Tomlinson-Harashima Precoding in Space-Time Transmission for Low-Rate Backward Channel

Robert F.H. Fischer, Christoph Windpassinger, Alexander Lampe, and Johannes B. Huber

Lehrstuhl für Nachrichtentechnik II, Friedrich–Alexander Universität Erlangen–Nürnberg,
Cauerstraße 7/NT, 91058 Erlangen, Germany, Email: {fischer,windpass,alampe,huber}@LNT.de

Abstract— **In this paper, Tomlinson-Harashima precoding, a nonlinear pre-equalization technique, is proposed for transmission over multiple-input/multiple-output channels. Instead of equalizing intersymbol interference (temporal equalization) here spatial equalization, i.e., equalization of multi-user interference is performed. If only a low-rate backward channel is available for communicating channel state information back from the receiver to the transmitter, a compromise precoder setting, calculated from (medium-term) average channel knowledge in combination with linear residual equalization at the receiver side is proposed. Compared to an optimal adjustment of the precoder, i.e., perfect channel state information at the transmitter, only small losses have to be accepted.**

Keywords— **BLAST, decision-feedback equalization, MIMO channel, space-time coding, Tomlinson-Harashima precoding**

I. Introduction

THE demand for high-rate wireless transmission significantly increased over the last years and stimulated the interest in so-called multiple-input/multiple-output (MIMO) systems. In general, a MIMO transmission scheme can be described by the basic relation $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}$. Here, $\boldsymbol{x}$ designates the transmit vector which comprises the transmit symbols of N_T parallel data streams. These streams can be due to a parallel (i.e., layered) encoding of a high-rate data signal, or they may belong to different and independent users. The vectors $\boldsymbol{y}$ and $\boldsymbol{n}$ of dimension N_R denote the vector of received symbols, and the vector of disturbances, respectively. The MIMO channel is characterized by its $N_\mathrm{R} \times N_\mathrm{T}$ channel matrix $\boldsymbol{H}$.

The main difficulty for transmission over MIMO channels is the separation or equalization of the parallel data streams, i.e., the recovery of the components of the transmitted vector $\boldsymbol{x}$ which interfere at the receiver side. The most obvious strategy for separating the data streams is linear (zero-forcing) equalization at the receiver side. Here, the decision vector is $\boldsymbol{r} = \boldsymbol{H}_{(l)}^{-1}\boldsymbol{y}$, where $\boldsymbol{H}_{(l)}^{-1}$ denotes the left (pseudo) inverse of the channel matrix $\boldsymbol{H}$. It is well-known that linear equalization suffers from noise enhancement and hence has a poor power efficiency. This disadvantage can be overcome by spatial decision-feedback equalization (DFE) [7], [16]. Unfortunately, error propagation may occur in DFE. Moreover, since immediate decisions are required, the application of channel coding requires some clever interleaving which in turn introduces significant delay.

If channel state information (CSI) is (partly) also available at the transmitter, the users can be separated by means of pre-equalization. Assuming perfect CSI at the transmitter, interference of the users at the receiver side can be completely avoided by a multiplication of the data vector $\boldsymbol{a}$ with the (pseudo) right inverse $\boldsymbol{H}_{(r)}^{-1}$ of the channel matrix $\boldsymbol{H}$ at the transmitter [1]. Instead of transmitting the data symbols $\boldsymbol{a}$ directly over the channel ($\boldsymbol{x} = \boldsymbol{a}$), the pre-distorted version $\boldsymbol{x} = \boldsymbol{H}_{(r)}^{-1}\boldsymbol{a}$ is fed into the channel. However, this (zero-forcing) linear pre-equalization suffers from the same loss in power efficiency as linear equalization at the receiver side. Here, instead of enhancing the noise, average transmit power is increased by the same factor.

Finally, the task of channel equalization can be split among transmitter and receiver. A popular strategy is based on the singular value decomposition (SVD) of the channel matrix: $\boldsymbol{H} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}$. Applying $\boldsymbol{V}^\mathsf{H}$ at the transmitter and $\boldsymbol{U}^\mathsf{H}$ at the receiver independent, parallel channels are generated [14], [13]. Here, in contrast to linear (pre-)equalization neither transmit power is increased, nor channel noise is enhanced.

Spatial equalization in MIMO systems (channel matrix $\boldsymbol{H}$) is tightly related to temporal equalization for single-input/single-output (SISO) transmission over intersymbol-interference (ISI) channels (channel transfer function $H(z)$). Each equalization strategy has its direct counterpart in the other domain. The analogies are depicted in Table I.

The correspondences for linear equalization at the receiver and linear pre-equalization at the transmitter are self-evident. Singular value decomposition for MIMO channels can be identified with orthogonal frequency-division multiplexing (OFDM) or discrete multitone (DMT) transmission over ISI channels. To be precise, SVD corresponds to a strategy called *vector coding*, e.g. [10], where blocks of consecutive symbols are processed at transmitter and receiver based on the eigenvectors of some channel matrix. In both cases, a partitioning of the underlying channel into parallel independent sub-channels is performed.

