ROBUST PRECODING FOR MULTI-USER MISO SYSTEMS WITH LIMITED-FEEDBACK CHANNELS

Paula M. Castro[1], Michael Joham[2], Luis Castedo[1], and Wolfgang Utschick[2]

1) University of A Coruña, Campus de Elviña s/n, A Coruña, Spain, 15071
Email: {pcastro,luis}@udc.es
2) Munich University of Technology, Theresienstr. 90, 80290 Munich, Germany
Email: {joham,utschick}@tum.de

ABSTRACT

In *Multi User Multiple Input Single Output* (MU-MISO) systems one centralized transmitter with multiple antennas serves several decentralized single-antenna receivers. Precoding is an attractive way to remove multiuser interferences because it reduces cost and power consumption in the user equipment. When implementing precoding, however, the base station should know the *Channel State Information* (CSI). In *Frequency Division Duplex* (FDD), this information is sent from the receivers by means of a feedback channel, whose data rate is often severely limited. In this paper we investigate the utilization of the *Karhunen-Loève* (KL) transform to compress the CSI sent through the feedback channel. We model the errors caused by channel estimation, truncation and quantization of the KL coefficients and feedback delay. This error modeling is the basic premise to design robust *Tomlinson Harashima Precoding* (THP) based on a sum *Mean Square Error* (MSE) criterion. Our results show that robust THP clearly outperforms conventional THP transmitting a minimum amount of information from the receivers.

1. INTRODUCTION

Dirty Paper Coding (DPC) [1] combined with a *Signal-to-Interference-plus-Noise Ratio* (SINR) criterion [2, 3] is an attractive signaling scheme for the downlink of *Multi User Multiple Input Single Output* (MU-MISO) systems. This is a point-to-multipoint channel which is referred to as broadcast channel in information theory [4]. In this work we focus on a low-cost and straightforward way to implement DPC, namely, *Tomlinson-Harashima Precoding* (THP) [5,6]. THP is suboptimal compared to DPC and suffers from shaping and modulo loss due to the modulo operations at the transmitter and the receivers [7]. As shown in [8], the SINR and the *Mean Square Error* (MSE) achievable regions for MU-MISO systems are

tightly related. Therefore, it is not surprising that the suboptimal sum MSE design of [9, 10] shows an excellent performance.

The design of DPC systems is already known for the ideal case where *Channel State Information* (CSI) at the transmitter is perfectly known. However, the situation is different for the case with erroneous CSI, where no scheme comparable to that of [1] has been proposed nor the capacity region has been found yet. Additionally, the application of the SINR criterion is questionable, since it is unclear up to now how to include the uncertainties in the SINR in a systematic way (see the discussion in [11] and the attempt in [12] for the case of statistical CSI). Consequently, it is inevitable to resort to an MSE criterion together with THP for the case of partial CSI, since a THP design based on the sum MSE criterion is possible, as demonstrated in [13, 14].

Most work on precoding with erroneous CSI has mainly focused on *Time Division Duplex* (TDD) systems. Contrarily, in this work we focus on the more extended case of *Frequency Division Duplex* (FDD) systems where the transmitter cannot obtain the CSI from the received signals, even under the assumption of perfect calibration, because the channels are not reciprocal. Instead, the receivers estimate their channels and send the CSI back to the transmitter by means of a feedback channel. This is a reasonable assumption since currently standardized wireless systems have control channels to implement adaptive transmit facilities such as power control or adaptive modulation. Since the data rate of the feedback channels is often limited [15], the CSI must be compressed to ensure that the tight scheduling constraints are satisfied. To limit the CSI sent to the transmitter we will use in this paper a truncation of the *Karhunen-Loève* (KL) decomposition which is optimum in the sense that it provides dimensionality reduction with the smallest possible MSE.

