizer combination should be chosen to minimize the MSE and not bit error rate (BER). The same design philosophy is used in the present paper.

In practice, to improve link reliability, modern systems are often assisted by the automatic repeat request (ARQ) mechanism. While extensive investigations have been presented for ARQ strategies primarily from the view point of FEC, the integrated effect of MIMO and ARQ has not been adequately addressed. To practically achieved capacity gain provided by MIMO systems, it has been proposed to decompose the MIMO channel into multiple orthogonal subchannels over which data stream can be sent without interchannel interference [13], [14]. By applying the similar idea, hybrid ARQ for the SISO system has been modified and proposed for the MIMO system [15]–[17]. Performance of different combining schemes for Type-I hybrid MIMO ARQ retransmissions has also been analyzed and compared in the [18].

In this paper, we consider the optimal linear precoding of MIMO ARQ retransmissions. We focus our precoder design on flat fading MIMO systems with perfect CSI knowledge at the transmitter and receiver. By sequentially optimizing the precoder for each ARQ transmission, we aim at minimizing the symbol MSE at any transmission stage without considering future retransmissions.

It should be remarked that sequential precoding optimization is a unique feature in ARQ systems as subsequent retransmission is only needed when previous transmissions failed to deliver satisfactory packet reception. As previous transmissions cannot be altered once a new ARQ request is sent and future retransmissions may not be needed, the transmitter can only process the current (re) transmission for optimum effects.

This paper is organized as follows. Section II introduces the system model and Section III summarizes the optimal precoder under the minimum MSE (MMSE) criterion for MIMO systems without ARQ. In Section IV, we extend the optimal precoder design to MIMO system with ARQ (re) transmissions. To minimize the MSE of the joint receiver output, we formulate the

Manuscript received January 16, 2005; revised April 14, 2005. This work was supported in part by the National Science Foundation under Grant ECS-0121469, Grant CCF-0515058, and Grant CNS-05020126, and in part by the U.S. Army Research Office under Grant W911NF-05-1-0382.

H. Sun and Z. Ding are with the Department of Electrical and Computer Engineering, University of California, Davis, CA 95616 USA (e-mail: sunht@ece.ucdavis.edu; zding@ece.ucdavis.edu).

J. H. Manton is with the Department of Information Engineering, Research School of Information Sciences and Engineering (RSISE), Australian National University, Canberra ACT 0200, Australia (e-mail: j.manton@ieee.org).

Digital Object Identifier 10.1109/JSAC.2005.862384

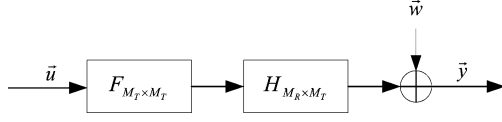


Fig. 1. Block diagram of linearly precoded flat fading MIMO communication systems.

optimal precoder design as a trace minimization problem under transmission power constraint. We then propose the optimal sequential precoder with proof given in the appendices. Simulation results are presented in Section V. Finally, Section VI provides concluding remarks.

Notations in this paper are mostly standard. Upper and bold case letters denote matrices and lower case letter with an arrow on top denote column vectors, such as  $\mathbf{A}$  and  $\vec{a}$ . Tabulated below is a list of the remaining notation symbols.

$\{\cdot\}^H$	Conjugate transpose operation.
$\mathbf{I}_N$	$N \times N$ identity matrix.
$\text{diag}(a_1, \dots, a_N)$	Diagonal matrix with diagonal elements $a_1, \dots, a_N$ .
$\det(\mathbf{A})$	Determinant of a square matrix $\mathbf{A}$ .
$\text{Tr}\{\mathbf{A}\}$	Trace of a square matrix $\mathbf{A}$ .
$E\{\cdot\}$	Expected value.
$\ \cdot\ $	2 norm.

## II. SYSTEM MODEL

Consider a flat fading MIMO wireless communication system with  $M_T$  transmit antennas and  $M_R$  receive antennas. In general, we assume that  $M_T \leq M_R$  to guarantee the symbol recoverability in presence of linear reception and single transmission. Let the source data vector be  $\vec{u}$  which is to be sent to the receiver using the MIMO channel under white noise. As in [5], [8], [9], and [19], we consider a transmitter that utilizes a linear precoding matrix  $\mathbf{F}$ . As shown in the simple diagram of Fig. 1, the received MIMO signal vector is given by

$$\vec{y} = \mathbf{H}\mathbf{F}\vec{u} + \vec{w} \quad (1)$$

in which  $\mathbf{H}$  is the  $M_R \times M_T$  channel matrix, while  $\vec{w}$  is an  $M_R \times 1$  independent white Gaussian vector representing noises from all  $M_R$  receive antennas.

The MMSE precoder design of  $\mathbf{F}$  in a single transmission problem has been investigated in the literature [5], [19]. The resulting MMSE design requires a power loading approach based on the singular value decomposition (SVD) of the channel matrix  $\mathbf{H}$ . In many wireless systems with appropriate feedback link, an ARQ mechanism can be incorporated where packet retransmissions are requested if detections of previous transmissions end in error. In order to improve transmission reliability, advanced receivers can benefit from joint demodulation and detection of the transmitted data vector from its multiple retransmissions. As shown in Fig. 2, a single integrated receiver can be used to extract the data vector  $\vec{u}$  from  $\{\vec{y}_i\}$ . Since diversity can lead to better performance, our approach, as shown in Fig. 2, is to apply different precoders  $\{\mathbf{F}_i\}$  against multiple MIMO channels  $\{\mathbf{H}_i\}$  at different retransmission epochs.

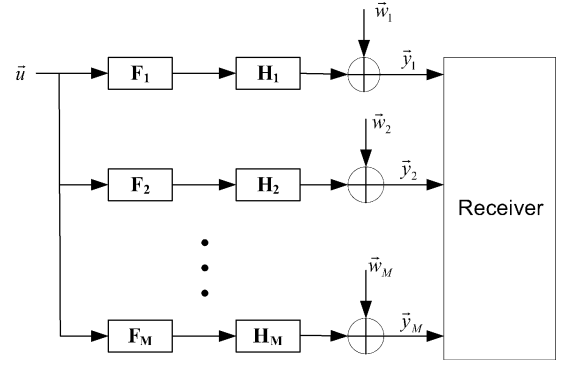


Fig. 2. Block diagram of  $M$  linearly precoded (re) transmissions of a packet through flat fading MIMO channel.

Using a quasi-static model, the MIMO channel matrix is assumed to be fixed during each individual transmission. We also assume that the same data packet will be transmitted during ARQ retransmissions, for instance, the same outer FEC encoder is used for all transmissions which simplifies the decoder design even though it is suboptimum. Hence, the received signal of the  $m$ -th ARQ transmission is modeled by

$$\vec{y}_m = \mathbf{H}_m \mathbf{F}_m \vec{u} + \vec{w}_m \quad (2)$$

where  $\mathbf{H}_m$  is the  $M_R \times M_T$  MIMO channel matrix,  $\mathbf{F}_m$  is the  $M_T \times M_T$  linear precoder matrix, and  $\vec{w}_m, m = 1, \dots, M$  is the independent identically distributed (i.i.d.) white Gaussian noise vector.

At the (integrated) receiver, after  $M$  ARQ transmissions of the same packet  $\vec{u}$ , the overall received signal can be expressed as

$$\vec{y} = \begin{bmatrix} \vec{y}_1 \\ \vec{y}_2 \\ \vdots \\ \vec{y}_M \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{H}_1 \mathbf{F}_1 \\ \mathbf{H}_2 \mathbf{F}_2 \\ \vdots \\ \mathbf{H}_M \mathbf{F}_M \end{bmatrix}}_{\mathbf{H}} \vec{u} + \underbrace{\begin{bmatrix} \vec{w}_1 \\ \vec{w}_2 \\ \vdots \\ \vec{w}_M \end{bmatrix}}_{\vec{w}} \quad (3)$$

$$= \mathbb{H}\vec{u} + \vec{w}.$$

Notice that the number  $M$  is not predetermined as it depends on the successes of the receiver given previous (re) transmissions.