Decision-feedback equalization is a non-linear equalization strategy at the receiver side. Its counterpart for MIMO channels is a matrix DFE for spatial equalization. In SISO transmission, the feedback part of the DFE can be transferred to the transmitter, leading to a scheme known as *Tomlinson-Harashima precoding (THP)*. It is well known that, neglecting a very small increase in average transmit power, the performance of (ideal feedback) DFE and THP is the same. But only for THP, which is a transmitter technique, error propagation at the receiver is avoided. Moreover, channel coding schemes can be applied in the same way as for the ideal additive white Gaussian noise (AWGN) or flat fading channel.

Finally, it has been shown (e.g. [5]) that the performance of DFE, which is a non-linear but time-invariant technique, is identical to that of multicarrier transmission (OFDM/DMT), which is a linear but time-variant strategy. Hence, in practice, the actual choice of transmission scheme depends on many points, such as complexity of implementation, inherent delay, or peak-to-average power ratio of the transmit signal.

TABLE I

CORRESPONDING EQUALIZATION STRATEGIES FOR ISI CHANNELS AND MIMO CHANNELS.

		ISI channel $H(z)$ (temporal equalization)	MIMO channel $\boldsymbol{H}$ (spatial equalization)
linear	at receiver	linear equalization via $1/H(z)$	linear equalization via $\boldsymbol{H}_{(l)}^{-1}$
	at transmitter	linear pre-equalization via $1/H(z)$	linear pre-equalization via $\boldsymbol{H}_{(r)}^{-1}$
	at transmitter/receiver	OFDM/DMT, vector coding	SVD
non-linear	at receiver	DFE	matrix DFE
	at transmitter/receiver	Tomlinson-Harashima precoding	**precoding for MIMO channels**

The aim of this paper is to present a non-linear precoding scheme for MIMO channels. Like in SISO transmission, it is the counterpart to matrix DFE and reflects the same duality to SVD as THP to multicarrier transmission. Moreover, we show that an average knowledge of the channel, i.e., its statistics, is sufficient at the precoder. Since—as explained above—spatial equalization in MIMO transmission has an analog in equalization of ISI in SISO transmission (temporal equalization), here the system concept proposed in [3] is adopted.

The paper is organized as follows: In Section II the channel model and the fundamentals of MIMO precoding for perfect channel knowledge at the transmitter are given. A compromise precoder setting and optimal linear residual equalization in case of mismatch of precoder and actual channel are derived in Section III. Numerical results illustrating the robustness of fixed precoding against channel mismatch are presented (Section IV). A summary and the conclusions are given in Section V.

II. CHANNEL MODEL AND TOMLINSON-HARASHIMA PRECODING

We consider the general model of transmission over a MIMO channel being described by (we restrict ourselves to $N_\mathrm{R} = N_\mathrm{T} = K$; the generalization is straightforward)

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n} \ , \tag{1}$$

where[1] $\boldsymbol{x} = [x_1, \ldots, x_K]^\mathsf{T}$ is the vector of the (complex) transmit symbols, each with variance $\mathrm{E}\{|x_i|^2\} = \sigma_x^2$. The transmit symbols x_k are calculated in some way for the data carrying symbols a_k drawn from a (finite) signal constellation. The $K \times K$ matrix $\boldsymbol{H}$ comprises the complex gain factors (flat fading channel) between the K transmit antennas and the K receive antennas. For simplicity, we restrict ourselves to square matrices $\boldsymbol{H}$, i.e., the same number of transmit and receive antennas is assumed. The K-dimensional vector $\boldsymbol{n}$ denotes additive white zero-mean complex noise with $\mathrm{E}\left\{\boldsymbol{n}\boldsymbol{n}^\mathsf{H}\right\} = \sigma_n^2 \boldsymbol{I}$, and $\boldsymbol{y} = [y_1, \ldots, y_K]^\mathsf{T}$ is the complex-valued vector of receive samples.

[1]Notation: $\boldsymbol{A}^\mathsf{T}$ is the transpose of matrix $\boldsymbol{A}$, $\boldsymbol{A}^\mathsf{H}$ the Hermitian (i.e., conjugate) transpose and $\boldsymbol{A}^{-\mathsf{H}}$ the inverse of the Hermitian transpose (for square matrix $\boldsymbol{A}$). $\boldsymbol{I}$ is the identity matrix.

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 $\boldsymbol{I}$ is desired—K parallel, independent channels should be generated. The decision vector $\boldsymbol{r} = \boldsymbol{a} + \boldsymbol{e}$ is then the initial data vector $\boldsymbol{a}$ plus some distortion $\boldsymbol{e}$, accounting for the channel noise and residual multi-"user" interference. Instead of performing non-linear equalization at the receiver side, non-linear pre-equalization at the transmitter, in particular a MIMO version of Tomlinson-Harashima precoding, is used in this paper. The transmission scheme applying MIMO precoding is depicted in Figure 1.

