Precoding and Signal Shaping for Transmission over MIMO Channels

Robert F.H. Fischer, Clemens Stierstorfer, Christoph Windpassinger
Lehrstuhl für Informationsübertragung, Friedrich-Alexander Universität Erlangen-Nürnberg,
Cauerstraße 7/LIT, 91058 Erlangen, Germany, Email: {fischer, clemens, windpass}@LIT.de

Abstract — Combined precoding and signal shaping for flat and intersymbol interference multiple-input/multiple-output channels is presented. Instead of equalizing the interference at the receiver side, non-linear preequalization together with a shaping algorithm is performed. Thereby, near-Gaussian distributed channel symbols are generated and a reduction in average transmit power compared to uniform signaling is achieved. Various approaches are discussed and compared. It turns out that shaping without scrambling, initially developed for linearly distorting channels, is also an attractive scheme for spatial/temporal precoded/shaped transmission over MIMO channels.

I. INTRODUCTION

Over the last years the interest in high-rate wireless transmission has significantly increased. It has been shown, e.g., [7, 13], that using systems with multiple transmit and receive antennas, the spectral efficiency and hence the supported data rate can be raised dramatically. The transmission has significantly increased. It has been shown, e.g., that using systems with multiple transmit and receive antennas, the spectral efficiency and hence the supported data rate can be raised dramatically. The main difficulty when transmitting over so-called multiple-input/multiple-output (MIMO) channels is the separation or equalization of the data streams transmitted in parallel.

Since spatial equalization in MIMO systems is tightly related to temporal equalization for single-input/single-output (SISO) transmission over intersymbol-interference (ISI) channels, each well known equalization strategy from the latter field can be directly applied to MIMO channels. Of particular interest is Tomlinson-Harashima precoding, which can be derived from decision-feedback equalization (DFE) [7, 14] by transferring its feedback part to the transmitter. The application of Tomlinson-Harashima precoding to MIMO channels, briefly “MIMO precoding”, and related techniques have been proposed in [4, 2, 8, 11].

In this paper we further want to enhance performance by combining MIMO precoding with signal shaping, a topic not yet present in literature. Using again the correspondences between schemes for SISO ISI channels and (flat) MIMO channels, we discuss the application of coding/shaping algorithms to the current setting. Various variants are analyzed and compared, especially with respect to the achievable shaping gain (reduction in average transmit power) and respective error rates. We show that shaping without scrambling [5] is an attractive scheme for combined spatial/temporal precoding/shaping over flat or ISI MIMO channels.

In Section II the MIMO channel model is introduced and MIMO precoding is reviewed. Various approaches for combined precoding and shaping are discussed in Section III. Section IV presents numerical results illustrating the advantages and disadvantage of the schemes presented. A brief summary is given in Section V.

II. CHANNEL MODEL AND MIMO PRECODING

For the moment, we restrict the discussion to flat fading MIMO channels and an equal number of $K$ transmit and receive antennas; the extension to intersymbol interference (ISI) MIMO channels is given later. Assuming flat fading channels, since no temporal interferences occur, each time interval can be processed on its own. Hence, the input/output relationship for each time instant simply reads

$$y = Hx + n,$$

where $x = [x_1, \ldots, x_K]^T$ is the vector of the (complex) transmit symbols, each with variance $E\{|x|^2\} = \sigma_x^2$. The $K \times K$ matrix $H = [h_{kl}]$ comprises the complex gain factors $h_{kl}$ between transmit antenna $l$ and receive antenna $k$. The $K$-dimensional vector $n$ denotes additive white zero-mean complex noise with $E\{|n|^2\} = \sigma_n^2 I$, and $y = [y_1, \ldots, y_K]^T$ is the complex-valued vector of receive samples.

The data symbols transmitted in parallel over the antennas interfere at the receiver side. These interferences caused by the MIMO channel have to be equalized, i.e., an end-to-end transfer matrix $I$ is desired — $K$ parallel, independent channels should be generated. Instead of performing non-linear equalization at the receiver side as done in the BLAST system, or SVD based equalization, non-linear preequalization at the transmitter, in particular a MIMO version of Tomlinson-Harashima precoding, is considered in this paper.

