

Linear and Nonlinear Preequalization/Equalization for MIMO Systems With Long-Term Channel State Information at the Transmitter

O. Simeone, Y. Bar-Ness, and U. Spagnolini

Abstract—A transceiver structure for frequency-flat multiple-input multiple-output (MIMO) systems that comprises linear/nonlinear preequalization/equalization is optimized according to the minimum mean square error (MMSE) criterion under the assumption that only long-term channel state information (i.e., correlation matrices of fading channel and noise) is available at the transmitter. The structure generalizes different techniques known from the literature, such as BLAST, linear preequalization and equalization, and Tomlinson-Harashima precoding (THP). Simulations show that relevant benefits can be obtained by exploiting the long term channel state information at the transmitter in both dense multipath channels with relatively large correlation at the transmitter side and in sparse multipath channels.

Index Terms—Channel state information (CSI), linear preequalization, multiple-input multiple-output (MIMO), Tomlinson-Harashima precoding (THP), V-BLAST.

I. INTRODUCTION

IN order to achieve the high spectral efficiencies promised by the information theory over a radio link with multiple antennas at both the transmitter and the receiver [i.e., multiple-input multiple-output (MIMO) link], different approaches have been proposed (e.g., space-time codes [1] and V-BLAST [2]). In this letter, we are concerned with the optimization of the transceiver structure shown in Fig. 1, that operates over a frequency-flat MIMO channel, under the constraint that the channel state information (CSI) at the transmitter is limited to the second-order statistics of channel and noise [3]–[5] (long-term CSI or, in short, LT-CSI). This assumption is of crucial relevance for systems in which the fading channel is sufficiently fast-varying to make the condition of instantaneous CSI (I-CSI) at the transmitter not realistic. The LT-CSI can be acquired by the transmitter either directly from measurements of the opposite link [6] or by feedback from the receiver. On the other hand, in the design of the receiver, the instantaneous realization of the channel matrix \mathbf{H} (i.e., I-CSI) is assumed known (effects of channel estimation errors are studied by means of simulations).

Manuscript received July 31, 2002; revised November 13, 2002 and December 20, 2002; accepted January 24, 2003. The editor coordinating the review of this letter and approving it for publication is A. Swami. This work was supported in part by NSF Grant CCR-0085846.

O. Simeone is with the Center for Communications and Signal Processing Research (CCSPR), New Jersey Institute of Technology (NJIT), University Heights, Newark, NJ 07102 USA, on leave from the Dipartimento di Elettronica e Informazione, Politecnico di Milano, I-20133 Milan, Italy (e-mail: simeone@elet.polimi.it).

Y. Bar-Ness is with the Center for Communications and Signal Processing Research (CCSPR), New Jersey Institute of Technology (NJIT), University Heights, Newark, NJ 07102 USA (e-mail: barness@yegal.njit.edu).

U. Spagnolini is with the Dipartimento di Elettronica e Informazione, Politecnico di Milano, I-20133 Milan, Italy (e-mail: spagnoli@elet.polimi.it).

Digital Object Identifier 10.1109/TWC.2003.821139

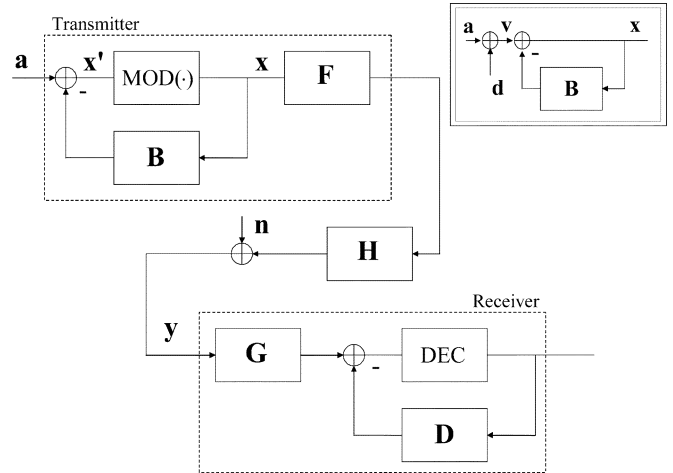


Fig. 1. Block diagram of the generalized transceiver.

Linear and nonlinear preequalization (or precoding) and equalization (or decoding) are considered in the scheme of Fig. 1. The structure reduces to known systems for specific constraints on the matrices $\{\mathbf{B}, \mathbf{F}, \mathbf{G}, \mathbf{D}\}$. For instance, imposing $\mathbf{B} = \mathbf{0}$ and $\mathbf{F} = \mathbf{I}$ the scheme reduces to a decision feedback equalizer (or equivalently to the V-BLAST receiver without optimal ordering [2], [7]); for $\mathbf{B} = \mathbf{0}$ and $\mathbf{D} = \mathbf{0}$ we have the linear preequalization-linear equalization structure (LP-LD), studied in [8], [9], under the assumption of I-CSI at both the transmitter and the receiver; for $\mathbf{F} = \mathbf{I}$ and $\mathbf{D} = \mathbf{0}$ the Tomlinson-Harashima precoding (THP) structure proposed in [10] and studied for LT-CSI at the transmitter in [4] is obtained.

