MMSE DFE FOR HIGH–RATE MIMO TRANSMISSION OVER CHANNELS WITH ISI

Robert F.H. Fischer

Lehrstuhl für Informationsübertragung, Friedrich–Alexander–Universität Erlangen–Nürnberg,
Cauerstrasse 7/LIT, 91058 Erlangen, Germany, Email: fischer@LNT.de

Keywords: MMSE decision-feedback equalisation, MIMO channel, intersymbol interference, multiuser interference

Abstract

Minimum mean-squared error decision-feedback equalisation (DFE) for joint spatial/temporal equalisation is derived. Here, no finite length assumptions are made but asymptotic results, i.e., ultimate limits are studied. The performance of MMSE DFE for transmission over MIMO ISI channels is assessed by means of numerical simulations for typical high-rate wireless scenarios.

1 Introduction

Using antenna arrays in transmitter and receiver, thus creating a *multiple-input/multiple-output (MIMO)* channel, spectral efficiency of a communication system can be increased dramatically. But transmitting a number of (independent) data streams in parallel, a superposition of the transmit signals is present at each receive antenna; the signals interfere and *multiuser interference (MUI)* is present. Task of the receiver in MIMO transmission schemes is to separate, i.e., equalise, the data streams and suppress or cancel MUI.

For the most part, flat fading MIMO channels are presumed in theoretical studies and in the design of practical transmission schemes. But if higher and higher data rates are desired, larger and larger bandwidths have to be used. As a result, the frequency selectivity of the channels becomes more and more important and has to be taken into account. Hence, due to the dispersive nature of the channels, *intersymbol interference (ISI)* is additionally present between pairs of transmit and receive antennas. Future high-rate transmission schemes have to deal with such MIMO ISI channels.

In this paper we derive *minimum mean-squared error (MMSE) decision-feedback equalisation (DFE)* for space-time transmission (short MIMO MMSE DFE), which performs joint spatial/temporal equalisation, as opposed, e.g., to *MIMO OFDM (orthogonal frequency division multiplexing)* systems which treat MUI (spatial equalisation) and ISI (temporal equalisation) separately. Contrary to the few derivations up to now known in literature [1], no finite-length assumptions are made but asymptotic, ultimate limits are studied.

In Section 2, the channel model is explained. The optimisation of the required matrix filters is done in Section 3. The performance of MMSE DFE for transmission over MIMO ISI channels is assessed by means of numerical simulations for typical high-rate wireless scenarios in Section 4. The MMSE DFE approach is compared to ZF (zero-forcing) DFE and to linear (ZF) equalisation. The degradation due to error propagation in the feedback loop, as well as the impact of an optimised processing order of the parallel data streams on the performance are discussed. It is shown that (MMSE) DFE is an attractive equalisation scheme for future multiple antenna systems transmitting over ISI channels.

2 Channel Model

Point-to-point transmission over a MIMO ISI channel with N transmit and $M \geq N$ receive antennas is assumed. Over each antenna a sequence $\langle a_\mu[k] \rangle$ of T-spaced data symbols $a_\mu[k]$, $k \in \mathbb{Z}$, $\mu = 1, 2, \ldots, N$, taken from a signal constellation $\mathcal{A}$ (variance $\sigma_a^2 \stackrel{\text{def}}{=} \mathrm{E}\{|a_\mu[k]|^2\}$, $\forall \mu$), is transmitted. The data symbols are combined into the vector[1] $\boldsymbol{a}[k] = [a_1[k], \ldots, a_N[k]]^{\mathsf{T}}$.

From literature [6], cf. [8, 9], it is known that the optimum receiver for transmission over MIMO ISI channels consists of a matrix of matched filters for the cascade of transmit pulse shaping filters and channel impulse responses, followed by T-spaced sampling and discrete-time processing. Hence, as an intermediate step, a discrete-time model, including transmit filters, matched filter front-end and T-spaced sampling, can be set up.[2] This end-to-end model is characterised by the $N \times N$ Hermitian matrix polynomial (i.e., $\boldsymbol{H}_{\mathrm{o}}(z) = \boldsymbol{H}_{\mathrm{o}}^{\mathsf{H}}(z^{-*})$)

$$\boldsymbol{H}_{\mathrm{o}}(z) = \sum_k \boldsymbol{H}_{\mathrm{o},k}\, z^{-k} \,. \tag{1}$$

Assuming white channel noise, the autocorrelation matrix of the additive noise $\boldsymbol{n}[k]$ in the end-to-end model (after matrix matched filtering) reads $\mathrm{E}\{\boldsymbol{n}[k+\kappa]\boldsymbol{n}^{\mathsf{H}}[k]\} \circ\!\!-\!\!\bullet\, \boldsymbol{\Phi}_{nn}(\mathrm{e}^{\mathrm{j}2\pi fT}) = \sigma_n^2\, \boldsymbol{H}_{\mathrm{o}}(\mathrm{e}^{\mathrm{j}2\pi fT})$, with $\sigma_n^2 = N_0/T$ (cf. the respective channel model including matched filter front-end in the single-input/singe-output case [3]).

Fig. 1 sketches the block diagram of the discrete-time end-to-end channel model.

[1] Notation: $\boldsymbol{A}^{\mathsf{T}}$: transpose of matrix $\boldsymbol{A}$; $\boldsymbol{A}^{\mathsf{H}}$: Hermitian (i.e., conjugate) transpose; $\boldsymbol{A}^{-\mathsf{H}}$: inverse of the Hermitian transpose of a square matrix $\boldsymbol{A}$. $\boldsymbol{I}$: Identity matrix. $\mathrm{E}\{\cdot\}$: expectation. z^{-*} abbreviates $(z^{-1})^*$. $\circ\!\!-\!\!\bullet$ symbolises a Fourier pair (time-domain signal and its Fourier transform). Matrices are printed in boldface uppercase letters, vectors in boldface lower case letters. For discrete-time vector/matrix signals, time domain and frequency domain are distinguished by the respective arguments, where $\boldsymbol{x}[k]$ denotes the time-domain signal, whereas $\boldsymbol{x}(z)$ denotes the corresponding frequency-domain signal.

