EURASIP Journal on Wireless Communications and Networking 2004:2, 344–356

Design of FIR Precoders and Equalizers for Broadband MIMO Wireless Channels with Power Constraints

Yongfang Guo

Department of Electrical and Computer Engineering, University of California, Davis, CA 95616-5294, USA
Email: yguo@ece.ucdavis.edu

Bernard C. Levy

Department of Electrical and Computer Engineering, University of California, Davis, CA 95616-5294, USA
Email: levy@ece.ucdavis.edu

Received 20 June 2003; Revised 13 July 2004; Recommended for Publication by Tho Le-Ngoc

This paper examines the optimum design of FIR precoders or equalizers for multiple-input multiple-output (MIMO) frequency-selective wireless channels. For the case of a left-coprime FIR channel, which arises generically when the number n_T of transmit antennas is larger than the number n_R of receive antennas, the Bezout matrix identity can be employed to design an FIR MIMO precoder that equalizes exactly the channel at the transmitter. Similarly, for a right-coprime FIR channel, the Bezout identity yields an FIR zero-forcing MIMO equalizer. Unfortunately, Bezout precoders usually increase the transmit power, and Bezout equalizers tend to amplify the noise power. To overcome this problem, we describe in this paper a convex optimization technique for the optimal synthesis of MIMO FIR precoders subject to transmit power constraints, and of MIMO FIR equalizers with output noise power constraints. The synthesis problem reduces to the minimization of a quadratic objective function under convex quadratic inequality constraints, so it can be solved by employing Lagrangian duality. Instead of solving the primal problem, we solve the lower-dimensional dual problem for the Lagrange multipliers. When an FIR MIMO precoder has already been selected, we also describe a technique for adding a vector shaping sequence to the transmitted signal in order to reduce the transmit power. The selection of effective shaping sequences requires a search over a trellis of large dimensionality, which can be accomplished suboptimally by employing reduced-complexity search techniques.

Keywords and phrases: precoder, equalizer, Bezout identity, quadratic optimization, Lagrangian duality, shaping sequence.

1. INTRODUCTION

The increasing demand for high data rates communication and the lack of wireless spectrum have prompted the consideration in recent years of multiantenna wireless communication systems that can support much higher data rates [1, 2, 3] than traditional single-input single-output wireless channels. However, for the case of frequency-selective multiple-input multiple-output (MIMO) wireless channels, the task of channel equalization becomes challenging, since the vector nature of the channel precludes the use of trellis-based equalization techniques due to the excessively large number of states required, and since MIMO decision-feedback equalizers (DFEs) [4, 5] have a high computational complexity. As a consequence, most broadband MIMO wireless system architectures address the equalization issue by purposefully trading off some amount of system performance against a lower implementation complexity. Usually, this is accomplished either by using an orthogonal frequency division multiplexing (OFDM) modulation format [6, 7] or, for single-carrier systems, by performing equalization in the frequency domain [8, 9]. In both cases, the transmitted data is divided in blocks and a cyclic prefix is appended to each block, which has the effect of ensuring that a block is not affected by intersymbol interference (ISI) from the previous block. This simplifies greatly the equalization problem, but at the price of reducing the overall transmission rate by the overhead represented by the cyclic prefix insertion. Similarly, the class of Bezout precoders or equalizers introduced in [10] for MIMO frequency-selective channels can be viewed as simplifying the equalization problem through the introduction of redundancy in space. These precoders exploit the fact that, whereas square filter impulse response (FIR) systems (systems with the same number of inputs and outputs) have an IIR inverse, rectangular FIR systems (systems with more inputs than outputs or vice-versa) admit FIR inverses which are obtained by using the Bezout identity [11] of right- or left-coprime FIR systems. Thus the intentional unbalancing of MIMO wireless systems through the introduction of additional transmit or receive antennas has the effect of simplifying the equalization problem. In contrast to Bezout precoders or equalizers which employ a single rectangular FIR filter, MIMO DFEs [4] or Tomlinson-Harashima (TH) precoders [12] require both feedforward and feedback FIR filters, as well as the solution of a matrix spectral-factorization problem, which typically requires the solution of a matrix Riccati equation. In addition to making equalization easier, the additional antennas introduced by making the system rectangular improve the performance of the system by increasing space diversity. Finally, since the order of Bezout precoders/equalizers tends to be about the same as the length of the MIMO FIR channel to be equalized, their computational complexity is smaller than that of OFDM systems or frequency-domain equalizers which require FFTs and inverse FFTs of about ten times the channel length [8].

Unfortunately, for MIMO systems with deep frequency fades, Bezout precoders have the tendency to increase the transmitted signal power, and similarly Bezout equalizers increase the noise power at the equalizer output. To reduce this problem, it was proposed in [10] to use the available degrees of freedom existing in Bezout precoders (resp., equalizers) to reduce the transmit signal power (resp., the equalizer output noise power). However, even with improvements of this type, by performing simulations that sample the space of MIMO channels at random, it is not difficult to verify that the occurrence of bad channels impacts significantly the average performance of Bezout precoders or equalizers. To overcome this problem, we explore here the tradeoff existing between the power of the residual ISI for the equalized channel and transmit power constraints or output noise power constraints. For MIMO systems with more transmit than receive antennas, a convex optimization technique is proposed for the synthesis of FIR MIMO precoders minimizing the power of the residual channel ISI subject to various power constraints. The constraints we examine involve either the power used by each user, the total power for the system of transmit antennas, or the power used by individual antennas. In all cases, the synthesis problem reduces to the minimization of a quadratic objective function under convex quadratic inequality constraints, so it can be solved by Lagrangian duality [13]. Specifically, since the number of power constraints is much smaller than the MIMO precoder coefficients, instead of solving the primal problem, we solve the lower-dimensional dual problem for the Lagrange multipliers. For systems with more receive than transmit antennas, we also discuss briefly the optimum synthesis of FIR MIMO equalizers minimizing the residual ISI power under various equalizer-output noise power constraints.

In this respect, it is worth noting that a wide range of results have already been derived concerning the separate or joint optimization of precoders and/or equalizers for frequency-selective channels. Unfortunately, with the exception of [14], [15, Section IV], most of the existing designs do not employ constant FIR filters. They either use IIR precoders and equalizers [16, 17, 18, 19] or divide the data in blocks and perform the design of precoders or equalizers on a blockwise basis [20, 21]. However, the block approach requires the use of guard symbols to prevent interblock ISI, and even when the channel is constant over a given block, the precoders and equalizers do not have a Toeplitz structure, so they cannot be implemented by time-invariant filters, which makes them unattractive. Also, whereas earlier precoder or equalizer design techniques consider only a total transmit power constraint, the FIR design methodology we propose can handle a wider range of such constraints, which includes for example constraints on each transmit antenna power, or on the power allocated to each user. Note that the results presented are equally applicable to single-user and multiuser MIMO systems, but in the multiuser case, cooperation is required between the different users to estimate the channel and then compute and implement the optimum MIMO precoder/equalizer.

For the precoder case, the design technique we propose can be viewed as an attempt to attain for Bezout precoders some of the features of TH precoders [12, 22] which solve the transmit power amplification problem through the introduction of a modulo operation in the transmitter feedback path at the expense of a small increase in transmit power (about 1.25 dB for QPSK transmission) and in BER due to the introduction of additional neighbors in the periodic extension of the signal set. After a MIMO precoder has been selected, we also describe a signal shaping technique similar to the one used for transmit power reduction in TH precoding. This method uses constellation expansion and consists of adding a vector shaping sequence to the transmitted signal in order to reduce the transmit power. The selection of effective shaping sequences requires a search over a trellis of large dimensionality, which can be accomplished suboptimally by employing reduced-complexity search techniques.

The paper is organized as follows. The MIMO wireless channel model and transmit/receive antenna system are described in Section 2. Section 3 considers the design of precoders for MIMO channels with more transmit than receive antennas. The minimization of the ISI under transmit power constraints yields a convex optimization problem which is solved by using Lagrangian duality. In Section 4, a method is described for generating shaping sequences that decrease the transmit power. The case of multiantenna systems with more receive than transmit antennas is examined in Section 5 where a design technique is presented for minimizing the residual ISI while ensuring that noise power constraints for the output signal are satisfied. Simulation results are presented in Section 6 and Section 7 draws conclusions and identifies further research issues.

Notation

In this paper, all boldface letters indicate vectors (lower case) or matrices (upper case). The superscripts T and H denote matrix transposition and the Hermitian transpose, respectively, and $\mathrm{tr}(\mathbf{A})$ and $|\mathbf{A}|$ represent, respectively, the trace and determinant of a matrix $\mathbf{A}$. Continuous-time vector signals are written as $\mathbf{a}(t)$, and discrete-time vector sequences as $\mathbf{a}(n)$. The vec-operator is defined as $\mathrm{vec}(x_i,\ 1 \le i \le q) = [x_1^H\ x_2^H\ \cdots\ x_q^H]^H$. The matrix derivative is defined as $(\partial f(\mathbf{A})/\partial \mathbf{A})_{ij} = \partial f(\mathbf{A})/\partial a_{ij}$.