The letter is organized as follows. The transceiver scheme depicted in Fig. 1 is described in Section II along with the basic assumptions. The optimal linear/non linear preequalization/equalization operators $\{\mathbf{B}, \mathbf{F}, \mathbf{G}, \mathbf{D}\}$ under the assumption of LT-CSI at the transmitter are derived and discussed in the context of the existing literature in Sections III and IV compares different structures obtained from the generalized scheme of Fig. 1 in terms of uncoded symbol error probability (SER) by means of simulation.

II. SIGNAL MODEL

We focus on a MIMO wireless link with an equal number of transmit and receive antennas N . The $N \times 1$ data vector \mathbf{a} (the time dependence of all the variables is implied) is composed of complex symbols taken from the M -QAM constellation, i.e., each entry a_i ($i = 1, \dots, N$) belongs to the set $\mathcal{A} = \{a^I + ja^Q | a^I, a^Q \in \{\pm 1, \pm 3, \dots, \pm \sqrt{M} - 1\}\}$. The data vector is passed through the nonlinear part of the precoder defined by the $N \times N$ strictly upper triangular matrix \mathbf{B} (i.e., $B_{ii} = 0$). In order to stabilize the precoder, or equivalently

to limit the dynamic range of the precoded sequence, a non-linear modulo-arithmetic operation (*MOD*) is introduced, as it is done in THP (see, e.g., [11], [12]). This operation performs a periodic mapping (or modulo reduction) of its input \mathbf{x}' on the square region of the complex plane that contains \mathcal{A} and has side length $2\sqrt{M}$, i.e., $\mathcal{R} = \{x^I + jx^Q | x^I, x^Q \in (-\sqrt{M}, \sqrt{M}]\}$. In other words, $x_i = \text{MOD}(x'_i) = x'_i + 2\sqrt{M}k_i$, where the real and imaginary parts of k_i are integers chosen to reduce $x_i \in \mathcal{R}$. Notice that there is only one k_i that satisfies this condition. It follows that the nonlinear part of the precoder can be equivalently redrawn by deleting the block *MOD* and adding at the input an input-dependent vector \mathbf{d} (see box in Fig. 1) such that $d_i = 2k_i\sqrt{M}$. Therefore, the effective symbols input to the non-linear precoder are $\mathbf{v} = \mathbf{a} + \mathbf{d}$. After linear preequalization with the $N \times N$ matrix \mathbf{F} and propagation through the $N \times N$ radio channel \mathbf{H} , the received signal can be written as

$$\mathbf{y} = \mathbf{H}\mathbf{F}\mathbf{x} + \mathbf{n} \quad (1)$$

where the circularly symmetric Gaussian noise has correlation $E[\mathbf{nn}^H] = \mathbf{R}_n$ and $E[\mathbf{xx}^H] = \sigma_x^2\mathbf{I}$. Notice that the latter assumption, also made in [10] and [4] to make the problem tractable, is not rigorously satisfied when $\mathbf{B} \neq \mathbf{0}$ since in this case $E[\mathbf{xx}^H]$ is a function of the unknown (i.e., design target) matrix \mathbf{B} . Furthermore, the power transmitted from each antenna is assumed to be independent on N since the scalability of the system as a function of the number of transmitting antennas is not of concern of this correspondence (as opposed to, e.g., [2]). In optimizing the scheme of Fig. 1, the power constraint $E[\|\mathbf{F}\mathbf{x}\|^2] = \sigma_x^2\text{tr}\{\mathbf{F}\mathbf{F}^H\} = \sigma_x^2N$ will be imposed. This condition is clearly satisfied when no linear preequalization is employed ($\mathbf{F} = \mathbf{I}$).

According to the geometry of array and scatterers, we consider two different models for the channel matrix \mathbf{H} .

A. Diversity Model

The channel matrix \mathbf{H} is assumed to be zero-mean (Rayleigh fading) circularly symmetric complex Gaussian distributed with a separable spatial correlation function [13]. It follows that the correlation between the channel gains H_{ij} and $H_{\ell m}$, i.e., $E[H_{ij}H_{\ell m}^*]$, is given by the product of the spatial correlation at the receiver $\gamma_R(i, \ell)$ and the spatial correlation at the transmitter $\gamma_T(j, m)$ so that

$$\mathbf{H} = \mathbf{R}_R^{H/2}\mathbf{H}_w\mathbf{R}_T^{1/2} \quad (2)$$

where the correlation matrices \mathbf{R}_R and \mathbf{R}_T are defined as $R_{R,i\ell} = \gamma_R(i, \ell)$ and $R_{T,jm} = \gamma_T(j, m)$ and \mathbf{H}_w is a matrix of independent identically distributed circularly symmetric complex Gaussian variables with unit power. The notation $(\cdot)^{1/2}$ defines the Cholesky factorization. We set the Frobenius norm of the channel matrix as $E[\|\mathbf{H}\|^2] = \text{tr}\{\mathbf{R}_T\}\text{tr}\{\mathbf{R}_R\} = N$ to account for the linear increase of the SNR as a function of the number of receive antennas. For simplicity, the numerical evaluation of the performance of the presented algorithms will be carried out for an AR(1) model of the spatial correlation: $\gamma_R(i, \ell) = 1/\sqrt{N}\rho_R^{|i-\ell|}$ and $\gamma_T(j, m) = 1/\sqrt{N}\rho_T^{|j-m|}$, where $0 \leq \rho_R, \rho_T \leq 1$ are the correlation coefficients at the receiver

and transmitter side, respectively. For later use, we remark that the correlation matrix of the channel (that is related to the LT-CSI available at the transmitter, see Section III) is

$$E[\mathbf{H}^H\mathbf{H}] = \mathbf{R}_T \cdot \text{tr}\{\mathbf{R}_R\} = \sqrt{N}\mathbf{R}_T. \quad (3)$$

The diversity model is appropriate if the interelement spacing at both sides of the link is sufficiently large to decorrelate the (impinging or transmitted) wavefront and/or if the scatterers are modeled as distributed sources (as opposed to point sources).

