Precoding in Multi-Antenna and Multi-User Communications

Christoph Windpassinger, Robert F.H. Fischer,
Tomáš Vencel, Johannes B. Huber

Abstract— **In this paper, Tomlinson-Harashima precoding for multiple-input/multiple-output systems including multiple-antenna and multi-user systems is studied. It is shown that nonlinear preequalization offers significant advantages over linear preequalization which increases average transmit power. Moreover, it outperforms decision-feedback equalization at the receiver side which is applicable if joint processing at the receiver side is possible, and which suffers from error propagation. A number of aspects of practical importance are studied. Loading, i.e., the optimum distribution of transmit power and rate is discussed in detail. It is shown that the capacity of the underlying MIMO channel can be utilized asymptotically by means of non-linear precoding.**

Keywords— **MIMO channels, Space-time coding, Decision-Feedback Equalization, Tomlinson-Harashima precoding, Code-Division Multiple Access**

I. Introduction

OVER the last years, the demand for high data rates in wireless transmission has increased significantly. This has led to a strong interest in so-called multiple-input/multiple-output (MIMO) systems which promise extremely high spectral efficiencies. Foschini and Gans [24] showed that the spectral efficiency can (for large signal-to-noise ratios) in principle grow linearly with the minimum over the number of transmit (N_T) and receive (N_R) antennas. The BLAST system (Bell Laboratories Layered Space-Time) is the most prominent example for high-rate wireless communication schemes.

Usually, a MIMO transmission scheme is 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 (flat-fading) channel is characterized by its $N_\mathrm{R} \times N_\mathrm{T}$ channel matrix $\boldsymbol{H}$.

Noteworthy, BLAST-like systems are not the only ones that can be described by a MIMO input-output relation of this type. In fact, many multi-user transmission schemes that transmit multiple, possibly independent, data streams in parallel over linear channels can be described in this way. In particular consider direct-sequence code-division multiple access (DS-CDMA) systems, where *users* with their individual spreading sequence are treated rather than *antennas*. Considering the transmission of K users in parallel, the same channel model as given above arises.

Besides information theoretic studies of MIMO channels [3], [33], [34], [37], [41], numerous work has been done to propose feasible receivers for various practical scenarios. 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 $\boldsymbol{x}$, of which superpositions are present at the receiver side. In the context of multiple access systems, this task is called multi-user detection [40].

In the present paper, precoding, i.e., non-linear predistortion of the transmit signals, for MIMO system is discussed. Processing at the transmitter side requires the channel to be (approximately) constant over some time, in order to be able to gain channel state information (CSI) at the transmitter. Hence, we concentrate on *block fading* channels which are assumed as constant over a reasonably large blocklength.

Moreover, in view of the demand of high transmission rates, we focus on high spectral efficiencies, in particular non-binary modulation alphabets and correspondingly we assume high signal-to-noise ratios.

The paper is organized as follows: After discussing the channel model in Section II, we review the concepts of spatial equalization and precoding in Section III. An important aspect for implementation are the different loading strategies, discussed in detail in Section IV. In conclusion, a number of aspects important in practice and some extensions are studied in Section V.

II. Channel Model

In this section we show that DS-CDMA downlink transmission as well as the multi-antenna point-to-point channel can be described (in complex baseband notation) by the same matrix equation. Hence, similar methods can be considered for receiver or transmitter signal processing.

A. Multi-Antenna Scenario

Assuming that between each pair of transmit and receive antenna a flat, i.e., non-dispersive, fading channel constant over some transmission burst is present, the discrete time equivalent complex baseband model can be written as

$$\boldsymbol{y}[\nu] = \boldsymbol{H}[\nu]\,\boldsymbol{x}[\nu] + \boldsymbol{n}[\nu]\,. \tag{1}$$

Here $\boldsymbol{x}[\nu]$ is the vector of complex channel input symbols $x_k[\nu]$, $\boldsymbol{H}[\nu]$ the matrix of complex channel coefficients $h_{kl}[\nu]$ constant during one transmission burst, $\boldsymbol{n}[\nu]$ the ad-

The authors are with Lehrstuhl für Informationsübertragung, Universität Erlangen–Nürnberg, Cauerstraße 7/NT, D-91058 Erlangen, Germany, email: {windpass,fischer,vencel,huber}@LNT.de

ditive white (complex) Gaussian noise and $\boldsymbol{y}[\nu]$ the vector of received signals $y_k[\nu]$ at time instant ν.

For simplicity we restrict ourselves to systems with an equal number of transmit and receive antennas, say K, hence vectors are of length K, and $\boldsymbol{H}[\nu] = \big[h_{kl}[\nu]\big]$, a complex $K \times K$ matrix, where $h_{kl}[\nu]$ is the fading coefficient between transmit antenna l and receive antenna k.

The noise samples $n_k[\nu]$ are assumed to be zero-mean Gaussian random variables, mutually uncorrelated, i.e.,

$$\mathrm{E}\{n_k[\nu+\mu]n_l^*[\nu]\} = \begin{cases} \sigma_n^2 & \text{if } k = l \text{ and } \mu = 0 \ , \\ 0 & \text{else} \ . \end{cases} \tag{2}$$

B. Multi-User Downlink Scenario

In this scenario we assume transmission from a central transmitter (base station) to users (mobile equipment) scattered over the service area. In particular, we focus on the downlink of a DS-CDMA system with a single transmit antenna at the base station and single receive antennas at the mobile terminals.

In DS-CDMA systems the transmit signal is the superposition of the users' spreading sequences multiplied by the transmit symbols $x_k[\nu]$. The interchip interference caused by frequency-selective fading results in non-orthogonal effective spreading sequences $\boldsymbol{s}_k[\nu]$. We assume the effective spreading sequences to be normalized to unit norm — the additional flat fading component is described by the channel weight $w_k[\nu]$. Here we neglect intersymbol interference, an idealization which is justified if the spreading sequences are substantially longer than the maximum delay spread of the channel. The receivers are expected to use matched filters $\boldsymbol{s}_k^{\mathsf{H}}[\nu]$ and the noise after the matched filter is denoted by $n_k[\nu]$. It is reasonable to model the noise as in the multi-antenna case as white mutually uncorrelated complex Gaussian processes, cf. (2).

Writing $\boldsymbol{S}[\nu] = \big[\boldsymbol{s}_1[\nu], \ldots, \boldsymbol{s}_K[\nu]\big]$ and $\boldsymbol{W}[\nu] = \mathbf{diag}(w_1[\nu], \ldots, w_K[\nu])$, and collecting the transmit symbols $x_k[\nu]$ and the received symbols $y_k[\nu]$ in vectors $\boldsymbol{x}[\nu]$ and $\boldsymbol{y}[\nu]$, respectively, we can express the transmission by

$$\boldsymbol{y}[\nu] = \big(\boldsymbol{W}[\nu]\boldsymbol{S}^{\mathsf{H}}[\nu]\boldsymbol{S}[\nu]\big)\boldsymbol{x}[\nu] + \boldsymbol{n}[\nu] \ . \tag{3}$$

Defining $\boldsymbol{H}[\nu] \stackrel{\text{def}}{=} \boldsymbol{W}[\nu]\boldsymbol{S}^{\mathsf{H}}[\nu]\boldsymbol{S}[\nu]$ the same MIMO channel model as above (eq. (1)) is present.

Noteworthy, the components of the receive vector are present at physically separate pieces of mobile equipment, and are not accessible jointly. Hence, only individual further processing of the receive symbols is possible. In terms of vectors and matrices, these processing steps are characterized by diagonal matrices.

C. Multi-Antenna, Multi-User Downlink Scenario

The combination of these two models to DS-CDMA downstream transmission using multiple antennas is immediately possible by choosing the channel matrix $\boldsymbol{H}$ suitably. In this case, processing of the receive vector is limited to *block diagonal* matrices.

For brevity, in this paper, we restrict ourselves to the above two "extreme cases"; however, the results are easily extendable to the general multi-antenna multi-user downlink scenario.

D. Signal-to-Noise Ratio

The instantaneous receive signal-to-noise ratio (SNR) is defined as

$$\mathrm{SNR}[\nu] = \frac{\mathrm{E}_{\boldsymbol{x}}\left\{\mathrm{tr}\left(\boldsymbol{H}[\nu]\boldsymbol{x}[\nu]\boldsymbol{x}^{\mathsf{H}}[\nu]\boldsymbol{H}^{\mathsf{H}}[\nu]\right)\right\}}{K\sigma_n^2} \tag{4}$$

which calculates for i.i.d. channel input symbols with variance $\sigma_x^2 \stackrel{\text{def}}{=} \mathrm{E}\{|x_k|^2\}$, $\forall k$, to

$$= \frac{\sigma_x^2}{K\sigma_n^2}\sum_{k=1}^{K}\sum_{l=1}^{K}|h_{kl}|^2 \ . \tag{5}$$

If the elements h_{kl} of the channel matrix are identically distributed zero-mean random variables with variance $\sigma_{h_{kl}}^2$, the SNR averaged over the channel realization (i.e., averaging over the bursts) is given by

$$\mathrm{SNR} \stackrel{\text{def}}{=} \mathrm{E}_{\boldsymbol{H}}\{\mathrm{SNR}[\nu]\} = \frac{\sigma_x^2}{K\sigma_n^2}\sum_{k=1}^{K}\sum_{l=1}^{K}\sigma_{h_{kl}}^2 \ . \tag{6}$$

It is common to express power efficiency in terms of the average received energy per information bit $\bar{E}_{\mathrm{b}}$ divided by the one-sided noise power spectral density N_0. For complex-valued signaling the following relation to the signal-to-noise ratio holds

$$\frac{\bar{E}_{\mathrm{b}}}{N_0} = \frac{\mathrm{SNR}}{R_{\mathrm{m}}} \ , \tag{7}$$

where R_{m} denotes the (average) number of information bits transmitted per channel symbol $x_l[\nu]$. For channel matrices with i.i.d. unit variance elements ($\sigma_{h_{kl}}^2 = 1$, $\forall k, l$) the signal-to-noise ratio simplifies to $\frac{\bar{E}_{\mathrm{b}}}{N_0} = \frac{K\sigma_x^2}{R_{\mathrm{m}}\sigma_n^2}$.