2. SYSTEM MODEL

The system we consider has n_T transmit and n_R receive antennas. Let $\{\mathbf{x}(n) \in \mathcal{S}^{n_U},\ n \in \mathbb{Z}\}$ denote the vector sequence of independent complex M-ary symbols to be transmitted, where $\mathcal{S}$ represents the symbol set and $n_U = \min(n_T, n_R)$. The components of $\mathbf{x}(n)$ can be from the different users or from the same user after serial-to-parallel conversion. This vector sequence is assumed to have zero-mean and covariance matrix $\sigma_x^2 \mathbf{I}_{n_U}$.

For frequency-selective channels, let $h_{qp}(k, n)$ denote the discrete-time channel impulse response (CIR) (including the transmit filter and receive filter) from the pth transmit antenna to the qth receive antenna at time instant nT due to an impulse applied at $nT - kT$, where $1/T$ is the symbol signalling rate. We assume it admits an L-path complex model of the form

$$h_{qp}(k, n) = \sum_{l=1}^{L} a_{lqp}(n) c\big(kT - \tau_{lqp}(nT)\big), \tag{1}$$

where $a_{lqp}(n)$ denotes the complex reflection coefficient specifying the amplitude and phase of the lth path from transmit antenna p to receive antenna q, $\tau_{lqp}(nT)$ represents the corresponding time delay, and $c(t) = r(t) * f(t)$ denotes the pulse obtained by convolving the transmit and receive filters ($*$ denotes the convolution operation). In the following, $\mathbf{H}(k, n)$ denotes the $n_R \times n_T$ matrix CIR with entries $h_{qp}(k, n)$, where $1 \leq q \leq n_R$ and $1 \leq p \leq n_T$. By neglecting the small taps in the discretization of the raised-cosine FIR $c(t)$, we obtain a time-varying CIR $\mathbf{H}_n(z) = \sum_{k=0}^{d} \mathbf{H}(k, n) z^{-k}$ with finite length $d + 1$ at time index n.

For the optimal design of precoders or equalizers with power constraints, we assume that the CIR is known at the transmitter or receiver. In practice, a training sequence, or a semiblind/blind channel estimation scheme can be used to obtain the CIR at the receiver. The CIR at the transmitter can be obtained either by means of a feedback channel from the receiver to the transmitter or, for the case of time-duplexed channels, from previous measurements due to the channel reciprocity property. Accordingly, the applicability of this approach is limited to slowly fading channels such as those arising for fixed broadband wireless systems of the type considered by the IEEE 802.16 working group [23]. In this case, the overhead associated to the introduction of a feedback channel is small. Furthermore, the data can be divided into blocks whose length is selected such that the CIR is approximately constant over the length of a block. The time dependence of the channel, which is represented here by the subscript n, can be dropped over a given block so that the CIR is denoted as $\mathbf{H}(z)$. For this channel model we can then design an optimal constant MIMO FIR precoder or equalizer that will be applied to the block we consider. For each successive block, the channel will need to be reestimated, and the optimum precoder/equalizer will need to be updated based on the new CIR.

Since the optimal precoders or equalizers we consider rely in part on the Bezout identity of linear system theory, we start by reviewing this identity. See [10, 11] for more details. Given a rectangular FIR matrix $\mathbf{H}(z)$ of dimension $n_R \times n_T$ with $n_R < n_T$ (or $n_R > n_T$), that is, a "fat" (or "tall") matrix, there exists an FIR matrix $\mathbf{F}(z)$ of dimension $n_T \times n_R$ (or $\mathbf{G}(z)$ of dimension $n_R \times n_T$) such that

$$\mathbf{H}(z)\mathbf{F}(z) = \operatorname{diag}\{z^{-k_i},\ 1 \leq i \leq n_R\}, \tag{2}$$

or

$$\mathbf{G}(z)\mathbf{H}(z) = \operatorname{diag}\{z^{-k_i},\ 1 \leq i \leq n_T\}, \tag{3}$$

where the integer delays z^{-k_i} with $k_i \geq 0$ are arbitrary if and only if the transfer function $\mathbf{H}(z)$ is left coprime (or right coprime). The expressions (2) or (3) form the Bezout identity. In this respect, recall that a polynomial matrix $\mathbf{H}(z)$ is left coprime (or right coprime) if it has full row rank (or full column rank) for all finite complex values of z^{-1}. The applicability of the Bezout identity to random wireless channels stems from the fact that rectangular FIR systems with $n_T \neq n_R$ are generically (i.e., with probability one) left or right coprime depending on $n_R < n_T$ or $n_R > n_T$. Finally, observe that the left or right coprimeness condition of $\mathbf{H}(z)$ used above to ensure the existence of identity (2) or (3) may be weakened slightly by requiring that $\mathbf{H}(z)$ should be delay-permissive left or right coprime. $\mathbf{H}(z)$ is delay-permissive left (right) coprime if it has full row (column) rank for $z^{-1} \neq 0$. In this case, the diagonal delay matrix appearing on the right-hand side of (2) or (3) is not arbitrary and must be a multiple of the greatest common left (right) divisor of $\mathbf{H}(z)$. Precoders or equalizers with this structure are called Bezout precoders and have been examined in [10]. According to the Bezout identity, when the number n_T of transmit antennas is larger than the number n_R of receive antennas, an $n_T \times n_R$ FIR precoder matrix $\mathbf{F}(z) = \sum_{k=0}^{r} \mathbf{F}(k) z^{-k}$ with order d is used to equalize the channel at the transmitter. The precoded system is shown in Figure 1. In this case, the sampled received signal takes, therefore, the form

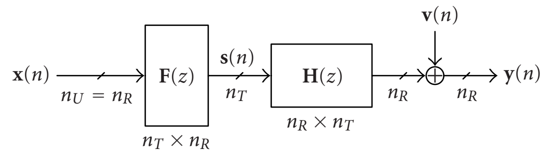

Figure 1: Precoded system.

$$\mathbf{y}(n) = \mathbf{H}(n) * \mathbf{F}(n) * \mathbf{x}(n) + \mathbf{v}(n), \tag{4}$$

where $\mathbf{v}(n)$ is a vector complex circular AWGN sequence independent of $\mathbf{x}(n)$, with zero-mean and variance $\sigma_v^2 I_{n_R}$. Similarly, when $n_R > n_T$, the equalized system is shown in Figure 2. In this case, the received sampled signal that passed through an $n_T \times n_R$ matrix FIR equalization filter $\mathbf{G}(z) = \sum_{k=0}^{r} \mathbf{G}(k) z^{-k}$ of order r can then be written as

$$\mathbf{y}'(n) = \mathbf{G}(n) * \mathbf{H}(n) * \mathbf{x}(n) + \mathbf{G}(n) * \mathbf{v}(n). \tag{5}$$

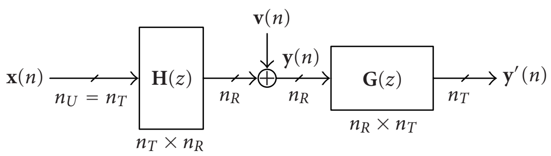

Figure 2: Equalized system.

Note that when identities (2) or (3) are satisfied, the expressions (4) and (5) become, respectively,

$$\mathbf{y}(n) = \operatorname{vec}\{x_i(n-k_i),\ 1 \le i \le n_R\} + \mathbf{v}(n) \tag{6}$$

or

$$\mathbf{y}'(n) = \operatorname{vec}\{x_i(n-k_i),\ 1 \le i \le n_T\} + \mathbf{G}(n) * \mathbf{v}(n), \tag{7}$$

so that perfect zero-forcing (ZF) equalization with adjustable delays is achieved in both instances. In spite of their apparent simplicity, Bezout precoders and equalizers have a significant defect. In the precoder case, the Bezout precoder may increase the transmit power significantly to overcome deep fades in the singular values of the channel matrix, while for the equalizer case, deep fades are compensated by amplifying the noise power. Since for sufficiently large order r, Bezout precoders or equalizers are not unique, it was suggested in [10] to use the additional degrees of freedom, as well as an optimum selection of the delays z^{-k_i}, to minimize the transmit power or the noise power at the receiver, respectively. Unfortunately, even with this additional flexibility, imposing an exact Bezout precoder or equalizer structure is still too constraining in some instances, and we explore below a more systematic tradeoff between the power of the residual vector ISI signal and transmit power constraints (resp., noise power constraints) in the design of MIMO FIR precoders (resp., equalizers). The optimal tradeoff can be obtained by convex optimization as discussed in the next section.