B. Beamforming Model

The channel matrix \mathbf{H} is described by d paths, each characterized by an angle of departure $\alpha_i^{(T)}$, an angle of arrival $\alpha_i^{(R)}$ and a complex amplitude β_i (Rayleigh fading, i.e., $\beta_i \sim \mathcal{CN}(0, \Omega_i)$) so that

$$\mathbf{H} = \mathbf{A}(\boldsymbol{\alpha}^{(R)})\text{diag}(\boldsymbol{\beta})\mathbf{A}(\boldsymbol{\alpha}^{(T)})^T \quad (4)$$

where $\boldsymbol{\alpha}^{(R)} = [\alpha_1^{(R)} \dots \alpha_d^{(R)}]^T$ ($\boldsymbol{\alpha}^{(T)}$ and $\boldsymbol{\beta}$ are similarly defined) and $\mathbf{A}(\boldsymbol{\alpha})$ denotes the $N \times d$ steering matrix, i.e., for a uniform linear array with interelement spacing equal to half wavelength, its i th column reads $[1 \exp(-i\pi \sin \alpha_i) \exp(-i2\pi \sin \alpha_i) \dots \exp(-i(N-1)\pi \sin \alpha_i)]^T$. The key assumptions underlying (4) are that the scatterers can be modeled as point sources and that the inter-element spacing at both sides of the link is sufficiently small to make the (impinging or transmitted) wavefront fully correlated. The rank of $\mathbf{A}(\boldsymbol{\alpha}^{(R)})$ (or $\mathbf{A}(\boldsymbol{\alpha}^{(T)})$), i.e., $r_R = \text{rank}(\mathbf{A}(\boldsymbol{\alpha}^{(R)}))$ (or $r_T = \text{rank}(\mathbf{A}(\boldsymbol{\alpha}^{(T)}))$) measures the number resolvable angles of arrival (or departure). We set the Frobenius norm of the channel matrix as $E[\|\mathbf{H}\|^2] = N^2 \sum_{i=1}^d \Omega_i = N$ by constraining the different paths to have the same power, $\Omega_i = 1/(Nd)$. The latter assumption is considered for mathematical convenience and appears to be realistic in a frequency-flat scenario, where all the paths are likely to experience the same path loss and shadowing. For later use, we remark that the correlation matrix of the channel is

$$E[\mathbf{H}^H\mathbf{H}] = \frac{1}{d}\mathbf{A}(\boldsymbol{\alpha}^{(T)})^* \mathbf{A}(\boldsymbol{\alpha}^{(T)})^T. \quad (5)$$

At the receiver side, the signal is linearly processed by the $N \times N$ matrix \mathbf{G} , passed through the feedback loop defined by the $N \times N$ strictly upper triangular matrix \mathbf{D} and modulo reduced into \mathcal{R} if $\mathbf{B} \neq \mathbf{0}$ (not shown in Fig. 1).

III. MMSE-BASED PREEQUALIZATION AND EQUALIZATION

Here we optimize the general transceiver scheme of Fig. 1 by minimizing the mean square error (MSE) between the variables at the input of the decision device and the effective data symbols $\mathbf{v} = \mathbf{a} + \mathbf{d}$ [10]. As previously discussed, we constrain the design of the operators at the transmitter side, i.e., of the matrices \mathbf{F} and \mathbf{B} , to be based only on LT-CSI, represented by the second-order statistics of channel and noise. In particular, the transmitter is given only the correlation matrix $E[\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{H}]$. On the other hand, the operators \mathbf{G} and \mathbf{D} at the receiver side are allowed to depend directly on the I-CSI, i.e., on the channel matrix \mathbf{H} . Furthermore, we will assume perfect error recovery at

the output of the decision device as it is usually done in the literature on decision feedback equalization. The effect of error propagation will be investigated in Section IV through simulations. To conclude the set of hypotheses, we recall that the transmitted power is constrained as $E[\|\mathbf{F}\mathbf{x}\|^2] = \sigma_x^2 \text{tr}\{\mathbf{F}\mathbf{F}^H\} = \sigma_x^2 N$.

We now proceed with the derivation of the optimum preequalization and equalization matrices. Since the vector at the input of the decision device can be written as $\mathbf{G}\mathbf{y} - \mathbf{D}\mathbf{v}$ (recall the assumption of no error propagation made earlier), the design problem can be stated as

$$\{\mathbf{B}, \mathbf{F}, \mathbf{G}, \mathbf{D}\} = \arg \min_{\{\mathbf{B}, \mathbf{F}, \mathbf{G}, \mathbf{D}\}} E[\|\mathbf{G}\mathbf{y} - \mathbf{D}\mathbf{v} - \mathbf{v}\|^2] \quad (6)$$

we will show below how we take into account the different types of CSI's at the transmitter and the receiver. From Fig. 1 one can easily show that $\mathbf{v} = \mathbf{C}\mathbf{x}$ where $\mathbf{C} = \mathbf{I} + \mathbf{B}$, so that

