Minimum Mean Square Error Vector Precoding

David A. Schmidt, Michael Joham, Wolfgang Utschick
Institute for Circuit Theory and Signal Processing, Munich University of Technology, 80333 Munich, Germany
Phone: +49 (89) 289-28510, Email: Michael.Joham@nws.ei.tum.de

Abstract— **We derive the *minimum mean square error* (MMSE) solution to vector precoding for frequency flat multiuser scenarios with a centralized multi-antenna transmitter. The receivers employ a modulo operation, giving the transmitter the additional degree of freedom to choose a *perturbation vector.* Similar to existing vector precoding techniques, the optimum perturbation vector is found with a closest point search in a lattice. The proposed MMSE vector precoder does not, however, search for the perturbation vector resulting in the lowest transmit energy, as proposed in all previous contributions on vector precoding, but finds an optimum compromise between noise enhancement and residual interference. We present simulation results showing that the proposed technique outperforms existing vector precoders, as well as the MMSE *Tomlinson-Harashima precoder.***

I. Introduction

In the case of *non-cooperative receivers* (*broadcast channel*), the only way to achieve the maximum possible performance is through the application of *precoding*. *Linear precoding* (see e. g. [1], [2], [3]) is attractive due to its simplicity, i. e. the data signal is linearly transformed at the transmitter and the received signal is only weighted with a scalar before quanitization. However, it is clearly outperformed by the nonlinear *Tomlinson-Harashima precoding* (THP), which was employed for MIMO channels with non-cooperative receivers in [4], [5], [6], [7]. The advantage of THP compared to linear precoding results from the introduction of a modulo operator inside the feedback loop at the transmitter, limiting the amplitude of the transmit signal. To counteract the operation of the modulo operator at the transmitter, the receivers also have to apply modulo operators before quantization.

The modulo operators at the receivers allow for a more general choice of the additive *perturbation signal* than done by THP, as was highlighted by Hochwald, Peel, and Swindlehurst in [8]. They proposed to use a linear transformation at the transmitter, whose input is the data signal superimposed with a perturbation signal with properties known from THP, i. e. its entries are integer multiples of the modulo constant. First, the linear transformation is chosen and kept fixed, e. g. following a zero-forcing criterion resulting in the weighted channel pseudoinverse. In a second heuristic step, the perturbation signal is optimized to minimize the transmit power [8]. Since the algorithm to find the perturbation signal is closely related to the *sphere decoder* (e. g. [9]), the algorithm was named *sphere encoder* [8]. However, we follow [10] and simply call the scheme *vector precoding*. We also note that vector precoding is *spatial shaping without scrambling* (cf. e. g. [11]).

To circumvent the computational complexity necessary for vector precoding, two suboptimum approaches have been proposed. In [10], Windpassinger et al. replaced the sphere decoder necessary for finding the perturbation signal by the respective *lattice reduction aided detector*, a technique known from the receiver side [12]. Meurer et al. [13] proposed to split the symbols into groups to reduce the dimensionality of the problem for the sphere decoder.

It is well researched and understood that the *zero-forcing* type is always outperformed by the *minimum mean square error* (MMSE) type for linear precoding (see [1], [3]) and for THP (see [7]). We can expect the same for vector precoding. Therefore, a variant with a *regularized* pseudoinverse as the linear transformation at the transmitter was proposed in [8]. Again, the perturbation signal is found by heuristically minimizing the transmit power. For the application of this scheme to frequency selective MIMO systems see [14]. As noted in [8], the choice of the regularization in the pseudoinverse is an open question and the results were obtained by a trial and error procedure. In [15], an SINR criterion was used to find the regularization, but the power of some intermediate signal was minimized instead of the total transmit power.