Assuming for the moment $\boldsymbol{H}$ to be known perfectly at receiver *and* transmitter, the optimal setting of feedforward matrix $\boldsymbol{F}$ at the receiver and feedback matrix $\boldsymbol{B}$ at the transmitter can be given. For adjusting them according to the minimum mean-squared error (MMSE) criterion

$$\mathrm{E}\{\boldsymbol{e}^\mathsf{H}\boldsymbol{e}\} \to \min \tag{2}$$

we have to perform a *Cholesky factorization* [8]

$$\boldsymbol{H}^\mathsf{H}\boldsymbol{H} + \zeta\boldsymbol{I} = \boldsymbol{S}^\mathsf{H}\boldsymbol{S} \ , \tag{3}$$

where $\boldsymbol{S} = [s_{kl}]$ is a lower left triangular matrix, i.e., $s_{kl} = 0$ for $k < l$. As usual for MMSE solutions, the parameter $\zeta \stackrel{\mathrm{def}}{=} \sigma_n^2/\sigma_x^2$ takes the signal-to-noise ratio into account. Equalization according to the zero-forcing criterion is obtained for $\zeta = 0$.

For the matrix $\boldsymbol{S}$, we define the scaling diagonal matrix as $\boldsymbol{V} = \mathbf{diag}(1/s_{11}, \ldots, 1/s_{KK})$, and a lower triangular matrix $\boldsymbol{C} = \boldsymbol{V}\boldsymbol{S}$ with unit diagonal. Then, the feedforward and feedback matrices are given as [8]

$$\boldsymbol{F} = \boldsymbol{S}^{-\mathsf{H}}\boldsymbol{H}^\mathsf{H} \ , \qquad \boldsymbol{B} = \boldsymbol{C} - \boldsymbol{I} \ . \tag{4}$$

The operation of MIMO precoding for M-ary square constellations

$$\mathcal{A} = \left\{ a_\mathrm{I} + \mathrm{j}a_\mathrm{Q} \mid a_\mathrm{I}, a_\mathrm{Q} \in \{\pm 1, \pm 3, \ldots, \pm(\sqrt{M} - 1)\} \right\} \ , \tag{5}$$

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

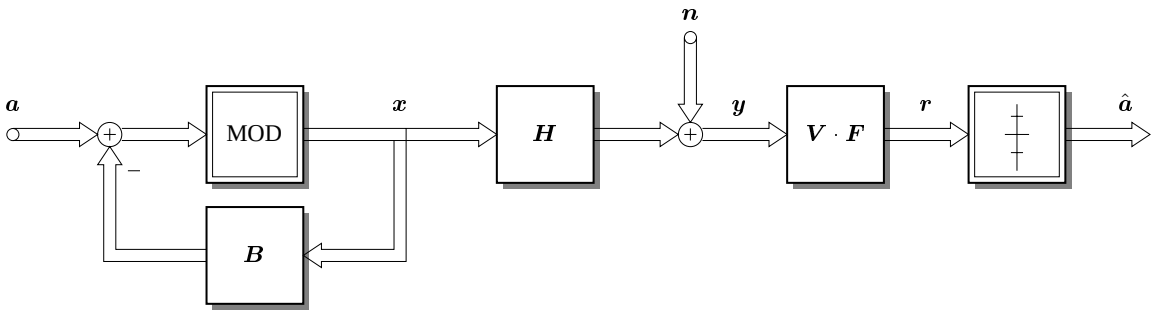

Fig. 1. Tomlinson-Harashima precoding for MIMO channels.

recursively calculated as ($\boldsymbol{B} = [b_{kl}]$ is the feedback matrix used in the precoder)

$$\begin{aligned} x_k &= \mathrm{MOD}\left\{a_k - \sum\nolimits_{l=1}^{k-1} b_{kl} x_l\right\} \\ &\stackrel{\mathrm{def}}{=} a_k + d_k - \sum\nolimits_{l=1}^{k-1} b_{kl} x_l \ , \quad k = 1, \ldots, K \ . \end{aligned} \tag{6}$$

The precoding symbols $d_k \in \left\{2\sqrt{M} \cdot (d_\mathrm{I} + \mathrm{j} d_\mathrm{Q}) \mid d_\mathrm{I}, d_\mathrm{Q} \in \mathbb{Z}\right\}$ 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 pre-distortion $\boldsymbol{C}^{-1} = (\boldsymbol{B} + \boldsymbol{I})^{-1}$ which is implemented by the feedback structure, *effective data symbols* $v_k \stackrel{\mathrm{def}}{=} a_k + d_k$ 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 that unique point is selected for transmission which results in a channel symbol with minimum magnitude of real and imaginary part [6].