$$\begin{aligned} MSE(\mathbf{B}, \mathbf{F}, \mathbf{G}, \mathbf{D}) &= E[\|\mathbf{G}\mathbf{y} - \mathbf{E}\mathbf{C}\mathbf{x}\|^2] \\ &= E[\|\mathbf{G}\mathbf{y} - \mathbf{K}\mathbf{x}\|^2] \end{aligned} \quad (7)$$

with $\mathbf{E} = \mathbf{I} + \mathbf{D}$. The upper triangular (with unit diagonal) feedback matrices \mathbf{C} and \mathbf{E} (or equivalently \mathbf{B} and \mathbf{D}) can not be independently identified using the minimum mean square error (MMSE) criterion. In the following we will thus set $\mathbf{K} = \mathbf{E}\mathbf{C}$ and restate (7) as $MSE(\mathbf{F}, \mathbf{G}, \mathbf{K})$. We remark that \mathbf{K} is still an upper triangular matrix with unit diagonal.

From the standard theory of Wiener linear filtering, we get $\mathbf{G} = \mathbf{K}\mathbf{E}[\mathbf{x}\mathbf{y}^H]E[\mathbf{y}\mathbf{y}^H]^{-1}$ and after algebraic manipulations

$$\begin{aligned} \mathbf{G} &= \mathbf{K}\mathbf{F}^H \tilde{\mathbf{H}}^H \left(\tilde{\mathbf{H}}\mathbf{F}\tilde{\mathbf{H}}^H + \frac{1}{\sigma_x^2} \mathbf{I} \right)^{-1} \cdot \mathbf{R}_n^{-H/2} \\ \mathbf{G} &= \mathbf{K}\mathbf{Q}^{-1} \mathbf{F}^H \tilde{\mathbf{H}}^H \cdot \mathbf{R}_n^{-H/2} \end{aligned} \quad (8)$$

where $\mathbf{Q} = \mathbf{F}^H \tilde{\mathbf{H}}^H \tilde{\mathbf{H}} \mathbf{F} + 1/\sigma_x^2 \mathbf{I}$ and $\tilde{\mathbf{H}} = \mathbf{R}_n^{-H/2} \mathbf{H}$. The result (8) states that the optimum linear filter at the front-end of the receiver performs the whitening of the received signal and then applies a linear operator that has the classical Wiener structure. Substituting (8) into the expression of $MSE(\mathbf{F}, \mathbf{G}, \mathbf{K})$ we obtain

$$MSE(\mathbf{F}, \mathbf{K}) = \text{tr}\{\mathbf{K}\mathbf{Q}^{-1}\mathbf{K}^H\}. \quad (9)$$

In our framework, minimizing (9) with respect to \mathbf{K} leads to different results depending on the way the matrix \mathbf{K} is factorized into the transmitter (\mathbf{C}) and receiver (\mathbf{E}) part. Here, we consider two cases:

A. Nonlinear Equalization ($\mathbf{C} = \mathbf{I} \Rightarrow \mathbf{K} = \mathbf{E}$)

Since the receiver has access to the I-CSI \mathbf{H} , the feedback matrix $\mathbf{K} = \mathbf{E}$ is obtained as

$$\mathbf{E} = \mathbf{V}\mathbf{Q}^{1/2} \quad (10)$$

where \mathbf{V} is a diagonal matrix that scales to unity the elements on the main diagonal of \mathbf{K} .

B. Nonlinear Preequalization [4] ($\mathbf{E} = \mathbf{I} \Rightarrow \mathbf{K} = \mathbf{C}$)

Since the transmitter is given only the LT-CSI $E[\tilde{\mathbf{H}}^H \tilde{\mathbf{H}}] = E[\mathbf{H}^H \mathbf{R}_n^{-1} \mathbf{H}]$ we can not minimize (9). Instead, it is reasonable to consider $E[\text{tr}\{\mathbf{K}\mathbf{Q}^{-1}\mathbf{K}^H\}]$ as the loss function. It can be

shown that $E[\text{tr}\{\mathbf{K}\mathbf{Q}^{-1}\mathbf{K}^H\}] \geq \text{tr}\{\mathbf{K}\bar{\mathbf{Q}}^{-1}\mathbf{K}^H\}$ where $\bar{\mathbf{Q}} = \mathbf{F}^H E[\tilde{\mathbf{H}}^H \tilde{\mathbf{H}}] \mathbf{F} + 1/\sigma_x^2 \mathbf{I}$ (see Appendix). Therefore, similarly to the approach of [3] and [4], we minimize the lower bound $\overline{MSE}(\mathbf{F}, \mathbf{K}) = \text{tr}\{\mathbf{K}\bar{\mathbf{Q}}^{-1}\mathbf{K}^H\}$ obtaining

$$\mathbf{C} = \mathbf{V}\bar{\mathbf{Q}}^{1/2} \quad (11)$$

where \mathbf{V} is the scaling matrix as in (10).

