Precoder Design for MIMO Systems with Iterative Equalization

Junjie Ma¹, Xiaojun Yuan², and Li Ping³
1Department of Electronic Engineering, City University of Hong Kong, Hong Kong
2Institute of Network Coding, the Chinese University of Hong Kong, Hong Kong

Abstract—This paper is concerned with precoder design for multiple-input multiple-output (MIMO) systems with iterative equalization. We first consider the case of no channel state information at the transmitter (CSIT). Based on evolution analysis, we derive the optimized precoder that minimizes the bit error rate (BER) of the system. We show that, with the optimized precoder, the linear precoding and iterative equalization scheme can achieve a genie-aided performance upper bound at high signal-to-noise ratio (SNR). We further consider the precoder design with perfect CSIT. We show that the precoder design problem reduces to a convex power-allocation problem that can be efficiently solved using standard convex programming tools. Numerical results are provided to demonstrate the performance advantages of the proposed scheme over its counterparts.

I. INTRODUCTION

Spatial-multiplexing multiple-input multiple-output (MIMO) communication systems [1]-[3] have attracted tremendous interest in the past decades. Such systems offers significant throughput increase without additional bandwidth or increased transmit power. In MIMO systems, the received signal is a superposition of data streams transmitted from different antennas. The receiver needs to mitigate the cross-antenna interference and separate the data streams from each other.

For this purpose, many equalization techniques have been developed in the literature [3]-[7]. Basic linear equalizers include zero-forcing (ZF) and linear minimum mean-square error (LMMSE) equalizers. It is known that linear precoding and decision feedback (DF) techniques can be employed to improve the performance of those linear equalizers. For example, to exploit the channel state information at the transmitter (CSIT), the authors in [5], [6] considered the optimization of linear precoder for ZF-DF and LMMSE-DF schemes. They showed that geometric-mean decomposition (GMD) based linear precoding is optimal for ZF-DF, while uniform channel decomposition (UCD) based linear precoding optimal for LMMSE-DF, both in the sense of minimizing BER. Both schemes utilize successive interference cancellation (SIC) in DF equalization, which may suffer from the error propagation and power loss problems.

To overcome these disadvantages, the authors in [7] proposed for MIMO systems the energy spreading transform (EST) at the transmitter (seen as a special case of linear precoding) and iterative equalization at the receiver. The EST precoder can be further optimized using the so-called root mean-square decomposition (RMSD). It is shown that RMSD-EST can considerably outperform the GMD and UCD based schemes. In our parallel work [8], we proposed a new iterative equalization algorithm based on the message-passing principle. It was shown that the proposed scheme in [8] can further outperform RMSD-EST.

In this paper, we focus on the design of the linear precoder for the iterative equalization scheme proposed in [8]. We first consider the case of no CSIT. Based on evolution analysis, we derive the optimized precoder that minimizes BER. We show that, with the optimized precoder, the linear precoding and iterative equalization scheme can achieve a genie-aided performance upper bound at high SNR. We further study the case of perfect CSIT. We show that, with perfect CSIT, the precoder design problem reduces to a convex power allocation problem which can be efficiently solved using standard convex programming tools [14]. Numerical results are provided to demonstrate the performance advantages of the proposed scheme over existing ones.

![Diagram of the proposed transceiver structure for a generic linear system](image)

Fig. 1. The diagram of the proposed transceiver structure for a generic linear system.

II. ITERATIVE EQUALIZATION AND EVOLUTION ANALYSIS

To begin with, we briefly outline the FFT precoding and iterative equalization scheme proposed in [8].

A. System Description

Consider the following generic linear system

\[ r = Az + \eta \]

where \( r \) is the output vector of the system, \( A \) the transfer matrix, \( z \) the input vector, and \( \eta \) the additive white Gaussian noise (AWGN) with variance \( \sigma^2 \). Without loss of generality, we denote by \( J \) the length of \( z \), where \( J \) is an integer. Throughout this paper, we assume that perfect CSI is available at the receiver, i.e., the receiver knows \( A \) perfectly.