In this work we study the following sources of errors in the proposed precoding design: channel estimation, truncation of the KL transform, quantization of the KL coefficients, and feedback channel delay. With the obtained error model, we develop robust THP that takes into account the statistical
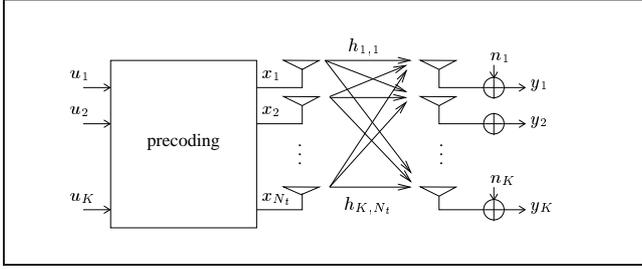
Fig. 1. Multiuser system model with precoding over flat MISO channels.

properties of the errors in the fed back CSI. Our simulation results show that using robust THP designs a considerable compression rate is possible without sacrificing performance.

This paper is organized as follows. Sections 2 and 3 describe the signal and channel models, respectively, and in Section 4, the models for the CSI error sources are developed. Section 5 presents our robust THP design. Illustrative computer simulations are presented in Section 6 and some concluding remarks are made in Section 7.

2. SYSTEM MODEL

We consider a MU-MISO system with N_{t} transmit antennas and K single antenna receivers as depicted in the Fig. 1. The precoder generates the transmit signal x from all data symbols $\{u_1, \ldots, u_K\}$ belonging to the different users $1, \ldots, K$. The signal x_ℓ from transmit antenna ℓ propagates over the channel with the coefficient $h_{k,\ell}$ to the k-th receiver, superimposes with the signals of the other transmit antennas, and is perturbed by the additive white Gaussian noise n_k with variance σ_n^2, i.e.,

$$y_k = \sum_{\ell=1}^{N_{\mathrm{t}}} h_{k,\ell} x_\ell + n_k = \boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{x} + n_k \qquad (1)$$

where $(\bullet)^{\mathrm{T}}$ denotes transpose and $\boldsymbol{h}_k = [h_{k,1}, \ldots, h_{k,N_{\mathrm{t}}}]^{\mathrm{T}} \in \mathbb{C}^{N_{\mathrm{t}} \times 1}$ represents the flat fading vector channel corresponding to the k-th user.

Figure 2 plots the block diagram of a MU-MISO system with THP. As you can see in the figure, the received signal y_k is then multiplied by a common gain factor $1/\beta$ that acts as an automatic gain control. In THP, a modulo operator is applied to the weighted received signal to remove the ambiguities introduced by the precoder (see e.g., [16]). The resulting estimate of u_k is denoted by $\hat{u}_k$.

At the transmitter, the feedforward filter $\boldsymbol{F}^{\mathrm{H}}$ suppresses parts of the interference linearly, whereas the feedback loop with the strictly lower triangular feedback filter $\boldsymbol{B}^{\mathrm{H}} - \boldsymbol{I}$ subtracts the remaining interferences non-linearly. Note that $(\bullet)^{\mathrm{H}}$ indicates conjugate transpose. Since the order of precoding has an effect on performance, the data signal $\boldsymbol{u}$ is reordered

$$
\boxed{
\begin{aligned}
&\boldsymbol{v} = \boldsymbol{P}\boldsymbol{u} \\
&\text{for } i = 1, \ldots, K \\
&\qquad \boldsymbol{v}(i) = \mathcal{M}(2\boldsymbol{v}(i) - \boldsymbol{B}^{\mathrm{H}}(:, i)\boldsymbol{v}) \\
&\boldsymbol{x} = \boldsymbol{F}^{\mathrm{H}}\boldsymbol{v}
\end{aligned}
}
$$

Table 1. Computing the feedback loop output from the permuted data.

by means of the permutation filter $\boldsymbol{P} = \sum_{i=1}^{K} \boldsymbol{e}_i \boldsymbol{e}_{k_i}^{\mathrm{T}}$, where $\boldsymbol{e}_i$ is the i-th column of the $K \times K$ identity matrix and k_i is the index of the i-th data stream to be precoded [10]. The signal $\boldsymbol{P}\boldsymbol{u}$ is passed through the feedback loop:

$$\boldsymbol{v} = \mathcal{M}\left(\boldsymbol{P}\boldsymbol{u} - \left(\boldsymbol{B}^{\mathrm{H}} - \boldsymbol{I}\right)\boldsymbol{v}\right) \qquad (2)$$