![Fig. 1: MIMO precoding (the slicer takes the modulo congruence into account).](image-url)

The transmission scheme applying MIMO precoding is depicted in Fig. 1. It consists of a feedforward filter $F$ and a scaling (gain) matrix $G$ at the receiver and a feedback structure employing the matrix $B$ at the transmitter side, by which the transmit symbols $x_k$ are calculated for the data carrying symbols $a_k$ drawn from a (finite) signal constellation.

Given the channel matrix $H$, the optimal setting of the matrices $F$, $G$, and $B$ at the transmitter can be given, e.g., [8, 6]. For adjusting them according to the zero-forcing (ZF) criterion a QR-type factorization of the

$^1$Notation: $A^\dagger$ denotes the transpose of matrix $A$, $A^H$ the Hermitian (i.e., conjugate) transpose and $A^{-1}$ the inverse of the Hermitian transpose (for square matrix $A$). $I$ is the identity matrix.
channel matrix $H$ is performed, i.e.

$$H = F^H S ,$$

where $F$ is the unitary feedforward matrix and $S = [s_{ij}]$ a lower left triangular matrix.

For convenience we define a diagonal scaling matrix $G = \text{diag}(s_1, \ldots, s_K)$, such that $B = [b_k] \equiv GS = GFH$ is a unit-diagonal lower triangular matrix. Noteworthy, since $F$ is a unitary matrix, we have $H^H H = S^H F^H F S = S^H S$. Hence, $S$ can be obtained by a modified Cholesky factorization of $H^H H$.

The operation of MIMO precoding for $M$-ary square constellations \(^1\)

$$A = \{ a_1 + j a_Q \mid a_1, a_Q \in \{ \pm 1, \pm 3, \ldots, \pm (\sqrt{M} - 1) \} \} ,$$

$a_k \in A$ (the generalization to other signal sets is easily possible, see [6]), is as follows: Given the data symbols $a_k$ (combined into the vector $a = [a_1, \ldots, a_K]^T$), the channel symbols are successively calculated as $[b_k]$.

$$x_k = \text{MOD} \{ a_k - f_k ,\} , \quad f_k = \sum_{i=1}^{k-1} b_{i+k} ,$$

(3) The preceeding symbols $d_k \in \{ 2\sqrt{M} \cdot (d_1 + j d_Q) \mid d_1, d_Q \in Z \}$ are chosen implicitly by a simple modulo reduction of real and imaginary part of the channel symbols $x_k$ into the (half-open) interval $[-\sqrt{M}, \sqrt{M})$.

Hence, instead of feeding the data symbols $a_k$ into the linear predistortion $B^{-1}$ which is implemented by the feedback structure, effective data symbols $v_k \equiv a_k + d_k$ are used—the initial signal constellation is extended periodically. All points, spaced in real and imaginary part by integer multiples of $2\sqrt{M}$ are congruent and represent the same data. From these equivalent points, symbol-by-symbol the unique point is selected for transmission which results in a channel symbol with minimum magnitude of real and imaginary part [6].

Since the linear predistortion $B^{-1}$ equalizes the cascade $GFH$, after prefiltering ($F$) and scaling ($G$) at the receiver side, the effective data symbols $v_k$ corrupted by additive noise (error) are visible: $r = v + n$. Using a slice which takes the modulo congruence into account an estimate for the data symbols can be generated. In summary, disregarding the modulo congruence, precoding transforms the MIMO channel into $K$ parallel, independent unit-gain AWGN channels with individual noise variance $\sigma^2_n, k = \sigma^2_n/s_{kk} ^2$.