Since most processing will only consider a single time instant ν, we will only use the time index if we want to stress time dependency, and otherwise omit it for brevity.

III. Equalization Strategies

In MIMO channels, superpositions (interference) of the components of the transmitted vector $\boldsymbol{x}$ are present at the receiver side. In particular, the receive symbol at the kth antenna/of the kth user is given by

$$y_k = h_{kk}x_k + \sum_{l=1, l\neq k}^{K} h_{kl}x_l + n_k \ , \tag{8}$$

where the second term on the right-hand side gives the "spatial" interference from input symbols x_l, $l \neq k$. In order to recover data, some kind of separation or equalization of the parallel data streams is required.

In this section, equalization strategies for multi-antenna and multi-user transmission are studied, see Table I. In each case it is indicated whether it is a receiver side

TABLE I

SUMMARY OF EQUALIZATION STRATEGIES FOR MULTI-ANTENNA AND MULTI-USER TRANSMISSION. IF THE EQUALIZATION STRATEGY IS APPLICABLE FO THE PARTICULAR SCENARIO THE NUMBER OF THE RESPECTIVE SUBSECTION IS INDICATED.

	Multi-Antenna Scenario	Multi-User Scenario	
		Downlink	Uplink
Linear Equalization (Receiver-Side Technique)	III.A	—	III.A
Linear Preequalization (Transmitter-Side Technique)	III.A	III.A	—
DFE (Receiver-Side Technique)	III.B	—	III.B
Precoding (Transmitter-Side Technique)	III.C	III.C	—
SVD (Both-Sides Technique)	III.D	—	—

or transmitter side technique, and whenever a particular equalization strategy is applicable to the multi-antenna scenario, the multi-user downlink scenario, or the multi-user uplink transmission, a scenario dual to the multi-user downlink, the number of the respective subsection is given.

A. Linear Equalization

The straightforward strategy for separating the data streams is linear equalization at the receiver side, also known as a decorrelating detector [40]. Here, the decision vector is generated by $\boldsymbol{r} = \boldsymbol{H}_{(l)}^{-1}\boldsymbol{y}$, where $\boldsymbol{H}_{(l)}^{-1}$ denotes the left (pseudo) inverse of the channel matrix $\boldsymbol{H}$. This strategy is applicable in the multi-antenna scenario, as well as in multi-user uplink transmission, i.e., transmission from users scattered over the service area to the central base station.

If channel state information (CSI) is available at the transmitter, the users can be separated by means of linear preequalization (sometimes calles "predistorted"). Such schemes are applicable in multi-antenna and multi-user downlink scenarios. Interference of the users at the receiver side can be completely avoided by 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 and feeding the preequalized version $\boldsymbol{x} = \boldsymbol{H}_{(r)}^{-1}\boldsymbol{a}$ into the channel [2], [42]. For further linear transmitter preprocessing strategies see, e.g., [5], [35], [12].

It is well-known that linear equalization suffers from noise enhancement and hence has poor power efficiency. The same is true for linear preequalization of the transmit signals, where transmit power is boosted.

B. Matrix DFE, V-BLAST

The disadvantage of linear equalization at the receiver side can be overcome by spatial or matrix decision-feedback equalization (DFE) [44], [11], [1], which is also used [27] in the (V-)BLAST [23], [25], [43] system for multi-antenna transmission. In the context of multi-user detection this strategy is known as "successive cancellation" [40].

The block diagram of decision-feedback equalization, which is applicable if joint processing at the receiver side is possible, is shown in Figure 1. As in DFE for intersymbol-interference (ISI) channels, the receiver consists of a feedforward matrix $\boldsymbol{F}$ (cascaded, for convenience, with a diagonal scaling matrix $\boldsymbol{G} = \mathbf{diag}(g_1, \ldots, g_K)$), and a feedback matrix $\boldsymbol{B}$.

The task of the feedforward matrix is to guarantee (i) spatially white noise at the decision devices (decorrelation of the noise) and (ii) spatial causality: The decision symbols with index l should only be disturbed by symbols with index $k < l$. Hence, the end-to-end matrix $\boldsymbol{B} = [b_{k\kappa}] \stackrel{\text{def}}{=} \boldsymbol{GFH}$ has to be lower triangular ($b_{k\kappa} = 0$, $k < \kappa$). The scaling matrix $\boldsymbol{G}$ is chosen for unit-gain signal transmission, i.e., the main diagonal elements of $\boldsymbol{B}$ are one ($b_{kk} = 1$). Via the feedback matrix $\boldsymbol{B}$, the interference caused by already detected symbols is canceled. Since $\boldsymbol{B} - \boldsymbol{I}$ is strictly lower left a causal loop for successive detection and interference cancellation results.

Assuming white channel noise, the calculation of the filters in BLAST [25] can be interpreted as performing a *QL-factorization* of a permuted version of the channel matrix, corresponding to a relabeling of the transmit antennas. Subsequently we assume that this relabeling is already included in the matrix $\boldsymbol{H}$. Then, the factorization reads

$$\boldsymbol{H} = \boldsymbol{F}^{\mathsf{H}}\boldsymbol{S} \ , \tag{9}$$

where $\boldsymbol{F}$ is the unitary feedforward matrix (preserving whiteness of the noise) and $\boldsymbol{S} = \left[s_{ij}\right]$ a lower left triangular matrix. The scaling matrix is then given as $\boldsymbol{G} = \mathbf{diag}(s_{11}^{-1}, \ldots, s_{KK}^{-1})$, and the feedback matrix reads $\boldsymbol{B} = \boldsymbol{GFH} = \boldsymbol{GS}$.

The above factorization results in filters adjusted accord-

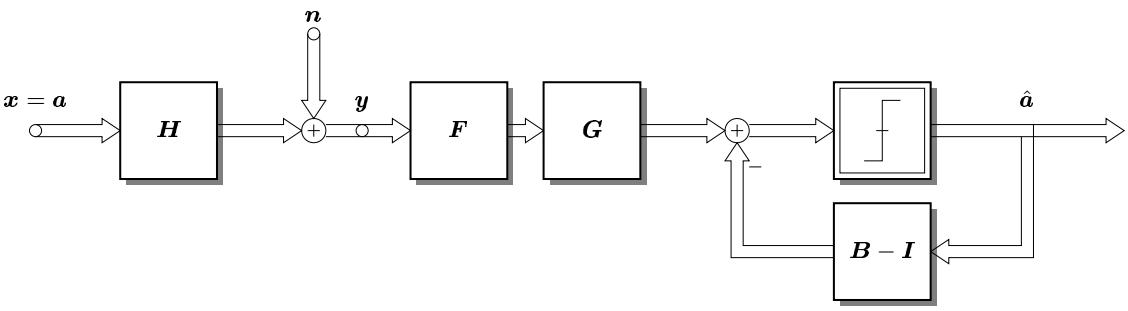

Fig. 1. Matrix DFE receiver.

ing to the *zero-forcing* (ZF) criterion. For low signal-to-noise ratios and, correspondingly, low transmission rates, some improvement may be possible by optimizing the matrices according to the *minimum mean-squared error* (MMSE) criterion. Here, we are interested in high rates and hence concentrate on the zero-forcing solution.

Due to the feedforward processing, the symbol at the decision device reads

$$y'_k = x_k + \sum_{l=1}^{k-1} b_{kl} x_l + n'_k \ , \tag{10}$$

with $n'_k = \tilde{n}_k / s_{kk}$ and $\tilde{\boldsymbol{n}} = \boldsymbol{F}\boldsymbol{n}$. Assuming error-free decisions, i.e, $\hat{x}_l = x_l$, $l = 1, \ldots, k-1$, the interference term is perfectly canceled and the decision variable for symbol x_k is only disturbed by white Gaussian noise $\tilde{n}_k$. That is, the MIMO channel is decomposed into K parallel AWGN channels with noise variances $\sigma_n^2 / |s_{kk}|^2$, $k = 1, \ldots, K$.

Since in practical systems erroneous decisions do occur, the BLAST algorithm [25] finds the optimum permutation of the receive antennas such that the decisions are taken starting from the most reliable signal to reduce the probability of error propagation. For simplicity we assume here that the antennas are already sorted in this optimum way.

Decision-feedback equalization is also popular in the multi-user uplink scenario, i.e., for joint detection at the base station. However, in the multi-user downlink scenario as described here no such equalization is possible. On the one hand, for complexity reasons, each receiver is only allowed to apply a single-user matched filter, and on the other hand, the components of $\boldsymbol{y}$ in the downlink channel model are not accessible jointly. Subsequently, we show how this problem can be solved by means of precoding.

C. Tomlinson-Harashima Precoding

In any communication system where channel state information is present at the transmitter, equalization can also be performed by a *precoding operation* on the transmit symbols.