3. OPTIMAL PRECODER DESIGN WITH POWER CONSTRAINTS

In this section, we assume that the number n_T of transmit antennas is larger than the number n_R of receive antennas, that is, $n_T > n_R$, and that the MIMO CIR is known at the transmitter. To measure the mismatch between the concatenated CIR $\mathbf{H}(n) * \mathbf{F}(n)$ and the ideal ZF impulse response

$$\mathbf{E}(n) \triangleq \operatorname{diag}\{\delta(n-k_i),\ 1 \le i \le n_R\} \tag{8}$$

with z transform

$$\mathbf{E}(z) \triangleq \operatorname{diag}\{z^{-k_i},\ 1 \le i \le n_R\}, \tag{9}$$

we employ the objective function

$$\begin{aligned} J(\mathbf{F}) &= \sum_{n=-\infty}^{\infty} \|\mathbf{E}(n) - \mathbf{H}(n) * \mathbf{F}(n)\|_F^2 \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \|\mathbf{E}(e^{j\theta}) - \mathbf{H}(e^{j\theta})\mathbf{F}(e^{j\theta})\|_F^2 d\theta, \end{aligned} \tag{10}$$

where

$$\|\mathbf{M}\|_F^2 = \sum_{i=1}^{n_R}\sum_{j=1}^{n_R} |m_{ij}|^2 \tag{11}$$

denotes the squared Frobenius norm of an $n_R \times n_R$ complex matrix M. Consider the $(d+r+1) \times (r+1)$ block Toeplitz matrix

$$\Gamma(\mathbf{H}) = \begin{bmatrix} \mathbf{H}(0) & 0 & \cdots & 0 \\ \mathbf{H}(1) & \mathbf{H}(0) & \ddots & \vdots \\ \vdots & \mathbf{H}(1) & \ddots & 0 \\ \mathbf{H}(d) & \vdots & \ddots & \mathbf{H}(0) \\ 0 & \mathbf{H}(d) & \ddots & \mathbf{H}(1) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \mathbf{H}(d) \end{bmatrix} \tag{12}$$

whose blocks have size $n_R \times n_T$, and the block column matrices

$$\mathbf{E} = \begin{bmatrix} \mathbf{E}(0) \\ \mathbf{E}(1) \\ \vdots \\ \mathbf{E}(d+r) \end{bmatrix}, \qquad \mathbf{F} = \begin{bmatrix} \mathbf{F}(0) \\ \mathbf{F}(1) \\ \vdots \\ \mathbf{F}(r) \end{bmatrix}, \tag{13}$$

where the $n_R \times n_R$ blocks $\mathbf{E}(i)$ represent the coefficients of $\mathbf{E}(z)$, and the blocks $\mathbf{F}(i)$ have size $n_T \times n_R$. Then the criterion $J(\mathbf{F})$ can be expressed in matrix form as

$$J(\mathbf{F}) = \operatorname{tr}\left[(\Gamma(\mathbf{H})\mathbf{F} - \mathbf{E})^H(\Gamma(\mathbf{H})\mathbf{F} - \mathbf{E})\right], \tag{14}$$

which is obviously quadratic in $\mathbf{F}$. For a Bezout precoder, the identity (6) can be rewritten as

$$\Gamma(\mathbf{H})\mathbf{F} = \mathbf{E} \tag{15}$$

so that in this case, $J(\mathbf{F}) = 0$.

Since the Hessian $\nabla^2 J(\mathbf{F}) = \Gamma(\mathbf{H})^H\Gamma(\mathbf{H})$ is positive semidefinite, the objective function J is convex. In the following subsections, we construct precoders which minimize J under several types of transmit power constraints. In all cases, the constraints are convex so that the resulting convex optimization problem can be solved by Lagrangian duality methods.

3.1. Power constraints on the precoder columns

We first consider the case where a power constraint is imposed on each column of the precoder. Specifically, let $\mathbf{f}^j(z)$ denote the jth column of the matrix filter $\mathbf{F}(z)$ so that

$$\mathbf{F}(z) = [\mathbf{f}^1(z)\ \ \mathbf{f}^2(z)\ \ \cdots\ \ \mathbf{f}^{n_R}(z)], \tag{16}$$

where $\mathbf{f}^j(z)$ has dimension $n_T \times 1$. If $\mathbf{f}^j(n)$ denotes the impulse response of $\mathbf{f}^j(z)$, the component of the transmitted signal $\mathbf{s}(n)$ due to the jth transmitted symbol sequence $x_j(n)$

is given by $\mathbf{f}^j(n) * x_j(n)$. We consider the case where a power constraint of the form

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\left\|\mathbf{f}^j(e^{j\theta})\right\|^2 d\theta = \left\|\mathbf{f}^j\right\|^2 \le P_j \tag{17}$$

is imposed on the column $\mathbf{f}^j(z)$, where $\mathbf{f}^j$ denotes the jth column of the block matrix $\mathbf{F}$. This constraint can be viewed as modeling a situation where each user j is guaranteed a fixed quality of service (QoS) in the form of an allocated power P_j. Note that the constraint (17) specifies a convex domain for $\mathbf{f}^j$.

Since the objective function $J(\mathbf{F})$ can be decomposed as

$$J(\mathbf{F}) = \sum_{j=1}^{n_R} J_j(\mathbf{f}^j) \tag{18}$$

with

$$J_j(\mathbf{f}^j) = \left\|\Gamma(\mathbf{H})\mathbf{f}^j - \mathbf{e}^j\right\|^2, \tag{19}$$

where $\mathbf{e}^j$ is the jth column of the block matrix $\mathbf{E}$, the minimization of $J(\mathbf{F})$ under the constraints (17) for $1 \le j \le n_R$ is equivalent to the separate minimization for each j of $J_j(\mathbf{f}^j)$ under the constraint (17). Since $J_j(\mathbf{f}^j)$ and the constraint (17) are both convex, this optimization problem can be solved either in primal space or in dual space [13]. Because of the large size of $\mathbf{f}^j$, the primal problem has a large dimension, whereas the dual problem reduces to a scalar optimization problem since there is only one constraint. Therefore, we solve the dual form of the optimization problem.

The Lagrangian associated with the maximization of $J_j(\mathbf{f}^j)$ under the constraint (17) takes the form

$$\begin{aligned} L_j(\mathbf{f}^j,\lambda_j) &= J_j(\mathbf{f}^j) + \lambda_j\left({\mathbf{f}^j}^H\mathbf{f}^j - P_j\right) \\ &= {\mathbf{f}^j}^H(\mathbf{M}+\lambda_j\mathbf{I})\mathbf{f}^j - {\mathbf{f}^j}^H\Gamma(\mathbf{H})^H\mathbf{e}^j \\ &\quad - {\mathbf{e}^j}^H\Gamma(\mathbf{H})\mathbf{f}^j - \lambda_j P_j + 1, \end{aligned} \tag{20}$$

with

$$\mathbf{M} \triangleq \Gamma(\mathbf{H})^H\Gamma(\mathbf{H}), \tag{21}$$

where the Lagrange multiplier $\lambda_j \ge 0$. As long as λ_j is such that the Hermitian matrix $\mathbf{M}+\lambda_j\mathbf{I}$ is positive definite, $L_j(\mathbf{f}^j,\lambda_j)$ is a positive-definite quadratic function of $\mathbf{f}^j$ and is convex. Since the constraint (17) specifies a compact domain for $\mathbf{f}^j$, the Lagrangian admits a saddle point $({\mathbf{f}^j}^{\text{opt}},\lambda_j^{\text{opt}})$ whose first component ${\mathbf{f}^j}^{\text{opt}}$ is the solution of the constrained optimization problem [13]. To find the saddle point, we first fix λ_j and minimize the Lagrangian over $\mathbf{f}^j$, yielding

$${\mathbf{f}^j}^{\text{opt}}(\lambda_j) = (\mathbf{M}+\lambda_j\mathbf{I})^{-1}\Gamma(\mathbf{H})^H\mathbf{e}^j. \tag{22}$$

The dual function is therefore given by

$$\begin{aligned} G_j(\lambda_j) &= L_j({\mathbf{f}^j}^{\text{opt}}(\lambda_j),\lambda_j) \\ &= -{\mathbf{e}^j}^H\Gamma(\mathbf{H})(\mathbf{M}+\lambda_j\mathbf{I})^{-1}\Gamma(\mathbf{H})^H\mathbf{e}^j + 1 - \lambda_j P_j \end{aligned} \tag{23}$$

over the domain

$$\mathcal{D}_j = \{\lambda_j : \lambda_j \ge 0,\ \mathbf{M}+\lambda_j\mathbf{I} > 0\}. \tag{24}$$

In order to obtain λ_j^{opt}, we only need to maximize the concave function $G_j(\lambda_j)$ over $\mathcal{D}_j$. This maximum λ_j^{opt} is unique and obeys the gradient condition

$$\nabla_{\lambda_j}G_j(\lambda_j) = {\mathbf{e}^j}^H\Gamma(\mathbf{H})(\mathbf{M}+\lambda_j\mathbf{I})^{-2}\Gamma(\mathbf{H})^H\mathbf{e}^j - P_j = 0. \tag{25}$$

However, solving the secular equation (25) is difficult, and it is easier to obtain λ_j^{opt} by employing the iteration

$$\lambda_j^{k+1} = \lambda_j^k + \alpha^k d^k \tag{26}$$

until $\nabla_{\lambda_j}G_j(\lambda_j^{k+1})$ is small enough. In this iteration, the descent direction d_k is obtained by a Newton (or quasi-Newton) technique and the step size α^k is selected to ensure that successive iterates stay inside the domain $\mathcal{D}_j$. Then, substituting the maximum λ_j^{opt} in (22) gives the optimal solution ${\mathbf{f}^j}^{\text{opt}}$.