Up to now, only flat fading channels have been considered, but MIMO precoding can be used straightforwardly for ISI MIMO channels, i.e., to perform joint spatial and temporal equalization. Assuming that the channel coefficients are constant over the transmission burst, the elements of the channel matrix are now impulse responses, rather than constant gain factors. Given the impulse responses $\{h_k[\mu]\} = \{h_k[0], h_k[1], \ldots\}$ from each transmit to each receive antenna and combining the $\mu$-th tap weights into the matrix $H[\mu] = [h_k[\mu]\mu = 0, 1, \ldots$, the received signal in time interval $\mu$ reads

$$y[\mu] = \sum_{\mu=0}^{\infty} H[\mu] z^{-\mu} = \sum_{\mu=0}^{\infty} H[\mu] z^{-\mu} .$$

(5) For calculating the optimum feedforward and feedback matrices, which—like the channel itself—are matrices of impulse responses, we define the $z$ transform of the channel matrix as

$$H(z) = \sum_{\mu=0}^{\infty} h_k[\mu] z^{-\mu} = \sum_{\mu=0}^{\infty} H[\mu] z^{-\mu} .$$

(6) Then, the required matrix filters can be calculated by spectral factorization, see, e.g. [4, 1], $z^{-1} \equiv (z^*)^{-1}$.

$$H^H(z) = S^H(z) $$

(7) which replaces the Cholesky factorization in the flat fading case. Here, the matrix filter $S(z) = \sum_{\mu=0}^{\infty} S[\mu] z^{-\mu}$ is strictly causal, i.e., $S[\mu] = 0, \mu < 0$, and the matrix $S[0] = [s_{k0}]$ is lower left triangular. From this, the feedback matrix is given as $F(z) = S^H(z) H^H(z)$, the scaling matrix is chosen as $G = \text{diag}(s_1^2, \ldots, s_K^2)$, and the feedback matrix reads $B(z) = GS(z)$.

III. Signal Shaping in Space-Time Transmission

The aim of signal shaping is to generate signals which exhibit some desired properties. The most important shaping aim is to reduce average transmit power. This is possible by replacing a uniformly distributed signal by a Gaussian one [1]. Subsequently, shaping strategies for MIMO channels are discussed and compared.

A FLAT FADEING MIMO CHANNEL

First, we study shaping for the flat fading MIMO channel. Regarding shaping, two different basic approaches have to be distinguished. Fig. 2 shows these two strategies in the space-time plane.

A.1 Shaping in Time Direction As in the case of SISO transmission, a natural choice is to perform signal shaping along the time axis (left part of Fig. 2). For the flat fading MIMO channel, the induced parallel channels are intersymbol interference free. Hence, shaping techniques developed for the AWGN channel can be applied. However, a small modification is necessary: for channels $k = 2, 3, \ldots, K$ the presubtraction of the interference term has to be taken into account.

Applying trellis shaping [10], the operation of the combined precoding/shaping scheme is as follows. Given the binary data to be transmitted, for the first data stream ($k = 1$) conventional trellis shaping generates a block of channel symbols $x_1[n]$, taken from an expanded signal set. Assuming one redundant shaping bit, the expanded constellation contains $2M$ signal points, and, e.g., a rate-1/2 shaping convolutional code may be used [10, 6]. The interference of the symbols transmitted over the first antenna into the second receive antenna is calculated via the feedback filter $B$ in the transmitter: $f_2[\mu] = b_2 x_1[\mu]$.

Trellis shaping for the second data stream also operates as usual, but after mapping of binary data to signals points, their squared magnitude does not directly give

\(^2\)Note, as in the DFE scheme used in the BLAST scheme, the ordering of the parallel data streams is essential. This corresponds to a relabeling of the transmit antennas and a permutation of the columns of the channel matrix. We always assume that the channel matrix $H$ is already permuted suitably.