C.1 Multi-Antenna Transmission

Initially, Tomlinson-Harashima precoding (THP) [30], [38] was proposed for the equalization of intersymbol interference of severely linear distorting SISO channels, but can readily be extended to MIMO channels, e.g., [17], [20], see also [13], [6]. THP can be interpreted as moving the feedback part of DFE to the transmitter, cf. Figures 1 and 2. Thereby, the two main shortcomings of the DFE scheme can be overcome: (i) since the operation is performed at the transmitter where signals are perfectly known, no error propagation can occur, and (ii) for equalization, no immediate decisions are required. Hence, channel coding can be applied in a straightforward manner, i.e., as it is done for the intersymbol-interference-free AWGN channel.

The operation of THP is tightly connected to the used signal constellation $\mathcal{A}$. For the moment assume that in each of the parallel data streams an M-ary square constellation (M a square number) is employed, where the coordinates of the signal points are odd integers, i.e., $\mathcal{A} = \{a_{\mathrm{I}} + \mathrm{j}a_{\mathrm{Q}} \mid a_{\mathrm{I}}, a_{\mathrm{Q}} \in \{\pm 1, \pm 3, \ldots \pm (\sqrt{M} - 1)\}\}$. The constellation is bounded by the square region of width $2\sqrt{M}$.

If the modulo device at the transmitter were ignored, *linear* preequalization of the cascade $\boldsymbol{B} = \boldsymbol{G}\boldsymbol{F}\boldsymbol{H}$ would be present. Because of the triangular structure of the feedback matrix $\boldsymbol{B}$, the channel symbols x_k, $k = 1, \ldots, K$, are successively generated from the data symbols $a_k \in \mathcal{A}$:

$$x_k = a_k - \sum\nolimits_{l=1}^{k-1} b_{kl} x_l \ , \quad k = 1, \ldots, K \ . \tag{11}$$

Since this strategy would increase transmit power significantly, Tomlinson-Harashima precoding *modulo reduces* the transmit symbols into the boundary region of $\mathcal{A}$. Mathematically, integer multiples of $2\sqrt{M}$ are added to the real and imaginary part of x_k. Now, the channel symbols are given as

$$x_k = a_k + p_k - \sum\nolimits_{l=1}^{k-1} b_{kl} x_l \ , \quad k = 1, \ldots, K \ , \tag{12}$$

where $p_k \in \left\{ 2\sqrt{M} \cdot (p_{\mathrm{I}} + \mathrm{j}p_{\mathrm{Q}}) \mid p_{\mathrm{I}}, p_{\mathrm{Q}} \in \mathbb{Z} \right\}$. In other words, instead of feeding the data symbols a_k into the linear preequalization, the *effective data symbols* $v_k \stackrel{\text{def}}{=} a_k + p_k$ are passed into $\boldsymbol{B}^{-1}$, which is implemented by the feedback structure. That is, the initial signal constellation is extended periodically, i.e., the effective data symbols v_k may be taken from the expanded set $\mathcal{V} = \{v_{\mathrm{I}} + \mathrm{j}v_{\mathrm{Q}} \mid v_{\mathrm{I}}, v_{\mathrm{Q}} \in \{\pm 1, \pm 3, \ldots\}\}$. Since the *precoding symbols* p_k are matched to the boundary region of the initial signal constellation, the points in the expanded signal set are also taken from a regular two-dimensional grid [19].

All points, spaced in real and imaginary part by integer multiples of $2\sqrt{M}$ are congruent and represent the same

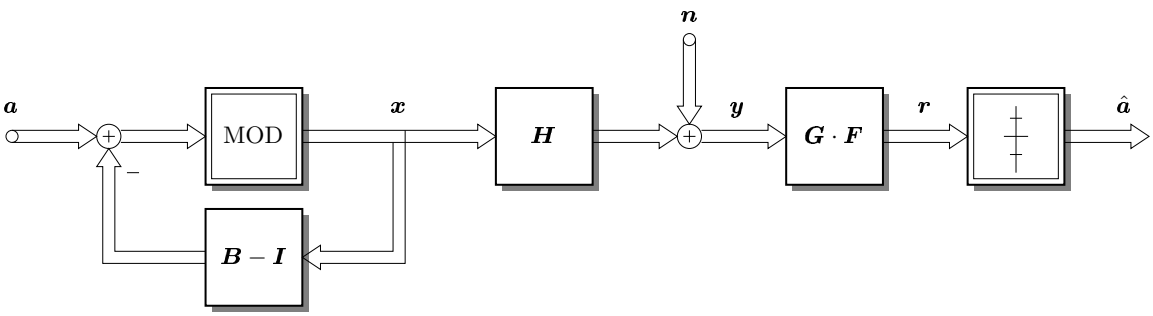

Fig. 2. Tomlinson-Harashima precoding for MIMO channels.

data. From these equivalent points, symbol-by-symbol that point is selected for transmission which results in a channel symbol falling into the boundary region of $\mathcal{A}$ [19]. Note that this choice is unique and done implicitly by the modulo operation.

Since the linear preequalization via $\boldsymbol{B}^{-1}$ equalizes the cascade $\boldsymbol{B} = \boldsymbol{GFH}$, after prefiltering and scaling, the effective data symbols v_k, corrupted by additive noise, are visible at the receiver, i.e., $\boldsymbol{r} = \boldsymbol{v} + \boldsymbol{n}'$. Here, $\boldsymbol{n}'$ denotes the filtered channel noise and $\boldsymbol{v} = [v_1, \ldots, v_K]^{\mathsf{T}}$. Using a slicer which takes the periodic extension into account an estimate for the data symbols (vector $\boldsymbol{a}$) can be generated. Alternatively, the received symbols r_k are first modulo reduced into the boundary region of the signal constellation $\mathcal{A}$. Then, a conventional slicer suffices. Hence, neglecting the modulo congruence, which is justified when using "large" QAM signal constellations (i.e., spectrally efficient modulation schemes), as in DFE unit-gain AWGN channels with signal-to-noise ratio $\frac{\sigma_x^2}{\sigma_n^2/|s_{kk}|^2}$ result.

Using Tomlinson-Harashima precoding, the channel symbols x_k are no longer taken from the signal constellation $\mathcal{A}$. In "classical" THP, the channel symbols are (approximately) uniformly distributed over the boundary region of $\mathcal{A}$. This leads to a somewhat increased transmit power compared to transmission of symbols taken from $\mathcal{A}$, quantified by the *precoding loss*, which for square QAM constellations and uniform distribution calculates to [19]

$$\gamma_{\mathrm{p,2D}}^2 = \frac{M}{M-1} \ . \tag{13}$$

Even for moderate sizes M this loss is negligible and vanishes as M increases. Moreover, the components x_k can be assumed to be mutually uncorrelated with variance $\sigma_x^2 = \mathrm{E}\{|x_k|^2\}$, $\forall k$.

When performing precoding with respect to the spatial dimension, each symbol interval is processed separately. As a consequence, the channel symbols take on more and more discrete levels when going from component x_1 to x_K. Since a uniform distribution is never achieved, the precoding loss in MIMO precoding is even lower than predicted by (13).

For illustration, Fig. 3 shows scatter plots of the channel symbols x_k and the received symbols r_k for a MIMO channel with $K = 4$ inputs and $K = 4$ outputs. In each of the parallel channels a 16-ary QAM constellation is used. From component 1 through 4, the channel symbols tend from the initial 16-QAM constellation to an almost uniform distribution over the boundary region. Simultaneously, the effective data symbols are taken from a more and more expanded signal set. The non-uniform distribution of the effective data symbols v_k can be seen. Moreover, the different noise variances which are effective for the different components are visible.

Tomlinson-Harashima precoding can work in combination with other than square constellations. The only requirement is that the precoding symbols are matched to the boundary region of the signal constellation such that the entire space is tiled with the periodic replicas, see [19].

Additionally we note that in case of low-rate transmission, precoding according to an MMSE criterion [13] and regarding the properties of so-called modulo channels [22] offers some additional gains. Moreover, going to higher-dimensional precoding lattices, the so-called shaping gap can be bridged. In [13] a scheme denoted by "inflated lattice" precoding is proved to be capacity-achieving. Again we note that in this paper we concentrate on high rates and hence the ZF solution.

C.2 Multi-User Transmission

The scheme presented above for the multi-antenna scenario was found by a direct translation of the concept of temporal (i.e., ISI) equalization to spatial equalization. However, in the present version it is not applicable to the multi-user scenario, since for distributed receivers the symbols y_k can not be processed jointly by the feedforward matrix $\boldsymbol{F}$.

Recall that the task of the feedforward matrix is to spatially whiten the channel noise and to force spatial causality. Since the channel noise is assumed to be white, only causality has to be achieved, which—in contrast to noise whitening—is also possible by a matrix at the transmitter side. Hence, the feedforward matrix $\boldsymbol{F}$ is moved to the transmitter, too, see Figure 4. The operation of the precoder remains the same as given above.