Without loss of generality, we assume that the symbols in the input vector  $\vec{u}$  are i.i.d. with zero mean. Hence, the signal vector covariance matrix is

$$\mathbf{R}_u = E\{\vec{u}\vec{u}^H\} = \sigma_u^2 \mathbf{I}_{M_T}.$$

In addition, the noises  $\{\vec{w}_i\}$  are also i.i.d. with zero mean, leading to covariance matrix

$$\mathbf{R}_w = E\{\vec{w}\vec{w}^H\} = \text{diag}(\mathbf{R}_{w_1}, \dots, \mathbf{R}_{w_M}) = \sigma_w^2 \mathbf{I}_{M \cdot M_R}$$

where

$$\mathbf{R}_{w_i} = E\{\vec{w}_i \vec{w}_i^H\} = \sigma_w^2 \mathbf{I}_{M_R}.$$

For convenience, we define the transmission signal-to-noise ratio (SNR) as  $\gamma = \sigma_u^2 / \sigma_w^2$ .

In order to design the optimal precoder, the optimality criterion has to be defined. In this paper, we consider the precoder

design under a well-established criterion of minimizing the MSE between the transmitted data and the receiver output sequence. This optimization is achieved by jointly designing the optimal linear precoder and decoder. For fairness, the precoder design must be under a resource constraint. We consider the MMSE design subject to the total transmit power constraint  $\text{Tr}\{\mathbf{F}_m \mathbf{F}_m^H\} = M_T$ .

Note that (3) is indeed a general model for linear systems. The case of flat fading MIMO channel is a special example. Relying on orthogonal frequency-division multiplexing (OFDM), our precoder design presented in this paper can be applied to the more general linear frequency-selective channels suffering from intersymbol interferences. To establish the performance benchmark and bound, we focus on the optimal precoder design when the transmitter and receiver have perfect knowledge of the CSI.

### III. LINEAR PRECODING FOR SINGLE TRANSMISSION

For single transmissions, optimal precoder design has been extensively studied in the literature [1], [5], [7], [9], [12], [11]. Similarly, when the number of transmissions is predetermined, the results are the same. In this section, we briefly summarize the design of single transmission linear precoders based on the MMSE criterion [5], [19], before investigating its nontrivial generalization to MIMO ARQ systems.

MMSE precoder design not only specifies the transmitter precoder, but at the same time it also determines the linear receiver that can achieve the minimum MSE for a given precoder. Hence, precoder design with the MMSE criterion requires the joint optimization of linear precoder  $\mathbf{F}_1$  at the transmitter and linear decoder  $\mathbf{G}_1$  at the receiver. The linear decoder generates an output as the estimated signal  $\hat{\mathbf{u}}$ . It is related to the original signal  $\mathbf{u}$  by

$$\hat{\mathbf{u}} = \mathbf{G}_1 \mathbf{H}_1 \mathbf{F}_1 \mathbf{u} + \mathbf{G}_1 \mathbf{w}_1. \quad (4)$$

Our objective is to find the optimal  $\mathbf{F}_1^{\text{opt}}$  and  $\mathbf{G}_1^{\text{opt}}$  that minimize symbol MSE under total transmit power constraint

$$\mathbf{F}_1^{\text{opt}}, \mathbf{G}_1^{\text{opt}} = \arg \min_{\mathbf{F}_1, \mathbf{G}_1} E \left\{ \|\hat{\mathbf{u}} - \mathbf{u}\|^2 \right\}$$

subject to

$$\text{Tr}\{\mathbf{F}_1 \mathbf{F}_1^H\} = M_T. \quad (5)$$

For a fixed precoder  $\mathbf{F}_1$  and channel matrix  $\mathbf{H}_1$ , the optimal linear decoder that minimize the symbol MSE is given by [20]

$$\begin{aligned} \mathbf{G}_1^{\text{opt}} &= \mathbf{R}_u \mathbf{F}_1^H \mathbf{H}_1^H (\mathbf{R}_{w_1} + \mathbf{H}_1 \mathbf{F}_1 \mathbf{R}_u \mathbf{F}_1^H \mathbf{H}_1^H)^{-1} \\ &= \gamma \mathbf{F}_1^H \mathbf{H}_1^H (\mathbf{I}_{M_R} + \gamma \mathbf{H}_1 \mathbf{F}_1 \mathbf{F}_1^H \mathbf{H}_1^H)^{-1}. \end{aligned} \quad (6)$$

Using this optimum  $\mathbf{G}_1^{\text{opt}}$ , the symbol MSE can be written as a function of precoder  $\mathbf{F}_1$  only

$$\begin{aligned} E \left\{ \|\hat{\mathbf{u}} - \mathbf{u}\|^2 \right\} &= \text{Tr} \left\{ \mathbf{R}_u (\mathbf{I}_{M_T} + \mathbf{F}_1^H \mathbf{H}_1^H \mathbf{R}_{w_1}^{-1} \mathbf{H}_1 \mathbf{F}_1 \mathbf{R}_u)^{-1} \right\} \\ &= \text{Tr} \left\{ \sigma_u^2 (\mathbf{I}_{M_T} + \gamma \mathbf{F}_1^H \mathbf{H}_1^H \mathbf{H}_1 \mathbf{F}_1)^{-1} \right\}. \end{aligned} \quad (7)$$

Given (7), the optimal precoder design problem is transformed into a trace minimization problem with total transmit power constraint as follows:

$$\mathbf{F}_1^{\text{opt}} = \arg \min_{\mathbf{F}_1} \text{Tr} \left\{ (\mathbf{I}_{M_T} + \gamma \mathbf{F}_1^H \mathbf{H}_1^H \mathbf{H}_1 \mathbf{F}_1)^{-1} \right\}$$

subject to

$$\text{Tr}\{\mathbf{F}_1 \mathbf{F}_1^H\} = M_T. \quad (8)$$

For convenience, define the singular value decomposition (SVD) of  $\mathbf{H}_1$  as  $\mathbf{H}_1 = \mathbf{U}_1 \mathbf{\Sigma}_1 \mathbf{V}_1^H$ . Without loss of generality, we assume that the diagonal entries of matrix

$$\mathbf{\Sigma}_1 = \text{diag}(\sqrt{\sigma_1}, \dots, \sqrt{\sigma_{M_T}})$$

are arranged in decreasing order, i.e.,  $\sigma_1 \geq \dots \geq \sigma_{M_T}$ . The solution of optimization problem (8) has been proved in [1] and [21] to be

$$\mathbf{F}_1^{\text{opt}} = \mathbf{V}_1 \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_{M_T}}) \quad (9)$$

in which the diagonal entries are determined by

$$\lambda_i = \frac{1}{\gamma \sigma_i} \left( \sqrt{\frac{\gamma \sigma_i}{\alpha}} - 1 \right)_+, \quad i = 1, 2, \dots, M_T \quad (10)$$

where  $(x)_+ = \max(0, x)$  is the positive cropping function and  $\alpha$  must be chosen to satisfy the power constraint  $\sum_{i=1}^{M_T} \lambda_i = M_T$ . This is the MMSE power loading precoder.

Note that a related objective of precoder design is the maximization of mutual information, or channel capacity, obtained by MIMO transmission. See, for example, the works by Telatar [7]. The maximization of mutual information results in a similar precoder structure but a different power loading strategy (through waterpouring).

### IV. PROGRESSIVE OPTIMIZATION OF MMSE PRECODING FOR MIMO ARQ

We are interested in the design and optimization of MMSE linear precoders when ARQ retransmission is incorporated in the MIMO system. Specifically, MMSE linear precoders under ARQ are different from those with predetermined time diversity because future retransmissions may not occur. For a given retransmission, the receiver can only utilize the current retransmission and the previously unsuccessful (re) transmissions of the same data packet. The ARQ transmitter must progressively find the best linear precoder based on the knowledge it has of its previous precoders. In particular, the current precoder is designed under the assumption that no further retransmissions will be required. In other words, the  $m$ th transmission precoder  $\mathbf{F}_m$  is optimum based on the current channel  $\mathbf{H}_m$  and the channels  $\{\mathbf{H}_i, i = 1, 2, \dots, m-1\}$  and the precoders  $\{\mathbf{F}_i, i = 1, 2, \dots, m-1\}$  in previous transmissions.