After substitution of (10) if $\mathbf{C} = \mathbf{I}$ or (11) if $\mathbf{E} = \mathbf{I}$ according to the two cases discussed earlier, we should in principle minimize $\overline{MSE}(\mathbf{F}, \mathbf{K}) = \text{tr}\{\mathbf{K}\bar{\mathbf{Q}}^{-1}\mathbf{K}^H\}$ with respect to the transmit preequalization matrix \mathbf{F} . To make the problem tractable and obtain a solution independent on \mathbf{K} , we minimize $\text{tr}\{\bar{\mathbf{Q}}^{-1}\}$ instead. In other words, the matrix \mathbf{F} is designed by assuming that neither nonlinear preequalization nor nonlinear equalization is included in the transceiver ($\mathbf{K} = \mathbf{I}$). Nonetheless, simulation results show that the so obtained linear precoder performs satisfactorily even for $\mathbf{K} \neq \mathbf{I}$ (see Section IV). The precoder \mathbf{F} can thus be obtained following the steps outlined in [8]. It is

$$\mathbf{F} = \mathbf{U}\Phi \quad (12)$$

where \mathbf{U} is obtained from the eigenvalue decomposition of the LT-CSI $E[\tilde{\mathbf{H}}^H \tilde{\mathbf{H}}] = \mathbf{U}\Lambda\mathbf{U}^H$ and Φ is a diagonal matrix such that

$$|\Phi_{ii}|^2 = \left(\frac{N + \sum_{n=1}^{\bar{N}} \lambda_{nn}^{-1}}{\sigma_x^2 \sum_{n=1}^{\bar{N}} \lambda_{nn}^{-1/2}} \lambda_{ii}^{-1/2} - \frac{1}{\lambda_{ii} \sigma_x^2} \right)^+ \quad (13)$$

where $(x)^+ = \max(x, 0)$ and $\bar{N} \leq N$ is such that $|\Phi_{nn}|^2 > 0$ for $n \in [1, \bar{N}]$ and $|\Phi_{nn}|^2 = 0$ for all other n .

Some remarks on the results of the optimization (8), (10)–(12) are in order. 1) In case we relax the assumption of LT-CSI at the transmitter, i.e., we allow the matrices \mathbf{C} and \mathbf{F} to depend on the instantaneous CSI, we get $\mathbf{K} = \mathbf{I}$ and \mathbf{F} and \mathbf{G} coincide with the results derived in [8]. In other words, if both sides of the link have access to the channel matrix $\tilde{\mathbf{H}}$ the setting that minimizes the MSE (6) results in linear preequalization and equalization. In this case, \mathbf{F} and \mathbf{G} are obtained from the SVD of the channel matrix $\tilde{\mathbf{H}}$ as it can be inferred from (8) and (12). 2) Setting $\mathbf{F} = \mathbf{I}$ and $\mathbf{E} = \mathbf{I}$ leads to the THP followed by a MMSE residual linear equalizer derived in 3) Following the approach of [8], the linear precoder \mathbf{F} can be obtained alternatively by minimizing $\text{tr}\{\bar{\mathbf{Q}}^{-1}\}$ subject to a peak power constraint or by maximizing the information rate, i.e., minimizing $\det\{\bar{\mathbf{Q}}^{-1}\}$. These alternatives will not be further pursued here.

IV. SIMULATION RESULTS

The performance of the precoder/decoder structure of Fig. 1 is first evaluated in terms of uncoded SER for a 16-QAM constellation ($M = 16$), $N = 8$ antennas and the diversity model. We compare the performance of the general setting with $\mathbf{K} = \mathbf{C}$ or $\mathbf{K} = \mathbf{E}$, referred to as NP-LE (nonlinear preequalization, linear equalization) and LP-NE (linear preequalization, nonlinear equalization) respectively, with the following special cases: 1) $\mathbf{B} = \mathbf{0}$, $\mathbf{F} = \mathbf{I}$: MMSE-V-BLAST

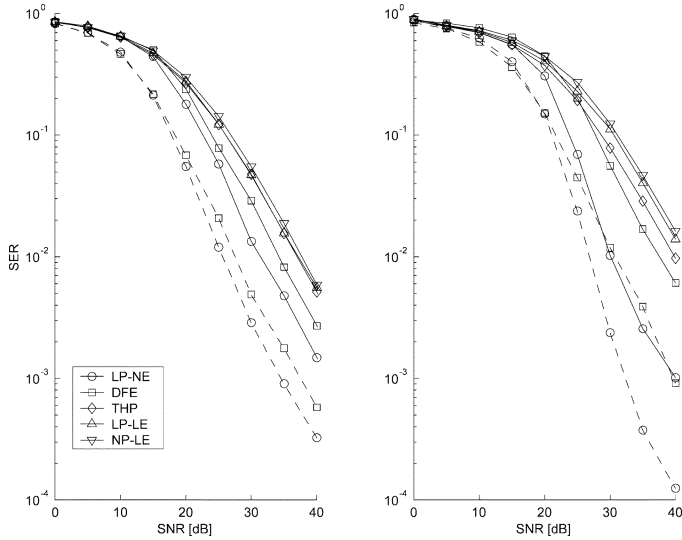


Fig. 2. SER versus SNR for $\rho_T = 0.4$ (left) and $\rho_T = 0.8$ (right) ($\rho_R = 0.4$, $N = 8$). Performance with perfect decision feedback is shown in dashed lines.

receiver (or MMSE-DFE); 2) $\mathbf{B} = \mathbf{0}$ and $\mathbf{D} = \mathbf{0}$: linear preequalization-linear equalization (LP-LE); 3) $\mathbf{F} = \mathbf{I}$ and $\mathbf{D} = \mathbf{0}$: THP with MMSE residual equalization [4]. We further limit the analysis to the spatially white noise case, i.e., $\mathbf{R}_n = \sigma_n^2 \mathbf{I}$ and the signal to noise ratio is defined as $SNR = \sigma_x^2 / \sigma_n^2$. Notice that appropriate scaling of the transmitted vector \mathbf{a} is performed to compensate for the power amplification due to nonlinear preequalization [11], [12] so that the performance comparison is based on equal total average transmitted power $N\sigma_x^2$ (see Section II). It is worth emphasizing again that all the schemes taken into account perform preequalization based on LT-CSI at the transmitter, except DFE that does not entail any processing at the transmitter side.