In the present situation, the required matrices can be calculated by decomposing the channel matrix according to

$$\boldsymbol{H} = \boldsymbol{SF}^{\mathsf{H}} \ , \tag{14}$$

where $\boldsymbol{F}$ is a unitary matrix, $\boldsymbol{S}$ is a lower left triangular matrix. Again, this is a well-known QR decomposition (here LQ) of the channel matrix, but note the difference to the

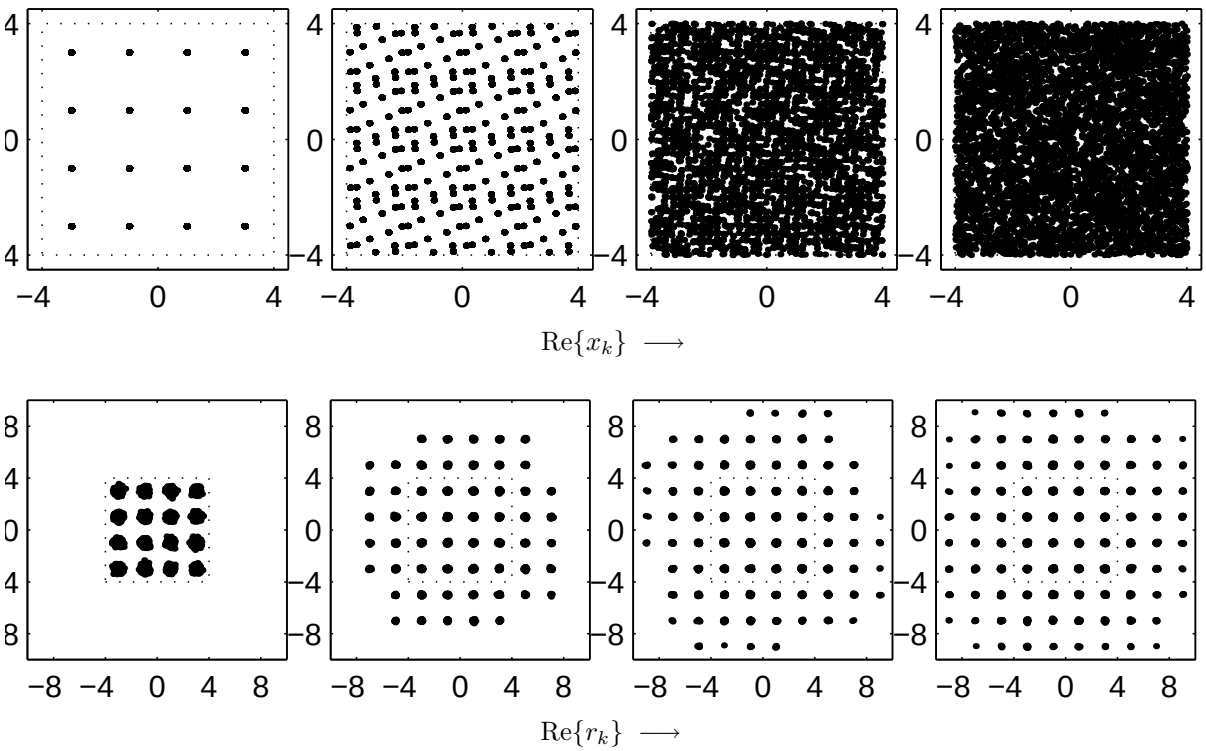

Fig. 3. Scatter plots of channel symbols x_k and received symbols r_k when using MIMO precoding. $K = 4$ in- and outputs. 16-QAM constellation. Left to right: Components $k = 1$ through $k = 4$.

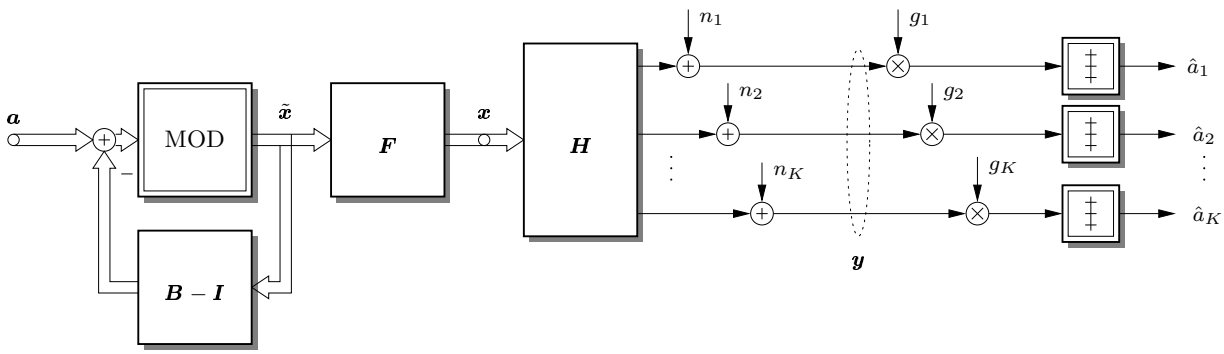

Fig. 4. Precoding for decentralized receivers.

factorization (9). $\boldsymbol{F}$ immediately gives the feedforward matrix, now applied at the transmitter side, the gain matrix reads as above $\boldsymbol{G} = \mathbf{diag}(s_{11}^{-1}, \ldots, s_{KK}^{-1})$, and the feedback matrix calcuates to $\boldsymbol{B} = \boldsymbol{GHF} = \boldsymbol{GS}$.

Note that in the context of multi-user broadcast transmission, precoding can be view as a scalar realization of so-called "Costa codes" [7]. This "dirty paper" result, which states that if interference is known at the transmitter side an encoding/decoding rule can be found that enables transmission as if the interference were not present, became very popular over the last few years.

C.3 QL versus LQ decomposition

The implementation of the feedforward matrix $\boldsymbol{F}$ at the transmitter side is indispensable in the multi-user scenario. In the multi-antenna setting, however, this strategy is an alternative to the scheme discussed above. While the QL decomposition can be seen as performing Gram-Schmidt orthogonalization [29] on the columns of $\boldsymbol{H}$, the LQ decomposition corresponds to orthogonalization along the rows of $\boldsymbol{H}$.

Although, for randomly chosen channel matrices $\boldsymbol{H}$, the statistical properties of the diagonal elements s_{kk}, which determine performance, will be the same for both strategies, regarding an actual realization of $\boldsymbol{H}$, a particular type of decomposition may show advantages compared to the other decomposition. Therefore one may consider an implementation of precoding where the position of the unitary "feed-forward" matrix is decided on-the-fly with regard to the current realization of $\boldsymbol{H}$.

Figure 5 shows simulated symbol error rates transmitting 16-QAM constellations over $K = 4$ transmit and receive antennas. The elements h_{kl} of the channel matrix are chosen i.i.d. complex Gaussian and the error rates are

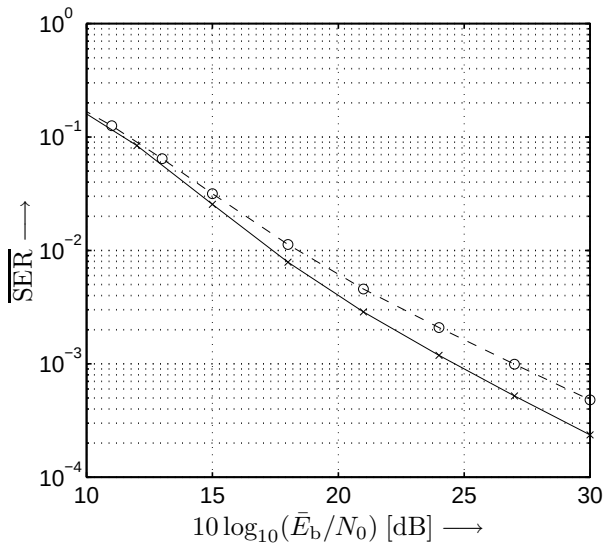

Fig. 5. Average symbole error rate for switching between QL and LQ decomposition (×) compared to only QL decomposition (○).

averaged over a large number of channel realizations. For each channel realization both factorizations are calculated and the "better" type of decomposition was selected. Since we are studying uncoded transmission, performance is governed by $|s_{kk,\min}|^2 \stackrel{\text{def}}{=} \min_k |s_{kk}|^2$. Hence, the factorization which results in the larger $|s_{kk,\min}|^2$ is selected. For comparison, (dashed line) the situation where $\boldsymbol{F}$ is always present at either receiver or transmitter is plotted, too. Using this additional degree of freedom in multi-antenna transmission with CSI available at the transmitter, further gains of more than 2 dB in signal-to-noise ratio are achievable.

D. Singular Value Decomposition

Another popular method for channel equalization in multi-antenna systems [36], [31] is based on the *singular value decomposition (SVD)* [29] of the channel matrix $\boldsymbol{H}$. Here, the task of channel equalization can be split among transmitter and receiver. Hence SVD-based equalization cannot be applied in the multi-user (either up- or downlink) scenario.

It is well known that any complex matrix $\boldsymbol{H}$ can be decomposed as

$$\boldsymbol{H} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathsf{H}} , \tag{15}$$

where $\boldsymbol{U}$ and $\boldsymbol{V}$ are unitary matrices containing the eigenvectors of $\boldsymbol{H}\boldsymbol{H}^{\mathsf{H}}$ and $\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H}$, respectively, and $\boldsymbol{\Sigma} = \mathbf{diag}(\varsigma_1, \ldots, \varsigma_K)$ is a real diagonal matrix of the positive square roots of the corresponding eigenvalues. Hence, applying $\boldsymbol{V}$ at the transmitter and $\boldsymbol{U}^{\mathsf{H}}$ at the receiver, K independent, parallel channels are present. Since $\boldsymbol{U}$ and $\boldsymbol{V}$ are unitary matrices, in contrast to linear (pre)equalization, neither an amplification of the channel noise, nor of transmit power occurs.