The transmitter structure is illustrated in the upper half of Fig. 1. For each transmission, the information bits are modulated to
form a J-by-1 vector \( x \). We assume that the symbols in \( x \) are taken from the signal constellation \( \mathcal{X} = \{ \alpha_1, \alpha_2, \ldots, \alpha_{Q_1} \} \) with equal probability, where \( |\mathcal{X}| \) denotes the cardinality of the set \( \mathcal{X} \). Without loss of generality, we assume that each entry of \( x \) has zero mean and power normalized to 1. The vector \( x \) is then multiplied by the length-\( J \) normalized discrete Fourier transform (DFT), yielding

\[
y = F_j x.
\]  

(2)

This process can be efficiently implemented using fast Fourier transform (FFT). The signal sequence \( y \) passes through the random interleaver \( \Pi \) to generate the system input vector \( z \).

The iterative receiver consists of an estimator and a soft detector, as illustrated in the lower half of Fig. 1. The estimator estimates \( z \) based on the channel observation \( r \) and the message vector \( \bar{z} \) from the detector, yielding a message vector \( \hat{z} \).

Now we consider the detection operation. As shown in Fig. 1, the input of the demodulator, denoted by \( \bar{y} \), is the inverse FFT \( y^\dagger \), i.e., \( \bar{y} = F_j y^\dagger \). Denote

\[
\rho = \left( \frac{1}{J} \sum_{j=1}^{J} v_j \right)^{-1}
\]  

(5a)

where \( v_j \) is calculated using in (4a). It was shown in [8] that \( \bar{y} \) can be modeled as

\[
\bar{y} = x + n^w
\]  

(5b)

where \( n^w \) is a white Gaussian noise vector with variance \( 1/\rho \). Let \( \bar{y}_i \) be the \( i \)th entry of \( \bar{y} \). Note that \( \rho \) is actually the input signal-to-interference-plus-noise ratio (SINR) of the detector. Given \( \bar{y}_i \), the conditional mean and variance of each \( x_i \) are respectively given by

\[
\bar{x}_i = \sum_{j=1}^{J} C_j \sum_{i=1}^{J} p(\bar{y}_i \mid x_i = \alpha_j) p(\bar{y}_i \mid x_i = \alpha_i) \]

(6a)

and

\[
u_i = \sum_{j=1}^{J} [\bar{y}_i - \bar{x}_i] \sum_{i=1}^{J} p(\bar{y}_i \mid x_i = \alpha_j) \]

(6b)

where \( p(\bar{y}_i \mid x_i = \alpha_j) = \mathcal{N}(\alpha_j, 1/\rho) \). Noting \( y = F_j x \), the a posteriori mean and covariance of \( y \) are thus given by

\[
\hat{y} = F_j \bar{y} \text{ and } U = F_j \text{diag}(u_1, u_2, \ldots, u_J) F_j^\dagger \]

(7)

Similarly to (4), we calculate the extrinsic mean and variance of \( y \) as

\[
v^{-1} = U^{-1} - \rho \]

(8a)

and

\[
\bar{\Sigma}_w \left( \frac{1}{v} \right) = \frac{\hat{y} - \bar{\Sigma}_w}{v^n} \rho \]

(8b)

where

\[
U = \frac{1}{J} \sum_{j=1}^{J} u_j.
\]  

(9)

Note that \( U \) actually gives the value of the diagonal elements of \( U \) (that are identical to each other).

Based on the above discussion, the overall iterative equalization algorithm is summarized as follows.

Initialization: \( \bar{y} = 0 \) and \( v = 1 \).

Step 1: Calculate (3)-(5) to yield \( \bar{y} \) and \( \rho \).

Step 2: Calculate (6)-(9) to yield \( \bar{y} \) and \( v \).

Repeat Steps 1 and 2 until convergence.

D. SINR-Variance Transfer Chart

We next describe an evolution technique to analyze the performance of the above iterative equalization algorithm [8], [11]. For the estimator, we use the input variance \( v \) to characterize the reliability of the input messages \( \bar{y} \), and the output SINR \( \rho \) to characterize the reliability of the estimator’s outputs. From (3b), (4a), and (5a), it can be readily shown that the transfer function of the estimator is given by

\[
\rho = \frac{1}{J} \sum_{j=1}^{J} a_j \bar{v} - v^n \to \phi(v)
\]  

(10)

where \( a_j \) denotes the \( j \)th column of \( A \).