#### A. General Formulation of MMSE Optimization

In a more formal representation of MMSE linear precoder design, we consider the precoder during the  $m$ th transmission and rewrite the system output (3) as

$$\begin{aligned} \vec{\mathbf{y}} &= \begin{bmatrix} \mathbf{H}_1 \mathbf{F}_1^{\text{opt}} \\ \vdots \\ \mathbf{H}_{m-1} \mathbf{F}_{m-1}^{\text{opt}} \\ \mathbf{H}_m \mathbf{F}_m \end{bmatrix} \mathbf{u} + \begin{bmatrix} \vec{\mathbf{w}}_1 \\ \vdots \\ \vec{\mathbf{w}}_{m-1} \\ \vec{\mathbf{w}}_m \end{bmatrix} \\ &= \mathbb{H} \mathbf{u} + \vec{\mathbf{w}} \end{aligned} \quad (11)$$

where  $\mathbf{F}_i^{\text{opt}}, i = 1, \dots, m-1$  are optimal precoders that have already been optimized using the same MMSE criterion and applied during the previous transmissions.  $\mathbb{H}$  is the effective

channel with dimension of  $mM_R$  by  $M_T$ . Following the progressive process (as in induction), our objective now is to design the current optimal precoder  $\mathbf{F}_m$  under the same MMSE criterion.

We first find the joint MMSE linear decoder  $\mathbf{G}_m^{\text{opt}}$  under given precoders  $\mathbf{F}_i$  and CSI  $\mathbf{H}_i$ . Similar to (6), the  $M_T$  by  $mM_R$  MMSE linear decoder is given by

$$\begin{aligned}\mathbf{G}_m^{\text{opt}} &= \mathbf{R}_u \mathbb{H}^H (\mathbf{R}_w + \mathbb{H} \mathbf{R}_u \mathbb{H}^H)^{-1} \\ &= \gamma \mathbb{H}^H (\mathbf{I}_{m \cdot M_R} + \gamma \mathbb{H} \mathbb{H}^H)^{-1}.\end{aligned}\quad (12)$$

Once the optimum joint receiver  $\mathbf{G}_m^{\text{opt}}$  is found, we rewrite the receiver output symbol MSE as a function of the unknown precoder  $\mathbf{F}_m$  only. We then determine the optimal linear precoder  $\mathbf{F}_m$  that minimizes the receiver output symbol MSE, as shown in (13) at the bottom of the page. The Hermitian matrix  $\sum_{i=1}^{m-1} (\mathbf{F}_i^{\text{opt}})^H \mathbf{H}_i^H \mathbf{H}_i \mathbf{F}_i^{\text{opt}}$  can be simply diagonalized by unitary matrix via

$$\sum_{i=1}^{m-1} (\mathbf{F}_i^{\text{opt}})^H \mathbf{H}_i^H \mathbf{H}_i \mathbf{F}_i^{\text{opt}} = \mathbf{U} \mathbf{\Lambda}_{m-1} \mathbf{U}^H.$$

As a result, the MSE of the receiver output can be simplified into

$$\begin{aligned}E\{\|\hat{u} - \bar{u}\|^2\} &= \sigma_u^2 \text{Tr} \left\{ (\mathbf{I}_{M_T} + \gamma \mathbf{U} \mathbf{\Lambda}_{m-1} \mathbf{U}^H + \gamma \mathbf{F}_m^H \mathbf{H}_m^H \mathbf{H}_m \mathbf{F}_m)^{-1} \right\} \\ &= \sigma_u^2 \text{Tr} \left\{ \mathbf{U}^H (\mathbf{I}_{M_T} + \gamma \mathbf{U} \mathbf{\Lambda}_{m-1} \mathbf{U}^H + \gamma \mathbf{F}_m^H \mathbf{H}_m^H \mathbf{H}_m \mathbf{F}_m)^{-1} \mathbf{U} \right\} \\ &= \sigma_u^2 \text{Tr} \left\{ (\mathbf{I}_{M_T} + \gamma \mathbf{\Lambda}_{m-1} + \gamma \mathbf{U}^H \mathbf{F}_m^H \mathbf{H}_m^H \mathbf{H}_m \mathbf{F}_m \mathbf{U})^{-1} \right\}.\end{aligned}\quad (14)$$

It is, therefore, clear that the  $m$ th progressive MMSE precoder should be

$$\arg \min_{\mathbf{F}_m} \sigma_u^2 \text{Tr} \left\{ (\mathbf{I}_{M_T} + \gamma \mathbf{\Lambda}_{m-1} + \gamma \mathbf{U}^H \mathbf{F}_m^H \mathbf{H}_m^H \mathbf{H}_m \mathbf{F}_m \mathbf{U})^{-1} \right\}$$

subject to

$$\text{Tr}\{\mathbf{F}_m \mathbf{F}_m^H\} \leq M_T. \quad (15)$$

Given the orthogonality of  $\mathbf{U}$ , we can simplify this problem by denoting

$$\tilde{\mathbf{F}}_m = \mathbf{F}_m \mathbf{U}$$

to obtain

$$\arg \min_{\tilde{\mathbf{F}}_m} \sigma_u^2 \text{Tr} \left\{ (\mathbf{I}_{M_T} + \gamma \mathbf{\Lambda}_{m-1} + \gamma \tilde{\mathbf{F}}_m^H \mathbf{H}_m^H \mathbf{H}_m \tilde{\mathbf{F}}_m)^{-1} \right\}$$

subject to

$$\text{Tr}\{\tilde{\mathbf{F}}_m \tilde{\mathbf{F}}_m^H\} \leq M_T. \quad (16)$$

This optimization problem for the  $m$ th transmission turns out to be equivalent to finding the optimum precoder for the second transmission. Hence, without loss of generality, we start with the design of optimal precoder  $\mathbf{F}_2^{\text{opt}}$  for the second transmission.

This will provide a general methodology. For the first transmission, optimal precoder  $\mathbf{F}_1^{\text{opt}}$  as described in the previous section diagonalizes the channel correlation matrix  $\mathbf{H}_1^H \mathbf{H}_1$

$$(\mathbf{F}_1^{\text{opt}})^H \mathbf{H}_1^H \mathbf{H}_1 \mathbf{F}_1^{\text{opt}} = \text{diag}(\lambda_1 \sigma_1, \dots, \lambda_{M_T} \sigma_{M_T}) = \mathbf{\Lambda}_1.$$

To design the precoder for the second transmission, the optimization problem becomes

$$\mathbf{F}_2^{\text{opt}} = \arg \min_{\mathbf{F}_2} \text{Tr} \left\{ (\mathbf{I}_{M_T} + \gamma \mathbf{\Lambda}_1 + \gamma \mathbf{F}_2^H \mathbf{H}_2^H \mathbf{H}_2 \mathbf{F}_2)^{-1} \right\}$$

subject to

$$\text{Tr}\{\mathbf{F}_2 \mathbf{F}_2^H\} \leq M_T. \quad (17)$$

Compared with the first transmission precoder design (8), the first term  $\mathbf{I}_{M_T} + \gamma \mathbf{\Lambda}_1$  now becomes an arbitrary diagonal matrix instead of identity matrix. The approaches taken in designing the MMSE precoder  $\mathbf{F}_1$  for the first transmission no longer apply here. This new optimization problem, in general, requires the optimal pairing of the singular values of the current channel correlation matrix with those of the previous transmission.

Let the SVD of  $\mathbf{H}_2^H \mathbf{H}_2$  be  $\mathbf{V}_2 \mathbf{\Sigma}_2 \mathbf{V}_2^H$ . We define  $\mathbf{P}_1$  and  $\mathbf{P}_2$  as permutation matrices to rearrange the order of the diagonal entries in  $\mathbf{I}_{M_T} + \gamma \mathbf{\Lambda}_1$  and  $\mathbf{\Sigma}_2$  into reverse order. Specifically

$$\begin{aligned}\tilde{\mathbf{\Sigma}}_2 &= \mathbf{P}_1 \mathbf{\Sigma}_2 \mathbf{P}_1^H \\ \tilde{\mathbf{\Lambda}}_1 &= \mathbf{P}_2 (\mathbf{I}_{M_T} + \gamma \mathbf{\Lambda}_1) \mathbf{P}_2^H\end{aligned}$$

in which  $\tilde{\mathbf{\Sigma}}_2 = \text{diag}(\tilde{\Sigma}_1^{(2)}, \dots, \tilde{\Sigma}_{M_T}^{(2)})$  whose diagonal entries are arranged in a nondecreasing order  $\tilde{\Sigma}_1^{(2)} \leq \dots \leq \tilde{\Sigma}_{M_T}^{(2)}$ . Similarly,  $\tilde{\mathbf{\Lambda}}_1 = \text{diag}(\tilde{\Lambda}_1^{(1)}, \dots, \tilde{\Lambda}_{M_T}^{(1)})$  with diagonal entries ordered decreasingly,  $\tilde{\Lambda}_1^{(1)} \geq \dots \geq \tilde{\Lambda}_{M_T}^{(1)}$ .