[2] When starting from a discrete-time (causal) channel model with impulse response of length L, $\boldsymbol{H}_{\mathrm{C}}(z) = \sum_{k=0}^{L-1} \boldsymbol{H}_k\, z^{-k}$, rather than a continuous-time one, the end-to-end model is given by $\boldsymbol{H}_{\mathrm{o}}(z) = \boldsymbol{H}_{\mathrm{C}}^{\mathsf{H}}(z^{-*})\boldsymbol{H}_{\mathrm{C}}(z)$.

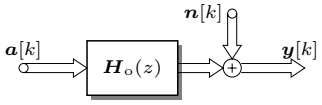

Fig. 1. Discrete-time end-to-end channel model.

Due to the dispersive nature of the channel, in addition to multiuser interference, which is present in each MIMO communication (given by the tap matrix $\boldsymbol{H}_{\mathrm{o},0}$), also intersymbol interference is present (quantified by the tap matrixes $\boldsymbol{H}_{\mathrm{o},k}$, $k \neq 0$).

3 MIMO MMSE DFE

An attractive equalisation strategy for transmission over dispersive channels is *decision-feedback equalisation (DFE)*, where already detected symbols assist in the detection of subsequent symbols. DFE is the realization of the chain rule of information theory [4]; it is known as successive cancellation in multiuser detection [5] and is the basic principle behind the BLAST detection algorithm [11].

The block diagram of DFE is shown in Fig. 2. First, the

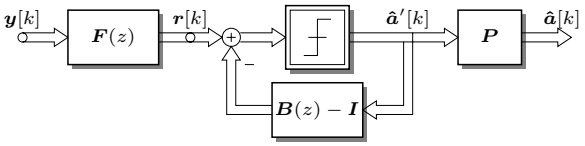

Fig. 2. Block diagram of spatial/temporal decision-feedback equalisation.

samples $\boldsymbol{y}[k]$ are passed through a (matrix) *feedforward filter* $\boldsymbol{F}(z)$. This filter does not equalise the signal, but shapes the end-to-end signal transfer function

$$\boldsymbol{H}(z) = \sum\nolimits_k \boldsymbol{H}_k\, z^{-k} \stackrel{\text{def}}{=} \boldsymbol{F}(z)\boldsymbol{H}_{\mathrm{o}}(z)\ . \tag{2}$$

Additionally, the noise is filtered ($\boldsymbol{n}'[k]$) and now has noise power spectral density

$$\begin{aligned} \boldsymbol{\Phi}_{n'n'}(\mathrm{e}^{\mathrm{j}2\pi fT}) &= \boldsymbol{F}(\mathrm{e}^{\mathrm{j}2\pi fT})\boldsymbol{\Phi}_{nn}(\mathrm{e}^{\mathrm{j}2\pi fT})\boldsymbol{F}^{\mathsf{H}}(\mathrm{e}^{\mathrm{j}2\pi fT}) \quad (3)\\ &= \sigma_n^2\, \boldsymbol{F}(\mathrm{e}^{\mathrm{j}2\pi fT})\boldsymbol{H}_{\mathrm{o}}(\mathrm{e}^{\mathrm{j}2\pi fT})\boldsymbol{F}^{\mathsf{H}}(\mathrm{e}^{\mathrm{j}2\pi fT})\ . \end{aligned}$$

Having correct decisions $\hat{\boldsymbol{a}}'[k-\kappa]$, $\kappa = 1, 2, \ldots$, on previous data symbols already available, their contribution to the ISI and MUI can be subtracted out. This is done by the *feedback filter* $\boldsymbol{B}(z) - \boldsymbol{I}$. Usually, detection in the natural order 1 through N does not give the best performance, cf. e.g., [5, 12]. The detection order is given by the permutation matrix $\boldsymbol{P}$ (non-singular matrix where each row and column contains a single one, $\boldsymbol{P}^{\mathsf{T}}\boldsymbol{P} = \boldsymbol{I}$), which corresponds to a relabelling of the transmit antennas. After decision, via $\boldsymbol{P}$, the original ordering is reestablished.

We assume the feedforward filter $\boldsymbol{F}(z)$ to be two-sided. Since, for the derivation, we admit noncausal infinite impulse responses, without loss of optimality, no decision delay has to be taken into account, i.e., decisions are based on the coefficient matrix $\boldsymbol{H}_0$ of the overall signal transfer function. The feedback filter $\boldsymbol{B}(z) - 1$ may have also infinite length, but, of course, is strictly causal. In *zero-forcing (ZF)* DFE $\boldsymbol{H}(z)$ is shaped to have only postcursors which are completely cancelled by the feedback filter, i.e., $\boldsymbol{B}(z) = \boldsymbol{H}(z)$ holds. Subsequently, we are interested in the generalisation to *minimum mean-squared error (MMSE)* DFE. The derivation follows that in the scalar case, cf. [3, 7].