Since the linear pre-distortion via $\boldsymbol{C}^{-1}$ equalizes the cascade $\boldsymbol{VFH}$, after pre-filtering ($\boldsymbol{F}$) and scaling ($\boldsymbol{V}$) at the receiver side, the effective data symbols v_k, corrupted by additive noise (error) are visible: $\boldsymbol{r} = \boldsymbol{v} + \boldsymbol{e}$. Using a slicer, and taking the modulo congruence into account, an estimate for the data symbols can be generated.

Compared to signaling using the data symbols a_k, Tomlinson-Harashima precoding slightly increases average transmit power, i.e., $\sigma_x^2 = \mathrm{E}\{|x_l|^2\} > \mathrm{E}\{|a_l|^2\} = \sigma_a^2$. This *precoding loss* is only relevant for "small" signal sets and vanishes completely as the number M of signal points goes to infinity. Additionally, the number of nearest neighbor points increases slightly due to the periodic extension of the constellation.

Using MIMO precoding, the major drawbacks of decision-feedback equalization are avoided. No error propagation may occur since the feedback filter is located at the transmitter where the signals are perfectly known. Additionally, channel coding can be applied in a straightforward manner. Using precoding, no immediate decisions at the receiver are required for equalization, which is irreconcilable with the principle of channel coding.

III. Fixed Precoder and Linear Residual Equalization

The MIMO precoding scheme described above assumes the transmitter to be informed about the actual realization of the channel matrix $\boldsymbol{H}$. Whenever the channel changes—even if it changes only slightly—it has to be communicated to the transmitter and the optimal precoder has to be recalculated.

Only in time-division duplex (TDD) schemes over a reciprocal channel, the channel can be estimated at each end from the received signals and immediately be used for the reverse direction [13]. Unfortunately, in most other transmission scenarios the channel parameters have to be estimated at the receiver side and then have to be communicated to the transmitter over a backward channel.

This back transmission of channel state information poses two problems: First, estimating the channel and transferring it back takes some time. Hence, in fast time-varying scenarios, the channel may have changed before the measured channel state is available at the transmitter. Consequently, even if the channel may be estimated perfectly at the receiver side and communicated to the transmitter without any error, channel information utilizable at the precoder and actual channel do not match.

Second, the transmission of channel state information itself requires high data rates. Hence, the backward channel suffers a loss in payload data rate. Obviously, it is desirable to transmit as little channel information as possible. One strategy is to supply only a quantized, and hence imperfect, version of the channel coefficients to the transmitter. This was proposed for space-time coding or beam forming e.g. in [12]. Like above, this approach leads to a mismatch between channel information available at the transmitter and the actual (instantaneous) channel.

However, if the channel varies only slowly over time, there is no need for instantaneous channel feedback. If precoder setting and actual channel differ only slightly, *residual linear equalization* at the receiver (cf. [3]) is a possibility to compensate for a mismatch. For that, the actual channel matrix is required *only* at the receiver.

Hence, we propose to adjust the precoder only to an average channel or an average covariance matrix, respectively, which has to be known at the transmitter (compromise precoder). The (medium to long-term) average $\mathrm{E}\{\boldsymbol{H}^\mathsf{H}\boldsymbol{H}\}$ has to be sent to the transmitter only once in a while. In turn, a low-rate backward channel is sufficient. Any mismatch will be compensated by residual linear equalization at the receiver. It has been shown that average knowledge at the transmitter may cause only a small loss in capacity, cf. e.g. [11], if the channel is sufficiently constant over a transmission burst.

TABLE II

COMPARISON OF PRECODER CALCULATION STRATEGIES.

	Precoder adjusted according to	
	actual channel $\boldsymbol{H}$	average $\mathrm{E}\{\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H}\}$
Factorization problem	$\boldsymbol{S}^{\mathsf{H}}\boldsymbol{S} = \boldsymbol{H}^{\mathsf{H}}\boldsymbol{H} + \zeta\boldsymbol{I}$	$\bar{\boldsymbol{S}}^{\mathsf{H}}\bar{\boldsymbol{S}} = \mathrm{E}\{\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H}\} + \zeta\boldsymbol{I}$
Receiver matrix	$\boldsymbol{F} = \boldsymbol{S}^{-\mathsf{H}}\boldsymbol{H}^{\mathsf{H}}$	$\bar{\boldsymbol{F}} = \bar{\boldsymbol{S}}^{-\mathsf{H}}\sqrt{\mathrm{E}\{\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H}\}}$
Precoder matrix	$\boldsymbol{C} = (\mathbf{diag}(\boldsymbol{S}))^{-1}\boldsymbol{S}\,, \qquad \boldsymbol{B} = \boldsymbol{C} - \boldsymbol{I}$	$\bar{\boldsymbol{C}} = \left(\mathbf{diag}(\bar{\boldsymbol{S}})\right)^{-1}\bar{\boldsymbol{S}}\,, \qquad \bar{\boldsymbol{B}} = \bar{\boldsymbol{C}} - \boldsymbol{I}$