Our contributions in this paper are:

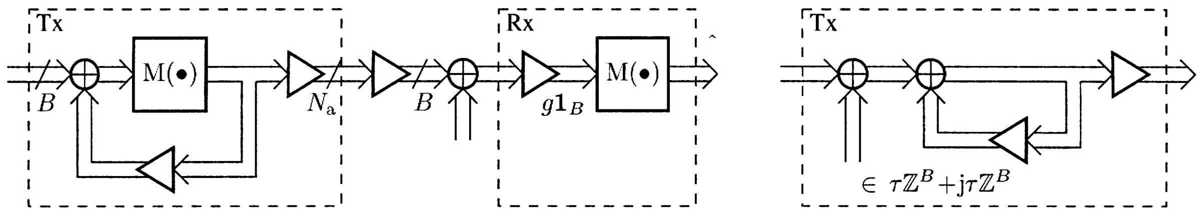
1) We base vector precoding on *one optimization* instead of the state-of-the-art two step optimization. Motivated by Forney [16, Remark 8], we use the *mean square error* (MSE) as the figure of merit.
2) We derive MMSE vector precoding, i. e. we find a closed form solution for the necessary regularization in the pseudoinverse.
3) We show that the minimization of the transmit power is not optimum in the MMSE sense.
4) By including a zero-forcing constraint, we find that the scheme of [8] is the solution for zero-forcing vector precoding. From this observation it is clear that our new MMSE vector precoding is superior to the variant in [8], since MMSE vector precoding does not have to fulfill the zero-forcing constraint.

Notation: Throughout the paper, we will denote vectors and matrices by lower and upper case bold letters, respectively. We use $\mathrm{E}[\bullet]$, $(\bullet)^*$, $(\bullet)^{\mathrm{T}}$, $(\bullet)^{\mathrm{H}}$, $\mathrm{tr}(\bullet)$, and $\mathrm{Re}(\bullet)$ for expectation, the complex conjugate, transposition, conjugate transposition, the trace of a matrix, and the real part, respectively. The M-dimensional zero vector is $\mathbf{0}_M$ and the $N \times N$ identity matrix is $\mathbf{1}_N$. We refer to the imaginary unit as j.

II. State of the Art Vector Precoding

We begin by reviewing the principle of THP for frequency flat MIMO channels (block diagram in Fig. 1(a),

978-3-8007-2909-8/05/$20.00 ©2005 IEEE



(a) Transmission Block Diagram

(b) Linear Representation of the Transmitter

Fig. 1. Spatial Tomlinson-Harashima Precoding

[5], [7]). We assume N_a antennas at the transmitter and B non-cooperative single-antenna receivers with different data streams. The transmitter uses the linear feedforward filter to give the effective channel desired characteristics. The data symbols belonging to the B streams are precoded as follows: the first symbol is transmitted unaltered; the second symbol is transmitted taking into account and subtracting the interference that will be caused by the first symbol, and so on, until the B-th symbol is transmitted with compensation for all other symbols. This successive interference cancelation can be expressed with the feedback filter matrix , which must have lower triangular structure with a zero main diagonal. Obviously, if the effective channel is lower triangular, interference-free transmission is possible.

Furthermore, every precoded symbol passes the modulo operation M(•), which maps both the real and the imaginary part of the symbol to the interval $[-\tau/2; \tau/2)$ by adding integer multiples of τ, where τ is the *modulo constant.* Let us consider, for instance, 4QAM symbols with real and imaginary parts that can have the values -1 and $+1$. We then might choose $\tau = 4$, so that the modulo output is always between -2 and $+2$. An input of $2.2 - 3.1\mathrm{j}$ would consequently be mapped to $-1.8+0.9\mathrm{j}$. In this case, the modulo operator added $-4 + 4\mathrm{j}$ to the symbol.

The receivers must apply the same modulo operation after scaling the received signal with g, to 'reverse' the effect of the transmitter modulo operation. In our example, the receiver modulo operator would add $4 - 4\mathrm{j}$, resulting in the desired data symbol, assuming that the interference was perfectly pre-subtracted and not taking into consideration the additive noise.

If we think of the modulo operation on the B data streams as the addition of an auxiliary vector signal , we can equivalently let it be located in front of the feedback loop, leading to the block diagram of the transmitter in Fig. 1(b). We must keep in mind, however, that the values of the auxiliary vector signal are determined by the modulo operation during the process of limiting the amplitude of the successively precoded signals . Nonetheless, the feedback filter and the feedforward filter now are both linear operations and can be combined into a single filter matrix $(\mathbf{1}_B - \ )^{-1}$. For *zero-forcing* THP (ZF-THP), where all interference is canceled, the resulting filter evaluates to $g^{-1}\ {}^{\mathrm{H}}(\ {}^{\mathrm{H}})^{-1}$, i. e. the scaled pseudoinverse of the channel [5], [7].