Now consider the detector. For a sufficiently large \( J \), the average a posteriori variance of \( x \) given \( \bar{y} \) (or equivalently, \( \bar{x} \) ) can be calculated by

\[
U = \frac{1}{J} \sum_{j=1}^{J} u_j = \gamma(\rho).
\]

(11)

where \( \gamma(\rho) \triangleq \mathbb{E}[\|x - \text{E}[x]\|^2] \). Here, \( x \) is uniformly taken over \( \mathcal{X} \), and \( n \sim \mathcal{C}(0, 1/\rho) \) is independent of \( x \). Then, the output variance of the detector in (8a) is given by

\[
v^{-1} - \rho^{-1} \to \psi(\rho).
\]

(11)
The behavior of the iterative equalizer can be characterized by \( \rho = \psi(v) \) and \( v = \varphi(\rho) \), i.e., the iterative process of the estimator and the detector can be tracked by a recursion of \( \rho \) and \( v \). Specifically, let \( q \) be the iteration number. Then, we have \( \rho^q = \psi(v^{q-1}) \) and \( v^q = \varphi(\rho^q) \), \( q = 1, 2, \ldots \)

The recursion continues and converges to a point \( v' \) satisfying \( \varphi(v') = \psi^{-1}(v') \) and \( \varphi(v) > \psi^{-1}(v') \), for \( v \in (v', 1] \).

where \( \psi^{-1}(\cdot) \) is the inverse of \( \psi(\cdot) \), which exists since \( \psi(\cdot) \) is continuous and monotonic. The system performance (e.g., BER or frame error rate) is uniquely determined by the convergence point \( v' \). This implies that we can predict the system performance using the above evolution technique.

## III. PRECODER DESIGN WITHOUT CSIT

In the preceding section, we described an FFT-precoding and iterative equalization scheme for the generic linear system in (1). In this section, we apply this scheme to MIMO channels. We discuss the design of linear precoding to improve the system performance under the assumption of no CSIT.

### A. Linearly Precoded MIMO Channel

Consider a quasi-static flat-fading MIMO channel with \( N \) transmit antennas and \( M \) receive antennas. Assume that the channel remains unchanged in a transmission frame consisting of \( L \) channel uses, but varies independently from frame to frame.

The received signal at the \( i \)th channel use can be modeled as

\[
r_i = H z_i + \eta_i, \quad \text{for } i = 1, 2, \ldots, L
\]

where \( r_i \) is the received signal vector at the \( i \)th channel use, \( H \) an \( M \times N \) channel matrix, \( z_i \) the transmitted signal, and \( \eta_i \) the additive noise. The entries of \( H \) are distributed as \( \mathcal{CN}(0, 1/N) \).

The received signal for the overall frame can be written as

\[
r = (I_L \otimes H) z + \eta
\]

where \( r = [r_1, r_2, \ldots, r_L] \), \( z' = [z_1, z_2, \ldots, z_L] \), \( \eta = [\eta_1, \eta_2, \ldots, \eta_L] \), \( I_L \) is the \( L \times L \) identity matrix, and \( \otimes \) represents the Kronecker product. The system input \( z \) in (1) is related to \( z' \) by a linear precoder \( P \) as \( z = P z' \). Then

\[
r = (I_L \otimes H) P z' + \eta.
\]

It is readily seen that (14) is a special case of the generic linear system in (1) by letting \( A = (I_L \otimes H) P \) (which implies \( J = NL \)). Therefore, the FFT-precoding and iterative equalization algorithm described in Section II can be directly applied to the system in (14). In what follows, we focus on the optimization of the precoder \( P \) to improve the system performance. We assume no CSIT in this section. The case of perfect CSIT will be addressed in Section IV.