\(^3\)Throughout the paper we assume operation in the high SNR region, where all parallel data streams use the same signal constellation and the same average transmit power, since waterfilling there tends to a uniform power distribution. Moreover, in this region the ZF approach is (almost) optimal and no gains by using the MMSE criterion can be obtained.
the metric for the shaping Viterbi algorithm but the interference \( f_k[\mu] \) is first presubtracted and the resulting signal is modulo reduced into the region of the expanded signal constellation. Note that this procedure resembles trellis precoding \([3]\), however, here the interference term is predetermined and not dependent on prior choices in the current shaping procedure. Hence, for trellis shaping a conventional Viterbi algorithm is sufficient and no per-survivor processing as in trellis precoding is required.

Using the described procedure—calculation of the interference \( f_k[\mu] = \sum_{l=1}^{K} b_{kl} x_l[\mu]; \) Trellis shaping with presubtraction of the interference and modulo reduction—the transmit symbols \( x_k[\mu] \) are successively generated off-line for all transmit antennas and then transmitted in succession.\(^4\)

Other techniques than trellis shaping are possible, too. In particular, Voronoi shaping \([10, 6]\) is an attractive alternative. Here, shaping is not done by a path search algorithm but is based on a high-dimensional lattice. The modulo device in MIMO precoding, operating on two-dimensional channel symbols is replaced by a modulo reduction into the Voronoi region of a 2\(N\)-dimensional lattice, treating blocks of \( N \) consecutive symbols. It has been shown \([2, 16]\) that using a lattice whose normalized second moment approaches that of a hypersphere, \(1/(2\pi e)\), the capacity of an AWGN channel without interference can be approached.

The main disadvantages of the above shaping strategies are the large delay introduced and error propagation at the receiver. In order to recover data, in trellis shaping an inverse syndrome former \([10]\), i.e., a linear dispersive device. For shaping, an algorithm determines the effect of the transmit or receive signal can propagate through this filter and take effect multiple times \([6]\). This in turn leads to a loss and the gross shaping gain (reduction in average transmit power) is reduced to the net shaping gain. Especially on flat Rayleigh fading channels this is highly undesirable, since doubling of the symbol error rate would lead to a loss of about 3 dB. The same is true for Voronoi shaping, where one erroneous symbol may cause errors in the entire decoded block. Fortunately, for all strategies, due to the preequalization, errors propagate only within the respective data stream and do not cause errors in the other parallel channels.

A.2 Shaping in Space Direction The above problems are overcome by performing shaping for each time step separately along the antennas, see right part of Fig. 2. In particular, since processing is limited to one time interval, no additional delay is introduced. As in this setting the choice of transmit symbol \( x_l \) influences the selection of subsequent symbols \( k > l \), combined shaping/precoding algorithms for ISI channels can be applied.

One possible candidate is trellis precoding, as proposed in \([15]\) for the broadcast channel; in the same way trellis precoding is applicable for the considered multi-antenna scenario. However, like in trellis shaping, a dispersive system is required to recover data and hence the shaping gain is lost in parts.

A more attractive candidate is a combined precoding/shaping scheme called shaping without scrambling \([5, 6]\), shown in Fig. 3. In MIMO precoding, the precoding symbols \( d_k \) and hence the effective data symbols \( v_k \) are (implicitly) selected symbol-by-symbol by the modulo device. For shaping, an algorithm determines the effective data symbols in the long run. Using this approach, almost any property of the transmit or receive signal can be influenced, cf. \([9]\).

Due to the flat fading channels, in each time interval the precoder starts from the all-zero state. Hence, the number of possible precoding sequences is finite and an exhaustive search—in contrast to the sub-optimum one used in \([5]\)—might be possible. For a large number of transmit antennas reduced complexity algorithms are preferable—in particular the M-algorithm, a member of the family of sequential decoding algorithms. Starting from the all-zero state, a fixed number of best precoding sequences \( (d_1, d_2, \ldots, d_K) \) are examined. In \([5, 6]\) it is shown that it is sufficient to regard only two alternatives in each shaping step. Hence, a simple binary tree has to be decoded. Since the last symbol \( x_K = a_K + d_K - f_K \) does not influence further selections, the precoding symbols are chosen such that \( x_K \) has minimum magnitude. In summary, the decoding tree has \( 2^{K-1} \) branches, which allows a full search for up to \( K = 8 \) or even more antennas.