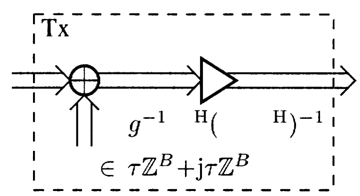


Fig. 2. Vector Precoding with Channel Inversion

This leads us to the idea of vector precoding, as proposed in [8] (cf. Fig. 2): instead of letting the successive procedure of interference cancelation and modulo operation determine the additive 'perturbation' signal, the vector precoder optimizes directly. The goal is to minimize the unscaled transmit energy, or equivalently, to minimize the noise enhancement factor g for a given transmit energy:

$$_{\mathrm{VP}} = \underset{a\in\tau\mathbb{Z}^B+\mathrm{j}\tau\mathbb{Z}^B}{\arg\min} \left\|\ {}^{\mathrm{H}}(\ {}^{\mathrm{H}})^{-1}(\ +\ )\right\|_2^2. \quad (1)$$

The optimization is equivalent to a closest point search in a lattice, of which the real valued representation can be solved with the methods in [9].

Hochwald et al. were able to further improve the performance of vector precoding by regularizing the matrix inverse in the linear pre-filter [8]. Finding the optimum regularization coefficient, however, was noted to be an open problem. When we look at the MMSE solution to THP, which does not cancel interference completely, but finds the optimum compromise between noise enhancement and interference, we notice that the combined feedforward filter resulting from the filters and in Fig. 1(b) can be expressed as $g^{-1}\ {}^{\mathrm{H}}(\ {}^{\mathrm{H}} + \xi\mathbf{1}_B)^{-1}$, where ξ denotes the inverse *signal-to-noise ratio* (SNR, see Eq. 6) [7]. The obvious heuristic approach to MMSE vector precoding would be to simply employ this regularized feedforward filter and to otherwise use the same procedure of minimizing the transmit energy as in (1), see [8]:

$$_{\mathrm{regVP}} = \underset{a\in\tau\mathbb{Z}^B+\mathrm{j}\tau\mathbb{Z}^B}{\arg\min} \left\|\ {}^{\mathrm{H}}(\ {}^{\mathrm{H}} + \xi\mathbf{1}_B)^{-1}(\ +\ )\right\|_2^2.$$

Simulations indeed show a performance improvement in most cases (cf. [8],[14]).

Surprisingly though, for low SNRs, MMSE-THP performs better than this regularized vector precoding technique

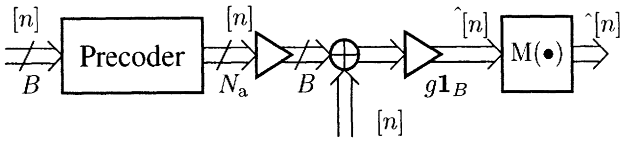


Fig. 3. System Model

(cf. Section IV), allowing the conclusion that there exists a better vector precoder, of which MMSE-THP is a constrained, and therefore suboptimum, variant. The problem with the regularized vector precoder is the fact that it does not take into account that interference between the data streams is not completely suppressed, and that different perturbation vectors result in different interferences. The additive signal that leads to the lowest unscaled transmit energy, and thus to the lowest noise gain g, might entail considerably higher interference at the receiver, even though the SNR is maximized.

In the following section, we introduce a vector precoder that minimizes the MSE, finding an optimum compromise between noise enhancement and residual interference for each symbol.

III. MMSE Vector Precoding

A. System Model

In our system (cf. Fig. 3), a base station with N_a transmit antennas serves B decentralized users with independent data streams. The data symbols of the B streams are collected in the vector

$$[n] = [s_1[n], \ldots, s_B[n]]^{\mathrm{T}} \in \mathbb{C}^B.$$

We consider the transmission of one block of data symbols of length N_B, during which symbols are transmitted with a constant scaling that is chosen to fulfill a certain average transmit energy. Based on the data symbols $[1], \ldots, [N_B]$, the precoder chooses the transmit signal vectors $[1], \ldots, [N_B]$, where

$$[n] = [y_1[n], \ldots, y_{N_a}[n]]^{\mathrm{T}} \in \mathbb{C}^{N_a}$$

contains the transmit signals at each antenna element.