Nonlinear preequalization (or equalization) causes the N transmitted data streams to have different error rates. This problem can be tackled by, e.g., coding across the different streams or using more powerful codes on weaker streams. In this correspondence, we limit the analysis to uncoded transmission, leaving the issues raised by the introduction of coding in the considered scheme (e.g., soft/hard equalization, horizontal/vertical layering) to further investigations. In the following, the SER is thus averaged over the N transmitted data streams.

In Fig. 2, the SER is plotted as a function of SNR for $\rho_T = 0.4$ (left) and $\rho_T = 0.8$ (right) ($\rho_R = 0.4$). Apart from the expected performance degradation due to the decreased spatial diversity, it can be seen that the benefits (if any) of preequalization based on LT-CSI compared to the DFE receiver are more relevant for increasing values of ρ_T . This is intuitively clear since for $\rho_T = 0$ the LT-CSI $E[\hat{\mathbf{H}}^H \hat{\mathbf{H}}] = \sqrt{N}/\sigma_n^2 \mathbf{R}_T = \sqrt{N}/\sigma_n^2 \mathbf{I}$ does not bring any side information that can be exploited by the transmitter to improve the performance of the link. In this case, it is $\mathbf{F} = \mathbf{I}$ and $\mathbf{B} = \mathbf{0}$ from (12) and (11), respectively. Furthermore, it can be concluded that the LP-NE gives the best performance in terms of uncoded SER. Simulations show that similar gain can be obtained even for $SNR > 30$ dB (not shown in the figure). To study the effect of error propagation, the performance of genie-aided (i.e., perfect past decisions) DFE and LP-NE are

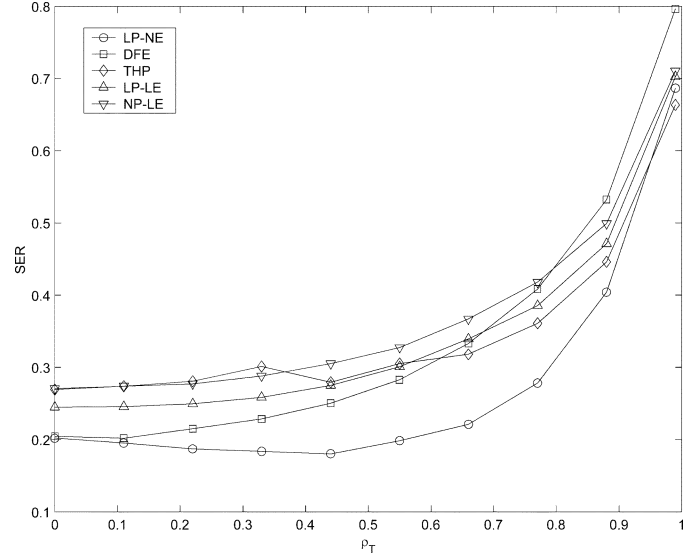


Fig. 3. Effect of spatial correlation at the transmitter side on the uncoded SER ($SNR = 20$ dB, $N = 8$, $\rho_R = 0.4$).

shown as dashed lines. It is important to remark that in comparing the performance of the schemes of interest other considerations, apart from the SER, should be taken into account. For instance, it is well known that nonlinear preequalization at the transmitter causes a relevant increase of the dynamic range at the input of the decision device that can limit its feasibility [15].

To have a clearer understanding of the role of the spatial correlation at the transmitter side on the performance of different schemes, Fig. 3 plots the uncoded SER as a function of ρ_T for $SNR = 20$ dB, $N = 8$ and $\rho_R = 0.4$. In accordance with the previous discussion, all preequalization schemes outperform the DFE for ρ_T large enough. Moreover, the LP-NE structure shows the lowest SER except for very high values of ρ_T , where it is slightly outperformed by the THP scheme.

We now want to assess the effects of an imperfect I-CSI at the receiver. To this end, we assume that for the design of \mathbf{G} (8) and \mathbf{D} (10) only a noisy version of the channel matrix is available. A conventional LS estimate of the channel is carried out at the receiver. The training sequences from all transmit antennas are assumed to be mutually orthogonal. Therefore, the estimate is unbiased and the estimated channel gains (conditioned on \mathbf{H}) are i.i.d. variables with variance $1/SNR \cdot 1/N_t$, where N_t is the length of the training sequence [16]. The SER is plotted as a function of SNR in Fig. 4 for $\rho_T = 0.4$ (left) and $\rho_T = 0.8$ (right) respectively ($\rho_R = 0.4$, $N_t = N = 8$). The same considerations discussed for perfect I-CSI apply also to the case in which the channel estimation error is taken into account, except for the performance degradation due to the imperfect I-CSI. In particular, LP-NE still gives the best performance uniformly with respect to the SNR.