### B. Upper Bound of the Estimator’s Output SINR

Without loss of generality, let the singular value decomposition (SVD) of \( P \) be \( P = U_P V_P^H \), where \( U_P \) and \( V_P \) are unitary matrices and \( D_P \) is a diagonal matrix for power allocation. Without CSIT, there is usually no reason to give priority to any directions in power allocation, and hence \( D_P \) is chosen to be a scaled identity matrix. This implies that it suffices to consider a unitary \( P \) in precoder optimization.

We have the following upper bound on the output SINR of the estimator.

**Proposition 1**: For any unitary matrix \( P, \phi(v) \) with \( A = (I_L \otimes H) P \) is upper bounded by

\[
\phi(v) \leq \left( \frac{1}{N} \text{tr} \left( H^H \left( v H H^H + \sigma^2 I \right)^{-1} H \right) \right)^{-1} - v^{-1}
\]

**Proof**: Let \( \Omega \equiv A^H (v A A^H + \sigma^2 I)^{-1} A \). Then

\[
\phi(v) = \left( \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\Omega(i,i)} - v \right) \leq \left( \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\Omega(i,i)} - v \right) = \left( v - \frac{1}{N} \text{tr} \left( H^H (v H H^H + \sigma^2 I)^{-1} H \right) \right)^{-1} - v^{-1}
\]

where \( \Omega(i,i) \) represents the \( i \)th diagonal element of \( \Omega \) step (a) follows from the arithmetic-harmonic-mean inequality, and step (b) can be verified using simple algebra.

Letting \( v \to 0 \) (implying that perfect a priori information is available at the estimator), the right hand side of (15) becomes

\[
\frac{1}{N \sigma^2} \text{tr} \left( H H^H \right).
\]

We refer to an AWGN channel with SNR given by (17) as the effective AWGN channel of the MIMO channel in (12). This AWGN bound will be used as a benchmark in evaluating the system performance.

### C. Precoder Design

We next study the design of the precoder \( P \) to achieve the upper bound in (15). Recall that the equality in (17) holds if and only if the matrix \( \Omega \) has identical diagonal elements given by

\[
\Omega(i,i) = \frac{1}{N} \text{tr} \left( H^H (v H H^H + \sigma^2 I)^{-1} H \right), \text{ for } i = 1, \ldots, J.
\]

Thus, our objective is to find such a unitary \( P \) that is independent of the channel \( H \) (since \( H \) is not known at the transmitter) while the equalities in (18) hold.

To meet the above requirements, our proposed precoder has the following structure:

\[
P = \Sigma (I_N \otimes F_L)
\]

where \( \Sigma \) is an \( J \)-by-\( J \) permutation matrix, \( I_N \) denotes the \( N \)-by-\( N \) identity matrix, and \( F_L \) is the \( L \)-by-\( L \) normalized DFT matrix. An illustration of the precoder structure can be found in Fig. 2. What remains is to determine \( \Sigma \).

Let \( s = (I_N \otimes F_L) z \) and \( \Sigma = \Sigma(s) \) (c.f., Fig. 2). We divide \( s \) into \( N \) blocks as \( s = [s_1^T, s_2^T, \ldots, s_N^T] \), and \( t \) into \( L \) blocks as \( t = [t_1^T, t_2^T, \ldots, t_L^T] \). To meet the conditions in (18), we impose the following criteria on \( \Sigma \):

- For each index \( i \), no two entries of \( s_i \) are transmitted at a same channel use;
- No two entries of \( s_i \) are transmitted at a same antenna.

The permutation matrix that meets the above requirements is not unique and can be easily constructed. Fig. 2 shows one particular example for \( N = L = 4 \).
Proposition 2: With the proposed precoder \( \mathbf{P} \) in (19) and \( \Sigma \) satisfying the two criteria above, the equality in (15) holds.

To prove Proposition 2, we only need to verify that, with \( \mathbf{P} \) in (19) and \( \Sigma \) satisfying the two criteria, the conditions in (18) hold. Due to space limitation, we omit the details here.

Note that the proposed precoder maximizes the SNR of the AWGN channel in (5b). This implies that the precoder is optimal in minimizing BER for the case of no CSIT.