The B users scale the received signal with $g \in \mathbb{R}^+$, yielding the estimates $\hat{}[n] = g \quad [n] + g \ [n] \in \mathbb{C}^B$, where $\in \mathbb{C}^{B \times N_a}$ is the frequency flat channel matrix and $[n]$ is a stationary zero-mean noise signal with the spatial covariance matrix

$$\mathrm{E}[\ [n] \ ^{\mathrm{H}}[n]] = \ _{\eta}.$$

The receivers apply the modulo operation M(•) (cf. Section II) to $\hat{}[n]$, yielding the estimate $\hat{}[n]$ of the data vector $[n]$. The modulo constant τ obviously must be large enough for both the real and imaginary part of every possible data symbol $s_i[n]$ to have a magnitude less than $\tau/2$.

B. Mean Square Error Optimization

The precoder has to perform two tasks: First, it must choose the symbols

$$[n] = \ [n] + \ [n],$$

of which the estimates at the receivers will be $\hat{}[n]$, where $[n] \in \tau\mathbb{Z}^B + \mathrm{j}\tau\mathbb{Z}^B$. As was noted in the end of Section II, the choice of the perturbation vector $[n]$ in general has an influence on the interference, except for the zero-forcing solution, where interference is always completely canceled. Second, the precoder must determine the transmit vectors $[n]$, so that the average transmit symbol energy in the block equals E_{tr} (transmit energy constraint), and the $\hat{}[n]$ are as close as possible to $[n]$. Inherent in the choice of the transmit symbols is the choice of the gain factor g.

In the following we will derive the joint MMSE or *Wiener filter* (WF) solution of $[n]$, $[n]$, and g, for a given block of data symbols $[n]$, $n = 1, \ldots, N_B$. We will then use this solution to describe the MMSE vector precoder.

We define the MSE for a given block of data symbols $[1], \ldots, [N_B]$ as

$$\varepsilon(\ [n], \ [n], g) = \frac{1}{N_B} \sum_{n=1}^{N_B} \mathrm{E}\left[\left\| \hat{}[n] - \ [n] \right\|_2^2 \ \middle| \ [n] \right].$$

Then

$$\begin{aligned} &\varepsilon(\ [n], \ [n], g) = \\ &= \frac{1}{N_B} \sum_{n=1}^{N_B} \Big(\ ^{\mathrm{H}}[n] \ [n] - 2g \operatorname{Re}(\ ^{\mathrm{H}}[n] \quad [n]) \\ &\quad + g^2 \ ^{\mathrm{H}}[n] \ ^{\mathrm{H}} \quad [n] + g^2 \operatorname{tr}(\ _{\eta}) \Big). \end{aligned} \quad (2)$$

Our optimization problem, together with the transmit energy constraint, reads as

$$\{\ _{\mathrm{WF}}[n], \ _{\mathrm{WF}}[n], g_{\mathrm{WF}}\} = \underset{\{a[n], y[n], g\}}{\operatorname{argmin}} \ \varepsilon(\ [n], \ [n], g)$$

$$\text{s.t.:} \quad \frac{1}{N_B} \sum_{n=1}^{N_B} \| \ [n] \|_2^2 = E_{\mathrm{tr}}.$$

In order to find the solution to this optimization problem, we form the Lagrangian function

$$\begin{aligned} &L(\ [n], \ [n], g, \lambda) = \\ &= \varepsilon(\ [n], \ [n], g) - \lambda \left(\frac{1}{N_B} \sum_{n=1}^{N_B} \ ^{\mathrm{H}}[n] \ [n] - E_{\mathrm{tr}} \right) \end{aligned}$$

and set the derivatives with respect to $[n]$ and g to zero:

$$\begin{aligned} \frac{\partial L(\ldots)}{\partial \ [n]} &= \frac{1}{N_B} \Big(-g \ ^{\mathrm{T}} \ ^*[n] + g^2 \ ^{\mathrm{T}} \ ^* \ ^*[n] - \lambda \ ^*[n] \Big) \\ &= \mathbf{0}_{N_a}, \end{aligned} \quad (3)$$

$$\begin{aligned} \frac{\partial L(\ldots)}{\partial g} &= \frac{1}{N_B} \sum_{n=1}^{N_B} \Big(-2 \operatorname{Re}(\ ^{\mathrm{H}}[n] \quad [n]) \\ &\quad + 2g \ ^{\mathrm{H}}[n] \ ^{\mathrm{H}} \quad [n] \Big) + 2g \operatorname{tr}(\ _{\eta}) = 0. \end{aligned} \quad (4)$$