Let us now consider the beamforming model. According to (5) the spatial correlation at the transmitter side is mainly related to the number of resolvable angle of departure r_T . To complement the analysis carried out in Fig. 3 for the diversity model, Fig. 5 shows the uncoded SER against r_T for $SNR = 20$ dB, $N = 8$, $r_R = 8$ and the beamforming model. Increasing r_T produces two simultaneous effects: increasing the rank of

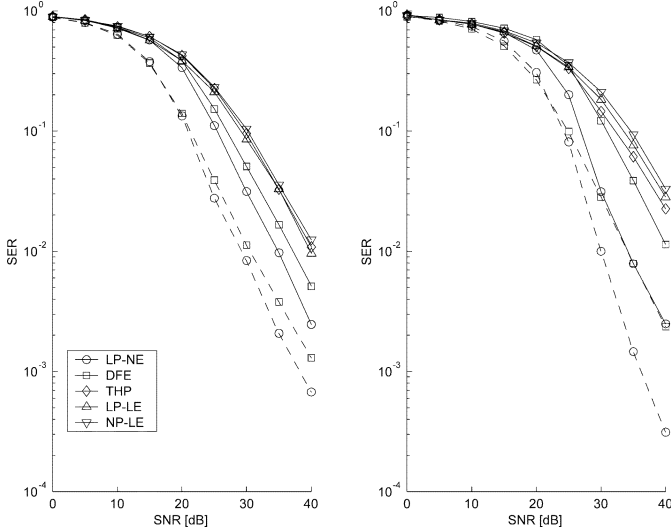


Fig. 4. SER versus SNR for $\rho_T = 0.4$ (left) and $\rho_T = 0.8$ (right) in case of imperfect I-CSI at the receiver ($\rho_R = 0.4$, $N = N_t = 8$). Performance with perfect decision feedback is shown in dashed lines.

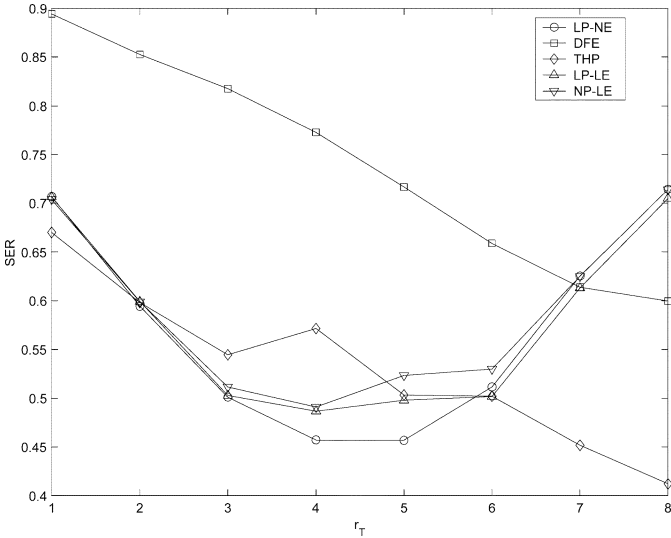


Fig. 5. SER versus r_T for the beamforming model ($SNR = 20$ dB, $N = 8$, $r_R = 8$).

the channel matrix ($\text{rank}(\mathbf{H}) = \min(r_T, r_R) = r_T$) and decreasing the spatial correlation at the transmitter side (i.e., increasing the spatial diversity). The first effect tends to reduce the SER (multiplexing gain, see, e.g., [2]) whereas the latter tends to lessen the benefits of preequalization. Accordingly the SER of DFE is monotonically decreasing while the SER of the different preequalization schemes, with the only exception of THP, presents a U-shape. For a wide range of values of r_T , LP-NE results in the lowest SER as for the diversity case. Nevertheless, for $r_T = 1$ (high spatial correlation) and $r_T \geq 7$, THP presents the best performance.

In summary, preequalization with long-term CSI appears to be advantageous in dense multipath channels (as for diversity model) with relatively large correlation at the transmitter ($\rho_T \geq 0.2$) or in sparse multipath channels (as for beamforming model). Moreover, the experimental results show that the most promising scheme is LP-NE, also considering the practical

limitations of nonlinear preequalization [15]. The preferred scheme essentially adds a linear precoder to a modified BLAST receiver, where the feedforward filter \mathbf{G} and the feedback filter \mathbf{D} are designed by taking into account the precoder \mathbf{F} according to (8) and (10).

V. CONCLUSION

A transceiver structure for frequency nonselective MIMO channels that includes linear/nonlinear preequalization/equalization has been studied under the assumption that the state information available at the transmitter is limited to the second-order statistics of channel and noise. Simulations have shown that relevant benefits can be obtained by exploiting the long-term CSI at the transmitter in both dense multipath channels with relatively large correlation at the transmitter side and in sparse multipath channels. Moreover, the preferred scheme essentially adds a linear precoder to a modified BLAST receiver.