Noteworthy, the capacity of the MIMO channel is calculated via the SVD approach, e.g., [37]. Assuming high signal-to-noise ratios ($\sigma_x^2/\sigma_n^2 \gg 1$), the capacity achieving power distribution over the parallel channels tends to a flat one, and the capacity of the MIMO channel, utilizable by the SVD scheme (assuming all dimensions are active[1]) is well approximated by

$$\begin{aligned} C_{\text{SVD}} &= \log\left|\boldsymbol{I} + \frac{\sigma_x^2}{\sigma_n^2}\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H}\right| \\ &\approx \log\left|\frac{\sigma_x^2}{\sigma_n^2}\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H}\right| = \log\left|\frac{\sigma_x^2}{\sigma_n^2}\boldsymbol{\Sigma}^{\mathsf{H}}\boldsymbol{\Sigma}\right| \\ &= \log\left(\prod_{k=1}^{K}\frac{\sigma_x^2}{\sigma_n^2}\varsigma_k^2\right) = K\log\frac{\sigma_x^2}{\sigma_n^2} + \log\prod_{k=1}^{K}\varsigma_k^2 . \end{aligned} \tag{16}$$

We note that $\prod_{k=1}^{K}\varsigma_k^2$ is just the determinant $|\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H}| = |\boldsymbol{\Sigma}^{\mathsf{H}}\boldsymbol{\Sigma}|$.

This capacity is to be compared to that usable with precoding. As shown above, decision-feedback equalization (assuming error-free previous decisions) and Tomlinson-Harashima precoding, result in parallel and independent unit-gain AWGN channels with noise variances $\sigma_n^2/|s_{kk}|^2$. Hence, total capacity calculates to

$$\begin{aligned} C_{\text{THP}} &= \sum_{k=1}^{K}\log\left(1 + \frac{\sigma_x^2}{\sigma_n^2}|s_{kk}|^2\right) \\ &\approx \sum_{k=1}^{K}\log\left(\frac{\sigma_x^2}{\sigma_n^2}|s_{kk}|^2\right) = \log\left(\prod_{k=1}^{K}\frac{\sigma_x^2}{\sigma_n^2}|s_{kk}|^2\right) \\ &= K\log\frac{\sigma_x^2}{\sigma_n^2} + \log\prod_{k=1}^{K}|s_{kk}|^2 . \end{aligned} \tag{17}$$

Since $\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H} = \boldsymbol{S}^{\mathsf{H}}\boldsymbol{F}\boldsymbol{F}^{\mathsf{H}}\boldsymbol{S} = \boldsymbol{S}^{\mathsf{H}}\boldsymbol{S}$ and $\boldsymbol{S}$ is triangular, we see that $\prod_{k=1}^{K}|s_{kk}|^2$ is just the determinant $|\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H}| = |\boldsymbol{S}^{\mathsf{H}}\boldsymbol{S}|$, and consequently for high SNRs we have

$$C_{\text{THP}} = C_{\text{SVD}} , \qquad \text{for} \quad \sigma_x^2/\sigma_n^2 \gg 1 . \tag{18}$$

In words, using non-linear precoding, the entire capacity of the underlying MIMO channel can be used. Hence, asymptotically, from a practical point of view, precoding is an optimal approach for channel equalization. In the next chapter, we discuss strategies how to best use this potential. This equivalence is similar to that of single- and multi-carrier modulation over ISI channels, see [16].

However, in each case for approaching capacity, the application of channel coding is indispensable. In particular, in precoding schemes as well as for the SVD approach, channel coding techniques developed for the AWGN channel may be employed cascaded with the equalization strategy. Since the performance of the codes in precoding

[1]For brevity of presentation, we assume that all singular values contribute to the transmission. If some singular values are (almost) zero and the respective dimensions do not contribute to the transmission, these dimensions simply have to be excluded from further considerations, studying a channel matrix of reduced dimension. The sums and products then only run over the active dimensions k.

From the theory of parallel Gaussian channels with noise variances N_k and fixed total power P, the transmit power of channel k is given by $P_k = (P + \sum_\kappa N_\kappa)/K - N_k$, see, e.g., [26], [8]. For high signal-to-noise ratios, i.e., $P \gg N_k$, $\forall k$, this tends to $P_k = (P + \sum_\kappa N_\kappa)/K - N_k \approx P/K = \text{const}$.

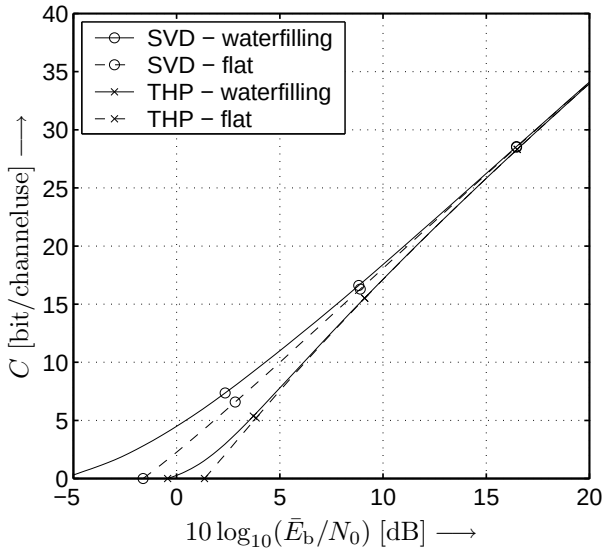

Fig. 6. Achievable rates for SVD-based equalization and for MIMO precoding. Solid: waterfilling result; dashed: capacity for flat power distribution.

schemes is essentially the same as on the AWGN channel [21], using "strong" codes the capacity of the channel can be approached asymptotically.

For transmission using $K = 4$ transmit and receive antennas the capacities (in bits per K-dimensional transmit vector) using the SVD approach and that of THP or DFE are plotted in Figure 6. Again, the elements of the channel matrix are chosen i.i.d. complex Gaussian and the results are averaged over a large number of channel realizations. For both equalization trategies, the waterfilling result and the capacity for a flat power distribution are given.

It can be seen that for high signal-to-noise ratios all curves merge. In this region the waterfilling power distribution tends to a flat one and the capacity utilizable by precoding approaches that of the underlying MIMO channel.

E. MIMO Channels with Intersymbol Interference

In the above discussion, only flat fading MIMO channels have been considered. One main advantage of DFE/precoding over the SVD approach is that they can straightforwardly be extended to channels which produce intersymbol interference. Here, joint spatial and temporal equalization has to be performed. Assuming again that the channel is (almost) constant over one transmission burst, the MIMO channel is now characterized by a channel matrix whose elements are impulse responses, rather than pure gain factors.

Denoting the matrix of (causal) impulse responses as $\langle \boldsymbol{H}[\nu] \rangle = [\langle h_{k\kappa}[\nu] \rangle]$, $\boldsymbol{H}[\nu] = [h_{k\kappa}[\nu]]$, $\langle h_{k\kappa}[\nu] \rangle = \langle h_{k\kappa}[0] h_{k\kappa}[1] \ldots \rangle$, the received signal in time interval ν reads

$$\boldsymbol{y}[\nu] = \sum_{\mu=0}^{\infty} \boldsymbol{H}[\mu] \boldsymbol{x}[\nu - \mu] + \boldsymbol{n}[\nu] \; . \tag{19}$$

For calculating the optimum feedforward and feedback matrices, we define the z-transform of the channel matrix as

$$\boldsymbol{H}(z) \stackrel{\text{def}}{=} \left[\sum_{\nu=0}^{\infty} h_{k\kappa}[\nu] z^{-\nu} \right] = \sum_{\nu=0}^{\infty} \boldsymbol{H}[\nu] z^{-\nu} \; . \tag{20}$$

Then, the factorization problem (9) $\boldsymbol{H} = \boldsymbol{F}^{\mathsf{H}} \boldsymbol{S}$ or $\boldsymbol{H}^{\mathsf{H}} \boldsymbol{H} = \boldsymbol{S}^{\mathsf{H}} \boldsymbol{F} \boldsymbol{F}^{\mathsf{H}} \boldsymbol{S} = \boldsymbol{S}^{\mathsf{H}} \boldsymbol{S}$ for a central receiver (multi-antenna scenario; $\boldsymbol{F}$ implemented at the receiver side) has to be replaced by a *spectral factorization* (cf. [45], [10])

$$\boldsymbol{H}^{\mathsf{H}}(z^{-*}) \boldsymbol{H}(z) = \boldsymbol{R}^{\mathsf{H}}(z^{-*}) \boldsymbol{R}(z) \; . \tag{21}$$

The time-domain matrix $\boldsymbol{R}[\nu]$ corresponding to $\boldsymbol{R}(z) = \sum_{\nu=0}^{\infty} \boldsymbol{R}[\nu] z^{-\nu}$ is strictly causal, i.e., $\boldsymbol{R}[\nu] = \boldsymbol{0}$, $\nu < 0$, and the matrix $\boldsymbol{R}[0] = [r_{k\kappa}[0]]$ additionally is lower triangular.

If $\boldsymbol{F}$ should be implemented at the transmitter side (e.g., in the multi-user case), factorization (14) is replaced by

$$\boldsymbol{H}(z) \boldsymbol{H}^{\mathsf{H}}(z^{-*}) = \boldsymbol{R}(z) \boldsymbol{R}^{\mathsf{H}}(z^{-*}) \; . \tag{22}$$

Having $\boldsymbol{R}(z)$, feedforward matrix is given as $\boldsymbol{F}(z) = \boldsymbol{R}^{\mathsf{H}}(z^{-*}) \boldsymbol{H}^{\mathsf{H}}(z^{-*})$, and the feedback matrix reads $\boldsymbol{B}(z) = \left(\mathbf{diag}(s_{11}[0], \ldots, s_{KK}[0]) \right)^{-1} \boldsymbol{S}(z)$.

The operation of combined spatial/temporal precoding is exactly the same as for pure spatial precoding—only the matrices have to be replaced by matrix filters. Moreover, again K parallel, independent channels are generated, the SNRs of which are given by the diagonal elements of $\boldsymbol{R}[0]$. Subsequently we return to flat fading channels but keep in mind that the results are immediately extendable to MIMO ISI channels.