Combining (3) and (4) with the transmit energy constraint and applying the matrix inversion lemma (e. g. [17]), we find that

$$_{\mathrm{WF}}[n] = g_{\mathrm{WF}}^{-1} \ ^{\mathrm{H}}(\quad ^{\mathrm{H}} + \xi \mathbf{1}_B)^{-1} \ [n], \quad (5)$$

$$g_{\mathrm{WF}} = \sqrt{\frac{\sum_{n=1}^{N_B} \ ^{\mathrm{H}}[n] \ (\ ^{\mathrm{H}} \quad + \xi \mathbf{1}_{N_a})^{-2} \ ^{\mathrm{H}} \ [n]}{E_{\mathrm{tr}} N_B}},$$

where

$$\xi = \frac{\mathrm{tr}(\ \ _{\eta})}{E_{\mathrm{tr}}}. \tag{6}$$

It is important to point out that the optimum transmit vector is a linearly filtered version of the desired symbol vector $[n]$ and that only the scaling g_{WF}, but not the structure of the linear filter, depends on the perturbation vector $[n]$.

Assuming that the optimum transmit vector and gain factor are employed, and making use of the matrix inversion lemma, the MSE in (2) simplifies to

$$\varepsilon(\ [n],\ _{\mathrm{WF}}[n], g_{\mathrm{WF}}) =$$
$$= \frac{\xi}{N_{\mathrm{B}}}\sum_{n=1}^{N_{\mathrm{B}}}(\ [n]+\ [n])^{\mathrm{H}}(\quad ^{\mathrm{H}}+\xi \mathbf{1}_B)^{-1}(\ [n]+\ [n]).$$

To find the optimum series of perturbation vectors $_{\mathrm{WF}}[n]$ we can minimize each summand of the MSE separately. With any matrix that fulfills

$$(\quad ^{\mathrm{H}}+\xi \mathbf{1}_B)^{-1} = \quad ^{\mathrm{H}}\ ,$$

which can be obtained e. g. via Cholesky factorization, we can rewrite the problem as

$$\begin{aligned} _{\mathrm{WF}}[n] &= \underset{\boldsymbol{a}[n]\in\tau\mathbb{Z}^B+\mathrm{j}\tau\mathbb{Z}^B}{\arg\min}\ \varepsilon(\ [n],\ _{\mathrm{WF}}[n], g_{\mathrm{WF}}) \\ &= \underset{\boldsymbol{a}[n]\in\tau\mathbb{Z}^B+\mathrm{j}\tau\mathbb{Z}^B}{\arg\min}\ \|\ (\ [n]+\ [n])\|_2^2. \end{aligned} \tag{7}$$

Obviously, the optimum choice of $[n]$ is again the solution to a closest point search in a lattice (cf. Section II). In this case, the lattice is generated by τ , and we are looking for the integer vector that corresponds to the lattice point closest to $-\ \ [n]$. Note that in general this is not the point resulting in the lowest unscaled transmit energy. Lattice searches have been shown to be NP-hard problems that grow exponentially in complexity with the dimensionality [9].

We can now specifiy the Wiener filter vector precoder as follows: for every symbol in the block, determine the perturbation vector with (7), and filter the resulting desired symbol with the regularized pseudoinverse of the channel $^{\mathrm{H}}(\quad ^{\mathrm{H}}+\xi\mathbf{1}_B)^{-1}$. Then, scale the whole block, so that the transmit energy constraint is fulfilled. The procedure is described in detail in Table I.