3.1 Optimisation

From Figure 2, assuming correct decisions ($\hat{\boldsymbol{a}}' = \boldsymbol{a}' \stackrel{\text{def}}{=} \boldsymbol{P}^{\mathsf{T}}\boldsymbol{a}$), the error sequence (expressed by its z transform) reads

$$\begin{aligned} \boldsymbol{e}(z) &= \boldsymbol{F}(z)\boldsymbol{y}(z) - (\boldsymbol{B}(z) - \boldsymbol{I})\boldsymbol{a}'(z) - \boldsymbol{a}'(z)\\ &= \boldsymbol{F}(z)\boldsymbol{y}(z) - \boldsymbol{B}(z)\boldsymbol{a}'(z)\ . \end{aligned} \tag{4}$$

In the first step, we assume the matrix feedback filter $\boldsymbol{B}(z)$ to be given. The matrix feedforward filter $\boldsymbol{F}(z)$ is then chosen such that the mean-squared error is minimised. The orthogonality principle (cf, e.g., [13]) states that this is achieved for the present problem if the error sequence $\boldsymbol{e}(z)$ is orthogonal to the observation $\boldsymbol{y}(z)$, i.e., the cross-correlation matrix has to be the zero matrix[3]

$$\boldsymbol{\Phi}_{ey}(z) = \boldsymbol{F}(z)\boldsymbol{\Phi}_{yy}(z) - \boldsymbol{B}(z)\boldsymbol{\Phi}_{a'y}(z) \stackrel{!}{=} \boldsymbol{0}\ . \tag{5}$$

Here, the definition $\boldsymbol{\Phi}_{uv}(z) \bullet\!\!-\!\!\circ\, \mathrm{E}\{\boldsymbol{u}[k+\kappa]\boldsymbol{v}^{\mathsf{H}}[k]\}$ for any signals $\boldsymbol{u}[k]$ and $\boldsymbol{v}[k]$ has been used.

Because of the matched-filter front-end and assuming an i.i.d. data sequence $\boldsymbol{a}(z)$ ($\boldsymbol{\Phi}_{aa}(z) = \sigma_a^2\boldsymbol{I}$), the cross PSDs calculate to

$$\boldsymbol{\Phi}_{a'y}(z) = \sigma_a^2\boldsymbol{P}^{\mathsf{T}}\boldsymbol{H}_{\mathrm{o}}(z) \tag{6a}$$

$$\boldsymbol{\Phi}_{ay}(z) = \sigma_a^2\boldsymbol{H}_{\mathrm{o}}(z) \tag{6b}$$

and

$$\begin{aligned} \boldsymbol{\Phi}_{yy}(z) &= \sigma_a^2\boldsymbol{H}_{\mathrm{o}}^2(z) + \sigma_n^2\boldsymbol{H}_{\mathrm{o}}(z)\\ &= \sigma_a^2\boldsymbol{H}_{\mathrm{o}}(z)\cdot\underbrace{(\boldsymbol{H}_{\mathrm{o}}(z) + \tfrac{\sigma_n^2}{\sigma_a^2}\boldsymbol{I})}\\ &\stackrel{\text{def}}{=} \sigma_a^2\boldsymbol{H}_{\mathrm{o}}(z)\cdot \qquad \boldsymbol{\Phi}_{ff}(z)\ . \end{aligned} \tag{6c}$$

Solving (5) for $\boldsymbol{F}(z)$ using (6a) and (6c) we have[4]

$$\begin{aligned} \boldsymbol{F}(z) &= \boldsymbol{B}(z)\boldsymbol{\Phi}_{a'y}(z)\boldsymbol{\Phi}_{yy}^{-1}(z)\\ &= \boldsymbol{B}(z)\boldsymbol{P}^{\mathsf{T}}\boldsymbol{H}_{\mathrm{o}}(z)\boldsymbol{\Phi}_{ff}^{-1}(z)\boldsymbol{H}_{\mathrm{o}}^{-1}(z)\\ &= \boldsymbol{B}(z)\boldsymbol{P}^{\mathsf{T}}\boldsymbol{\Phi}_{ff}^{-1}(z)\ , \end{aligned} \tag{7}$$

and the error sequence is given by:

$$\begin{aligned} \boldsymbol{e}(z) &= \boldsymbol{B}(z)\left[\boldsymbol{P}^{\mathsf{T}}\boldsymbol{\Phi}_{ff}^{-1}(z)\boldsymbol{y}(z) - \boldsymbol{P}^{\mathsf{T}}\boldsymbol{a}(z)\right]\\ &\stackrel{\text{def}}{=} \boldsymbol{B}(z)\boldsymbol{P}^{\mathsf{T}}\boldsymbol{e}'(z)\ . \end{aligned} \tag{8}$$

For the PSD of the newly defined error sequence $\boldsymbol{e}'[k] \circ\!\!-\!\!\bullet\, \boldsymbol{e}'(z)$, we obtain (note that $\boldsymbol{H}_{\mathrm{o}}(z)$ and $\boldsymbol{\Phi}_{ff}(z)$ are Hermitian matrix polynomials)

$$\begin{aligned} \boldsymbol{\Phi}_{e'e'}(z) &= \boldsymbol{\Phi}_{ff}^{-1}(z)\boldsymbol{\Phi}_{yy}(z)\boldsymbol{\Phi}_{ff}^{-\mathsf{H}}(z^{-*}) - \boldsymbol{\Phi}_{ff}^{-1}(z)\boldsymbol{\Phi}_{ya}(z)\\ &\qquad - \boldsymbol{\Phi}_{ay}(z)\boldsymbol{\Phi}_{ff}^{-\mathsf{H}}(z^{-*}) + \sigma_a^2\boldsymbol{I}\\ &= \sigma_a^2\big(\boldsymbol{I} - \boldsymbol{H}_{\mathrm{o}}(z)\boldsymbol{\Phi}_{ff}^{-1}(z)\big)\\ &= \sigma_a^2\big(\boldsymbol{\Phi}_{ff}(z) - \boldsymbol{H}_{\mathrm{o}}(z)\big)\boldsymbol{\Phi}_{ff}^{-1}(z)\\ &= \sigma_n^2\boldsymbol{\Phi}_{ff}^{-1}(z)\ . \end{aligned} \tag{9}$$

[3] For $\boldsymbol{q}(z) \stackrel{\text{def}}{=} \boldsymbol{F}(z)\boldsymbol{y}(z) \bullet\!\!-\!\!\circ\, \boldsymbol{q}[k] = \boldsymbol{F}[k] * \boldsymbol{y}[k]$ straightforward calculations give $\mathrm{E}\{\boldsymbol{q}[k+\kappa]\boldsymbol{a}^{\mathsf{H}}[k]\} \circ\!\!-\!\!\bullet\, \boldsymbol{F}(z)\boldsymbol{\Phi}_{ya}(z) = \boldsymbol{\Phi}_{qa}(z)$.