A. Calculation of Precoder Setting

When only the (medium-term) average $\mathrm{E}\{\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H}\}$ is available at the transmitter side, a natural choice for calculating the relevant matrices is to replace the instantaneous value $\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H}$ by its expectation $\mathrm{E}\{\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H}\}$ in the Cholesky factorization problem (3). Table II compares the calculation[2] of the matrices $\boldsymbol{B}$, $\bar{\boldsymbol{B}}$ (precoder) and $\boldsymbol{F}$, $\bar{\boldsymbol{F}}$ (receiver front-end) with knowledge of the actual channel matrix or the (medium-term) expectation ("$\bar{\cdot}$"), respectively. Thus, the transmission scheme (precoder matrix $\bar{\boldsymbol{B}}$ and receiver matrix $\bar{\boldsymbol{F}}$) is designed only for a medium-term average.

B. Optimal Linear Residual Equalization

We assume that the precoder, i.e., the feedback matrix $\bar{\boldsymbol{B}}$, and the feedforward matrix $\bar{\boldsymbol{F}}$ at the receiver are fixed. If this setting does not match the actual realization $\boldsymbol{H}$ of the channel, significant unequalized multi-"user" interference may be present. This residual interference between the components of the receive vector $\boldsymbol{r}$ can be mitigated—and hence performance can be improved—by applying an additional linear equalization matrix $\boldsymbol{G}^{\mathsf{H}}$ at the receiver subsequent to $\bar{\boldsymbol{F}}$. In fact, an equivalent feedforward matrix $\boldsymbol{F} = \boldsymbol{G}^{\mathsf{H}}\bar{\boldsymbol{F}}$ is adjusted suitably. In practice, this will be done by means of some adaptive algorithm.

The optimal setting of the matrix $\boldsymbol{G}$ according to the MMSE criterion (minimum variance of filtered channel noise plus unequalized interference) is readily derived by applying standard techniques. Optimally, it is given as

$$\boldsymbol{G}_{\mathrm{opt}} = \bar{\boldsymbol{F}}^{-\mathsf{H}}\boldsymbol{H}\left(\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H} + \zeta\boldsymbol{I}\right)^{-1}\bar{\boldsymbol{C}}^{\mathsf{H}}\,. \tag{7}$$

with $\bar{\boldsymbol{C}} = \bar{\boldsymbol{B}} + \boldsymbol{I}$, and the correlation matrix of the residual error $\boldsymbol{e}$ at the decision device reads

$$\Phi_{ee} = \sigma_n^2\bar{\boldsymbol{C}}\left(\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H} + \zeta\boldsymbol{I}\right)^{-1}\bar{\boldsymbol{C}}^{\mathsf{H}}\,. \tag{8}$$

The noise variances $\sigma_{n,k}^2$ of the K parallel, unit-gain AWGN channels are given by the main-diagonal elements of $\boldsymbol{\Phi}_{ee}$.

IV. NUMERICAL EXAMPLES

In order to assess and compare the performance of the precoding schemes proposed above, as well as the influence of the imperfect actual channel knowledge at the transmitter, numerical examples are presented.

[2] The square root of a Gram matrix $\boldsymbol{M}$ is well defined and it holds: $\boldsymbol{X} = \sqrt{\boldsymbol{M}} \Leftrightarrow \boldsymbol{X}\boldsymbol{X} = \boldsymbol{M}$.

A. Channel

The underlying MIMO channel is assumed to be composed of two components: one, $\hat{\boldsymbol{H}}$ (almost) constant during one transmission burst, and one fast, $\Delta\boldsymbol{H}$, which can be seen as a deviation from the mean $\hat{\boldsymbol{H}}$. The actual channel $\boldsymbol{H}$ is hence given as

$$\boldsymbol{H} = \hat{\boldsymbol{H}} + \Delta\boldsymbol{H}\,. \tag{9}$$

The components δ_{ij} of $\Delta\boldsymbol{H}$ are assumed to be i.i.d. zero-mean Gaussian, mutually uncorrelated and with variance σ_Δ^2, i.e.,

$$\mathrm{E}\{\delta_{ij}\delta_{lm}^*\} = \begin{cases} \sigma_\Delta^2, & i = l \text{ and } j = m \\ 0, & \text{else} \end{cases}\,. \tag{10}$$

Thus, the medium-term average correlation matrix in a specific burst is

$$\mathrm{E}_{\Delta\boldsymbol{H}}\{\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H}\} = \hat{\boldsymbol{H}}^{\mathsf{H}}\hat{\boldsymbol{H}} + \sigma_\Delta^2\boldsymbol{I}\,. \tag{11}$$

This matrix is fed back to the transmitter, where it is used for adjusting the compromise precoder.