TABLE I

THE WIENER FILTER VECTOR PRECODER

factorize $(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}+\xi\mathbf{1}_B)^{-1} = \boldsymbol{L}^{\mathrm{H}}\boldsymbol{L}$
for $n = 1,\ldots,N_{\mathrm{B}}$:
$\quad\boldsymbol{a}_{\mathrm{WF}}[n] \leftarrow \underset{\boldsymbol{a}[n]\in\tau\mathbb{Z}^B+\mathrm{j}\tau\mathbb{Z}^B}{\arg\min} \|\boldsymbol{L}(\boldsymbol{s}[n]+\boldsymbol{a}[n])\|_2^2$
$\quad\boldsymbol{y}[n] \leftarrow \boldsymbol{H}^{\mathrm{H}}(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}+\xi\mathbf{1}_B)^{-1}(\boldsymbol{s}[n]+\boldsymbol{a}_{\mathrm{WF}}[n])$
$g_{\mathrm{WF}} \leftarrow \sqrt{\frac{1}{E_{\mathrm{tr}}N_{\mathrm{B}}}\sum_{n=1}^{N_{\mathrm{B}}}\boldsymbol{y}^{\mathrm{H}}[n]\boldsymbol{y}[n]}$
for $n = 1,\ldots,N_{\mathrm{B}}$:
$\quad\boldsymbol{y}_{\mathrm{WF}}[n] \leftarrow g_{\mathrm{WF}}^{-1}\boldsymbol{y}[n]$

C. Optimization with Zero-Forcing Constraint

Now we use the same method to derive the *zero-forcing* (ZF) vector precoder. We only need to include complete interference cancelation as an additional constraint in the optimization, which now reads as

$$\{\ _{\mathrm{ZF}}[n],\ _{\mathrm{ZF}}[n], g_{\mathrm{ZF}}\} = \underset{\{\boldsymbol{a}[n],\boldsymbol{y}[n],g\}}{\arg\min}\ \varepsilon(\ [n],\ [n], g)$$
$$\text{s.t.:}\ \frac{1}{N_{\mathrm{B}}}\sum_{n=1}^{N_{\mathrm{B}}}\|\ [n]\|_2^2 = E_{\mathrm{tr}} \quad \text{and}$$
$$g\quad [n] = \ [n],\quad n=1,\ldots,N_{\mathrm{B}}.$$

The MSE in (2) then simplifies to

$$\varepsilon(\ [n],\ [n], g) = g^2\,\mathrm{tr}(\ \ _{\eta}).$$

The method of Lagrangian multipliers leads to the solution

$$_{\mathrm{ZF}}[n] = g_{\mathrm{ZF}}^{-1}\ \ ^{\mathrm{H}}(\quad ^{\mathrm{H}})^{-1}\ [n] \quad \text{and}$$
$$g_{\mathrm{ZF}} = \sqrt{\frac{1}{E_{\mathrm{tr}}N_{\mathrm{B}}}\sum_{n=1}^{N_{\mathrm{B}}}\ ^{\mathrm{H}}[n](\quad ^{\mathrm{H}})^{-1}\ [n]}.$$

With $_{\mathrm{ZF}}[n]$ and g_{ZF}, the MSE is

$$\varepsilon(\ [n],\ _{\mathrm{ZF}}[n], g_{\mathrm{ZF}}) =$$
$$= \frac{\xi}{N_{\mathrm{B}}}\sum_{n=1}^{N_{\mathrm{B}}}(\ [n]+\ [n])^{\mathrm{H}}(\quad ^{\mathrm{H}})^{-1}(\ [n]+\ [n]).$$

We can therefore find the optimum perturbation vectors with

$$_{\mathrm{ZF}}[n] = \underset{\boldsymbol{a}[n]\in\tau\mathbb{Z}^B+\mathrm{j}\tau\mathbb{Z}^B}{\arg\min}\left\|\ \ ^{\mathrm{H}}(\quad ^{\mathrm{H}})^{-1}(\ [n]+\ [n])\right\|_2^2.$$

For blocklength $N_{\mathrm{B}} = 1$, this solution is identical to vector precoding as proposed in [8] (cf. Section II and Eq. 1).