[4] Assuming $\boldsymbol{H}_{\mathrm{o}}^{-1}(z)$ exists, it is easy to show that the following holds: $\boldsymbol{H}_{\mathrm{o}}(z)\,(\boldsymbol{H}_{\mathrm{o}}(z) + \zeta\boldsymbol{I})^{-1}\,\boldsymbol{H}_{\mathrm{o}}^{-1}(z) = (\boldsymbol{H}_{\mathrm{o}}(z) + \zeta\boldsymbol{I})^{-1}$.

Following prediction theory, in the optimum, $\boldsymbol{e}(z)$ is a white sequence. Hence, regarding (8), $\boldsymbol{B}(z)\boldsymbol{P}^{\mathsf{T}}$ has to be the matrix whitening filter for $\boldsymbol{e}'(z)$ (or $\boldsymbol{B}(z)$ has to be the matrix whitening filter for $\boldsymbol{P}^{\mathsf{T}}\boldsymbol{e}'(z)$). For this, we decompose $\boldsymbol{\Phi}_{ff}(z)$ by using a spectral factorisation according to (cf. [17, 18])

$$\begin{aligned}\boldsymbol{P}^{\mathsf{T}}\boldsymbol{\Phi}_{ff}(z)\boldsymbol{P} &= \boldsymbol{P}^{\mathsf{T}}\big(\boldsymbol{H}_{\mathrm{o}}(z)+\tfrac{\sigma_n^2}{\sigma_a^2}\boldsymbol{I}\big)\boldsymbol{P}\\ &= \boldsymbol{S}^{\mathsf{H}}(z^{-*})\cdot\boldsymbol{\Sigma}\cdot\boldsymbol{S}(z)\;, \qquad (10)\end{aligned}$$

where $\boldsymbol{\Sigma}=\mathbf{diag}(\varsigma_1,\ldots,\varsigma_N)$ is a real-valued diagonal matrix, $\boldsymbol{P}$ the permutation matrix, and the matrix polynomial $\boldsymbol{S}(z)=\sum_{k\geq 0}\boldsymbol{S}_k\,z^{-k}$ is causal and minimum-phase, i.e., $\det(\boldsymbol{S}(z))\neq 0$, $|z|\geq 1$. Moreover, $\boldsymbol{S}_0$ has to be lower triangular with unit main diagonal ("monic" matrix polynomial $\boldsymbol{S}(z)$). With this additional constraint, the factorisation (10) is unique and it exists as long as $\log\big(\det\big(\boldsymbol{H}_{\mathrm{o}}(\mathrm{e}^{\mathrm{j}2\pi fT})\big)\big)$ is absolutely integrable [18]. Efficient algorithms for solving this factorisation task exists, e.g., [14, 9, 10].

Since the feedback filter in (MMSE) DFE should be a monic matrix polynomial like $\boldsymbol{S}(z)$, we have

$$\boldsymbol{B}(z) = \boldsymbol{S}(z)\;, \qquad (11a)$$

which, using (7) and (10), gives the corresponding matrix feedforward filter

$$\begin{aligned}\boldsymbol{F}(z) &= \boldsymbol{S}(z)\boldsymbol{P}^{\mathsf{T}}\boldsymbol{\Phi}_{ff}^{-1}(z)\\ &= \boldsymbol{S}(z)\boldsymbol{P}^{\mathsf{T}}\,\boldsymbol{P}\boldsymbol{S}^{-1}(z)\boldsymbol{\Sigma}^{-1}\boldsymbol{S}^{-\mathsf{H}}(z^{-*})\boldsymbol{P}^{\mathsf{T}}\\ &= \boldsymbol{\Sigma}^{-1}\boldsymbol{S}^{-\mathsf{H}}(z^{-*})\boldsymbol{P}^{\mathsf{T}}\;. \qquad (11b)\end{aligned}$$

Using these matrix filters, the PSD of the error sequence $\boldsymbol{e}(z)$ at the input of the slicers calculates to

$$\begin{aligned}\boldsymbol{\Phi}_{ee}(z) &= \boldsymbol{B}(z)\boldsymbol{P}^{\mathsf{T}}\boldsymbol{\Phi}_{e'e'}(z)\boldsymbol{P}\boldsymbol{B}^{\mathsf{H}}(z^{-*})\\ &= \boldsymbol{S}(z)\boldsymbol{P}^{\mathsf{T}}\sigma_n^2\boldsymbol{\Phi}_{ff}^{-1}(z)\boldsymbol{P}\boldsymbol{S}^{\mathsf{H}}(z^{-*})\\ &= \sigma_n^2\boldsymbol{\Sigma}^{-1}\;. \qquad (12)\end{aligned}$$

It is noteworthy that the results for MMSE DFE are almost identical to that of ZF-DFE [9]. The spectral factorisation is only amended by the additional term $(\sigma_n^2/\sigma_a^2)\boldsymbol{I}$. Hence, for fixed channel realization and high SNRs ($\sigma_n^2/\sigma_a^2\to 0$) the results for the MMSE criterion tend to the ZF solution.