3.2. *Total power constraint*

Next, consider the case where a constraint is imposed on the total power used by all transmit antennas. This constraint can be expressed as

$$\operatorname{tr}\left[\frac{1}{2\pi}\int_{-\pi}^{\pi}\mathbf{F}^H(e^{j\theta})\mathbf{F}(e^{j\theta})d\theta\right] = \operatorname{tr}(\mathbf{F}^H\mathbf{F}) \le P_T. \tag{27}$$

The Lagrangian corresponding to the minimization of $J(\mathbf{F})$ under the constraint (27) can be expressed as

$$\begin{aligned} L(\mathbf{F},\lambda) &= J(\mathbf{F}) + \lambda(\operatorname{tr}(\mathbf{F}^H\mathbf{F}) - P_T) \\ &= \operatorname{tr}\left[\mathbf{F}^H(\mathbf{M}+\lambda\mathbf{I})\mathbf{F} - \mathbf{E}^H\Gamma(\mathbf{H})\mathbf{F} - \mathbf{F}^H\Gamma(\mathbf{H})^H\mathbf{E}\right] \\ &\quad + n_R - \lambda P_T. \end{aligned} \tag{28}$$

We use the fact that $\partial\text{tr}(\mathbf{AXB})/\partial\mathbf{X} = \mathbf{A}^H\mathbf{B}^H$ and $\partial\text{tr}(\mathbf{AX^HB})/\partial\mathbf{X} = \mathbf{BA}$ [24], which gives

$$\mathbf{F}^{\text{opt}}(\lambda) = (\mathbf{M}+\lambda\mathbf{I})^{-1}\mathbf{\Gamma}(\mathbf{H})^H\mathbf{E}. \tag{29}$$

The dual function is, therefore, given by

$$\begin{aligned} G(\lambda) &= L(\mathbf{F}^{\text{opt}}(\lambda),\lambda) \\ &= \operatorname{tr}\left[-\mathbf{E}^H\mathbf{\Gamma}(\mathbf{H})(\mathbf{M}+\lambda\mathbf{I})^{-1}\mathbf{\Gamma}(\mathbf{H})^H\mathbf{E}\right] + n_R - \lambda P_T \end{aligned} \tag{30}$$

over the domain

$$\mathcal{D} = \{\lambda : \lambda \ge 0,\ \mathbf{M}+\lambda\mathbf{I} > 0\}. \tag{31}$$

For the concave function $G(\lambda)$, the optimal value λ^{opt} satisfies the gradient condition

$$\nabla_\lambda G(\lambda) = \operatorname{tr}\left[\mathbf{E}^H\mathbf{\Gamma}(\mathbf{H})(\mathbf{M}+\lambda\mathbf{I})^{-2}\mathbf{\Gamma}(\mathbf{H})^H\mathbf{E}\right] - P_T = 0. \tag{32}$$

This secular equation is again difficult to solve, and it is preferable to compute the optimal Lagrange multiplier λ^{opt} numerically by maximizing $G(\lambda)$ over $\mathcal{D}$ with a Newton or quasi-Newton method. Substitution of λ^{opt} inside (29) then yields the optimal precoder $\mathbf{F}^{\text{opt}}$.

3.3. Power constraint on each transmit antenna

In practical situations, the power amplifier for each transmit antenna has a region of linear operation. Thus instead of imposing a total power constraint on the system of antennas, it is more realistic to impose a power constraint on each transmit antenna, or equivalently on each row of the precoder. Let $\mathbf{f}_i(z)$ denote the ith row of the precoder with $1 \leq i \leq n_T$, that is,

$$\mathbf{F}(z) = \begin{bmatrix} \mathbf{f}_1(z) \\ \mathbf{f}_2(z) \\ \vdots \\ \mathbf{f}_{n_T}(z) \end{bmatrix}. \tag{33}$$

Then, denoting the maximum operation power for antenna i as P_i, the power constraint on this antenna can be expressed as

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \mathbf{f}_i(e^{j\theta})\mathbf{f}_i^H(e^{j\theta})\,d\theta = \operatorname{tr}\left(\mathbf{F}^H\mathbf{C}_i\mathbf{F}\right) \leq P_i \tag{34}$$

with

$$\mathbf{C}_i = \mathbf{I}_{r+1} \otimes \left(\mathbf{u}_i\mathbf{u}_i^H\right), \tag{35}$$

where $\mathbf{u}_i$ denotes the unit vector of length n_T with all zero entries, except for its ith entry, which equals one, and $\otimes$ represents the Kronecker product. The constraint (34) is quadratic positive semidefinite and thus convex. The Lagrangian associated to the minimization of $J(\mathbf{F})$ under the constraints (34) for $1 \leq i \leq n_T$ can be written as

$$L(\mathbf{F}, \boldsymbol{\lambda}) = J(\mathbf{F}) + \boldsymbol{\lambda}^T\mathbf{c}(\mathbf{F}), \tag{36}$$

with

$$\begin{aligned} \boldsymbol{\lambda} &= \left[\lambda_1 \;\; \lambda_2 \;\; \cdots \;\; \lambda_{n_T}\right]^T, \\ \mathbf{c}(\mathbf{F}) &= \left[c_1(\mathbf{F}) \;\; c_2(\mathbf{F}) \;\; \cdots \;\; c_{n_T}(\mathbf{F})\right]^T, \end{aligned} \tag{37}$$

where

$$c_i(\mathbf{F}) = \operatorname{tr}\left(\mathbf{F}^H\mathbf{C}_i\mathbf{F}\right) - P_i. \tag{38}$$

The computation of the optimum precoder $\mathbf{F}^{\text{opt}}$ then proceeds along the same lines as in the previous two subsections, except that the dual function depends now on a vector $\boldsymbol{\lambda}$ of Lagrange multipliers instead of a scalar. The Lagrangian can be expressed as

$$\begin{aligned} L(\mathbf{F}, \boldsymbol{\lambda}) = \operatorname{tr}\Bigg[&\mathbf{F}^H\left(\mathbf{M} + \sum_{i=1}^{n_T}\lambda_i\mathbf{C}_i\right)\mathbf{F} - \mathbf{E}^H\Gamma(\mathbf{H})\mathbf{F} \\ &- \mathbf{F}^H\Gamma(\mathbf{H})^H\mathbf{E}\Bigg] \\ &+ n_R - \sum_{i=1}^{n_T}\lambda_i P_i, \end{aligned} \tag{39}$$

which is minimized by

$$\mathbf{F}^{\text{opt}}(\boldsymbol{\lambda}) = \left(\mathbf{M} + \sum_{i=1}^{n_T}\lambda_i\mathbf{C}_i\right)^{-1}\Gamma(\mathbf{H})^H\mathbf{E}. \tag{40}$$

The domain of the dual function

$$\begin{aligned} G(\boldsymbol{\lambda}) &= L\left(\mathbf{F}^{\text{opt}}, \boldsymbol{\lambda}\right) \\ &= \operatorname{tr}\left[-\mathbf{E}^H\Gamma(\mathbf{H})\left(\mathbf{M} + \sum_{i=1}^{n_T}\lambda_i\mathbf{C}_i\right)^{-1}\Gamma(\mathbf{H})^H\mathbf{E}\right] \\ &\quad + n_R - \sum_{i=1}^{n_T}\lambda_i P_i \end{aligned} \tag{41}$$

is then given by

$$\mathcal{D} = \left\{\boldsymbol{\lambda} \in \mathbb{R}^{n_T} : \lambda_i \geq 0,\; \mathbf{M} + \sum_{i=1}^{n_T}\lambda_i\mathbf{C}_i > 0\right\}. \tag{42}$$

Then using a Newton technique to maximize $G(\boldsymbol{\lambda})$ over the convex domain $\mathcal{D}$ gives the optimal Lagrange multiplier vector $\boldsymbol{\lambda}^{\text{opt}}$, which in turn yields the optimal precoder after substitution inside $\mathbf{F}^{\text{opt}}(\boldsymbol{\lambda})$.

Up to this point, we have explored the tradeoff existing between ISI and ICI minimization and transmit power constraints in precoder design. However, even after a precoder has been selected, it is possible to reduce further the required transmit power by adding a signal shaping sequence to the precoder input.

4. SIGNAL SHAPING SEQUENCE SELECTION

The selection of a signal shaping sequence to minimize the transmit power is similar conceptually to the selection of a shaping sequence for TH precoding [22], except that it is easier to implement since the MIMO precoder has an FIR structure, whereas the TH precoder has a feedback structure which complicates the selection of a shaping sequence.

The basic idea is to perform a constellation expansion by adding to the vector input sequence $\mathbf{x}(n)$ a shaping sequence $\mathbf{d}(n)$, yielding the transmitted sequence

$$\mathbf{z}(n) = \mathbf{x}(n) + \mathbf{d}(n). \tag{43}$$

The full system is shown in Figure 3. For simplicity, we consider the case of QPSK modulation so that each entry $x_k(n)$ of the input sequence admits a complex representation of the

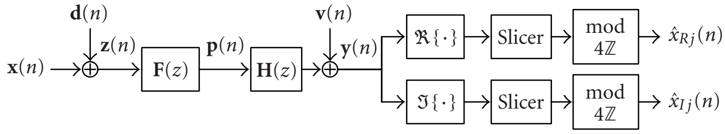

Figure 3: Transmission system with the shaping sequence $\mathbf{d}(n)$.

form $x_k(n) = x_{Rk}(n) + jx_{Ik}(n)$ with $x_{Rk}(n), x_{Ik}(n) \in \{1, -1\}$. Each entry of the shaping sequence takes the form $d_k(n) = d_{Rk}(n) + jd_{Ik}(n)$ with $d_{Rk}(n), d_{Ik}(n) \in \{0, \pm 4\}$. The transmit signal $\mathbf{p}(n)$ is then given by

$$\begin{aligned}\mathbf{p}(n) &= \sum_{l=0}^{r} \mathbf{F}(l)\mathbf{z}(n-l) \\ &= \sum_{l=0}^{r} \left[\mathbf{F}(l)\mathbf{x}(n-l) + \mathbf{F}(l)\mathbf{d}(n-l)\right].\end{aligned} \tag{44}$$

For a fixed input sequence $\mathbf{x}(n)$, the first term in (43) is known. The shaping sequence $\mathbf{d}(n)$ appearing in the second term is selected such that the transmit power is minimized. That is, the optimal shaping sequence is given by

$$\mathbf{d}^{\text{opt}}(n) = \arg\min_{\mathbf{d}(n)} \left(\sum_{n=0}^{L_s} \|\mathbf{p}(n)\|^2 \right), \tag{45}$$

where L_s denotes a fixed block length.