IV. Loading in MIMO Systems

In systems where channel state information is available at the transmitter the question of so-called *loading* arises. Since the transmitter has knowledge of how the transmitted symbols are attenuated and disturbed by the channel, it may adjust transmit rate and/or power in an optimized way. Moreover, in the multi-antenna scenario, the MIMO channel might not have full rank, i.e., offer as many independent paths as the dimensionality of the channel matrix suggests. As a result, some of the equivalent parallel channels might have very bad transmission properties or not be present at all.

Loading schemes are well known from multicarrier transmission, e.g., [9], [15], where total transmit power and target transmission rate are distributed over the individual carriers. In MIMO transmission this translates to "spatial" loading; however, for fading channels, a second dimension should be included in the consideration: Loading may also be performed with respect to time [28], [41]. Hence, for optimum performance when transmitting over MIMO fading channels, two-dimensional rate and power loading, i.e., "space-time loading" should be performed. This concept is illustrated in Fig. 7, where different signal constellations are used to carry different transmission rates in the subchannels (index k). The shading of the blocks illustrates the channel states in the particular subchannels (assuming

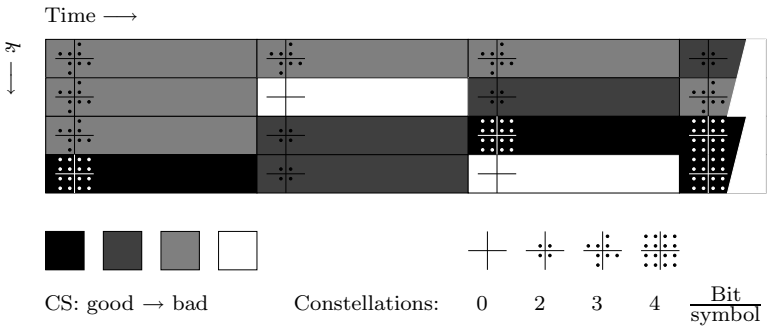

Fig. 7. Illustration of combined spatial and temporal loading.

block fading channels), and the "zero" constellation marks subchannels that are not used at all.

Regarding the parameters "space", "time", "rate", and "power", in different scenarios different loading strategies may be of particular interest. Subsequently, we discuss in more detail the specific requirements of multi-antenna and multi-user systems, respectively, and contrast various approaches for loading.

Time: When performing loading with respect to time, rate and/or power are adjusted according to the current realization $\boldsymbol{H}[\nu]$ for maximum average reliability for a given average transmit power and rate. Our focus is on precoding, which is only applicable if the channel is slowly time varying (in order to be able to gain CSI at the transmitter). In this situation, loading over time will introduce long-term variations of the actual transmit power or the instantaneous transmission rate. Hence, in systems which should support a constant data rate, large buffers at transmitter and receiver are required, which in turn introduces large delays, unacceptable in a great number of applications.

Space: The diagonal entries s_{kk} in Equation (9) show very dissimilar statistics for different k. In fact, in the multi-antenna scenario, entry $|s_{kk}|^2$ is distributed according to a χ^2-distribution of $2(K-k+1)$ degrees of freedom [23]. Due to these asymmetries, spatial loading is of particular interest.

In the multi-user scenario, however, "space" has the meaning of user. While in the multi-antenna transmission total rate per symbol period is of relevance, in the multi-user case variable rates for one subchannel would mean variable transmission rates for the corresponding user. Hence, "spatial" loading in the multi-user setting should focus on the adjustment of transmit powers for the subchannels at constant rates. Only if (large) delays are acceptable, variable-rate transmission may be an option.

Please note that in the present paper we concentrate on uncoded transmission. Here it is natural that optimum loading results in a balancing of bit error probabilities in the parallel channels. However, considering coded transmission, from the analogy to the present situation to OFDM it is known that loading is not necessary in order to achieve capacity [4]. Instead, a single, long codeword is arranged over space and time. Then, the code itself averages over the different fading conditions of the subchannels. However, even in coded systems loading is favorable from a practical point of view. Especially for low-delay channel coding schemes a balancing of the predecoding error probabilities is rewarding. For uncoded transmission, loading may offer significant gains.

Subsequently, we discuss more in detail loading strategies for the two scenarios multi-antenna and multi-user transmission and present simulation results.

A. Multi-Antenna Scenario

As the baseline for comparison, Figure 8 gives simulated symbol error rates for a multi-antenna scenario with $K = 4$. The elements of the channel matrix $\boldsymbol{H}$ are chosen i.i.d. complex Gaussian and the error rates are averaged over a large number of channel realizations. Here *no power or rate loading* is active. Each of the parallel data streams employs 16-QAM constellations with equal average power, and the optimal detection ordering according to [25] is used.

Noticeably, the SVD based transmission scheme (curve "SVD", cf. Section III-D) suffers from the highly unequal distribution of the singular values ς_k. With good approximation, the worst subchannel (smallest singular value) dominates the error rate. For high SNRs the performance of the SVD scheme approaches that of linear preequalization (cf. "Linear"). It is easy to see that the performance of linear preequalization is also dominated by the smallest singular value.

Tomlinson-Harashima precoding (curve labeled "THP") for the MIMO channel shows a clear advantage over linear preequalization and also some advantage over the V-BLAST system ("VBLAST") which uses a DFE receiver. Comparing the error rate of V-BLAST with the curve labeled "gVBLAST", which is valid for *genie-aided* DFE, i.e., using perfect decisions for the feedback, the effect of error propagation at the receiver becomes evident. Moreover, relative to genie-aided DFE, THP exhibits a small loss of approximately 0.2 dB due to the increased transmit power (precoding loss) and an additional factor of $\approx \frac{4}{3}$ in error rate due to the increase in the number of nearest neighbors by extending the signal set periodically (4 nearest neighbors instead of 3 in 16-ary QAM).

In the present scenario, a non-uniform distribution of transmission rates as illustrated in Fig. 7 is possible since both transmitter and receiver have access to all components of the vector channel. As described above, since we consider block fading channels, loading over time is of mi-

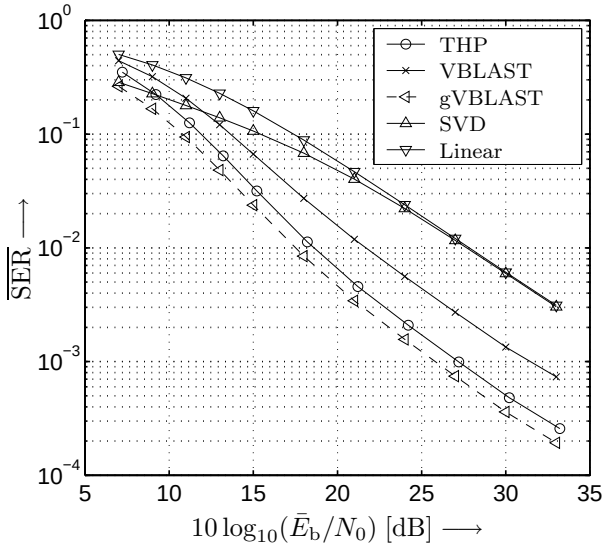

Fig. 8. Average symbol error rates for different equalization strategies without loading. Multi-antenna scenario, $K = 4$.

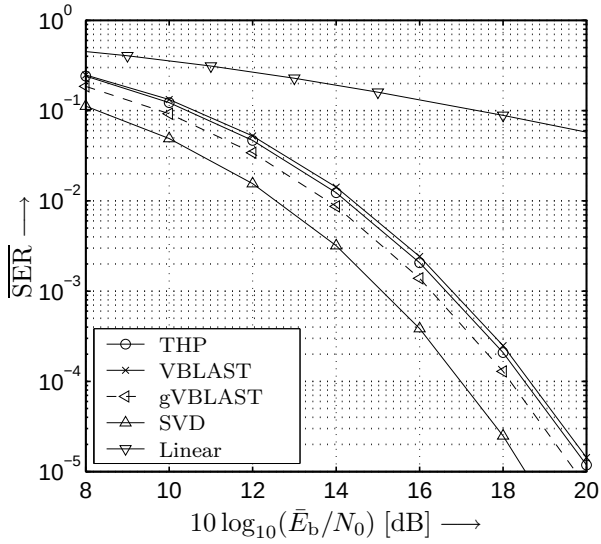

Fig. 9. Average symbol error rates for different equalization strategies. Multi-antenna scenario, $K = 4$; optimum rate and power loading.

nor interest and the most interesting option is spatial rate loading while keeping the sum rate constant.

From the OFDM literature, several algorithms are known to find the optimum rate and power distribution for discrete modulation alphabets, e.g., [9], [15]. The algorithm [15] distributes a fixed sum rate over the parallel channels such that the overall symbol error rate is minimized. Denoting, as above, the gain factors of the K parallel AWGN channels as s_{kk}, total rate R is distributed according to

$$R_k = \frac{R}{K} + \frac{1}{K} \cdot \log_2 \left(\frac{|s_{kk}|^{2K}}{\prod_{l=l}^{K} |s_{ll}|^2} \right) . \tag{23}$$

These rates are quantized to integers in an allowed range. Ideally, total power has to be distributed uniformly over the subchannels. In order to compensate for the quantization to integer rates and to guarantee the same error rate within the parallel subchannels some fine-adjustment is performed [15].