3.2 Unbiased MIMO MMSE DFE

In summary, applying the MMSE-DFE matrix feedforward filter, the filtered receive sequence (in z domain) calculates to

$$\begin{aligned}\boldsymbol{r}(z) &= \boldsymbol{F}(z)\boldsymbol{y}(z)\\ &= \boldsymbol{\Sigma}^{-1}\boldsymbol{S}^{-\mathsf{H}}(z^{-*})\boldsymbol{P}^{\mathsf{T}}\left(\boldsymbol{H}_{\mathrm{o}}(z)\boldsymbol{a}(z)+\boldsymbol{n}(z)\right)\\ &= \boldsymbol{\Sigma}^{-1}\boldsymbol{S}^{-\mathsf{H}}(z^{-*})\boldsymbol{P}^{\mathsf{T}}\big(\big(\boldsymbol{P}\boldsymbol{S}^{\mathsf{H}}(z^{-*})\boldsymbol{\Sigma}\boldsymbol{S}(z)\boldsymbol{P}^{\mathsf{T}}\\ &\qquad -\tfrac{\sigma_n^2}{\sigma_a^2}\boldsymbol{I}\big)\boldsymbol{a}(z)+\boldsymbol{n}(z)\big)\\ &= \boldsymbol{S}(z)\boldsymbol{P}^{\mathsf{T}}\boldsymbol{a}(z)+\boldsymbol{\Sigma}^{-1}\boldsymbol{S}^{-\mathsf{H}}(z^{-*})\boldsymbol{P}^{\mathsf{T}}\boldsymbol{n}(z)\\ &\qquad \underbrace{-\tfrac{\sigma_n^2}{\sigma_a^2}\boldsymbol{\Sigma}^{-1}\boldsymbol{S}^{-\mathsf{H}}(z^{-*})\boldsymbol{P}^{\mathsf{T}}\boldsymbol{a}(z)}_{\boldsymbol{e}(z)}\;. \qquad (13)\end{aligned}$$

The filtered receive sequence is thus decomposed into three parts: First, the data sequence, permuted with $\boldsymbol{P}^{\mathsf{T}}$ and filtered with $\boldsymbol{S}(z)$; second, an additive Gaussian noise sequence; and third, the residual, anticausal intersymbol interference. Since $\boldsymbol{S}^{\mathsf{H}}(z^{-*})$ is monic and anticausal, its inverse $\boldsymbol{S}^{-\mathsf{H}}(z^{-*})$ also has these properties. Thus, the biasing term $-(\sigma_n^2/\sigma_a^2)\boldsymbol{\Sigma}^{-1}$ is present at the decision point. In order to get an unbiased receiver, we rewrite (13) as

$$\begin{aligned}\boldsymbol{r}(z) &= \big(\boldsymbol{S}(z)-\tfrac{\sigma_n^2}{\sigma_a^2}\boldsymbol{\Sigma}^{-1}\big)\boldsymbol{P}^{\mathsf{T}}\boldsymbol{a}(z)\\ &\qquad +\tfrac{\sigma_n^2}{\sigma_a^2}\boldsymbol{\Sigma}^{-1}\big(\boldsymbol{I}-\boldsymbol{S}^{-\mathsf{H}}(z^{-*})\big)\boldsymbol{P}^{\mathsf{T}}\boldsymbol{a}(z)\\ &\qquad +\boldsymbol{\Sigma}^{-1}\boldsymbol{S}^{-\mathsf{H}}(z^{-*})\boldsymbol{P}^{\mathsf{T}}\boldsymbol{n}(z)\\ &= \big(\boldsymbol{I}-\tfrac{\sigma_n^2}{\sigma_a^2}\boldsymbol{\Sigma}^{-1}\big)\boldsymbol{S}'(z)\boldsymbol{P}^{\mathsf{T}}\boldsymbol{a}(z)\\ &\qquad +\tfrac{\sigma_n^2}{\sigma_a^2}\boldsymbol{\Sigma}^{-1}\big(\boldsymbol{I}-\boldsymbol{S}^{-\mathsf{H}}(z^{-*})\big)\boldsymbol{P}^{\mathsf{T}}\boldsymbol{a}(z)\\ &\qquad +\boldsymbol{\Sigma}^{-1}\boldsymbol{S}^{-\mathsf{H}}(z^{-*})\boldsymbol{P}^{\mathsf{T}}\boldsymbol{n}(z)\;, \qquad (14)\end{aligned}$$

with

$$\begin{aligned}\boldsymbol{S}'(z) &= \big(\boldsymbol{I}-\tfrac{\sigma_n^2}{\sigma_a^2}\boldsymbol{\Sigma}^{-1}\big)^{-1}\cdot\big(\boldsymbol{S}(z)-\tfrac{\sigma_n^2}{\sigma_a^2}\boldsymbol{\Sigma}^{-1}\big)\\ &= \big(\boldsymbol{\Sigma}-\tfrac{\sigma_n^2}{\sigma_a^2}\boldsymbol{I}\big)^{-1}\cdot\big(\boldsymbol{\Sigma}\boldsymbol{S}(z)-\tfrac{\sigma_n^2}{\sigma_a^2}\boldsymbol{I}\big) \qquad (15)\end{aligned}$$

Like $\boldsymbol{S}(z)$, $\boldsymbol{S}'(z)$ is causal, minimum-phase[5] and monic. Since only the coefficient at $k=0$ is changed, the feedback part of the DFE remains unchanged; it cancels the term $(\boldsymbol{S}(z)-\boldsymbol{I})\boldsymbol{P}^{\mathsf{T}}\boldsymbol{a}(z)$. Hence, from the first term in (14), only the nondelayed contribution $\boldsymbol{I}-(\sigma_n^2/\sigma_a^2)\boldsymbol{\Sigma}^{-1}$ remains. Residual interference (second term in (14)) is strictly anti-causal. To compensate for the bias, the signal at the slicer input is scaled by the diagonal scaling matrix

$$\boldsymbol{G}=\big(\boldsymbol{I}-\tfrac{\sigma_n^2}{\sigma_a^2}\boldsymbol{\Sigma}^{-1}\big)^{-1}=\mathbf{diag}\left(\tfrac{\varsigma_\mu}{\varsigma_\mu-\sigma_n^2/\sigma_a^2}\right)\;. \qquad (16)$$

The unbiased MIMO MMSE DFE is shown in Figure 3.