The components $\hat{h}_{ij}$ of the locally constant part $\hat{\boldsymbol{H}}$ are assumed to be i.i.d. zero-mean Gaussian, mutually uncorrelated and with variance 1.

Unless otherwise stated, for all numerical examples a MIMO channel with $K = 4$ transmit and receive antennas is considered. Noteworthy, the symbols are detected in random ordering—the optimal sorting of [9] is not applied here.

B. Performance Criteria

In each case, the variances $\sigma_{n,k}^2$, $k = 1, 2, \ldots, K$, of the additive disturbance (noise + residual multi-user interference) in the parallel channels are calculated. From these, performance measures for three different scenarios are derived:

(A) Uncoded transmission: Here the maximum over the noise variances of the parallel channels is meaningful, and we use as performance measure

$$\sigma_{n,\max}^2 \stackrel{\mathrm{def}}{=} \max_k \sigma_{n,k}^2\,.$$

(B) Coded transmission without use of channel state information: It can easily be seen that in this scenario the average noise variance (arithmetic mean) is (approximately) active, and the performance is assessed by

$$\sigma_{n,\mathrm{avg}}^2 \stackrel{\mathrm{def}}{=} \frac{1}{K}\sum_{k=1}^{K}\sigma_{n,k}^2\,.$$

(C) Optimal rate and power assignment: From optimal loading in multicarrier systems it is known that here only the geometric mean over the noise variances takes effect, and we regard

$$\sigma^2_{n,\mathrm{geo}} \stackrel{\mathrm{def}}{=} \sqrt[K]{\prod_{k=1}^{K} \sigma^2_{n,k}} \; .$$

C. Results

Histograms of the loss when using a fixed precoder relative to an optimally adjusted one are plotted in Figure 2. Thereby, the loss is expressed as

$$\Delta\mathrm{SNR} \stackrel{\mathrm{def}}{=} \frac{\left.\frac{\sigma_x^2}{\sigma_{n,\bullet}^2}\right|_{\text{optimal adjusted}}}{\left.\frac{\sigma_x^2}{\sigma_{n,\bullet}^2}\right|_{\text{fixed precoder}}} = \frac{\left.\sigma_{n,\bullet}^2\right|_{\text{fixed precoder}}}{\left.\sigma_{n,\bullet}^2\right|_{\text{optimal adjusted}}} \; . \quad (12)$$

The histograms are taken over 1000 random medium-term channel realizations $\hat{\boldsymbol{H}}$, and for each channel 1000 random deviations (fast varying components $\Delta\boldsymbol{H}$) are generated. The variance is assumed to be $\sigma_\Delta^2 = 0.03$, the variance of the transmit symbols is $\sigma_x^2 = 1$, and the noise variance is $\sigma_n^2 = 0.01$.

From left to right, the situation for uncoded transmission (scenario (A), $\bullet = \max$), coding without use of channel state information (scenario (B), $\bullet = \mathrm{avg}$), and for optimal rate and power distribution (scenario (C), $\bullet = \mathrm{geo}$), respectively, is shown. For uncoded transmission, a fixed precoder and linear residual equalization causes almost no loss. With probability of 60% the loss is smaller than 0.1 dB. If the arithmetic mean of the noise variances is active, the loss very seldomly exceeds 1 dB. The performance gap between optimal and compromise precoder increases when loading rate and power, since an optimally adjusted precoder benefits most from an optimized signaling. In this case, however, best performance can be achieved, anyway.

Noteworthy, for a fixed precoder, adjusting the modulation alphabets of the parallel channels is possible, even if only a low-rate backward channel is available. Communicating back the rate distribution (K integer numbers) takes much less effort than sending the actual channel ($K \times K$ complex numbers).

The corresponding average loss

$$\overline{\Delta\mathrm{SNR}} \stackrel{\mathrm{def}}{=} \mathrm{E}\{\Delta\mathrm{SNR}\} \quad (13)$$

expressed in dB when using a fixed precoder is plotted over the signal-to-estimation-error ratio $1/\sigma_\Delta^2$ (in dB) in Figure 3. Noteworthy, the signal-to-noise ratio σ_x^2/σ_n^2 is adjusted to 20 dB. Hence, over a wide range the deviation from the actual channel is much larger than the variance of the channel noise. Nevertheless, the loss of compromise precoding is still in a tolerable range.

Finally, in Figure 4 the average loss $\overline{\Delta\mathrm{SNR}}$ is displayed over the dimension K of the MIMO channel, i.e., the number of transmit and receive antennas. The other simulation parameters are again as given above; in particular $\sigma_\Delta^2 = 0.03$ is assumed. Remarkably, even for large dimensions of the matrix channel, the average loss is moderate.

To conclude, the adjustment of the Tomlinson-Harashima precoder to an average scenario is an attractive strategy. Only a low-rate backward channel is required to communicate the (average) channel data to the transmitter. Residual multi-user interference is mitigated by means of linear residual equalization at the receiver. The loss of this procedure compared to an optimally, instantaneously adjusted precoder is tolerable for almost all situations in practice.