Since each entry $d_k(n)$ of the shaping sequence can take nine possible values (three values each for the real and imaginary parts), the minimization in (45) can be performed by applying standard trellis search techniques for a suitable state trellis. The state at time index n is

$$\left[\mathbf{d}^T(n-1) \ \ \mathbf{d}^T(n-2) \ \ \cdots \ \ \mathbf{d}^T(n-r)\right] \tag{46}$$

so that the trellis has 9^{rn_U} states. Note that this number can be rather large even for small values of n_U and r.

The Viterbi algorithm (VA) can be employed to solve the minimization problem (45). However, both the VA's computational complexity and storage complexity grow exponentially with the order r of the precoder and the value of n_U. To keep the computational burden manageable, one can employ suboptimal trellis search techniques. One such technique is the list Viterbi method described in [25, Chapter 5], which at any given time considers the most promising K paths or states in the trellis. The number K corresponds to the size of the list. At each depth, the most promising K paths are extended, instead of all for the VA. The beam VA [26] follows a similar philosophy, except that the number of retained paths varies instead of being fixed. Specifically, the retained paths are those whose metrics fall within a certain beam width or distance from the path with the lowest metric. The size of the beam can be decided based on different application needs. The list VA (LVA) and beam VA are only two instances of a wider class of reduced-complexity VAs, which includes other methods based on state partitioning. For the simulations presented in Section 6, we consider only the LVA. At each time index l, the K best paths are extended. There are 9^{n_U} subpaths extended from one selected state. The trellis metric is the transmit power $\sum_{n=0}^{l} \|\mathbf{p}(n)\|^2$ up to time l. The power reduction which is achievable through shaping depends on the size K of the list. In the theory, the larger K, the better the performance. When the order r of the precoder or the value of n_U increases, K needs to be larger. But in this case, the computational complexity and storage requirements of the LVA become prohibitive.

The VA and its variants can be viewed as performing a nonbacktracking breadth-first search over a state trellis. The sequential algorithms (SAs), such as the stack algorithm or Fano's algorithm [25, 27], follow instead a depth-first trellis search strategy, with some possible backtracking. The fundamental idea of these algorithms is that the search process should only explore the most promising path at any given time. If a path to a node looks "bad," we can discard all the paths stemming from that node without a significant loss in performance. The key problem for SAs is to find the Fano metric. In the VA, we always compare the metrics of paths with the same length. However, the path metrics in SAs must be adjusted according to the length of the path. The adjustment to the path metric based on the length of the path is called the Fano metric. For the problem we consider, we want to minimize the transmit power by selecting the shaping sequence $\mathbf{d}(n)$. Thus, the Fano metric can be defined as the transmit energy divided by the length of the path up to the current node, that is, $(\sum_{n=0}^{l} \|\mathbf{p}(n)\|^2)/l$. For the simulations presented in Section 6, we implement only the stack algorithm. In this algorithm, the new nodes are extended from the best node and the metrics of the new extended nodes are evaluated and sorted together with the metrics of the old nodes. After sorting, a new best node appears at the top of the stack, and the extension process is initiated again.

Once the shaping sequence has been selected and added to the transmitted input, since the real and imaginary parts of each entry of the transmitted signal $\mathbf{z}(n)$ belong to the expanded constellation $\{\pm 1, \pm 3, \pm 5\}$, at the receiver end, a slicer adapted to this expanded constellation is applied to the real and imaginary parts of each entry of the received signal $\mathbf{y}(n)$, followed by a reduction modulo $4\mathbb{Z}$, as shown in Figure 3.

5. OPTIMAL EQUALIZER DESIGN WITH OUTPUT NOISE POWER CONSTRAINTS

When $n_T < n_R$, provided the channel is right coprime, a Bezout equalizer can be used to recover the transmitted vector signal [10]. But, like all ZF equalizers, Bezout equalizers tend to amplify the received noise. In multimedia applications, several information streams, such as audio or video, are typically sent over different channels with various noise power constraints. For example, video has a smaller noise power requirement than audio. For such applications, it is natural, as shown in [20, 28], to introduce explicitly QoS considerations in the form of noise power constraints in the design of MIMO precoders and decoders. We consider below the design of equalizers that minimize the power of the remaining ISI component at their output while obeying output noise power constraints. This problem is of course dual to the precoder design with transmit power constraints considered earlier.

We assume that the orders of the channel and equalizer are d and r, respectively. If $\mathbf{G}(n)$ denotes the impulse response of the equalizer, the signal at the equalizer output is given by (4). Let

$$\mathbf{E}'(n) \triangleq \operatorname{diag}\{\delta(n-k_i),\ 1 \le i \le n_T\} \tag{47}$$

be the impulse response of the target equalized channel, and let

$$\mathbf{E}'(z) = \operatorname{diag}\{z^{-k_i},\ 1 \le i \le n_T\} \tag{48}$$

represent its z transform. Then the signal at the equalizer output can be decomposed into three components: the desired signal $\mathbf{E}'(n) * \mathbf{x}(n)$, the ISI component $[\mathbf{G}(n) * \mathbf{H}(n) - \mathbf{E}'(n)] * \mathbf{x}(n)$, and the noise component $\mathbf{G}(n) * \mathbf{v}(n)$. We design the MIMO equalizer $\mathbf{G}(n)$ to minimize the power of the ISI component while ensuring that the noise power is below a preset threshold. The power of the ISI component can be expressed as

$$\begin{aligned} P_{\text{ISI}} &= E\Big[\big|\big|[\mathbf{G}(n) * \mathbf{H}(n) - \mathbf{E}'(n)] * \mathbf{x}(n)\big|\big|^2\Big] \\ &= \sigma_x^2 J(\mathbf{G}), \end{aligned} \tag{49}$$

where

$$\begin{aligned} J(\mathbf{G}) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \|\mathbf{E}'(e^{j\theta}) - \mathbf{G}(e^{j\theta})\mathbf{H}(e^{j\theta})\|_F^2 d\theta \\ &= \operatorname{tr}\Big[\big(\mathbf{G}\Gamma(\mathbf{H}^H)^H - \mathbf{E}'\big)\big(\mathbf{G}\Gamma(\mathbf{H}^H)^H - \mathbf{E}'\big)^H\Big], \end{aligned} \tag{50}$$

$$\mathbf{G} = [\mathbf{G}(0)\ \ \mathbf{G}(1)\ \ \cdots\ \ \mathbf{G}(r)], \tag{51}$$

$$\mathbf{E}' = [\mathbf{E}'(0)\ \ \mathbf{E}'(1)\ \ \cdots\ \ \mathbf{E}'(d+r)]. \tag{52}$$

From (49), we see that minimizing the ISI power is equivalent to minimizing $J(\mathbf{G})$. Also, the noise power constraint is equivalent to

$$\operatorname{tr}\left[\frac{1}{2\pi}\int_{-\pi}^{\pi} \mathbf{G}(e^{j\theta})\mathbf{G}^H(e^{j\theta})d\theta\right] = \operatorname{tr}(\mathbf{G}\mathbf{G}^H) \le P_T. \tag{53}$$

This reduces the equalizer design to a problem of the same form as the one we considered in Section 3.2 for the design of a precoder with a total power constraint. By using the Lagrangian duality method, the closed-form solution of the optimal equalizer with the output noise power constraint is given by

$$\mathbf{G}^{\text{opt}} = \mathbf{E}'\Gamma(\mathbf{H}^H)(\mathbf{M}' + \lambda^{\text{opt}}\mathbf{I})^{-1}, \tag{54}$$

where $\mathbf{M}' = \Gamma(\mathbf{H}^H)^H\Gamma(\mathbf{H}^H)$. We see that expression (54) for the optimal equalizer coincides with the MMSE equalizer when $\lambda^{\text{opt}} = \sigma_v^2/\sigma_x^2$ and with the ZF equalizer when $\lambda^{\text{opt}} = 0$. Note that the ZF equalizer minimizes the residual ISI power, since it has zero ISI, and the MMSE equalizer minimizes the MSE, whereas the optimal equalizers we design minimize the residual ISI signal power under an output noise power constraint. Therefore, they define a family of equalizers among which the ZF and MMSE equalizers represent two special cases corresponding to $\lambda^{\text{opt}} = 0$ and $\lambda^{\text{opt}} = \sigma_v^2/\sigma_s^2$, respectively.