The complexity of the proposed scheme is dominant by the inverse of \( \mathbf{R} \) (c.f., (3)) and the length-\( J \) FFT operation. Note that \( \mathbf{R} \) is a \( J \)-by-\( J \) block-diagonal matrix with \( L \) blocks and each block of size \( N \)-by-\( N \). Thus, the complexity involved in \( \mathbf{R}^{-1} \) is \( O(\text{LN}^3) \). The complexity of length-\( J \) FFT is \( O(\text{LNlog}(LN)) \) by noting \( J = LN \). Therefore, the overall computational complexity is \( O(\text{LN}^3) + O(\text{JlogJ}) \) per iteration.

\[ \text{Fig. 2. An example of the precoder structure with } N = L = 4. \text{ The four blocks of } \mathbf{H} \text{ represents } L = 4 \text{ channel uses.} \]

D. Numerical Results

Numerical results are presented in Fig. 3 to demonstrate the performance of the proposed precoder for Rayleigh fading channels. In simulation, \( M = N = 8 \) and \( L = 2048 \). We assume that Gray-mapped QPSK modulation is employed. The SNR is defined as \( \text{SNR} \triangleq 1/\sigma^2 \).

From Fig. 3, we see that the proposed scheme without precoder optimization outperforms the other schemes, including ZF-SIC [3], MMSE-SIC [4] and RMSD-EST [7]. An extra gain of about 0.7 dB is achieved at \( \text{BER} = 10^{-5} \) by applying the proposed precoder in (19). We also see that the scheme with the proposed precoder can approach the effective AWGN bound (alternatively referred to as the genie-aided upper bound) at high SNR. Moreover, simulation results match well with the performance predicted by evolution analysis.

\[ \text{Fig. 3. The BER performances of various schemes without utilizing CSIT for quasi-static Rayleigh fading channel. } M = N = 8 \text{ and } L = 2048. \]

IV. PRECODER DESIGN WITH PERFECT CSIT

A. Precoder Design with Perfect CSIT

To exploit the potential benefit of the available CSIT, we propose to introduce an extra precoding matrix \( \mathbf{Q} \) between the original precoder in (19) and the physical channel, i.e., the proposed precoder now becomes

\[ \mathbf{P} = \left( \mathbf{I}_L \otimes \mathbf{Q} \right) \Sigma \left( \mathbf{I}_N \otimes \mathbf{F}_L \right). \]

The precoder structure is illustrated in Fig. 4. The matrix \( \mathbf{Q} \) adapts to the varying channel to improve system performance. Substituting \( \mathbf{A} = \left( \mathbf{I}_N \otimes \mathbf{Q} \right) \mathbf{P} \) with \( \mathbf{P} \) given above, the transfer function of the estimator in (10) becomes

\[ \phi(v) = \left( v^2 - \frac{v^2}{N} \text{tr}\left( \mathbf{Q}^H \mathbf{H}^2 \left( \mathbf{H}^H \mathbf{Q} \mathbf{Q}^H + \sigma^2 \mathbf{I} \right)^{-1} \mathbf{H} \mathbf{Q} \right) \right)^{-1} v^{-1}. \]  

We can therefore tune the system performance by adjusting the precoding matrix \( \mathbf{Q} \).

- We want to minimize the transmission power \( \text{tr}(\mathbf{QQ}^H) \) (per channel use).
- At the same time, we need to ensure that the iterative equalizer converges to a designated output variance \( \nu_0 \), where \( \nu_0 \) is determined by the target BER of the system.

This optimization problem can be formulated as

\[ \min_{\mathbf{Q}} \text{tr}(\mathbf{QQ}^H) \tag{21a} \]

s.t. \( \phi(v) > \nu^{-1}(v) \), for \( 1 \geq v > v_0 \). \tag{21b} \]

We next discuss how to solve the above problem. Without loss of generality, let the SVD of \( \mathbf{H} \) and \( \mathbf{Q} \) be

\[ \mathbf{H} = \mathbf{U}_H \mathbf{D}_H \mathbf{V}_H^H \]  

and \( \mathbf{Q} = \mathbf{U}_Q \mathbf{D}_Q \mathbf{V}_Q^H \) \tag{22} \]

where \( \mathbf{U}_H, \mathbf{V}_H, \mathbf{U}_Q \) and \( \mathbf{V}_Q \) are unitary matrices, and \( \mathbf{D}_H \) and \( \mathbf{D}_Q \) are diagonal matrices with non-negative diagonal elements arranged in the descending order.