The drawback of this shaping strategy is that the achievable shaping gain is limited to a much lower value than the ultimate shaping gain of 1.53 dB \([10, 6]\). The upper bound on this gain is given by the power saving of a 2\(K\)-dimensional hypersphere over a 2\(K\)-dimensional hypercube. For \( K = 4 \) this maximum gain evaluates to approximately 0.73 dB, for \( K = 16 \) to 1.17 dB, and approaches 1.53 dB as \( K \to \infty \). However, for the proposed scheme no error multiplication occurs at the receiver, as a simple modulo device eliminates the precoding sequence, and the entire shaping gain can be utilized.

\(^4\) Instead of performing shaping on blocks of symbols, trellis shaping can operate continuously. After some decoding delay, the channel symbols are known and the algorithm operating on the next data stream can start.
B  ISI MIMO Channel

As for ISI MIMO channels combined spatial/temporal equalization has to be performed, shaping also needs to operate over space and time. Since the choice of one channel symbol $x_k[n]$ has influence on the symbol $x_l[n]$, $k > l$, in the same time step, and on all symbols $x_{k'}[n']$, $n' > n$, of subsequent time steps, precoding/shaping schemes for the ISI channel may be transferred to the scenario at hand.

Again, trellis precoding is a possible choice. However, as in trellis shaping, error multiplication occurs in data recovery and part of the shaping gain is lost. Due to time diversity introduced by the ISI channel, the error rate curve will have a steeper slope than that for the flat fading MIMO channel. In turn, the loss will be somewhat smaller but still not negligible.

Shaping without scrambling, i.e., precoding/shaping solely based on the selection of the additive precoding sequence, is also preferable for ISI MIMO channels. Thereby, due to the lower triangular structure of $H[0]$, i.e., spatial causality, in each time step the symbols are processed consecutively from index 1 to $K$, then the next time instant is processed. That is, shaping operates in a zig-zag (serialized) fashion over space and time. In this way, each symbol has influence on subsequent symbols but not the other way round. However, the generation of the interference depends on the current antenna—in the serialized description, a periodically varying “impulse response” is present in the precoding/shaping algorithm. Except of this fact, the same situation as for shaping on ISI SISO channels is present, and shaping without scrambling can straightforwardly be applied. In particular, the shaping Viterbi algorithm can preferably operate on the trellis defined by an imaginary scrambler (see [5]).

At the receiver side, a simple modulo device eliminates the influence of the precoding symbols $d_k[n]$. Since decision is done symbol-by-symbol and no dispersive system is present at the receiver, the reduction in average transmit power is entirely utilizable.

IV. Numerical Examples

In order to assess the performance and achievable gains of signal shaping for MIMO channels, numerical examples are presented. Unless otherwise stated, we assume $K = 4$ transmit and receive antennas. Each of the 4 parallel data streams uses uncoded 16-QAM signaling; they are processed in the optimum ordering [12]. For shaping, the constellation is expanded to a 32-ary square constellation, rotated by $45^\circ$, cf. [3, 6].

A  Flat Fading MIMO Channel

We now study shaping for flat fading MIMO channels. The $K \times K$ channel matrix with i.i.d. complex Gaussian distributed unit variance entries is randomly selected. The results are averaged over a large number of channel realizations (bursty transmission over block-wise constant channel conditions).

First, shaping in temporal direction is considered. The symbol error rates of trellis shaping applied to each of the 4 parallel data streams as described above are given in Fig. 4 over the ratio of average received energy per information bit $E_b$ and one-sided noise power spectral density $N_0$. The 4-state rate-1/2 shaping code given in [10] and a Viterbi shaping decoder with path register length of 16 symbols are applied.