where the modulo operator $\mathcal{M}(\bullet)$ limits the amplitude of $\boldsymbol{v}$ and thus, the power of the transmit signal $\boldsymbol{x}$. Additionally, the entries of $\boldsymbol{v}$ have statistical properties which approximately only depend on the modulo constant τ (see e.g., [7]). In particular, $E[\boldsymbol{v}\boldsymbol{v}^{\mathrm{H}}] = \sigma_v^2 \boldsymbol{I}$ with $\sigma_v^2 = \tau^2/6$. The output $\boldsymbol{v}$ of the feedback loop can alternatively be computed with the pseudo code in Table 1. The signal $\boldsymbol{v}$ is transformed by the feedforward filter $\boldsymbol{F}^{\mathrm{H}}$ to get the transmit signal $\boldsymbol{x}$, which must satisfy an average total transmit power constraint, i.e., $E[\|\boldsymbol{x}(n)\|_2^2] = E_{\mathrm{tr}}$.

3. CHANNEL MODEL

We model the k-th user channel vector as a vector of zero-mean circularly symmetric complex Gaussian distributed random variables, i.e.,

$$\boldsymbol{h}_k \sim \mathcal{N}_{\mathbb{C}}\left(\boldsymbol{0}, \boldsymbol{C}_{h,k}\right) \qquad (3)$$

where $\boldsymbol{C}_{h,k}$ is the covariance matrix of the k-th user's channel. The channels of the different users are statistically independent.

In the q-th time slot, our model for the k-th user channel vector is

$$\boldsymbol{h}_k(q) = \boldsymbol{C}_{h,k}^{1/2} \boldsymbol{h}_{\mathrm{w},k}(q). \qquad (4)$$

with $\boldsymbol{h}_{\mathrm{w},k}(q)$ being a vector of stationary circularly symmetric complex white Gaussian processes (with unit variance elements) and where $(\bullet)^{1/2}$ represents the Cholesky decomposition. According to the modified Jakes model [17, 18] described in [19], temporal channel correlations are modeled by $\boldsymbol{h}_{\mathrm{w},k}(q)$ whereas the spatial correlations are introduced by the multiplication by $\boldsymbol{C}_{h,k}^{1/2}$ [20].

Notice that, according to our model, the channel $\boldsymbol{h}_k$ is stationary because $\boldsymbol{h}_{\mathrm{w},k}$ is stationary. Realistic channels are often non-stationary, e.g., either the location of the receiver or the scenario geometry can change. Thus, the channel covariance matrix has to be tracked in real situations. However, since the covariance matrix changes very slowly compared to the channel itself, it is realistic to assume that it is constant and perfectly known at both the receiver and the transmitter.
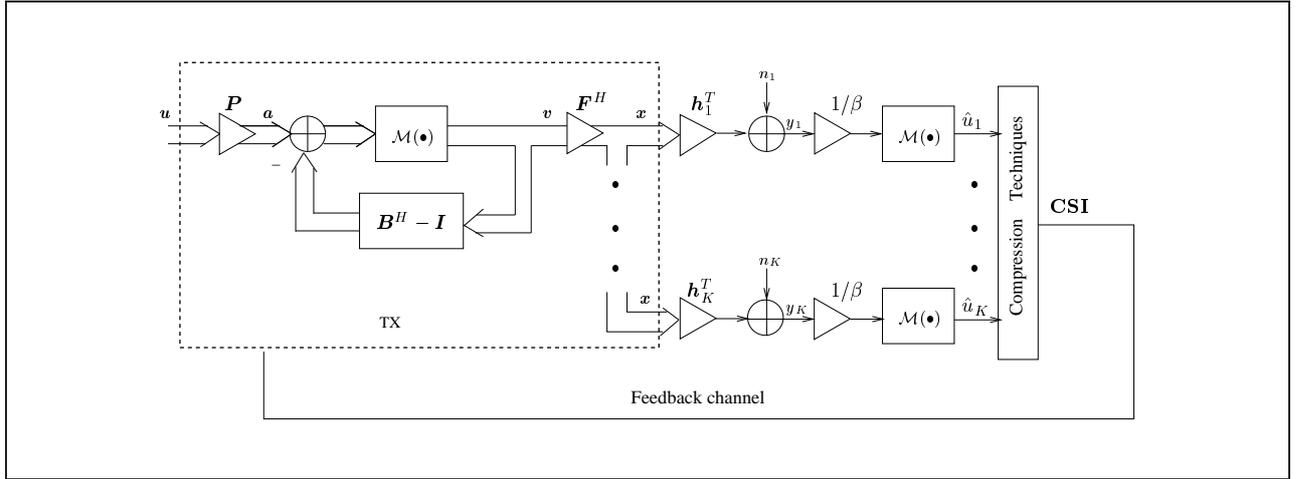
Fig. 2. Block diagram of Tomlinson-Harashima precoding.