Proposition 3: For the optimization problem in (21), the optimal \( \mathbf{U}_Q \) is \( \mathbf{U}_Q = \mathbf{V}_H \); the optimal \( \mathbf{V}_Q \) is arbitrary.
Proof: Substituting (22) into (20), the transfer function of the estimator can be shown to be
\[
\phi(v) = \left[ v^2 + \frac{1}{N} \left( \sigma^2 + 1 \right) \right]^{-1} - v^{-1} = \left( \frac{1}{N} \left( \sigma^2 + 1 \right) \right)^{-1} - v^{-1} \quad (23)
\]
where \( \hat{U} \equiv V^H U \). Then, it is easy to see that both the objective and the constraint of (21) are independent of \( V \). Therefore, the optimal \( V \) is arbitrary. What remains is to show that the optimal \( \hat{U} \) is \( \hat{U} = I \).

To this end, we first observe that the objective function \( \text{tr}(QQ^H) \) is independent of \( \hat{U} \). Thus, it suffices to show that \( \hat{U} = I \) maximizes the value of \( \phi(v) \) in (23).

Define \( E \equiv \frac{1}{\sigma} D_{\hat{U}} D_{\hat{U}}^H D_{\hat{U}}^H + \frac{1}{v} I \). The objective is equivalent to minimizing \( \text{tr}(E) \), which is a Schur-concave function of the diagonal elements of \( E \). It is proved in [12] (c.f., Theorem 1 in [12]) that the optimal \( E \) is diagonal, or equivalently \( \hat{U} = I \). This completes the proof.

With Proposition 3, the problem in (21) can be simplified to a power-allocation problem as
\[
\min_{d_{\hat{U}}} \sum_{i,j} d_{\hat{U}_{ij}} \quad \text{s.t. } \bar{\phi}(v) > \psi^{-1}(v), \quad 1 \geq v > v_0, \quad (24b)
\]
where \( \bar{\phi}(v) = \left( \frac{1}{N} \sum_{i,j} \frac{d_{\hat{U}_{ij}} d_{\hat{U}_{ij}}}{\sigma^2 + v^{-1}} \right)^{-1} - v^{-1} \) and \( d_{\hat{U}_{ij}} \) (resp. \( d_{\hat{U}_{ij}} \)) is the \( ij \)-th entry of \( D_{\hat{U}}^H D_{\hat{U}} \) (resp. \( D_{\hat{U}}^H D_{\hat{U}}^H \)). It can be shown that the problem in (24) is convex in \( \{ d_{\hat{U}_{ij}} \} \) [13]. Letting the constraint in (24b) hold on the discretized values of \( v \) in the range of \( (v_0, 1] \), we can solve (24) using standard convex optimization tools [14]. We omit the details due to space limitation.

B. Numerical Results

We provide numerical results to illustrate the performance of the proposed scheme in Fig. 5. System settings are the same as those in Fig. 3. The target BER is set to be 10^{-3} for the optimized precoder. It is interesting to see that our proposed scheme without exploiting CSIT (i.e., using the precoder in (19)) performs better than all the benchmark schemes with CSIT, including GMD in [5], UCD in [6], and RMSD-EST in [7]. The precoder optimization based on CSIT (by solving (24)) can provide a significant power gain of about 2 dB at BER = 10^{-3} or 3 dB at BER = 10^{-5}. We also see that the BER curve of the optimized precoder approaches that of the effective AWGN channel at high SNR, as expected.

V. CONCLUSIONS

In this paper, we studied precoder optimization for MIMO systems with iterative equalization. For the case of no CSIT, we showed that the proposed precoder can approach the effective AWGN bound at high SNR. For the case of perfect CSIT, we showed that the precoder design reduces to a convex power-allocation problem that can be efficiently solved. We showed that the optimized precoder can achieve a significant power gain compared with the one without using CSIT.

REFERENCES