In order to simplify the trace optimization problem, we introduce the following transformation based on  $\mathbf{P}_1$  and  $\mathbf{P}_2$

$$\mathbf{F}_2 = \mathbf{V}_2 \mathbf{P}_1^H \tilde{\mathbf{F}}_2 \mathbf{P}_2. \quad (18)$$

By substituting (18) into (17), we have a better formulated equivalent optimization problem

$$\min_{\tilde{\mathbf{F}}_2} \text{Tr} \left[ \left( \tilde{\mathbf{\Lambda}}_1 + \gamma \tilde{\mathbf{F}}_2^H \tilde{\mathbf{\Sigma}}_2 \tilde{\mathbf{F}}_2 \right)^{-1} \right]$$

subject to

$$\text{Tr}\{\tilde{\mathbf{F}}_2 \tilde{\mathbf{F}}_2^H\} \leq M_T. \quad (19)$$

## B. Solution to the MMSE Precoder Design

For the MMSE precoder, we rewrite using simpler notations, the general optimization problem

$$\mathbf{F}^{\text{opt}} = \arg \min_{\mathbf{F}} \text{Tr} \left\{ (\mathbf{\Lambda} + \mathbf{F}^H \mathbf{\Sigma} \mathbf{F})^{-1} \right\}$$

subject to

$$\text{Tr}\{\mathbf{F}^H \mathbf{F}\} \leq M_T \quad (20)$$

where matrices  $\mathbf{F}$ ,  $\mathbf{\Lambda}$  and  $\mathbf{\Sigma}$  have the same dimension of  $M_T \times M_T$ .  $\mathbf{\Lambda}$  is the diagonal matrix with positive diagonal entries  $\mathbf{\Lambda} =$

$$\begin{aligned}E\{\|\hat{u} - \bar{u}\|^2\} &= \text{Tr} \left\{ \mathbf{R}_u (\mathbf{I}_{M_T} + \mathbb{H}^H \mathbf{R}_u^{-1} \mathbb{H} \mathbf{R}_u)^{-1} \right\} \\ &= \sigma_u^2 \text{Tr} \left\{ \left( \mathbf{I}_{M_T} + \gamma \sum_{i=1}^{m-1} (\mathbf{F}_i^{\text{opt}})^H \mathbf{H}_i^H \mathbf{H}_i \mathbf{F}_i^{\text{opt}} + \gamma \mathbf{F}_m^H \mathbf{H}_m^H \mathbf{H}_m \mathbf{F}_m \right)^{-1} \right\}\end{aligned}\quad (13)$$

$\text{diag}(\Lambda_1, \dots, \Lambda_{M_T})$ , while  $\Sigma$  is also diagonal with nonnegative entries. Note that, without loss of generality, we can assume that the diagonal entries of  $\Lambda$  and  $\Sigma$  are paired in any appropriate way. In the following solution, we will automatically show that the optimal pairing is the reverse order pairing achieved by the optimal precoder  $\mathbf{F}^{\text{opt}}$ .

The Lagrangian associated with this constrained optimization problem is

$$f(\mathbf{F}, \alpha) = \text{Tr} \left\{ (\Lambda + \mathbf{F}^H \Sigma \mathbf{F})^{-1} \right\} + \alpha \text{Tr} \{ \mathbf{F}^H \mathbf{F} \} \quad (21)$$

where  $\alpha$  is Lagrangian multiplier. By setting derivative of (21) (with respect to  $\mathbf{F}$ ) to zero, we can see that optimal  $\mathbf{F}$  must satisfy

$$\Sigma \mathbf{F} = \alpha \mathbf{F} (\Lambda + \mathbf{F}^H \Sigma \mathbf{F})^2. \quad (22)$$

We assume that  $\Sigma$  has  $K$  distinct eigenvalues (diagonal entries)  $\Sigma_1 \geq \dots \geq \Sigma_K$  with multiplicity  $\nu_1, \dots, \nu_K$ , respectively. Without loss of generality, we can arrange the diagonal entries of  $\Sigma$  in nonincreasing order, hence,  $\Sigma$  is a block diagonal matrix and each block is a scalar of identity matrix

$$\Sigma = \begin{pmatrix} \Sigma_1 \mathbf{I}_{\nu_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \Sigma_2 \mathbf{I}_{\nu_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \Sigma_K \mathbf{I}_{\nu_K} \end{pmatrix}.$$

We now present two useful lemmas for the derivation of MMSE precoders. The proofs of the lemmas are given in the Appendix .

*Lemma 1:* For any matrix  $\mathbf{F}$  that satisfies (22),  $\mathbf{F}$  can be decomposed as

$$\mathbf{F} = \mathbf{U}_F \mathbf{D}_F \mathbf{Q}_F$$

$$\mathbf{U}_F = \begin{pmatrix} \mathbf{U}_{\nu_1}^1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_{\nu_2}^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{U}_{\nu_K}^K \end{pmatrix} \quad (23)$$

where  $\mathbf{U}_F$  is a unitary block diagonal matrix with the same block structure as  $\Sigma$ . In other words,  $\mathbf{U}_F^H \Sigma \mathbf{U}_F = \Sigma$ .  $\mathbf{D}_F$  is a diagonal matrix with nonnegative entries and  $\mathbf{Q}_F$  is a unitary matrix.  $\square$

*Lemma 2:* If (22) holds,  $\Lambda$  and  $\mathbf{F}^H \Sigma \mathbf{F}$  can be jointly diagonalized by a unitary matrix.  $\square$

Using Lemma 1, since  $\mathbf{U}_F$  does not change the cost function  $\Lambda + \mathbf{F}^H \Sigma \mathbf{F}$  and the trace constraint, without loss of generality, we can assume  $\mathbf{U}_F$  to be identity. Hence, the optimization problem (20) can be simplified into

$$(\mathbf{D}_F^{\text{opt}}, \mathbf{Q}_F^{\text{opt}}) = \arg \min_{\mathbf{D}_F, \mathbf{Q}_F} \text{Tr} \left\{ (\Lambda + \mathbf{Q}_F^H \Sigma \mathbf{D}_F^2 \mathbf{Q}_F)^{-1} \right\},$$

such that  $\text{Tr} \{ \mathbf{D}_F^2 \} \leq M_T$ . (24)

Applying Lemma 2, instead of optimizing over an arbitrary unitary matrix  $\mathbf{Q}_F$ , we can further constrain the  $\mathbf{Q}_F$  to be a permutation matrix denoted by  $\mathbf{P}_F$ . This gives us the optimization problem

$$(\mathbf{D}_F^{\text{opt}}, \mathbf{P}_F^{\text{opt}}) = \arg \min_{\mathbf{D}_F, \mathbf{P}_F} \text{Tr} \left\{ (\mathbf{P}_F \Lambda \mathbf{P}_F^H + \Sigma \mathbf{D}_F^2)^{-1} \right\},$$

such that  $\text{Tr} \{ \mathbf{D}_F^2 \} \leq M_T$ . (25)

In the optimization of (25),  $\mathbf{D}_F$  achieves power loading, while  $\mathbf{P}_F$  pairs and rearranges the order of diagonal entries of  $\Sigma$  and  $\Lambda$  so as to minimize the cost. We claim that pairing the diagonal entries of  $\Sigma$  and  $\Lambda$  in the reverse order gives the best power loading performance. Without loss of generality, it is sufficient to show this result for a  $2 \times 2$  case.

Consider  $a_2 \geq a_1 > 0$ ,  $b_2 \geq b_1 \geq 0$ , and two functions

$$f_1(z_1, z_2) = \frac{1}{a_1 + b_2 z_1} + \frac{1}{a_2 + b_1 z_2}$$

$$f_2(x_1, x_2) = \frac{1}{a_2 + b_2 x_1} + \frac{1}{a_1 + b_1 x_2}.$$

Both function  $f_1(z_1, z_2)$  and  $f_2(x_1, x_2)$  are minimized under identical constraint

$$z_1 + z_2 \leq M_T \quad z_1, z_2 \geq 0$$

and

$$x_1 + x_2 \leq M_T \quad x_1, x_2 \geq 0.$$

In order to show that  $f_1(z_1, z_2)$  can achieve smaller minimum, it is sufficient to prove that for any  $x_1, x_2 \geq 0$ , we can find a pair of nonnegative  $z_1, z_2$  such that

$$f_1(z_1, z_2) \leq f_2(x_1, x_2)$$

and

$$z_1 + z_2 \leq x_1 + x_2.$$

We only need to consider two cases.