According to [20], we have that

$$C_{h,k} = \frac{\sigma_{\mathrm{SF}}}{M} \sum_{m=1}^{M} r_{k,m} r_{k,m}^{\mathrm{H}} \qquad (5)$$

with

$$[r_{k,m}]_i = e^{j(i\pi \sin(\theta_{\mathrm{BS},k} + \theta_m))}, \quad i \in \{1, \ldots, N_t\} \qquad (6)$$

where $\theta_{\mathrm{BS},k} \in [0, 2\pi)$ is the angle of departure from the transmitter to the k-th user, θ_m (different constant values depending on the environment, see [20]) is the offset for the m-th sub-path with respect to $\theta_{\mathrm{BS},k}$ and σ_{SF} represents the log-normal shadowing, which is $8\,\mathrm{dB}$ for suburban macro-cell [20]. In [20], the number M of sub-paths per-path is 20.

4. IMPERFECT CSI

In realistic situations, the CSI that is available at the transmitter is not perfectly known. In this case, it is a matter of discussion what kind of information has to be sent to the transmitter and the way of recovering it from the receiver side. CSI can be obtained by different mechanisms in the receiver side which gives rise to more or less degradation in the final information sent through the feedback channel. In the system that we propose in this paper we start estimating the channel at the receivers using the observations of the pilot symbols, which requires knowledge about its covariance matrix. Then, we project the resulting channel estimation onto the eigenvectors of the channel covariance matrix to obtain the Karhunen-Loève transformation of the channel vector which optimally provides a dimensionality reduction with the smallest possible MSE. The coefficients of the truncated KL expansion are then quantized prior to transmission over the feedback channel which introduces a delay. We incorporate this delay in

our model considering a system without feedback delay but a delayed observation for the channel estimator. That partial CSI is then used at the transmitter to reconstruct the channel vector and to design the THP filters.

In the following subsections we describe this process with more detail and obtain the statistical description of the errors incurred at each step. Along this section we will assume that the signals and errors are uncorrelated.

4.1. Statistical model for channel estimation errors (Type A)

We use linear estimators at the receiver based on N_{tr} pilot symbols per time slot q to enable the channel vector estimation for the k-th user, i.e.,

$$\hat{h}_k(q) = W_k(q) y_k(q) \qquad (7)$$

where $y_k(q)$ reads as

$$y_k(q) = S h_k(q) + n_k(q) \qquad (8)$$

with $S \in \mathbb{C}^{N_{\mathrm{tr}} \times N_t}$ containing the training symbols [21] and $n_k(q) \in \mathbb{C}^{N_{\mathrm{tr}} \times 1}$ being the AWGN with variance σ_n^2.

We use the heuristic estimator [22] so that the least-square channel estimates are obtained when we consider the estimator $W_k(q) = S^{\dagger} = (S^{\mathrm{H}} S)^{-1} S^{\mathrm{H}}$. Therefore, the channel estimate is given by

$$h_{\mathrm{LS},k}(q) = S^{\dagger} y_k(q) = h_k(q) + S^{\dagger} n(q) = h_k(q) + n_{\mathrm{LS},k}(q) \qquad (9)$$

with

$$n_{\mathrm{LS},k}(q) \sim \mathcal{N}_{\mathbb{C}}(0, \sigma_n^2 (S^{\mathrm{H}} S)^{-1}). \qquad (10)$$