V. Summary and Conclusions

In this paper we have shown that MIMO precoding introduced in [8] is applicable even when perfect knowledge on the channel is not available at the transmitter. The precoder is designed according to the average of the channel correlation matrix, and the mismatch of precoder and actual channel is mitigated by (adaptive) residual linear equalization. Numerical results show that the loss of this transmission strategy compared to an optimal adjustment of the precoder is tolerable. The strategy presented requires only a low-rate backward channel for the infrequent update of the precoder in time-varying scenarios.

Although we have focused on spatial equalization for flat fading channels, the transmission scheme is immediately applicable to frequency selective channels. Via MIMO precoding, combined spatial and temporal equalization can then be realized.

References

[1] P.W. Baier, M. Meurer, T. Weber, H. Troeger. Joint Transmission (JT), an alternative rationale for the downlink of Time Division CDMA using multi-element transmit antennas. In *Proc. of IEEE Intern. Symp. on Spread Spectrum Techniques and Applications (ISSSTA)*, pp. 1–5, Parsippany, NJ, Sep. 2000.

[2] M. Bossert, A. Donder, V. Zyablov. OFDM-Übertragung über Mobilfunkkanäle: Bemerkungen zur Kanalkapazität, *2. OFDM-Fachgespräch*, Braunschweig, Sept. 1997. (In German).

[3] R. Fischer, W. Gerstacker, J. Huber. Dynamics Limited Precoding, Shaping, and Blind Equalization for Fast Digital Transmission over Twisted Pair Lines. *IEEE Journal on Selected Areas in Communications*, JSAC–13, pp. 1622–1633, December 1995.

[4] R. Fischer, J. Huber. A New Loading Algorithm for Discrete Multitone Transmission, *IEEE Global Telecommunications Conference (GLOBECOM '96)*, pp. 724–728, London, November 1996.

[5] R. Fischer, J. Huber. On The Equivalence of Single- and Multicarrier Modulation: A New View, *Proceedings of the IEEE International Symposium on Information Theory (ISIT '97)*, p. 197, Ulm, June/July 1997.

[6] R. Fischer. *Precoding and Signal Shaping for Digital Transmission.* Habilitationsschrift, Universität Erlangen–Nürnberg, 2001.

[7] G. Foschini, M. Gans. On limits of wireless communication in a fading environment when using multiple antennas. *Wireless Personal Communications*, pp. 311–335, March 1998.

[8] R. Fischer, C. Windpassinger, A. Lampe, J. Huber. Space-Time Transmission using Tomlinson-Harashima Precoding, To be presented at *4th International ITG Conference on Source and Channel Coding*, Berlin, Jan. 2002.

[9] G. Golden, G. Foschini, R. Valenzuela and P. Wolniasky. Detection algorithm and initial laboratory results using the V-BLAST space-time communication architecture, *Electronics Letters*, pp. 14–15, January 1999.

[10] S. Kasturia, J.T. Aslanis, J.M. Cioffi. Vector Coding for Partial Response Channels. *IEEE Trans. Inform. Theory*, pp. 741–762, July 1990.

[11] M. Ivrlac, T. Kurpjuhn, C. Brunner, J. Nossek. On Channel Capacity of Correlated MIMO Channels, *ITG Fokusprojekt: Mobilkommunikation "Systeme mit intelligenten Antennen"*, Ilmenau, Mar. 2001.

[12] G. Jöngren, M. Skoglund. Utilizing Quantized Feedback Information in Orthogonal Space-Time Block Coding, *Proceedings of the IEEE Global Telecommunications Conference (Globecom '00)*, Nov. 2000.

[13] R. Knopp and G. Caire. Power Control Schemes for TDD Systems with Multiple Transmit and Receive Antennas. *Proc. of IEEE Global Telecommunications Conference (Globecom)*, pp. 2326–2330, Rio de Janeiro, Nov. 1999.

[14] G. Raleigh, J. Cioffi. Spatio-Temporal Coding for Wireless Communication. *IEEE Trans. Commun.*, pp. 357–366, March 1998.

[15] E. Telatar. Capacity of multi-antenna Gaussian Channels, *European Transactions on Telecommunications (ETT)*, pp. 585–596, Nov. 1999.

[16] P. Wolniansky, G. Foschini, G. Golden, R. Valenzuela. V-BLAST: An Architecture for Realizing Very High Data Rates Over the Rich-Scattering Wireless Channel. In *Proc. ISSSE–98*, Pise, Italy, Sep. 1998.