IV. SIMULATION RESULTS

For the results in Fig. 4 and Fig. 5, we employed a channel model with i. i. d. unit variance Rayleigh fading coefficients and assumed perfect channel state information at the transmitter. $N_{\mathrm{B}} = 100$ 16QAM symbols were transmitted to each receiver per channel realization, the modulo constant τ was set to four times the distance between nearest neighbors in the symbol constellation. The Tomlinson-Harashima precoders employed the optimum precoding order of the data streams (cf. [7]). For comparison, we also included linear pre-equalization in our simulations.

The simulations show that the vector precoders (WF-VP and ZF-VP) outperform the respective Tomlinson-Harashima precoders (WF-THP and ZF-THP), especially for high SNRs, and that the WF-VP, in particular, is superior to all other schemes. For the low-dimensional case in Fig. 4, the *bit error rate* (BER) graphs of the vector precoders have a steeper slope than the respective THP BER plots. In general, the diversity order of WF-THP and ZF-THP is the same as that of the linear WF and ZF. However, we observed that with an increasing number of users the true diversity order of THP becomes visible only for very high SNRs. The vector precoder graphs,

978-3-8007-2909-8/05/$20.00 ©2005 IEEE

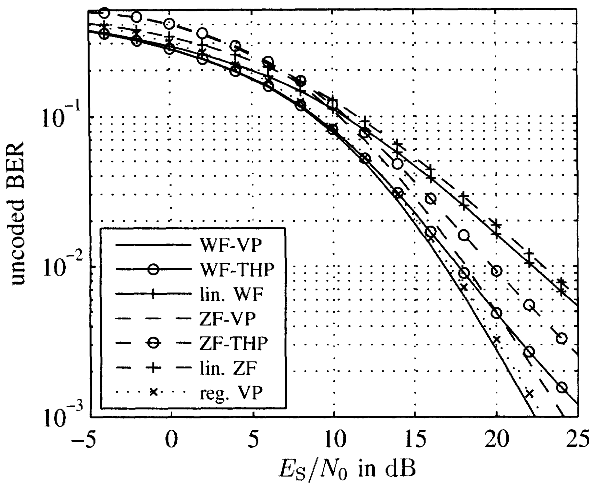


Fig. 4. $B = 2$ Users, $N_a = 2$ Transmit Antennas

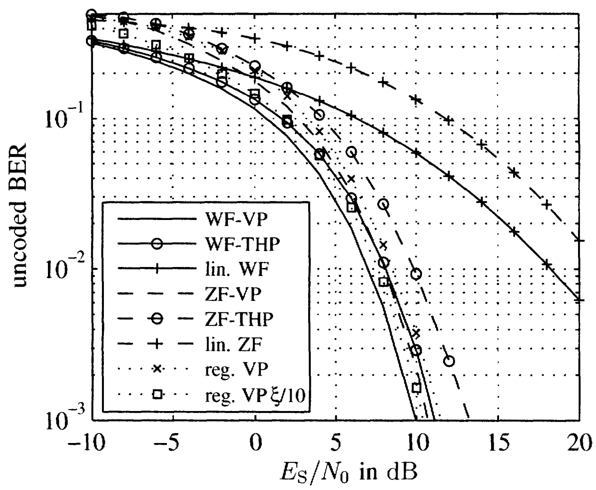


Fig. 5. $B = 10$ Users, $N_a = 10$ Transmit Antennas

on the contrary, do not exhibit an asymptotic limitation of the slope, i. e. the vector precoders achieve full diversity order. In the high-dimensional case in Fig. 5, the diversity order of THP is not visible in the shown SNR region. The VP and THP graphs run approximately parallel, with a performance difference between WF-VP and WF-THP of about 1 dB.

We can also see that the heuristic regularization of the feedforward filter as described in Section II (reg. VP) is clearly inferior to the WF-VP. In Fig. 5, the regularization even causes a performance degradation compared to ZF-VP. In [8], a regularization factor of $\xi/10$ was found to be 'optimum' for the 10 user scenario (reg. VP $\xi/10$), leading to a small improvement over ZF-VP. Note that for different numbers of users, other regularization factors perform better. The regularized vector precoders, however, seem always to perform worse than WF-THP for low SNRs.