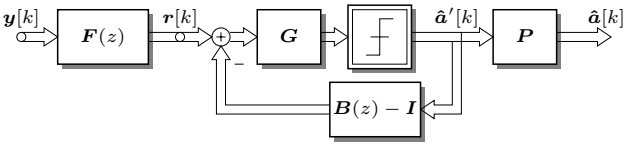

Fig. 3. Unbiased minimum mean-squared error MIMO decision-feedback equaliser.

By construction, the effective distortion sequence $\tilde{\boldsymbol{e}}(z)\stackrel{\text{def}}{=}\boldsymbol{e}(z)+(\sigma_n^2/\sigma_a^2)\boldsymbol{\Sigma}^{-1}\boldsymbol{P}^{\mathsf{T}}\boldsymbol{a}(z)$ after postcursor subtraction and prior to scaling is independent of $\boldsymbol{a}(z)$ ($\tilde{\boldsymbol{e}}(z)$ contains only past and future samples of the data sequence $\boldsymbol{a}(z)$). Due to this independence and taking $\boldsymbol{e}(z)=\tilde{\boldsymbol{e}}(z)-(\sigma_n^2/\sigma_a^2)\boldsymbol{\Sigma}^{-1}\boldsymbol{P}^{\mathsf{T}}\boldsymbol{a}(z)$ into account, the correlation matrix of the sequences are related by $\boldsymbol{\Phi}_{ee}(z)=\boldsymbol{\Phi}_{\tilde{e}\tilde{e}}(z)+\left(\sigma_n^2/\sigma_a^2\right)^2\boldsymbol{\Sigma}^{-1}\sigma_a^2\boldsymbol{\Sigma}^{-\mathsf{H}}$, and the correlation matrix of $\tilde{\boldsymbol{e}}(z)$ calculates to

$$\begin{aligned}\boldsymbol{\Phi}_{\tilde{e}\tilde{e}}(z) &= \sigma_n^2\boldsymbol{\Sigma}^{-1}-(\sigma_n^4/\sigma_a^2)\boldsymbol{\Sigma}^{-2}\\ &= \mathbf{diag}\left(\sigma_n^2(\varsigma_\mu^{-1}-\sigma_n^2/\sigma_a^2\cdot\varsigma_\mu^{-2})\right)\\ &= \mathbf{diag}\left((\sigma_n^2(\varsigma_\mu-\sigma_n^2/\sigma_a^2))/\varsigma_\mu^2\right)\;. \qquad (17)\end{aligned}$$

[5] If $\boldsymbol{S}'(z)$ would be non-minimum-phase, performance could be improved by introducing a "space-time all-pass filter". However, this filter could be accounted to the feedforward filter, which is already assumed to lead to the best performance. Hence, $\boldsymbol{S}'(z)$ is minimum-phase.

After scaling by $\boldsymbol{G} = \mathbf{diag}\left(\varsigma_\mu/(\varsigma_\mu - \sigma_n^2/\sigma_a^2)\right)$, the MSE for symbol transmitted over antenna μ is thus

$$\frac{\sigma_n^2(\varsigma_\mu - \sigma_n^2/\sigma_a^2)}{\varsigma_\mu^2} \cdot \frac{\varsigma_\mu^2}{(\varsigma_\mu - \sigma_n^2/\sigma_a^2)^2} = \frac{\sigma_n^2}{\varsigma_\mu - \sigma_n^2/\sigma_a^2} \,, \quad (18)$$

which corresponds to the signal-to-noise ratio

$$\begin{aligned} \mathrm{SNR}_\mu^{\mathsf{MMSE-DFE,u}} &= \frac{\sigma_a^2(\varsigma_\mu - \sigma_n^2/\sigma_a^2)}{\sigma_n^2} \\ &= \frac{\sigma_a^2 \varsigma_\mu}{\sigma_n^2} - 1 \\ &= \mathrm{SNR}_\mu^{\mathsf{MMSE-DFE}} - 1 \,. \end{aligned} \quad (19)$$

This result is consistent with the general statement on the relationship between unbiased and biased SNR, cf. [3, 7].

4 Numerical Results and Discussion

To assess the performance of (unbiased MMSE) MIMO DFE numerical simulations were performed. Throughout this section we assume $N = 4$ transmit and $M = 4$ receive antennas, i.e., a 4×4 MIMO ISI channel, and 16-QAM signalling over each antenna. Moreover, a block-fading MIMO ISI channel is assumed which reflects a bursty transmission over a block-wise constant radio channel. The numerical results are averaged over a large number of channel realisations. The MIMO ISI channel is normalised such that its average energy equals $N \cdot M$ (16 in the present situation). Hence its average energy equals that of a one-tap (i.e., flat-fading) MIMO channel with the usual assumption of unit-energy fading coefficients. In each case, spectral factorisation (10) is performed using the optimal permutation matrix $\boldsymbol{P}$ (for details see [10]).

First, a MIMO ISI channel with constant power-delay profile (pdp) (equal-gain test channel, T-spaced taps) is considered. The elements of the tap matrices of $\boldsymbol{H}_\mathrm{C}(z)$ are are chosen i.i.d. complex Gaussian with variance $1/L$. The length L of the impulse response is selected as $L = 2$, 4, and 8. From that, the end-to-end model is obtained as $\boldsymbol{H}_\mathrm{o}(z) = \boldsymbol{H}_\mathrm{C}^\mathsf{H}(z^{-*})\boldsymbol{H}_\mathrm{C}(z)$. Figure 4 shows the symbol error rate (SER) over the ratio of total (average) transmitted energy per information bit E_b and one-sided noise power spectral density N_0 of the channel. In the present situation $E_\mathrm{b}/N_0 = \sigma_a^2/(4\,\sigma_n^2)$ holds. For comparison, the lighter dashed curves are valid for a genie-aided receiver where perfect, i.e., error-free, feedback in the DFE loop is assumed.