APPENDIX

The inequality $E[\text{tr}\{\mathbf{K}\mathbf{Q}^{-1}\mathbf{K}^H\}] \geq \text{tr}\{\mathbf{K}E[\mathbf{Q}]^{-1}\mathbf{K}^H\} = \text{tr}\{\mathbf{K}\mathbf{Q}^{-1}\mathbf{K}^H\}$ directly follows from the Jensen's inequality once the function $\text{tr}\{\mathbf{K}\mathbf{Q}^{-1}\mathbf{K}^H\}$ is proved to be convex in the positive definite matrix \mathbf{Q} . To show the convexity of the function of interest, it is sufficient to demonstrate that $\mathbf{k}^H(\lambda\mathbf{Q}_1 + (1-\lambda)\mathbf{Q}_2)^{-1}\mathbf{k} \leq \lambda\mathbf{k}^H\mathbf{Q}_1^{-1}\mathbf{k} + (1-\lambda)\mathbf{k}^H\mathbf{Q}_2^{-1}\mathbf{k}$ where $0 \leq \lambda \leq 1$ and \mathbf{k} is any vector. The aforementioned condition can be stated as $(\lambda\mathbf{Q}_1 + (1-\lambda)\mathbf{Q}_2)^{-1} \leq \lambda\mathbf{Q}_1^{-1} + (1-\lambda)\mathbf{Q}_2^{-1}$ or equivalently as (th. 7.7.3 of [17])

$$\varrho((\lambda\mathbf{Q}_1 + (1-\lambda)\mathbf{Q}_2)(\lambda\mathbf{Q}_1^{-1} + (1-\lambda)\mathbf{Q}_2^{-1})) \geq 1 \quad (14)$$

where $\varrho(\cdot)$ denotes the spectral radius. After simple manipulations, we obtain

$$\varrho(\mathbf{Q}_1\mathbf{Q}_2^{-1} + \mathbf{Q}_2\mathbf{Q}_1^{-1}) \geq 2. \quad (15)$$

Since \mathbf{Q}_1 and \mathbf{Q}_2 are hermitian matrices (in particular, they are positive definite), we can find a nonsingular matrix \mathbf{Z} such that $\mathbf{Q}_1 = \mathbf{Z}\mathbf{Z}^H$ and $\mathbf{Q}_2 = \mathbf{Z}\mathbf{\Theta}\mathbf{Z}^H$, where $\mathbf{\Theta} = \text{diag}([\Theta_1 \dots \Theta_N])$ is diagonal and Θ_i is real and positive (th. 7.6.3 and 7.6.5 of [17]). It follows that $\mathbf{Q}_1\mathbf{Q}_2^{-1} + \mathbf{Q}_2\mathbf{Q}_1^{-1} = \mathbf{Z}(\mathbf{\Theta} + \mathbf{\Theta}^{-1})\mathbf{Z}^{-1}$, which implies that (15) becomes

$$\Theta_i + \frac{1}{\Theta_i} \geq 2 \quad \forall i = 1, 2, \dots, N \quad (16)$$

that is clearly satisfied $\forall \Theta_i > 0$.

REFERENCES

- [1] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar. 1998.
- [2] G. D. Golden, C. J. Foschini, R. A. Valenzuela, and P. W. Wolnianski, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture," *Electron. Lett.*, vol. 35, no. 1, pp. 14–16, Jan. 1999.
- [3] M. T. Ivrlac, T. P. Kurpjuhn, C. Brunner, and W. Utschick, "Efficient use of fading correlations in MIMO systems," in *VTC 2001*, vol. 4, 2001, pp. 2763–2767.

- [4] R. F. H. Fischer, C. Windpassinger, A. Lampe, and J. B. Huber, "Tomlinson-harashima precoding in space-time transmission for low-rate backward channel," *Int. Zurich Seminar on Broadband Commun.*, pp. 7_1–7_6, 2002.
- [5] H. Sampath and A. Paulraj, "Linear precoding for space-time coded systems," *IEEE Commun. Lett.*, vol. 6, pp. 239–241, June 2002.
- [6] B. M. Hochwald and T. L. Marzetta, "Adapting a downlink array from uplink measurements," *IEEE Trans. Signal Processing*, vol. 49, pp. 642–653, Mar. 2001.
- [7] G. Ginis and J. M. Cioffi, "On the relation between V-BLAST and the GDFE," *IEEE Commun. Lett.*, vol. 5, pp. 364–366, Sept. 2001.
- [8] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Signal Processing*, vol. 50, pp. 1051–1064, May 2002.
- [9] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers. Part I: Unification and optimal design," *IEEE Trans. Signal Processing*, vol. 47, pp. 1987–2006, July 1999.
- [10] R. F. H. Fischer, C. Windpassinger, A. Lampe, and J. B. Huber, "Space-time transmission using tomlinson-harashima precoding," in *Proc. 4 ITG Conf. Source and Channel Coding*, Jan. 2002, pp. 139–147.
- [11] H. Harashima and H. Miyakawa, "Matched-transmission technique for channels with intersymbol interference," *IEEE Trans. Commun.*, vol. 20, pp. 774–780, Aug. 1972.
- [12] M. Tomlinson, "New automatic equaliser employing modulo arithmetic," *Electron. Lett.*, pp. 138–139, March 1971.
- [13] S. D. G. J. Foschini and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, pp. 502–513, Mar. 2000.
- [14] C. Tidestav, A. Ahlen, and M. Sternad, "Realizable MIMO decision feedback equalizers: Structure and design," *IEEE Trans. Signal Processing*, vol. 49, pp. 121–133, Jan. 2001.
- [15] R. F. H. Fischer, R. Tzschoppe, and J. B. Huber, "Signal shaping for peak-power and dynamics reduction in transmission schemes employing precoding," *IEEE Trans. Commun.*, vol. 50, pp. 736–741, May 2002.
- [16] "Tech. Memo.," Bell Labs., Lucent Technol., 2000.
- [17] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, MA: Cambridge University Press, 1996.