- $a_1 + b_1 x_2 \leq a_2$ : It is easy to show that  $f_1(x_1 + x_2, 0) \leq f_2(x_1, x_2)$  since

$$\frac{1}{a_1 + b_2(x_1 + x_2)} - \frac{1}{a_2 + b_2 x_1} \leq \frac{1}{a_1 + b_1 x_2} - \frac{1}{a_2}.$$

Hence, in this case,  $z_1 = x_1 + x_2$ , while  $z_2 = 0$ .

- $a_1 + b_1 x_2 > a_2$ : we have  $f_1((a_2 - a_1)/b_2 + x_1, (a_1 - a_2)/b_1 + x_2) = f_2(x_1, x_2)$  and obviously

$$\frac{a_2 - a_1}{b_2} + \frac{a_1 - a_2}{b_1} + x_1 + x_2 \leq x_1 + x_2.$$

Hence,  $z_1 = (a_2 - a_1)/b_2 + x_1$  while  $z_2 = (a_1 - a_2)/b_1 + x_2$ .

By induction, we now generalize this optimal pairing to any size matrix  $\mathbf{D}_F$ . Let there be two arbitrarily ordered sequences with optimal power loading. Find two entries from each sequence that are in the same order. If we keep all other entries and corresponding power loading unchanged. By only reverse-ordering the two entries from one sequence and performing different power loading on these two entries, we have already shown that a lower cost value can be achieved without increasing the total power. Hence, the optimal pairing of the two sequences must be such that individual pairing of any two entries is in reverse order. This implies that the two sequences must be entirely paired in reverse order.

This optimum pairing of entries in  $\mathbf{P}_F \Lambda \mathbf{P}_F^H$  and  $\Sigma \mathbf{D}_F^2$  can be achieved by the permutation matrix  $\mathbf{P}_F$ . Once the optimum pairing is determined, the power constraint can be achieved according to optimal power loading such as the results in [5]. Hence, to summarize, the following theorem describes the progressive MMSE precoder design.

*Theorem 1:* The solution  $\tilde{\mathbf{F}}_2^{\text{opt}}$  of the optimization problem (19) is a diagonal matrix with the diagonal entries determined via

$$\tilde{\mathbf{F}}_2^{\text{opt}} = \text{diag} \left( \sqrt{\mathbf{F}_1^{(2)}}, \dots, \sqrt{\mathbf{F}_{M_T}^{(2)}} \right)$$

$$\mathbf{F}_i^{(2)} = \frac{1}{\gamma \tilde{\Sigma}_i^{(2)}} \left( \sqrt{\frac{\gamma \tilde{\Sigma}_i^{(2)}}{\alpha}} - \tilde{\Lambda}_i^{(1)} \right)_+, \quad i=1, 2, \dots, M_T \quad (26)$$

in which  $\alpha$  is chosen such that  $\text{Tr}(\tilde{\mathbf{F}}_2^{\text{opt}}(\tilde{\mathbf{F}}_2^{\text{opt}})^H) = M_T$ . Hence, the optimal sequential linear precoder for the second transmission is

$$\mathbf{F}_2^{\text{opt}} = \mathbf{V}_2 \mathbf{P}_1^H \tilde{\mathbf{F}}_2^{\text{opt}} \mathbf{P}_2. \quad (27)$$

□

*Remarks:* The optimal MMSE precoder of (26) and (27) consists of three important parts.

- 1) Unitary matrix  $\mathbf{V}_2$  indicating that the optimal transmission subspaces are the orthogonal eigenvectors determined by the channel correlation matrix  $\mathbf{H}_2^H \mathbf{H}_2$ .
- 2) Diagonal matrix  $\tilde{\mathbf{F}}_2^{\text{opt}}$  that implements the constrained power loading.
- 3) Permutation matrix  $\mathbf{P}_1$  and  $\mathbf{P}_2$  to achieve the optimal reverse order pairing of the singular values of  $\mathbf{H}_2$  with the singular values of previous precoded transmissions  $\mathbf{I}_{M_T} + \gamma \mathbf{\Lambda}_1$ , i.e., the smallest singular value of  $\mathbf{H}_2$  is paired with the largest diagonal element of  $\mathbf{I}_{M_T} + \gamma \mathbf{\Lambda}_1$ , the second smallest with the second largest, and so on.

This result can be directly applied to optimal sequential precoder design for multiple ARQ transmissions. By induction, we can assume that at the  $m$ th ARQ transmission, all the previous  $m-1$  precoders have been sequentially determined with each  $\mathbf{F}_i^{\text{opt}}$  diagonalizing the  $\mathbf{H}_i^H \mathbf{H}_i$ ,  $i=1, \dots, m-1$ . The optimal precoder  $\mathbf{F}_m^{\text{opt}}$  employs the similar structure as (26) and (27) with a minor change of indices

$$\mathbf{F}_m^{\text{opt}} = \mathbf{V}_m (\mathbf{P}_1^{(m)})^H \text{diag} \left( \sqrt{\mathbf{F}_1^{(m)}}, \dots, \sqrt{\mathbf{F}_{M_T}^{(m)}} \right) \mathbf{P}_2^{(m)} \quad (28)$$

in which

$$\mathbf{F}_i^{(m)} = \frac{1}{\gamma \tilde{\Sigma}_i^{(m)}} \left( \sqrt{\frac{\gamma \tilde{\Sigma}_i^{(m)}}{\alpha}} - \tilde{\Lambda}_i^{(m-1)} \right)_+, \quad i=1, 2, \dots, M_T$$

while  $\alpha$  is chosen such that  $\text{Tr}\{\mathbf{F}_m^{\text{opt}}(\mathbf{F}_m^{\text{opt}})^H\} = M_T$ . The SVD of  $\mathbf{H}_m^H \mathbf{H}_m$  is  $\mathbf{V}_m \mathbf{\Sigma}_m \mathbf{V}_m^H$ , and  $\mathbf{P}_1^{(m)}$  and  $\mathbf{P}_2^{(m)}$  are permutation matrices such that

$$\tilde{\mathbf{\Sigma}}_m = \mathbf{P}_1^{(m)} \mathbf{\Sigma}_m (\mathbf{P}_1^{(m)})^H = \text{diag}(\tilde{\Sigma}_1^{(m)}, \dots, \tilde{\Sigma}_{M_T}^{(m)})$$

and

$$\tilde{\mathbf{\Lambda}}_{m-1} = \mathbf{P}_2^{(m)} \left( \mathbf{I}_{M_T} + \gamma \sum_{i=1}^{m-1} (\mathbf{F}_i^{\text{opt}})^H \mathbf{H}_i^H \mathbf{H}_i \mathbf{F}_i^{\text{opt}} \right) (\mathbf{P}_2^{(m)})^H$$

$$= \text{diag}(\tilde{\Lambda}_1^{(m-1)}, \dots, \tilde{\Lambda}_{M_T}^{(m-1)})$$

with  $\tilde{\Sigma}_1^{(m)} \leq \dots \leq \tilde{\Sigma}_{M_T}^{(m)}$  and  $\tilde{\Lambda}_1^{(m-1)} \geq \dots \geq \tilde{\Lambda}_{M_T}^{(m-1)}$ .

Note that, similar precoder structure (28) characterized by the reverse order pairing, but with a different powering loading strategy, has also been proved to be optimal [22] for sequential precoder design under maximum mutual information criterion.