Instead of imposing an aggregate output noise power constraint, it may be preferable to require that each component of the vector signal at the equalizer output obeys a separate noise power constraint. Let $\mathbf{g}_i(z)$ and $\mathbf{e}_i'(z)$ denote the ith rows of the FIR matrices $\mathbf{G}(z)$ and $\mathbf{E}'(z)$, with $1 \le i \le n_T$. Then the noise power constraint for the ith component of the equalizer output can be expressed as

$$||\mathbf{g}_i||^2 \le P_i, \tag{55}$$

where $\mathbf{g}_i$ denotes the ith row of the matrix $\mathbf{G}$ specified by (51). Then, minimizing the ISI and ICI power under the noise power constraints $\|\mathbf{g}_i\|^2 \le P_i$ for $1 \le i \le n_T$ is equivalent to minimizing

$$J(\mathbf{G}) = \sum_{i=1}^{n_T} J_i(\mathbf{g}_i) \tag{56}$$

with

$$J_i(\mathbf{g}_i) = \left\|\mathbf{g}_i\Gamma(\mathbf{H}^H)^H - \mathbf{e}_i'\right\|^2, \tag{57}$$

where $\mathbf{e}_i'$ denotes the ith row of $\mathbf{E}'$. But this problem is clearly equivalent to the separate minimization of $J_i(\mathbf{g}_i)$ under the constraint (55), which is a problem of the same type as was considered in Section 3.1. The optimal equalizer $\mathbf{g}_i$ is given by

$$\mathbf{g}_i^{\text{opt}} = \mathbf{e}_i'\Gamma(\mathbf{H}^H)(\mathbf{M}' + \lambda_i^{\text{opt}}\mathbf{I})^{-1}. \tag{58}$$

Thus, for channels with more receive than transmit antennas, the design of ISI and ICI minimizing equalizers under output noise constraints gives rise to problems identical to those considered earlier for the dual case of more transmit than receive antennas. This is not a surprise since for MIMO channels that operate in time-division duplex, the channel is the same in both directions so that a precoder designed for transmission in one direction becomes automatically an equalizer when the transmission direction is reversed.

6. SIMULATION RESULTS

In this section, simulation results are presented to illustrate the performance of optimal precoders with power constraints. Since the same design technique is applicable to both equalizers and precoders, we only show results for the precoder case. In our simulations, a 4-input 2-output single-carrier frequency-selective wireless channel is considered. The transmit signal uses an uncoded QPSK constellation format and the combined response of the transmit and receive filters has a raised cosine spectrum with a roll-off factor of 0.2. The broadband transmission channel from input p to output q is described by a five-path fading model. We assume that the transmit antennas are located in the far field of the receive antennas so that for each propagation ray emanating from a fixed transmit antenna, the receive antennas have the same fading amplitudes, arriving angles, and multipath delays. Specifically, if we consider the complex CIR (1), we have

$$a_{lqp}(n) = a_{lp}(n) \exp\left(j\phi_{lqp}(n)\right), \quad \tau_{lqp}(nT) = \tau_{lp}(nT), \tag{59}$$

where for the case of a uniform linear receive antenna array, the interantenna phase factor $\phi_{lqp}(n)$ can be measured with respect to the first antenna and takes the form

$$\phi_{lqp}(n) = \frac{(q-1)2\pi d_s \sin\left(\theta_{lp}(n)\right)}{\lambda}, \tag{60}$$

where $d_s = 10\lambda$ denotes the interantenna spacing, λ is the carrier wavelength, and $\theta_{lp}(n)$ represents the angle of arrival of the lth ray from the pth transmit antenna measured with respect to the normal of the receiver array at time index nT. In the CIR model (1), the first term corresponds to a line-of-sight path with unit gain. The arrival angles and the multipath delays are assumed to be random and uniformly distributed over the intervals $[0, 2\pi]$ and $[0, 12T]$, respectively. The fading amplitudes exponentially decrease with the path delay according to the relation $a_{lp}(n) = -\tau_{lp}(nT)/T$ in dB. The length of the channel is truncated to 5 since the impulse response samples beyond the 5th are statistically very small. So in the following simulations, the signals are transmitted over a length-5 frequency-selective MIMO channel. Since optimal precoders are applicable only to fixed or slowly time-varying channels, we consider these two types of channels in our simulation. For the case of a fixed channel, the channel is assumed to be quasistationary, that is, it is stationary during the transmission of one block but changes independently from one block to another. This simulation procedure ensures that the results presented do not depend on one specific (good or bad) channel, but instead sample exhaustively the space of all possible channels. The simulation results shown below represent an average over 1000 random channels and the length of each block is 1000 symbols. The channel noise is simulated by adding independent complex circular white Gaussian noise sequences which zero-mean and variance σ_v^2 to each receive antenna signal. In all plots, the SNR (in dB) is defined as SNR $= \mathrm{tr}(\mathbf{F}\mathbf{F}^H)\sigma_x^2/\sigma_v^2$, where $\sigma_x^2 = 2$.

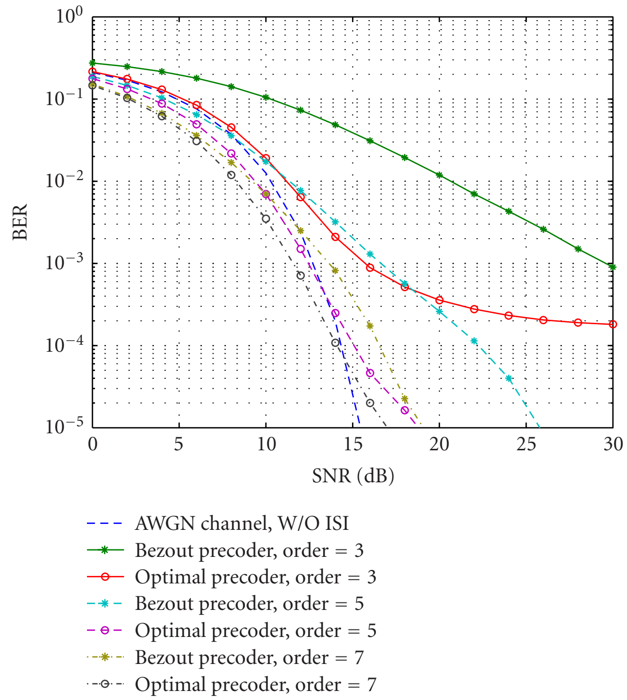

FIGURE 4: Comparison of the BER performance of Bezout precoders and optimal precoders with a total power constraint for orders 3, 5, and 7.

Figure 4 compares the BER performance of Bezout precoders and of optimal precoders with a total power constraint when channels are quasistationary and the CIR is perfectly known at the transmitter. Since only two signals are transmitted, we apply the power constraint $\mathrm{tr}(\mathbf{F}\mathbf{F}^H) \leq P_T = 2$ which ensures that the total transmit power for the two signals is not amplified by the 4×2 precoder $\mathbf{F}(z)$. The Bezout precoders are obtained by using the technique described in [10] for minimizing the transmit power. The figure shows the BER performance for Bezout and optimal precoders of orders 3, 5, and 7, respectively. As a benchmark, we also show the BER performance for QPSK signals transmitted over an AWGN channel. The figure indicates that optimal precoders with a total power constraint have a better BER performance than Bezout precoders and the BER performance gets progressively closer to the AWGN bound at mid-to-high SNR as the precoder order increases. For low SNR, the BER performance of optimal precoders is better than the AWGN bound because the SNR is defined as the transmit power over the noise power and the transmit power of the optimal precoder is less than or equal to $P_T\sigma_x^2 = 4$, whereas for AWGN channels, the transmit power is always 4. Therefore, for the same SNR, the noise added to a system with an optimal precoder is less than the corresponding AWGN channel noise, which results in a better BER performance. In contrast, at high SNR, the noise becomes very small, and the residual ISI and ICI of optimal precoders become the dominant factor controlling BER performance instead of noise. Figure 5 shows the BER performance

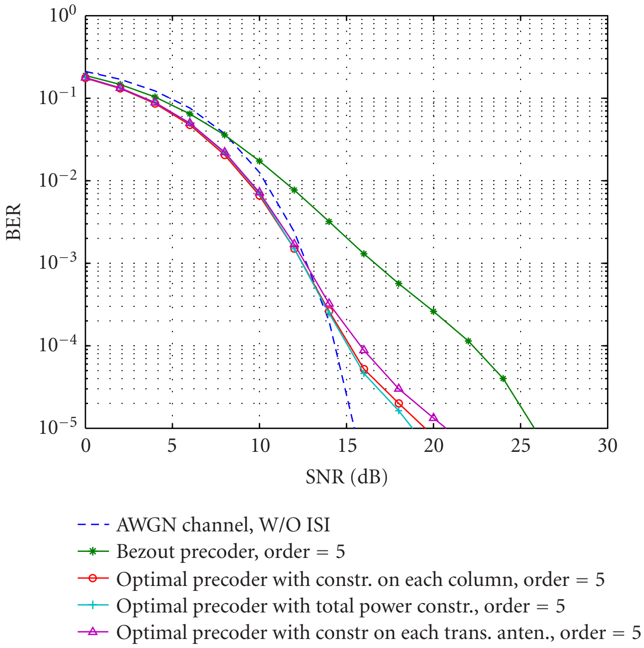

Figure 5: BER performance comparison for optimal precoders designed with a total power constraint, or with power constraints on the precoder columns or rows; order of precoders = 5.