Using shaping a power reduction of about 0.7 dB is achieved in the present situation. However, error multiplication occurs at the inverse syndrome former at the receiver side leading to an increased error rate. In total, a (small) loss compared to uniform signaling occurs when applying this version of shaping. Hence, especially for fading channels, error propagation at the receiver has to be avoided.

Next, Fig. 5 shows the achievable shaping gain (i.e., reduction in average transmit power compared to uniform signaling) over the number of antennas when shaping is performed in spatial direction only. Shaping without scrambling [5] using the M-algorithm with different numbers of active paths is applied. For comparison, the maximum achievable gain in $2K$ dimensions is included.

It is visible that only a moderate number of paths has to be regarded in the shaping algorithm. The achievable gain clearly follows the theoretical curve and ranges from 0.25 dB for 4 antennas to more than 1 dB for large numbers of antennas. It can be concluded that shaping in spatial direction is rewarding for 8 or more antennas. Noteworthy, since there is no error multiplication at the receiver, the shaping gain immediately gives the gain in power efficiency.
B  ISI MIMO Channel

We now turn to ISI MIMO channels. Since we are not aware of commonly accepted realistic models for ISI MIMO channels, some kind of test channel is used. The impulse responses have length $3$; $H[0]$ is required to be a lower left triangular matrix with unit diagonal, where all non-zero entries and all elements of $H[1]$ and $H[2]$ are independent and complex Gaussian distributed. The coefficients of matrix $H[0]$ have unit variance; those of $H[1]$ and $H[2]$ have variance $1/2$ and $1/3$, respectively (harmonically decreasing power-delay profile).

In spatial/temporal combined precoding/shaping for ISI MIMO channels, the gain is no longer dependent or limited by the number of antennas. Numerical simulations typically show gains of $0.8$ dB (short decoding delays) to more than $1$ dB (decoding delays of approx. $100$ symbols). This gain lies in the region of that achievable with the same algorithm over ISI SISO channels.

The bit error rates over the signal-to-noise ratio ratio $E_b/N_0$ of different precoding/shaping strategies are compared in Fig. 6. The three curves correspond to MIMO precoding, which is the reference ($0$ dB shaping gain), trellis precoding, applied similar as proposed in [15] for the broadcast channel, and spatial/temporal shaping without scrambling. Both shaping techniques use a Viterbi algorithm operating on a $16$-state trellis, hence the complexity is the same. Moreover, the decoding delays of trellis precoding ($100$ symbols) and shaping without scrambling ($12$ symbols) are chosen such that both strategies achieve the same shaping gain of about $0.8$ dB.

![Fig. 6: Bit error rates over $E_b/N_0$ in dB. ×: MIMO precoding; +: Trellis precoding; ≈: Shaping without scrambling.](image)

However, in trellis precoding, error multiplication occurs at the receiver. Consequently, part of the shaping gain is lost. For error rates above $10^{-7}$ even a loss compared to MIMO precoding occurs. In contrast, the gain of shaping without scrambling is usable in total—the error rate curve of MIMO precoding is simply shifted to the left by the shaping gain.

V. SUMMARY

In this paper we have discussed the application of combined precoding and signal shaping to MIMO channels. It turns out that shaping without scrambling, where shaping is solely based on the selection of the precoding sequence, and which has been initially proposed for shaping/precoding on ISI channels, is also an attractive scheme for MIMO channels. The receiver side is identical to that of MIMO Tomlinson-Harashima precoding: A simple modulo device eliminates the influence of the precoding sequence. In turn, contrary to other shaping techniques such as trellis shaping or trellis precoding, the entire shaping gain can be used. Moreover, using the appropriate metric in the shaping algorithm, besides a reduction of average transmit power almost any signal property of transmit or receive signals can be generated.

We have concentrated the discussion to the multi-antenna scenario. However, the results are in the same way valid for the broadcast channel, i.e., a downlink scenario from a central transmitter to distributed receivers. In this case, MIMO precoding or combined precoding/shaping, where the feedforward matrix $F$ is present at the transmitter, too, are the dual to multuser detection in uplink transmission.

REFERENCES