TABLE I  
OPTIMAL SEQUENTIAL CHANNEL DIRECTION ASSIGNMENT AND POWER LOADING WITH SNR  $\gamma = 10$  dB

	Chan. Direc. 1 (4.4478)	Chan. Direc. 2 (2.7104)	Chan. Direc. 3 (1.5754)	Chan. Direc. 4 (0.3515)
1st Tran.	Sym. 1 (0.5135)	Sym. 2 (0.6537)	Sym. 3 (0.8475)	Sym. 4 (1.6096)
2nd Tran.	Sym. 4 (1.0002)	Sym. 3 (1.1918)	Sym. 2 (1.2567)	Sym. 1 (0.0000)
3rd Tran.	Sym. 1 (1.2645)	Sym. 2 (1.4458)	Sym. 3 (0.5573)	Sym. 4 (0.0000)
4th Tran.	Sym. 3 (1.4647)	Sym. 4 (1.3619)	Sym. 2 (0.0000)	Sym. 1 (0.0000)

### C. Additional Considerations

In the sequential precoder design under the MMSE criterion, we assume that both input signal and noise are white. For cases where correlations exist between input symbols, i.e.,  $\mathbf{R}_u \neq \sigma_u^2 \mathbf{I}_{M_T}$ , a prewhitening operation can be performed prior to precoding, and the corresponding inverse operation can be applied after decoding. Cases with noise correlations can be similarly dealt with.

Note that we require the noise vectors  $\vec{w}_1, \dots, \vec{w}_M$  to be zero mean and independent. This is a reasonable assumption in practice since it is unlikely that the noise correlations span longer than the delay involved in requesting and transmitting retransmissions.

## V. SIMULATION RESULTS

To illustrate the performance improvement of optimal sequential linear precoding, specifically, the importance of reverse order alignment of singular values, we test over a MIMO system with four transmit and four receive antennas ( $M_R = M_T = 4$ ). The input bit stream is quadrature phase-shift keying (QPSK) modulated. The entries of  $\mathbf{H}$  matrix are zero-mean i.i.d. complex Gaussian random variables with unit variance in each dimension, and are generated independently for every packet. We consider a maximum of four transmissions per packet. To focus on the precoder performance with retransmissions, we assume that every packet will be retransmitted no matter whether it is correctly received or not during the previous transmissions.

### A. Subspace Pairing

We first present the simulation results that demonstrate the effect of reverse order pairing. The optimal precoder (28) transmits each symbol of the data vector  $\vec{u}$  via a different channel subspace indicated by the different column of  $\mathbf{V}_m$  with different power loading. The diagonal entries of  $\mathbf{\Sigma}_m$  quantify the quality of different channel subspaces. Reverse ordering of the singular values in effect tries to compensate for the unreliable symbols that are transmitted via the weaker channel subspaces in the previous transmissions by assigning them to the stronger channel subspaces in current transmission.

Table I shows the channel subspace assignment and power loading at SNR  $\gamma = 10$  dB. We randomly generate one channel matrix and keep it constant for subsequent retransmissions. Values given in parentheses show either the channel subspace quality or the loaded power for different symbols. In the first transmission, the first symbol is transmitted via the strongest channel subspace which gives the most reliable symbol de-

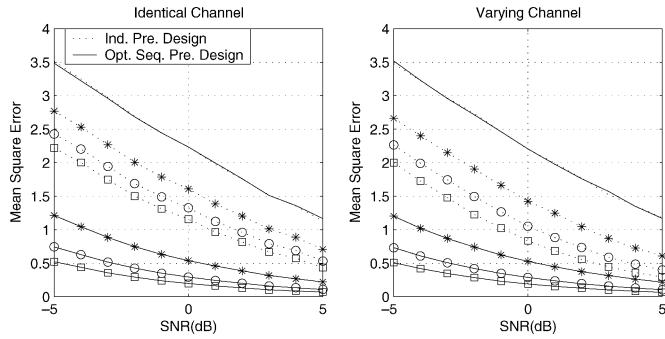


Fig. 3. Symbol MSE achieved using different precoding strategies, with varying and identical channels: 1) —○—: 1st tran.; 2) -\*-: 2nd tran.; 3) -○-: 3rd tran.; and 4) -□-: 4th tran.

tection, while the last symbol is transmitted via the weakest channel subspace and, hence, becomes the least reliable symbol. During the second transmission, in order to compensate for the unequal channel assignment and power loading, the last symbol will be transmitted via the best channel subspace, while the first symbol is transmitted using the worst channel subspace. However, since the first symbol is reliable enough during the first transmission, indeed, it is not transmitted during the second transmission (loaded with zero power). In summary, reverse singular value pairing achieves the optimal channel assignment, which, together with the optimal power loading, enables the sequential minimization of the symbol MSE.

This example represents a very special case in which MIMO channels are stationary. It is not surprising that the most reasonable and fair thing to do is to reverse pair the subspaces. Overall, our MMSE precoder design provides the optimum MMSE precoding design for the more general cases with fading channels.

### B. MSE and BER Comparison With Individually Optimized Precoding

We next present the MSE and BER simulation results to demonstrate the effectiveness of our proposed sequential optimal linear precoding. We compare two types of precoding: *optimal sequential precoding* derived by pairing the singular values in reverse order and *individually optimized* precoding without considering previous transmission. Note that for individually optimized precoder, there exists a design ambiguity because right-multiplying the optimal precoder (9) with any unitary matrix  $\mathbf{Q}$  ( $\mathbf{F}_1^{\text{opt}} \mathbf{Q}$ ) still achieves the optimality. While this unitary matrix  $\mathbf{Q}$  has no effects on a single transmission, it has significant impact during joint design. In our simulation of individually optimized precoding, we let  $\mathbf{Q}$  be the identity matrix. For all test cases, we compare situations when the channels  $\mathbf{H}_1, \dots, \mathbf{H}_M$  remain identical and when the channels vary independently for different retransmissions.

Fig. 3 illustrates the resulting MSE for the two different precoding strategies. For any given SNR, Monte Carlo simulations over 1000 packets are averaged to obtain the results.

The optimal sequential precoder clearly provides substantial improvement in terms of decoder MSE. It is interesting to note that, whether or not retransmission channels vary, the difference in performance improvement provided by optimal sequential precoding and individually optimized precoding remains nearly the same. As shown in the simulation, optimal pairing

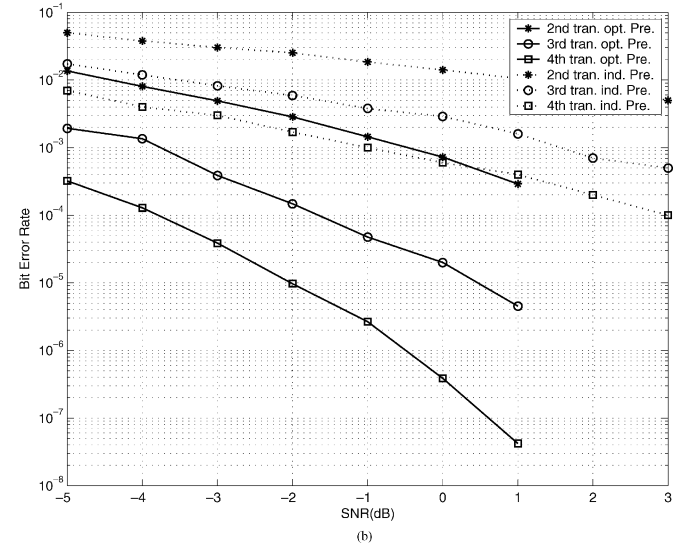
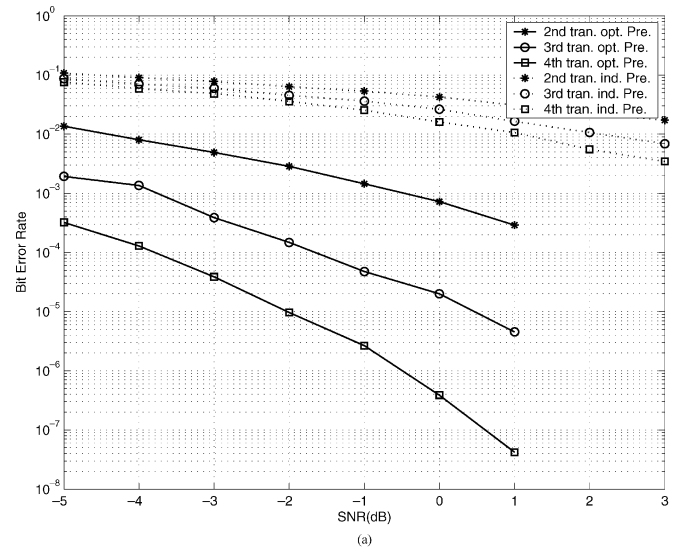


Fig. 4. BER performance improvement of optimal sequential precoding. (a) Compared with individually optimized precoding without permutation. (b) Compared with individually optimized precoding with random permutation.

of the singular values plays a critical role in minimizing the decoder output MSE. When ARQ precoder is independently optimized for each transmission without considering precoders and channels in previous transmissions, there is a substantial performance loss.