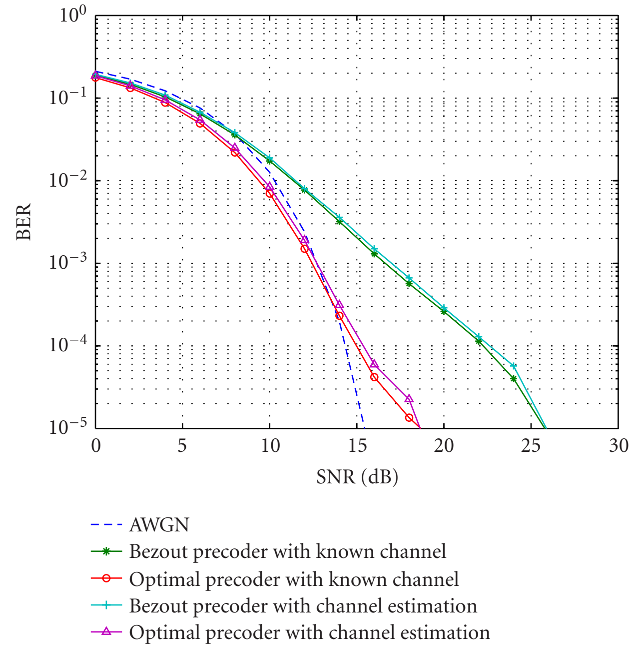

Figure 6: Comparison of the BER performance of optimal precoders with perfect channel knowledge and with channel estimation; order of precoders = 5.

for optimal precoders designed with different power constraints and perfect channel knowledge at the transmitter. The power constraint applied to each column of the precoder in (17) is $P_j = P_T/n_R$, where here $n_R = 2$. In other words, the total power is divided evenly among the two users. Similarly, the power constraint applied to each transmit antenna in (34) is P_T/n_T, where $n_T = 4$. Thus in this case, the total transmit power $P_T = 2$ is divided evenly across all 4 transmit antennas. In all cases, the precoder has order 5. Figure 5 shows that all three types of power constrained precoders exhibit a similar BER performance. Although the difference in performance between power constrained precoders is relatively small, it appears that the fewer the constraints, the better the performance. Thus, the precoder with a total power constraint performs best since it involves only one constraint, followed by the precoder with power constraints on the precoder columns, which has $q = 2$ constraints, and the precoder with a uniform power constraint for all transmit antennas performs the worst since it involves $p = 4$ constraints.

In real systems, channel estimation needs to be considered, so Figure 6 shows the effect of channel estimation on the performance of Bezout and optimal precoders. A vector training sequence of 200 symbols is employed, and the recursive least-squares algorithm is used to estimate the channel. The estimated CIR is then used for precoder design, instead of the true CIR. It is also of interest to evaluate the BER performance of optimal precoders for slowly time-varying channels. To simulate the channel time variations, we use, as Doppler spectrum,

$$S(f) = \begin{cases} 1 - 1.72 f_0^2 + 0.785 f_0^4, & |f_0| \leq 1, \\ 0, & |f_0| > 1, \end{cases} \tag{61}$$

according to the fixed broadband wireless channel model proposed by the IEEE 802.16 Working Group [23]. Here $f_0 = f/f_m$, and f_m is the maximum Doppler frequency. The Doppler spectrum (61) affects both the magnitude and phase of the CIR. How Doppler shifts affect the magnitude of the impulse response follows the simulation codes provided by the IEEE 802.16 Working Group [23]. When Doppler shifts are considered, the phase factor $\phi_{lqp}(n)$ in (60) becomes

$$\phi_{lqp}(n) = \frac{(q-1)2\pi d_s \sin\left(\theta_{lp}(n)\right)}{\lambda} + 2\pi f_d(n), \tag{62}$$

where $f_d(n)$ is the Doppler frequency characterized by the spectrum (61). The data rate is 10000 times the maximum Doppler frequency so that the channel is almost fixed during the transmission of a block of 1000 symbols. In our simulations, only one sequence of channel parameters is generated for the slowly time-varying channel, that is, we do not perform an average over multiple simulation runs. The length of the transmitted sequence is 100 000 symbols long and the optimal precoder is updated after each block of 1000 symbols. Figure 7 shows that optimal precoders are more robust than Bezout precoders for slowly time-varying channels since the matrix singular value decomposition which is used for the Bezout precoder design is sensitive to small errors.

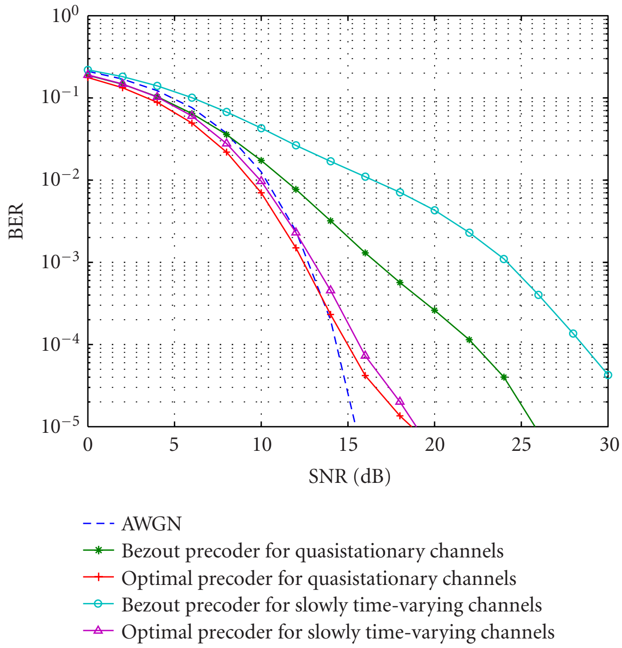

Figure 7: Comparison of the BER performance of optimal precoders for quasistationary channels and slowly time-varying channels; order of precoders = 5.

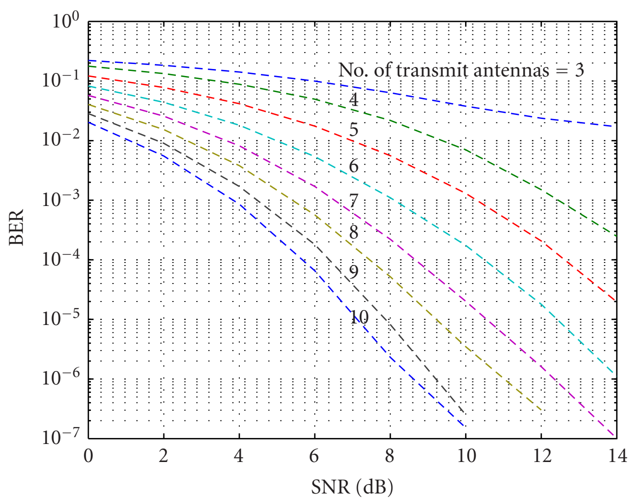

Figure 8: BER performance for a precoder with a total power constraint as the number of transmit antennas is increased from 3 to 10.

To illustrate the effect of increasing the number of transmit antennas while keeping the number of receive antennas fixed, Figure 8 shows the BER obtained by an optimal precoder of order 5 with a total transmit power constraint for a quasistationary channel with $n_R = 2$ receive antennas, as the number of transmit antennas is increased from $n_T = 3$ to 10. The figure shows that the performance for a system with only 3 transmit antennas is rather poor, but as the number of transmit antennas increases from 4 to 10, the performance improves steadily due to the increased spatial diversity of the system.

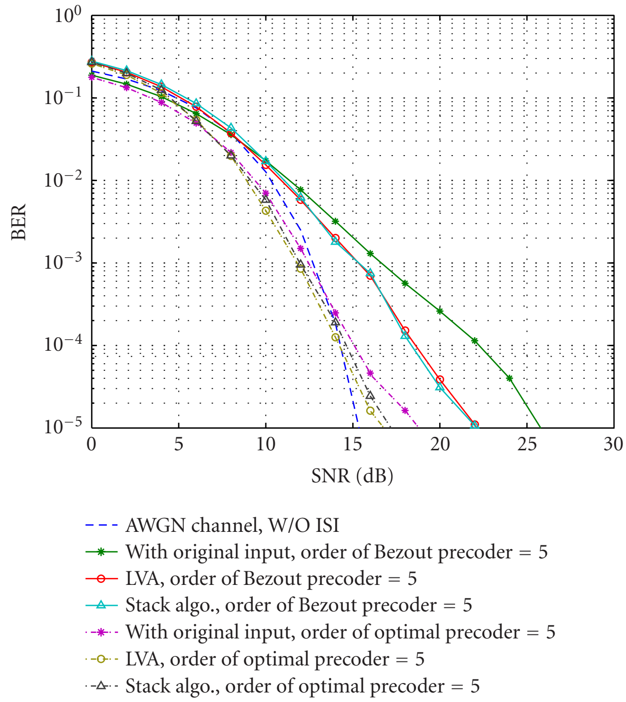

Figure 9: Effect of shaping on the BER performance of optimal precoders with a total power constraint and of Bezout precoders.