The simulation results shown in Figure 9 compare different equalization strategies. In each case the algorithm [15] is employed to find the optimum rate (and power) distribution. Thereby, in active subchannels transmission with integer rates ranging from 2 up to 8 bits is possible, corresponding to M-QAM signal sets with cardinality $M = 2^2, \ldots, 2^8$. The channel model is the same as in the previous figure, and total rate is again fixed to 16 bit per symbol interval. Please note that when using loading the sorting of the parallel subchannels according to [25] no longer is optimum, see [39]. In this paper, each time when using loading, the modified processing ordering according to [39] is applied.

Now, performance of the SVD scheme is superior to the other schemes; due to loading no longer the worst-case subchannel dominates. As a side effect of loading, error propagation in DFE is mitigated and Tomlinson-Harashima precoding performs almost the same as V-BLAST. Compared to genie-aided V-BLAST, both strategies show only a very small loss. Since linear preequalization always results in parallel, independent channels with the same SNR, loading is inactive and very poor performance is visible.

A comparison of Figures 8 and 9 shows the substantial performance improvements possible by spatial loading in precoding systems. Noticeably, loading leads to some diversity gain; the flat Rayleigh fading behavior present in Figure 8 is overcome and tends to that of an AWGN channel.

B. Multi-User Scenario

As already mentioned, in multi-user downlink scenarios an adjustment of the transmission rates of the subchannels, which correspond to different users, is undesirable in most practical schemes. The obvious strategy is hence to use equal rates in all subchannels, but allow for an adjustment of transmit power to account for different fading states seen by the users.

Since we are interested in a fair treatment of the users, in addition to equal data rates we demand equal error rates in the subchannels. This leads to a so-called channel inversion, since the goal is equivalent to achieving the same signal-to-noise ratio for all users. However, for averaging error rates high transmit powers are required for very bad channels (small fading gains). In turn, most of the total transmit power has to be assigned to these channels, leading to an unsatisfactory overall performance. In most situations it is advantageous not to use the worst-case channels, i.e., to discard the corresponding user from transmission, and to spend power for transmission of the "better" users. This strategy is tightly related to *truncated channel inversion* [28]. Thereby, as a new dimension, "outage" probabilities for the users would have to be taken into consideration. In this paper, no outage is allowed at the cost of an increased total error rate.

Figures 10 and 11 show simulation results for a multi-

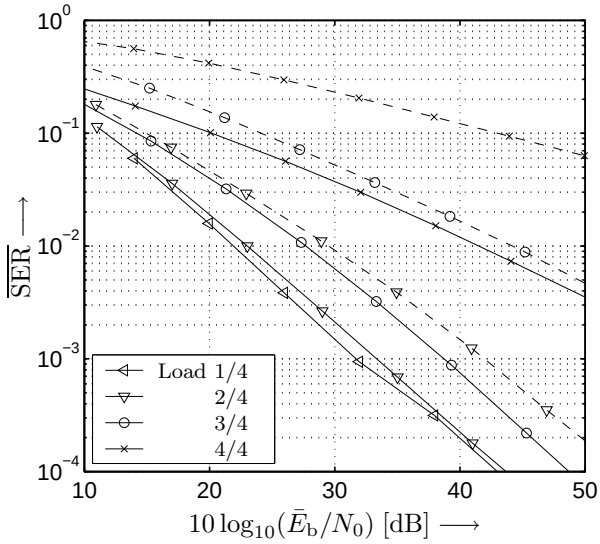

Fig. 10. Average symbol error rates for THP (–) and linear preequalization (- -). Multi-user scenario with equal rates and no power loading. System load $K/N = 1/4, \ldots, 4/4$.

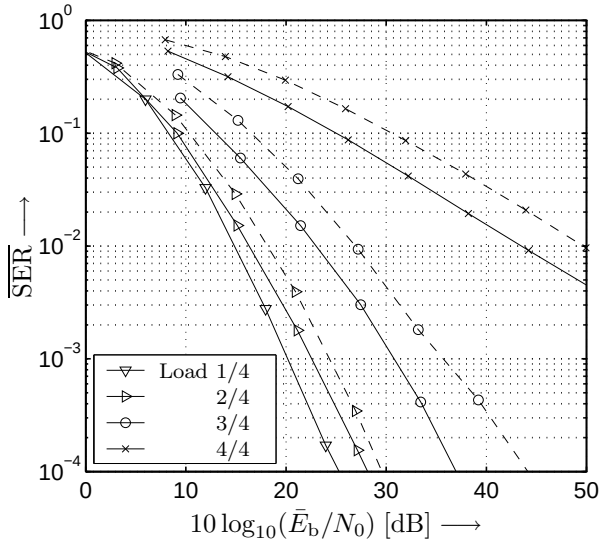

Fig. 11. Average symbol error rates for THP (–) and linear preequalization (- -). Multi-user scenario with equal rates and power loading. System load, $K/N = 1/4, \ldots, 4/4$.

user scenario with linear preequalization and Tomlinson-Harashima precoding, respectively. The matrix $\boldsymbol{S}[\nu]$ of effective spreading sequences is chosen as an $N \times K$ matrix of i.i.d. complex Gaussian random variables, where the columns are normalized to unit norm. The fading powers $|w_k|^2$ are assumed to be χ^2-distributed random variables with 4 degrees of freedom and unit variance. This closely reflects a multipath fading channel (incorporated into the effective spreading sequences) with two dominant Raleigh-fading taps. Again, the error rates are averaged over a large number of channel realizations. The different curves are valid for different system loads $\Gamma = K/N$.

Figure 10 gives the simulation results for no power adaptation, i.e., it corresponds to Figure 8 in the multi-antenna scenario, while Figure 11 shows results for channel inversion. In each case, 16-QAM constellations are applied and the spreading sequences are of length $N = 4$. Noticeably, especially for medium to high system loads, precoding offers considerable gains over linear preequalization. In the one user case ($K = 1$) both linear pre-equalization and THP degrade to single-user transmission over a flat-fading channel and hence perform the same.

V. Discussion and Extension

In this paper, Tomlinson-Harashima precoding for transmission over MIMO channels has been discussed. In a unified framework, BLAST-like multi-antenna systems and multi-user downlink communication, i.e., the downlink direction of DS-CDMA systems, were treated. The application of precoding at the transmitter side is always possible if the channel is slowly time-varying and channel state information is available at the transmitter. Moreover, in the DS-CDMA downlink case precoding allows to concentrate computational complexity at the base station and to use low-complexity mobile terminals.

As shown, precoding offers significant gains over linear (pre)equalization and over decision-feedback equalization. From a capacity point of view, precoding performs as well as the optimum singular-value decomposition approach. For optimal performance, rate and power should be distributed optimally over the independent, parallel channels induced by precoding. Some of such loading strategies have been addressed.

If perfect channel state information is not available at the transmitter, precoding may be adjusted to an average channel knowledge. A mismatch of precoder and actual channel realization can be compensated by means of linear adaptive residual equalization at the receiver. In realistic scenarios, the loss of this strategy compared to an optimally adjusted precoder is moderate. For details the reader is referred to [18].

Finally, the operation of precoding can further be enhanced by combining precoding with *signal shaping*. As proposed for ISI SISO channels [14], combined precoding/signal shaping for MIMO channels can be done by eliminating the modulo device and performing a long-term selection of the precoding symbols p_k by means of a shaping algorithm. Via shaping, a maximum gain in average transmit power of 1.53 dB is possible. However, besides creating transmit signals with lowest possible power, almost any shaping aim, e.g., to control the dynamics range of the effective data symbols or to reduce the peak-to-average power ratio of the continuous-time transmit signal (after the pulse shaping filter) [14], [19] can be achieved.

VI. Acknowledgements

The authors wish to acknowledge stimulating discussions with Dr. A. Lampe.

References

[1] N. Al-Dhahir and A. H. Sayed. The Finite-Length Multi-Input Multi-Output MMSE-DFE *IEEE Trans. Signal Processing*, pp. 2921–2936, October 2000.

[2] P. W. Baier, M. Meurer, T. Weber, and 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, September 2000.

[3] E. Biglieri, G. Caire, and G. Taricco. Limiting Performance of Block–Fading Channels with Multiple Antennas. *IEEE Transactions on Information Theory*, pp. 1273–1289, May 2001.

[4] M. Bossert, A. Donder, and V. Zyablov. OFDM–Übertragung über Mobilfunkkanäle: Bemerkungen zur Kanalkapazität (in German). In *2. OFDM–Fachgespräch*, Braunschweig, Germany, September 1997.

[5] M. Brandt-Pearce, A. Dharap. Transmitter-based multiuser interference rejection for the down-link of a wireless CDMA system in a multipath environment. *IEEE Journal on Selected Areas in Communications*, pp. 407-417, March 2000.

[6] G. Caire and S. Shamai. On the achievable throughput of a multi-antenna Gaussian broadcast channel. Submitted to *IEEE Transactions on Information Theory*, July 2001.

[7] M.H.M. Costa. Writing on dirty paper. *IEEE Transactions on Information Theory*, IT-29(5), pp. 439–441, May 1983.

[8] T.M. Cover and J.A. Thomas. *Elements of Information Theory*. John Wiley & Sons, Inc., New York, 1991.

[9] P. S. Chow, J. M. Cioffi, J. A. C. Bingham. A Practical Discrete Multitone Transceiver Loading Algorithm for Data Transmission over Spectrally Shaped Channels. *IEEE Transactions on Communications*, pp. 773–775, February 1995.

[10] M.C. Davis. Factoring the Spectral Matrix. *IEEE Transactions on Automatic Control*, pp. 296-305, October 1963.

[11] A. Duel-Hallen. A Family of Multiuser Decision-Feedback Detectors for Asynchronous Code-Division Multiple-Access Channels. *IEEE Transactions on Communications*, pp. 421–434, February/March/April 1995.