Since the MMSE precoder design specifies the joint decoder after each retransmission, we test the BER of the overall wireless communication system. Shown in Fig. 4, BER is an alternative performance metric. The BER results after the first transmission are not shown as they are identical for both precoders.

Evidently, optimal sequential precoding provides substantial BER improvement. The surprisingly poor performance of individually optimized precoding shown in Fig. 4(a) is attributed to the lack of *any* singular value pairing. As stated earlier, by fixing the ambiguity matrix  $\mathbf{Q}$  as identity, we always load the least amount of power on the last symbol (data stream). Hence, the resulting BER will be dominated by the low-power data transmitted on the last stream. In order to alleviate this problem, for individually optimized precoding, we randomly permute the stream order during each transmission. As shown in Fig. 4(b),

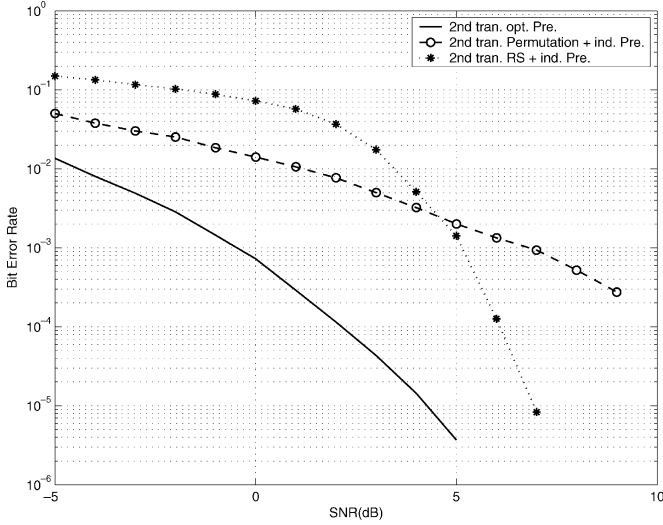


Fig. 5. BER performance improvement of optimal sequential precoding compared with joint type-II hybrid ARQ by using Reed-Solomon code and individually optimized precoding.

even though random permutation reduces the BER, compared with optimal pairing (permutation) implemented by the sequential precoding, it still suffers from significant performance loss.

### C. Comparison With Traditional Type-II Hybrid ARQ

As progressive MIMO ARQ design demonstrates substantial performance advantages, it is well known that practical systems may be in favor of implementing a Type-II hybrid ARQ. Type-II hybrid ARQ does not repeat the data packet in retransmissions. Instead, it can send parity bits of a FEC codeword involving the original data packet for better error correction.

For further comparison, we also consider a Type-II hybrid ARQ [23], [24] scheme by sending incremental redundancy during retransmissions. We use a (63, 31) Reed-Solomon code over  $GF(2^6)$  field to generate the parity bits for retransmission. During the first transmission, information bits will be precoded and transmitted through MIMO channel. If detection error occurs after the first transmission and retransmission is requested, instead of transmitting the same data again, parity bits generated by Reed-Solomon encoder will be precoded and transmitted. At the receiver side, after applying MMSE linear decoder on each transmission individually, hard Reed-Solomon decoding will be performed based on the concatenation of the two received data vectors. For simplicity of illustration, we only consider two maximum ARQ transmissions. Since different transmissions contain different data, the concept of sequential optimal linear precoding cannot be applied here. Instead, we individually optimize precoder for each transmission. Fig. 5 compares the joint detection BER performance for our proposed sequential linear precoder versus the joint Type-II hybrid ARQ by using Reed-Solomon code and individually optimized precoding. For reference, we also include the individually optimized precoding with random permutation. As shown in Fig. 5, for individually optimized precoding, Type-II hybrid ARQ outperforms random permutation at medium to high SNR (when BER is lower than  $10^{-3}$ ), since transmitting parity bits is much more efficient than sending the identical but randomly

permuted data. As shown from the simulation results, our proposed sequential precoding continue to outperform the tested Type-II hybrid ARQ scheme given low to moderate SNR. For instance, at  $BER = 10^{-4}$ , individually optimized precoding with Type-II hybrid ARQ achieves more than 3 dB SNR gain compared with random permutation, but suffers from about 4 dB SNR loss compared with the MMSE sequential optimal precoding. By increasing the number of retransmissions, more powerful (lower rate) channel code can be employed by Type-II hybrid ARQ, it is highly possible that the performance gap will become smaller.

### D. Discussions and Complexity

It should be noted, however, that the performance of Type-II hybrid ARQ can be further improved by using much more powerful channel codes, including longer RS codes, turbo codes, and low-density parity-check (LDPC) codes. Additional performance gain can also be achieved by employing more advance decoding strategies, such as soft decoding or iterative decoding. Nonetheless, compared with our simple linear detection, longer codes and iterative decoding require considerable increase in decoding delay and computational complexity. For static or quasi-static channel, the optimum linear precoder and decoder can be computed offline or updated very slowly. During each packet transmission, only simple matrix multiplication is required. The complexity of computing optimum linear precoder and decoder mainly consists of SVD and inversion of a Hermitian matrix which is roughly  $O(M_R^3)$ . While for Type-II hybrid ARQ, channel decoding has to be performed for every codeword and the complexity will increase at least linearly with the length of codeword which usually will be much larger than  $M_R$ .

## VI. CONCLUDING REMARKS

This paper investigates the progressive design of sequential linear MMSE precoders for MIMO wireless communications. Given the available ARQ diversity, properly precoded MIMO packet retransmissions can significantly improve the overall performance under wireless fading channels. Incorporating a joint receiver that utilizes multiple retransmissions, we design the optimal sequential linear precoders to minimize the receiver output MSE. The result shows that, in addition to the optimal power loading idea used in single transmission, optimal sequential precoding requires the reverse order pairing of the singular values of current MIMO channel with previous transmissions. Simulation results demonstrate significant gains in terms of receiver MSE and BER under optimum MMSE precoding.

## APPENDIX I PROOF OF LEMMA 1

### A. Proof

We multiply both sides of (22) on the right by  $\mathbf{F}^H$ . The resulting equality implies that  $\mathbf{\Sigma}\mathbf{F}\mathbf{F}^H$  is a hermitian matrix. Therefore,  $\mathbf{\Sigma}$  commutes with  $\mathbf{F}\mathbf{F}^H$ , i.e.,  $\mathbf{\Sigma}\mathbf{F}\mathbf{F}^H = \mathbf{F}\mathbf{F}^H\mathbf{\Sigma}$ . Considering the special block diagonal structure of  $\mathbf{\Sigma}$ ,  $\mathbf{F}\mathbf{F}^H$  must be a block diagonal matrix with the same structure which



can be unitarily diagonalized by a block diagonal matrix  $\mathbf{U}_F$ ,  $\mathbf{U}_F^H \mathbf{F} \mathbf{F}^H \mathbf{U}_F = \mathbf{D}_F^2$ . Let  $\tilde{\mathbf{F}} = \mathbf{U}_F^H \mathbf{F}$  and denote its SVD as  $\tilde{\mathbf{F}} = \mathbf{U}_{\tilde{F}} \mathbf{D}_{\tilde{F}} \mathbf{V}_{\tilde{F}}^H$ , we have  $\tilde{\mathbf{F}} \tilde{\mathbf{F}}^H = \mathbf{U}_{\tilde{F}} \mathbf{D}_{\tilde{F}}^2 \mathbf{U}_{\tilde{F}}^H = \mathbf{D}_{\tilde{F}}^2$ . Since  $\mathbf{D}_F$  takes all the nonnegative square roots of diagonal entries of  $\mathbf{D}_{\tilde{F}}^2$ ,  $\mathbf{D}_F$  and  $\mathbf{D}_{\tilde{F}}^2$  must share the same eigenvectors. Therefore,  $\mathbf{U}_{\tilde{F}}$  also diagonalizes  $\mathbf{D}_F$ ,  $\mathbf{U}_{\tilde{F}} \mathbf{D}_F \mathbf{U}_{\tilde{F}}^H = \mathbf{D}_F$ . Finally, we have

$$\begin{aligned} \mathbf{F} &= \mathbf{U}_F \tilde{\mathbf{F}} \\ &= \mathbf{U}_F \mathbf{U}_{\tilde{F}} \mathbf{D}_F \mathbf{V}_{\tilde{F}}^H \\ &= \mathbf{U}_F \underbrace{\mathbf{U}_{\tilde{F}} \mathbf{D}_F \mathbf{U}_{\tilde{F}}^H}_{\mathbf{D}_F} \underbrace{\mathbf{U}_{\tilde{F}} \mathbf{V}_{\tilde{F}}^H}_{\mathbf{Q}_F} \end{aligned}$$