Finally, we illustrate the vector shaping sequence technique of Section 4 for the case of QPSK (4-QAM) modulation and the simulation is run over quasi-stationary channels. The LVA and stack algorithm are employed to select a shaping sequence minimizing the transmit power. In our simulations, the size of the list or of the stack is only 10 to keep the computational load reasonable. Figure 9 compares the BER performance with and without shaping of an optimal precoder with a total transmit power constraint and a Bezout precoder. The total transmit power constraint is again $\text{tr}(\mathbf{FF}^H) \leq P_T = 2$ and the precoders have order 5. We see from the figure that because the optimal precoder has already decreased significantly the transmit power, the shaping gain is smaller than that of the Bezout precoder. The results show that systems with a shaping sequence performs better than those without shaping at high SNR (in our simulation, for a SNR > 8 dB). The fact that an improvement appears only at high SNR is due to the fact that after constellation expansion, the original 4-QAM system becomes equivalent to a 36-QAM system, whose performance at low SNR is poorer than a 4-QAM system. Finally, note from Figure 9 that the LVA and stack algorithm have a similar performance.

7. CONCLUSION

In this paper, to overcome a limitation of Bezout precoders or equalizers for frequency-selective MIMO channels, a technique has been presented for the optimal synthesis of precoders subject to transmit power constraints, or equalizers subject to output noise power constraints. This design technique is applicable to unbalanced channels with either $n_T > n_R$ or $n_R > n_T$, and assumes full knowledge of the CIR at either the transmitter or the receiver. It formulates the design of precoders or equalizers minimizing the ISI and ICI power under transmit power or noise power constraints as a positive-semidefinite quadratic minimization problem with convex positive-definite quadratic constraint. Lagrangian duality is employed to solve the lower-dimensional dual form of this problem. Due to the reciprocity principle of wave propagation, the design of optimal precoders and equalizers have an identical form. When a precoder has been chosen, a shaping technique based on constellation expansion is also presented for reducing the required transmit power. Since the selection of effective shaping sequences requires a search over a large trellis, the search is performed suboptimally by using the LVA or stack algorithms.

We have focused our attention here excursively on equalization. However, advanced MIMO wireless systems require the development of integrated coding and equalization systems [29]. While the design of the MIMO coding and precoder/equalizer components can certainly be kept separate, there exist interesting possibilities for joint design that might be worth exploring. For example, instead of using a channel formed by pure delays as a target for the equalized channel $\mathbf{H}(z)\mathbf{F}(z)$ in (6), one could just attempt to diagonalize approximately the MIMO channel and then use the resulting diagonal terms as inner codes for decoupled serially concatenated turbo-like codes for each transmit signal.

ACKNOWLEDGMENT

This research was supported by NSF Grant ECS-0121469.

REFERENCES

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, 1998.

[2] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[3] D. Gesbert, M. Shafi, D.-S. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: an overview of MIMO space-time coded wireless systems," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 3, pp. 281–302, 2003.

[4] A. Duel-Hallen, "Equalizers for multiple-input/multiple-output channels and PAM systems with cyclostationary input sequences," *IEEE Journal on Selected Areas in Communications*, vol. 10, no. 3, pp. 630–639, 1992.

[5] A. Lozano and C. Papadias, "Layered space-time receivers for frequency-selective wireless channels," *IEEE Trans. Communications*, vol. 50, no. 1, pp. 65–73, 2002.

[6] R. Van Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*, Artech House, Boston, Mass, USA, 2000.

[7] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Communications*, vol. 46, no. 3, pp. 357–366, 1998.

[8] D. Falconer, S. L. Ariyavisitakul, A. Benyamin-Seeyar, and B. Eidson, "Frequency domain equalization for single-carrier broadband wireless systems," *IEEE Communications Magazine*, vol. 40, no. 4, pp. 58–66, 2002.

[9] N. Al-Dhahir, "Single-carrier frequency-domain equalization for space-time block-coded transmissions over frequency-selective fading channels," *IEEE Communications Letters*, vol. 5, no. 7, pp. 304–306, 2001.

[10] S.-Y. Kung, Y. Wu, and X. Zhang, "Bezout space-time precoders and equalizers for MIMO channels," *IEEE Trans. Signal Processing*, vol. 50, no. 10, pp. 2499–2514, 2002.

[11] T. Kailath, *Linear Systems*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1980.

[12] C. Windpassinger, R. F. H. Fischer, T. Vencel, and J. B. Huber, "Precoding in multiantenna and multiuser communications," *IEEE Trans. Wireless Communications*, vol. 3, no. 4, pp. 1305–1316, 2004.

[13] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, Belmont, Mass, USA, 1999.

[14] H. S. Malvar, *Optimal pre- and post-filtering in noisy sampled-data systems*, Ph.D. thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Mass, USA, September 1986.

[15] M. L. Honig, P. Crespo, and K. Steiglitz, "Suppression of near- and far-end crosstalk by linear pre- and post-filtering," *IEEE Journal on Selected Areas in Communications*, vol. 10, no. 3, pp. 614–629, 1992.

[16] J. Salz, "Digital transmission over cross-coupled linear channels," *AT & T Technical Journal*, vol. 64, no. 6, pp. 1147–1159, 1985.

[17] J. Yang and S. Roy, "On joint transmitter and receiver optimization for multiple-input-multiple-output (MIMO) transmission systems," *IEEE Trans. Communications*, vol. 42, no. 12, pp. 3221–3231, 1994.

[18] J. Yang and S. Roy, "Joint transmitter-receiver optimization for multi-input-multi-output systems with decision feedback," *IEEE Transactions on Information Theory*, vol. 40, no. 5, pp. 1334–1347, 1994.

[19] D. P. Palomar and M. A. Lagunas, "Joint transmit-receive space-time equalization in spatially correlated MIMO channels: a beamforming approach," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 730–743, 2003.

[20] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Communications*, vol. 49, no. 12, pp. 2198–2206, 2001.

[21] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Signal Processing*, vol. 50, no. 5, pp. 1051–1064, 2002.

[22] R. F. H. Fischer, *Precoding and Signal Shaping for Digital Transmission*, John Wiley & Sons, New York, NY, USA, 2002.

[23] V. Erceg, K. V. S. Hari, M. S. Smith, and D. S. Baum, "Channel models for fixed wireless applications," IEEE 802.16 Broadband Wireless Access Working Group, contribution IEEE 802.16a-03/01, June 2003.

[24] M. Athans and F. C. Schweppe, "Gradient matrices and matrix calculations," Tech. note 1965-53, Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Mass, USA, November 1965.

[25] R. Johannesson and K. S. Zigangirov, *Fundamentals of Convolutional Coding*, IEEE Press, New York, NY, USA, 1999.

[26] F. Jelinek, *Statistical Methods for Speech Recognition*, MIT Press, Cambridge, Mass, USA, 1998.

[27] S. Lin and D. J. Costello Jr, *Error Control Coding*, Prentice-Hall, Upper Saddle River, NJ, USA, 2nd edition, 2004.

[28] D. P. Palomar, M. A. Lagunas, and J. M. Cioffi, "Optimum linear joint transmit-receive processing for MIMO channels with QoS constraints," *IEEE Trans. Signal Processing*, vol. 52, no. 5, pp. 1179–1197, 2004.

[29] B. A. Bjerke and J. G. Proakis, "Equalization and decoding for multiple-input multiple-output wireless channels," *EURASIP Journal on Applied Signal Processing*, vol. 2002, no. 3, pp. 249–266, 2002.

Yongfang Guo received the B.S. degree in electrical engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 1999, and the M.S. and Ph.D. degrees in electrical engineering from the University of California at Davis, Davis, Calif, in 2003 and 2004, respectively. She is currently a system engineer with Intel Corp., Santa Clara, Calif. Her general research interests include communications, signal processing, and networking. Her current research focuses on precoding and equalization for MIMO communication systems.

Bernard C. Levy received the Diplôme d'Ingénieur Civil des Mines from the Ecole Nationale Supérieure des Mines, Paris, France, in 1974, and the Ph.D. degree in electrical engineering from Stanford University in 1979. From July 1979 to June 1987, he was an Assistant and then Associate Professor in the Department of Electrical Engineering and Computer Science, MIT. Since July 1987, he has been with the University of California at Davis (UC Davis), where he is a Professor of electrical engineering and a member of the Graduate Group in Applied Mathematics. He served as a Chair of the Department of Electrical and Computer Engineering, UC Davis, from 1996 to 2000. He was a Visiting Scientist at the Institut de Recherche en Informatique et Systèmes Aléatoires (IRISA), Rennes, France, from January to July 1993, and at the Institut National de Recherche en Informatique et Automatique (INRIA), Rocquencourt, France, from September to December 2001. His research interests are in statistical signal processing, estimation, detection, multidimensional signal processing, and inverse problems. He currently serves as an Associate Editor of the IEEE Transactions on Circuits and Systems II and served as a Member of the Image and Multidimensional Signal Processing Technical Committee, IEEE Signal Processing Society, from 1992 to 1998. He is a Fellow of the IEEE, and a Member of SIAM and the Acoustical Society of America.