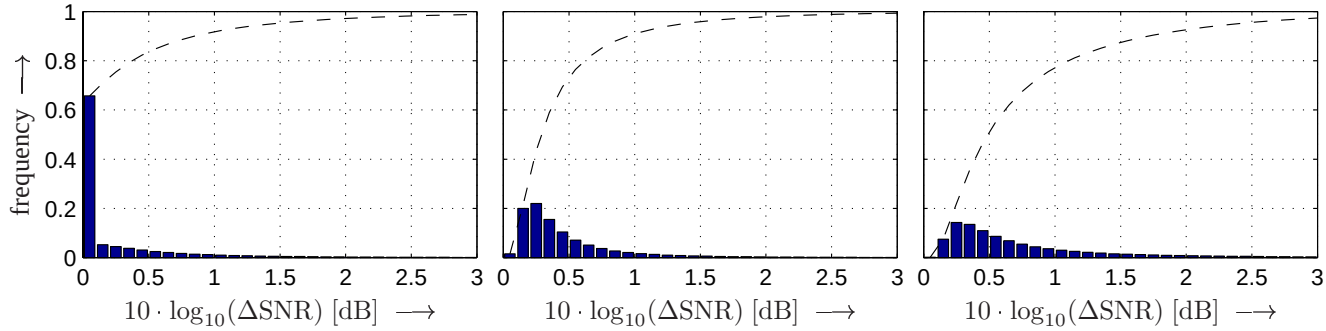

Fig. 2. Histograms of loss when using a fixed precoder relative to an optimal adjusted one. Left to right: uncoded transmission, coding without use of channel state information, optimal rate and power distribution. Dashed line: cumulative distribution.

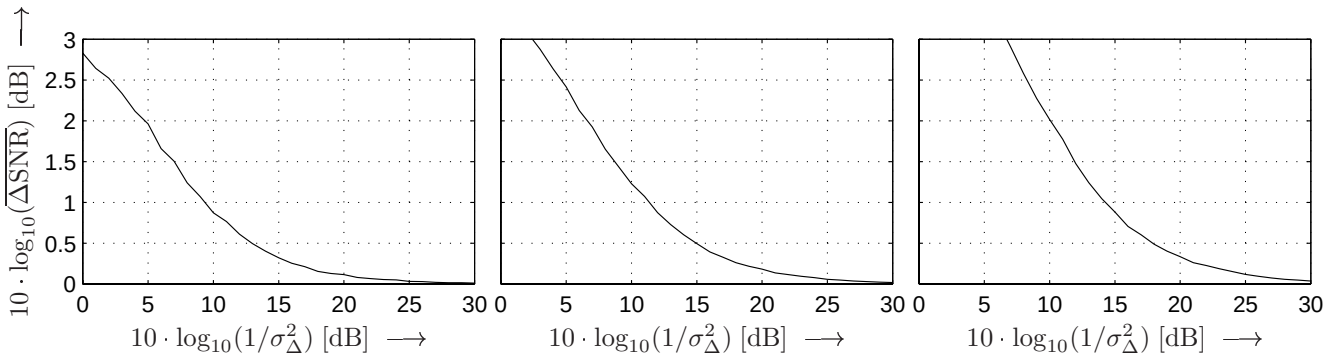

Fig. 3. Average loss when using a fixed precoder relative to an optimal adjusted one over the estimation variance in dB. Left to right: uncoded transmission, coding without use of channel state information, optimal rate and power distribution.

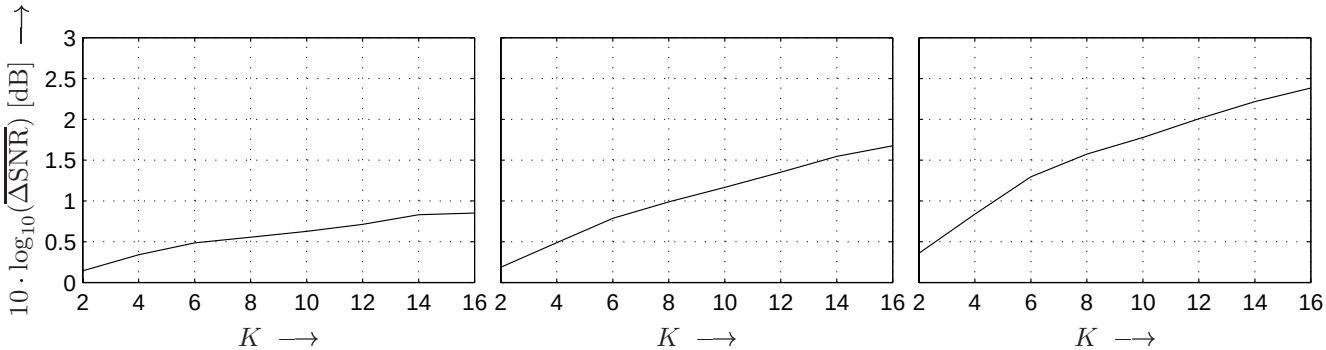

Fig. 4. Average loss when using a fixed precoder relative to an optimal adjusted one over the size K of the MIMO channel. Left to right: uncoded transmission, coding without use of channel state information, optimal rate and power distribution.