■

## APPENDIX II PROOF OF LEMMA 2

### A. Proof

If we multiply both sides of (22) on the left by  $\mathbf{F}^H \Sigma$ , similarly, we can show that  $\mathbf{F}^H \Sigma \mathbf{F} (\Lambda + \mathbf{F}^H \Sigma \mathbf{F})^2$  is a hermitian matrix. Hence,  $\mathbf{F}^H \Sigma \mathbf{F}$  commutes with  $(\Lambda + \mathbf{F}^H \Sigma \mathbf{F})^2$ . For two commutative hermitian matrices, there exists a unitary matrix  $\mathbf{V}$  that jointly diagonalizes both of them [25], i.e.,  $\mathbf{V} \mathbf{F}^H \Sigma \mathbf{F} \mathbf{V}^H = \mathbf{D}_1$ ,  $\mathbf{V} (\Lambda + \mathbf{F}^H \Sigma \mathbf{F})^2 \mathbf{V}^H = \mathbf{D}_2$ . Since  $(\Lambda + \mathbf{F}^H \Sigma \mathbf{F})$  is a hermitian positive definite matrix, we prove that  $\mathbf{V}$  must also diagonalize  $(\Lambda + \mathbf{F}^H \Sigma \mathbf{F})$ .

Consider a hermitian positive semidefinite matrix  $\mathbf{B}$ , denote its SVD as  $\mathbf{B} = \mathbf{U}_B \mathbf{D}_B \mathbf{U}_B^H$ , all the diagonal entries of  $\mathbf{D}_B$  are nonnegative. For a matrix  $\mathbf{V}$  that diagonalizes  $\mathbf{B}^2$ , we have

$$\mathbf{V} \mathbf{B}^2 \mathbf{V}^H = \mathbf{V} \mathbf{U}_B \mathbf{D}_B^2 \mathbf{U}_B^H \mathbf{V}^H = \mathbf{D}_B^2.$$

Note that  $\mathbf{D}_B$  contains all the nonnegative square roots of diagonal entries of  $\mathbf{D}_B^2$ , we have

$$\mathbf{V} \mathbf{B} \mathbf{V}^H = \mathbf{V} \mathbf{U}_B \mathbf{D}_B \mathbf{U}_B^H \mathbf{V}^H = \mathbf{D}_B.$$

By now, we prove that  $\mathbf{V}$  jointly diagonalizes the  $\mathbf{F}^H \Sigma \mathbf{F}$  and  $\Lambda + \mathbf{F}^H \Sigma \mathbf{F}$ , obviously,  $\mathbf{V}$  also diagonalizes the  $\Lambda$ . ■

## REFERENCES

- [1] H. Sampath and A. J. Paulraj, "Joint transmit and receive optimization for high data rate wireless communication using multiple antennas," in *Proc. 33rd Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, Oct. 1999, pp. 215–219.
- [2] H. Malvar and D. Staelin, "Optimal pre- and postfilters for multichannel signal processing," *IEEE Trans. Signal Process.*, vol. 36, no. 2, pp. 287–289, Feb. 1988.
- [3] J. Yang and S. Roy, "On joint transmitter and receiver optimization for multiple-input-multiple-output (MIMO) transmission systems," *IEEE Trans. Commun.*, vol. 42, pp. 3221–3231, Dec. 1994.
- [4] —, "Joint transmitter-receiver optimization for multi-input multi-output systems with decision feedback," *IEEE Trans. Inf. Theory*, vol. 40, pp. 1334–1347, Sep. 1994.
- [5] A. Scaglione, P. Stoica, S. Barbarossa, G. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1051–1064, May 2002.
- [6] A. Scaglione, S. Barbarossa, and G. B. Giannakis, "Filterbank transceivers optimizing information rate in block transmissions over dispersive channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 3, pp. 1019–1032, Apr. 2002.
- [7] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, pp. 585–596, Nov. 1999.
- [8] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2634–2639, Sep. 2001.
- [9] S. A. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1165–1175, Jul. 2004.
- [10] A. Goldsmith, S. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE Trans. Sel. Areas Commun.*, vol. 21, no. 5, pp. 684–702, Jun. 2003.
- [11] J. H. Manton, "Design and analysis of linear precoders under a mean square error criterion, Part I: Foundations and worst case designs," *Syst. Control Lett.*, vol. 49, no. 2, pp. 121–130, 2003.
- [12] —, "Design and analysis of linear precoders under a mean square error criterion, Part II: MMSE designs and conclusions," *Syst. Control Lett.*, vol. 49, no. 2, pp. 131–140, 2003.
- [13] J. C. Tu, "Theory, design and application of multi-channel modulation for digital communications," Ph.D. dissertation, Stanford Univ., Stanford, CA, 1991.
- [14] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Commun.*, vol. 46, no. 3, pp. 357–366, Mar. 1998.
- [15] T. Koike, H. Murata, and S. Yoshida, "Hybrid ARQ scheme suitable for coded MIMO transmission," in *Proc. IEEE Intl. Conf. Commun.*, vol. 5, Jun. 2004, pp. 2919–2923.
- [16] H. Zheng, A. Lozano, and M. Haleem, "Multiple ARQ processes for MIMO systems," in *Proc. 13th IEEE Intl. Pers., Indoor, Mobile Radio Commun.*, vol. 3, Sep. 2002, pp. 1023–1026.
- [17] Z. Ding and M. Rice, "Type-I hybrid-ARQ using MTCM spatio-temporal vector coding for MIMO systems," in *Proc. IEEE Intl. Conf. Commun.*, vol. 4, May 2003, pp. 2758–2762.
- [18] E. N. Onggosanusi, A. G. Dabak, Y. Hui, and G. Jeong, "Hybrid ARQ transmission and combining for MIMO systems," in *Proc. IEEE Intl. Conf. Commun.*, vol. 5, May 2003, pp. 3205–3209.
- [19] N. Khaled, G. Leus, C. Desset, and H. D. Man, "A robust joint linear precoder and decoder MMSE design for slowly time-varying MIMO channels," in *Proc. IEEE Intl. Conf. Acoustics, Speech, and Signal Process.*, vol. IV, May 2004, pp. 485–488.
- [20] A. Gelb, *Applied Optimal Estimation*. Cambridge, MA: MIT Press, 1974.
- [21] A. Scaglione, S. Barbarossa, and G. B. Giannakis, "Redundant filterbank precoders and equalizers Part I: Unification and optimal designs," *IEEE Trans. Signal Process.*, vol. 47, pp. 1988–2006, Jul. 1999.
- [22] H. Sun, H. Samra, Z. Ding, and J. Manton, "Constrained capacity of linear precoded ARQ in MIMO wireless systems," in *Proc. IEEE Intl. Conf. Acoustics, Speech, and Signal Process.*, Philadelphia, PA, Mar. 2005.
- [23] D. M. Mandelbaum, "An adaptive-feedback coding scheme using incremental redundancy," *IEEE Trans. Inf. Theory*, vol. 20, no. 3, pp. 388–389, May 1974.
- [24] Y.-M. Wang and S. Lin, "A modified selective-repeat type-II hybrid ARQ system and its performance analysis," *IEEE Trans. Commun.*, vol. 31, no. 5, pp. 593–608, May 1983.
- [25] R. A. Horn and C. R. Johnson, *Matrix Analysis*, U.K.: Cambridge Univ. Press, 1985.

**Haitong Sun** (S'03), photograph and biography not available at the time of publication.

**Jonathan H. Manton** (S'95–M'02–SM'03), photograph and biography not available at the time of publication.

**Zhi Ding** (S'88–M'90–SM'95–F'03), photograph and biography not available at the time of publication.