[12] R. Esmailzadeh, M. Nakagawa, E.A. Sourour. Time-Division Duplex CDMA Communications. *IEEE Personal Communications*, pp. 51-56, April 1997.

[13] U. Erez, S. Shamai, and R. Zamir. Capacity and Lattice-Strategies for Cancelling Known Interference. In *Proceedings of ISITA 2000*, Honolulu, Hawaii, November 2000.

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

[15] R. F. H. Fischer and J. B. Huber. "A New Loading Algorithm for Discrete Multitone Transmission," In *IEEE Global Telecommunications Conference*, pp. 724-728, London, November 1996.

[16] R. F. H. Fischer and J. B. 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.

[17] R. F. H. Fischer, C. Windpassinger, A. Lampe, and J. B. Huber. "Space-Time Transmission using Tomlinson-Harashima Precoding," *4th International ITG Conference on Source and Channel Coding*, pp. 139-147, Berlin, January 2002.

[18] R. F. H. Fischer, C. Windpassinger, A. Lampe, and J. B. Huber. "Tomlinson-Harashima Precoding in Space-Time Transmission for Low-Rate Backward Channel," *International Zurich Seminar on Broadband Communications (IZS '02)*, Zurich, Switzerland, February 2002.

[19] R. F. H. Fischer. Precoding and Signal Shaping for Digital Transmission, John Wiley & Sons, 2002.

[20] R. F. H. Fischer, C. Windpassinger, A. Lampe, and J. B. Huber. "MIMO Precoding for Decentralized Receivers," *International Symposium on Information Theory (ISIT '02)*, p. 496, Lausanne, Switzerland, July 2002.

[21] G.D. Forney and G. Ungerböck. Modulation and Coding for Linear Gaussian Channels. *IEEE Transactions on Information Theory*, pp. 2384–2415, October 1998.

[22] G.D. Forney, M.D. Trott, and S.-Y. Chung. Sphere–Bound–Achieving Coset Codes and Multilevel Coset Codes. *IEEE Transactions on Information Theory*, pp. 820–850, May 2000.

[23] G. J. Foschini. Layered Space–Time Architecture for Wireless Communication in a Fading Environment When Using Multiple Antennas. *Bell Laboratories Technical Journal*, pp. 41–59, Autumn 1996.

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

[25] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky. Simplified Processing for High Spectral Efficienty Wireless Communication Employing Multi-Element Arrays. *IEEE Journal on Selected Areas in Communications*, pp. 1841–1852, November 1999.

[26] R.G. Gallager. *Information Theory and Reliable Communication*. John Wiley & Sons, Inc., New York, London, 1968.

[27] G. Ginis and J.M. Cioffi. On the Relation between V-BLAST and the GDFE. *IEEE Communications Letters*, pp. 364–366, September 2001.

[28] A. J. Goldsmith and P. P. Varaiya. Capacity of Fading Channels with Channel Side Information *IEEE Transactions on Communications*, pp. 1986–1992, November 1997.

[29] G. H. Golub and C. F. Van Loan. *Matrix Computations*, 3rd Edition, The Johns Hopkins University Press, Baltimore, 1996.

[30] H. Harashima and H. Miyakawa. Matched-transmission technique for channels with intersymbol interference, *IEEE Transactions on Communications*, pp. 774–780, August 1972.

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

[32] S. Kasturia, J. T. Aslanis, and J. M. Cioffi. Vector Coding for Partial Response Channels. *IEEE Transactions on Information Theory*, pp. 741–762, July 1990.

[33] R. R. Müller, A. Lampe, and J. B. Huber. Gaussian multiple–access channels with weighted energy constraint. In *Proc. of IEEE Information Theory Workshop (ITW)*, pp. 106–107, Killarney, Ireland, June 1998.

[34] R. R. Müller. An Asymptotic Analysis of BLAST–like Systems. In *Proc. of IEEE International Conference on Communications (ICC)*, Helsinki, June 2001.

[35] A. Noll Barreto. Signal Pre-Processing in the Downlink of Spread-Spectrum Communications Systems. *VDI Fortschrittberichte*, no. 687, vol. 10, 2000.

[36] G. G. Raleigh and J. M. Cioffi. Spatio–Temporal Coding for Wireless Communication. *IEEE Transactions on Communications*, pp. 357–366, March 1998.

[37] E. Telatar. "Capacity of multi-antenna Gaussian Channels," *European Trans. Telecomm. (ETT)*, pp. 585–596, November 1999.

[38] M. Tomlinson. New automatic equaliser employing modulo arithmetic, *Electronics Letters*, pp. 138–139, March 1971.

[39] T. Vencel, C. Windpassinger, R. Fischer. Sorting in the V-BLAST Algorithm and Loading. In *Proceedings of Communication Systems and Networks (CSN 2002)*, pp. 304-309, Malaga, Spain, Sept. 2002.

[40] S. Verdú. *Multiuser Detection*. Cambridge University Press, New York, 1998.

[41] P. Viswanath, D. Tse, and V. Anantharam. Asymptotically Optimal Water–Filling in Vector Multiple–Access Channels. *IEEE Transactions on Information Theory*, pp. 241–267, January 2001.

[42] B. R. Vojčić and W. M. Jang. Transmitter Precoding in Synchronous Multiuser Communications *IEEE Transactions on Communications*, pp. 1346–1355, October 1998.

[43] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela. V-BLAST: An Architecture for Realizing Very High Data Rates Over the Rich-Scattering Wireless Channel. In *Proc. ISSSE–98*, invited paper, Pise, Italy, September 1998.

[44] J. Yang and S. Roy. Joint Transmitter-Receiver Optimization for Multi-Input Multi-Output Systems with Decision Feedback. *IEEE Transactions on Information Theory*, pp. 1334–1347, September 1994.

[45] D.C. Youla. On the Factorization of Rational Matrices. *IEEE Transactions on Information Theory*, pp. 172–189, July 1961.

Christoph Windpassinger (S'00) was born in Würzburg, Germany, in 1975. He received the Dipl.-Inf. (computer science) degree from the University of Erlangen-Nürnberg in 2000 and has since been working toward the Ph.D. degree at the Institute for Information Transmission of the University of Erlangen.
His research interests include detection and precoding in space-time transmission and related information theoretic questions.

Robert F. H. Fischer (S'93–M'99) received the Dipl.-Ing. degree in electrical engineering in 1992, the Dr.-Ing. degree in 1996, and the habilitation degree in 2001, all from the University of Erlangen-Nürnberg, Erlangen, Germany. The subject of his dissertation was multichannel and multicarrier modulation, that of his habilitation was precoding and signal shaping.
Form 1992 to 1996, he was a Research Assistant at the Telecommunications Institute, University of Erlangen-Nürnberg. During 1997, he was with the IBM Research Laboratory, Zürich, Switzerland. In 1998, he returned to the Telecommunications Institute II, University of Erlangen–Nürnberg. Currently, he teaches graduate courses in digital communications. His research concentrates on fast digital transmission including single- and multicarrier modulation techniques. His current interests are information theory, coded modulation, digital communications and signal processing, and especially precoding and shaping techniques for high-rate transmission schemes.
Dr. Fischer received the Dissertation Award from the Technische Fakultät, University of Erlangen-Nürnberg, in 1997, the Publication Award of the German Information Technology Society (ITG) in 2000, and the Wolfgang Finkelnburg habilitation award in 2002. He is author of the textbook *Precoding and Signal Shaping for Digital Transmission* (John Wiley & Sons, New York, 2002).

Tomáš Vencel was born in Nitra, Slovakia, on April 10, 1978. He received the Dipl.-Ing. degree in electrical engineering from the University of Erlangen-Nürnberg, Germany, in 2002. In his studies he focused on digital communications, with particular emphasis on equalization techniques.
After graduation he changed the field of activity and since 2002 has been with Framatome GmbH, Erlangen, Germany, and is now a project manager, working on safety and control systems for power plants.

Johannes B. Huber (M'87–SM'02) was born in Hausham, Germany, in 1951. He received the Dipl.-Ing. degree in electrical engineering from the Technische Universität, München in 1977. From 1977 to 1982 he was research assistant at the Lehrstuhl für Nachrichtentechnik of the Universität der Bundeswehr, München, from which he received the Dr.-Ing. degree with a thesis on coding for channels with memory.
From 1982 to 1991, he was an Akademischer Oberrat at the Universität der Bundeswehr, München. In 1991 he was visiting scientist at the IBM Research Laboratory, Zürich, Switzerland. Since autumn 1991, he is Professor at the Universität Erlangen-Nürnberg, Germany. His research interests are information and coding theory, modulation schemes, high rate baseband transmission, algorithms for signal detection and adaptive equalization for channels with severe intersymbol interference and channel coding and signaling for multiple-input multiple-output (MIMO) channels. Since 2001, he is Chairman of the Department of Electrical, Electronic and Communication Engineering of the University of Erlangen-Nuremberg.
In 1996, Prof. Huber was appointed member of the Committee 5.1 "Information und System Theory" of the German Information Technology Society (ITG). From 1999 to 2001 he served as chairman of this committee.
Since 1994 Prof. Huber has been on the Editorial Board of the International Journal on Electronics and Communications (AEÜ) and between 1997 and 2002 he was Editor-in-Chief of this journal. From 1996 to 1999 he was an Associate Editor of the IEEE Transactions on Communications. He was member of the board of governors of the IEEE Information Theory Society from January 1999 to December 2001. He has been Co-Chairman and member of the program committee of several international conferences on information theory and coding.
Prof. Huber received the Research Award of the German Information Technology Society (ITG) in 1988.