A comparison of the error rate curves shows the gains due to temporal diversity. Increasing the length of the impulse response, temporal diversity increases and the performance of (MMSE) DFE gets better. Moreover, the effect of error propagation in the DFE feedback loop is visible. Unfortunately, in combined spatial/temporal DFE errors may propagate over space (parallel data streams) and time. At low signal-to-noise ratios, this effect is more pronounced for channels with longer impulse responses.

Next, the MMSE DFE approach is compared to other equalisation strategies in Figure 5. Again, a channel with constant pdp and length $L = 4$ is assumed. The figure shows the error rate curves of DFE optimised according to the MMSE criterion (biased and unbiased version) and optimised according the zero-forcing (ZF) approach, as well as linear (ZF) equalisation.

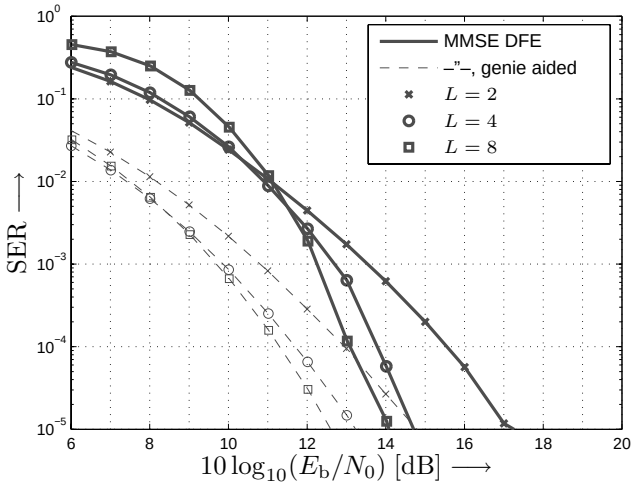

Fig. 4. Symbol error rate over signal-to-noise E_b/N_0 (in dB). Solid: (unbiased) MMSE DFE (solid); dashed: respective genie-aided DFE. $N = 4$ transmit antennas, $M = 4$ users, 16-QAM modulation. Constant power-delay profile with $L = 2$, 4, and 8 T-spaced taps; i.i.d. complex Gaussian elements of the tap matrices.

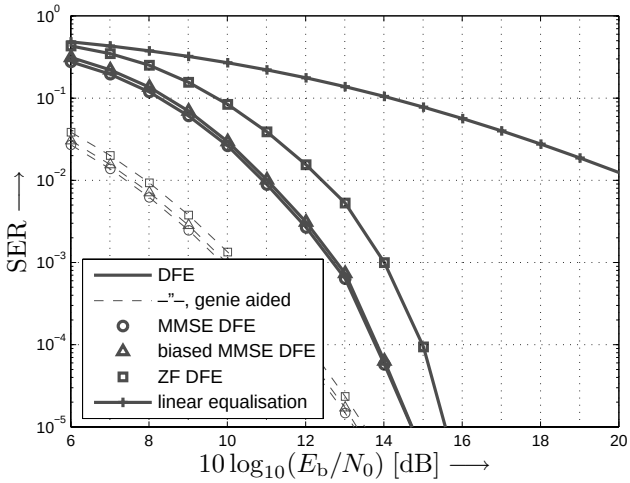

Fig. 5. Symbol error rate over signal-to-noise E_b/N_0 (in dB). Solid: biased and unbiased MMSE DFE, ZF DFE, and linear equalisation; dashed: respective genie-aided DFE. $N = 4$ transmit antennas, $M = 4$ users, 16-QAM modulation. Constant power-delay profile with $L = 4$ T-spaced taps; i.i.d. complex Gaussian elements of the tap matrices.

As expected, the MMSE approach outperforms the ZF solution. In particular, comparing the curves with the respective genie-aided curves, it is visible that the MMSE approach suffers from less error propagation than the ZF approach. Interestingly, even for high signal-to-noise ratio, MMSE and ZF curves do not converge. This can be explained as follows: for a fixed channel realisation the error rates of MMSE and ZF DFE converge as the signal-to-noise ratio tends to infinity (cf. the convergence of the respective filters (10)). When averaging over the channel realisation, the error rates are weighted with the relative frequency of occurrence (to

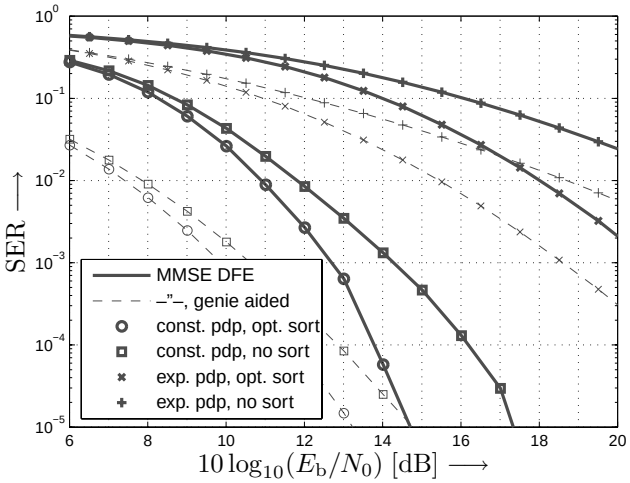

Fig. 6. Symbol error rate over signal-to-noise E_b/N_0 (in dB). Solid: MMSE DFE with different permutation matrices $\boldsymbol{P}$ in the filter calculation; dashed: respective genie-aided DFE. $N = 4$ transmit antennas, $M = 4$ users, 16-QAM modulation. Constant and exponentially decaying power-delay profiles with $L = 4$ T-spaced taps; i.i.d. complex Gaussian elements of the tap matrices.

be precise, weighted with the probability density function) of the particular channel. As a consequence, for each SNR value the average error rate is dominated by the high-error portion of the individual error rate curves. In this region, however, MMSE DFE has a clear advantage over ZF DFE. As long as averaging leads to a error rate curve with linear slope in double-logarithmic scale (as it is the case for fading channels with finite diversity order (cf., e.g., [15, 2]), the gain of MMSE at high error rates shows up over the entire SNR region. Hence, in the case of fading channels with finite diversity order—which is also true for MIMO ISI channels—the MMSE approach is rewarding compared to the ZF one for all signal-to-noise ratios and error rate curves do not converge.