V. Conclusion

Even though the performance gain over WF-THP that can be achieved with MMSE vector precoding seems modest—especially when considering the complexity involved in closest point searches in lattices—we believe our results to be of great relevance. On the one hand, the same approaches to complexity reduction are applicable to our method as to ZF vector precoding (e. g. [10]); we expect that a lattice-reduction-aided MMSE vector precoder will be able to achieve nearly optimum performance at a complexity comparable to THP.

On the other hand, we were able to show that our vector precoder is truly optimum in the MSE-sense. Following the arguments of [16], we therefore think that the proposed MMSE vector precoder is a strong candidate for the optimum precoder for decentralized modulo receivers.

References

[1] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A Vector-Perturbation Technique for Near-Capacity Multi-Antenna Multi-User Communication—Part I: Channel Inversion and Regularization," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 195–202, January 2005.

[2] B. W. Vojčić and W. M. Jang, "Transmitter Precoding in Synchronous Multiuser Communications," *IEEE Trans. Commun.*, vol. 46, no. 10, pp. 1346–1355, October 1998.

[3] M. Joham, W. Utschick, and J. A. Nossek, "Linear Transmit Processing in MIMO Communications Systems," 2005, *Accepted for Publication in IEEE Trans. Signal Processing.*

[4] G. Ginis and J. M. Cioffi, "A Multi-user Precoding Scheme achieving Crosstalk Cancellation with Application to DSL Systems," in *Proc. Asilomar Conference on Signals, Systems, and Computers*, vol. 2, October 2000, pp. 1627–1632.

[5] R. F. H. Fischer, C. Windpassinger, A. Lampe, and J. B. Huber, "MIMO Precoding for Decentralized Receivers," in *Proc. ISIT 2002*, June/July 2002, p. 496.

[6] J. Liu and A. Duel-Hallen, "Tomlinson-Harashima Transmitter Precoding for Synchronous Multiuser Communications," in *Proc. CISS'03*, March 2003.

[7] M. Joham, J. Brehmer, and W. Utschick, "MMSE Approaches to Multiuser Spatio-Temporal Tomlinson-Harashima Precoding," in *Proc. ITG SCC'04*, January 2004, pp. 387–394.

[8] B. M. Hochwald, C. B. Peel, and A. L. Swindlehurst, "A Vector-Perturbation Technique for Near-Capacity Multi-Antenna Multi-User Communication—Part II: Perturbation," *IEEE Trans. Commun.*, vol. 53, no. 3, pp. 537–544, March 2005.

[9] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest Point Search in Lattices," *IEEE Trans. Inform. Theory*, vol. 48, no. 8, pp. 2201–2214, August 2002.

[10] C. Windpassinger, R. F. H. Fischer, and J. B. Huber, "Lattice-Reduction-Aided Broadcast Precoding," *IEEE Trans. Commun.*, vol. 52, no. 12, pp. 2057–2060, December 2004.

[11] R. F. H. Fischer, *Precoding and Signal Shaping for Digital Transmission.* New York: John Wiley & Sons, Inc., 2002.

[12] H. Yao and G. W. Wornell, "Lattice-Reduction-Aided Detectors for MIMO Communications Systems," in *Proc. Globecom'02*, vol. 1, November 2002, pp. 424–428.

[13] M. Meurer, T. Weber, and W. Qiu, "Transmit Nonlinear Zero Forcing: Energy Efficient Receiver Oriented Transmission in MIMO CDMA Mobile Radio Downlinks," in *Proc. ISSSTA 2004*, September 2004, pp. 260–269.

[14] R. Habendorf, R. Irmer, W. Rave, and G. Fettweis, "Nonlinear Multiuser Precoding for Non-Connected Decision Regions," 2005, *Submitted to SPAWC 2005.*

[15] S. Shi and M. Schubert, "Precoding and Power Loading for Multi-Antenna Broadcast Channels," in *Proc. CISS 2004*, March 2004.

[16] G. D. Forney, Jr., "On the role of MMSE estimation in approaching the information-theoretic limits of linear Gaussian channels: Shannon meets Wiener," in *Proc. 2003 Allerton Conference*, October 2003.

[17] L. L. Scharf, *Statistical Signal Processing.* Addison-Wesley, 1991.