Even though the gains are relatively small, in MMSE DFE the bias should always be compensated by proper scaling (matrix $\boldsymbol{G}$) prior to threshold decision. Lastly, it is evident that DFE significantly outperforms linear channel equalisation.

Finally, Figure 6 compares the results for the channel with constant pdp with that of a channel with exponentially decaying pdp, borrowed from the "Pedestrian A" power-delay profile, cf. [16]. In addition, the effect of an optimally sorted spectral factorisation in the filter calculation on the performance of MMSE DFE is visualised, too. Spectral factorisation (10) is done (i) using the optimal permutation matrix $\boldsymbol{P}$ (for details see [9, 10]) and (ii) for $\boldsymbol{P} = \boldsymbol{I}$.

Since this channel with exponentially decaying pdp does not provide as much diversity as the equal-gain channel with $L = 4$, the slope of the error rate curves is significantly smaller and poorer performance is observed. This is in accordance with Figure 4. Moreover, the influence of an optimised detection order is clearly visible. As the channel provides more and more diversity, the elements ς_μ of $\boldsymbol{\Sigma}$ in the spectral factorisation (10) become very similar, and in turn the use of an optimised detection order is of minor importance. However, as the channel with exponentially decaying pdp offers only little temporal diversity, detection in an optimised order is of much more importance than for the channel with constant pdp, for which still significant gains (more than 2 dB) are possible.

5 Summary

In this paper, MMSE DFE for joint spatial/temporal equalisation (MIMO DFE) has been discussed. The required matrix filters have been derived and the performance for some selected channels has been studied by means of numerical simulations. The results reveal the advantage of the MMSE approach over zero-forcing DFE as well as over linear zero-forcing equalisation. The impact of an optimised processing order of the parallel data streams on the performance, as well as the degradation due to error propagation in the feedback loop have been discussed. In summary it can be stated that (MMSE) DFE is an attractive equalisation scheme for future multiple antenna systems transmitting over ISI channels.

References

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

[2] E. Biglieri, J. Proakis, S. Shamai (Shitz). Fading Channels: Information-Theoretic and Communications Aspects. *IEEE Trans. Inf. Theory.*, pp. 2619–2692, Oct. 1998.

[3] J.M. Cioffi, G.P. Dudevoir, M.V. Eyuboğlu, G.D. Forney. MMSE Decision-Feedback Equalizers and Coding — Part I: Equalization Results, Part II: Coding Results. *IEEE Trans. Comm.*, pp. 2582–2604, Oct. 1995.

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

[5] A. Duel-Hallen. A Family of Multiuser Decision-Feedback Detectors for Asynchronous Code-Division Multiple-Access Channels. *IEEE Trans. Comm.*, pp. 421–434, Feb./Mar./Apr. 1995.

[6] W. Van Etten. Maximum Likelihood Receiver for Multiple Channel Transmission Systems. *IEEE Trans. Comm.*, pp. 276–283, Feb. 1976.

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

[8] R.F.H. Fischer, J.B. Huber, C. Windpassinger. Signal Processing in Decision-Feedback Equalization of Intersymbol-Interference and Multiple-Input / Multiple-Output Channels: A Unified View. *ELSEVIER Signal Processing*, Vol. 83, No. 8, pp. 1633–1642, Aug. 2003.

[9] R.F.H. Fischer, J.B. Huber. Signal Processing in Receivers for Communication over MIMO ISI Channels. *IEEE International Symposium on Signal Processing and Information Technology (ISSPIT 2003)*, Darmstadt, Germany, Dec. 2003.

[10] R.F.H. Fischer, C. Stierstorfer, J.B. Huber. Precoding for Point-to-Multipoint Transmission over MIMO ISI Channels. *International Zurich Seminar (IZS '04)*, pp. 208–211, Zurich, Switzerland, Feb. 2004.

[11] G. Ginis, J.M. Cioffi. On the Relation Between V-BLAST and the GDFE. *IEEE Comm. Letters*, pp. 364–366, Sep. 2001.

[12] G.D. Golden, G.J. Foschini, R.A. Valenzuela, P.W. Wolniansky. Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture. *Electronics Letters*, pp. 14–15, Jan. 1999.

[13] S. Haykin. *Adaptive Filter Theory*. Prentice-Hall, Inc., Englewood Cliffs, NJ, 3. edition, 1996.

[14] V. Kučera. Factorization of Rational Matrices: A Survey of Methods. *IEE Intern. Conf. on Control*, pp. 1074–1078, Edinburgh, 1991.

[15] J.G. Proakis. *Digital Communications*. McGraw-Hill, New York, 4. edition, 2001.

[16] ETSI. *Selection Procedures for the Choice of Radio Transmission Technologies of the UMTS (UMTS 30.03 version 3.2.0)*, 1998.

[17] J. Yang, S. Roy. Joint Transmitter-Receiver Opimization for Multi-Input Multi-Output Systems with Decision Feedback. *IEEE Trans. Inf. Theory.*, pp. 1334–1347, Sep. 1994.

[18] D.C. Youla, N.N. Kazanjian. Bauer-Type Factorization of Positive Matrices and the Theory of Matrix Polynomials Orthogonal on the Unit Circle, *IEEE Trans. Circ. and Systems*, pp. 57–69, Feb. 1978.