4.2. Statistical model for Karhunen–Loève errors (Type B)

The eigenvalue decomposition of $C_{h,k}$ reads as

$$C_{h,k} = E[h_k(q)h_k^{\mathrm{H}}(q)] = \sum_{i=1}^{r_k} \lambda_{k,i} v_{k,i} v_{k,i}^{\mathrm{H}} = V'_k \Lambda_k V'^{\mathrm{H}}_k \tag{11}$$

where r_k is the rank of $C_{h,k}$ and $v_{k,i}$ and $\lambda_{k,i}$ are, respectively, the i-th eigenvector (or i–th column of the matrix V'_k) and the i-th eigenvalue of $C_{h,k}$ (or the i–th entry of the diagonal matrix Λ_k).

Applying the KL transform, the channel vector in the time slot q can be obtained as $V'_k c_k(q)$, where $c_k(q)$ are the coefficients of the KL transform given by $c_k(q) = V'^{\mathrm{H}}_k h_{\mathrm{LS},k}(q)$. No errors are added to our channel estimation if all the coefficients of the KL transform are employed. To compress the channel information and taking into account the good energy compaction properties of the KL decomposition, we can approximate the vector channels $h_k(q)$ by

$$\begin{aligned} h_{\mathrm{KL},k}(q) = V_k c_k(q) &= V_k V_k^{\mathrm{H}} h_{\mathrm{LS},k}(q) \\ &= V_k V_k^{\mathrm{H}} h_k(q) + V_k V_k^{\mathrm{H}} n_{\mathrm{LS},k}(q) \end{aligned} \tag{12}$$

where $V_k = [v_{k,1}, \ldots, v_{k,r}, 0_{N_t \times N_{t-r}}]$ and r denotes the number of KL coefficients sent from the receiver after truncation. The noise $V_k V_k^{\mathrm{H}} n_{\mathrm{LS},k}(q)$ and the signal $V_k V_k^{\mathrm{H}} h_k(q)$ lie in the same subspace spanned by the columns of V_k. Therefore, $h_{\mathrm{KL},k}(q)$ gives us no information about the properties of $h_k(q)$ lying in range$(V_k)^\perp$.

The resulting error contribution due to the KL truncation reads as

$$\begin{aligned} h''_k(q) &= (I - V_k V_k^{\mathrm{H}}) h_k(q) \\ &\sim \mathcal{N}_{\mathbb{C}}(0, (I - V_k V_k^{\mathrm{H}}) C_{h,k}(I - V_k V_k^{\mathrm{H}})). \end{aligned} \tag{13}$$

Note that $V_k V_k^{\mathrm{H}} n_{\mathrm{LS},k}(q)$ is orthogonal to $h''_k(q)$.

So, we have

$$h_k(q) = h'_k(q) + h''_k(q) \tag{14}$$

with

$$h'_k(q) = h_{\mathrm{KL},k}(q) + n_{\mathrm{KL},k}(q) \tag{15}$$

and $n_{\mathrm{KL},k}(q) \sim \mathcal{N}_{\mathbb{C}}(0, \sigma_n^2 V V^{\mathrm{H}} (S^{\mathrm{H}} S)^{-1} V V^{\mathrm{H}})$.

4.3. Statistical model for quantization errors (Type C)

Given that our channel covariance matrix does not depend on time [cf. (5)], the modal matrix obtained from its eigenvalue decomposition is also constant over time. With this assumption, only the KL coefficients have to be sent from the receiver to the transmitter due to the fast variations of the channel (so referred as *short–term* variations).

The uniform quantizer is the most common of the scalar quantizers whose principle is rather simple. With the normalization of the KL coefficients $c_{k,i}(q)$ given by $c_{k,i}(q)/\sqrt{\Lambda_{k,i}}$, where i denotes the i-th entry of the diagonal matrix Λ_k and of the KL coefficients vector $c_k(q)$, we get a standard Gaussian distribution for them. We can approximately assume that the input is bounded, with real and imaginary values independently quantized lying in the range included between $-1.96/\sqrt{2}$ and $1.96/\sqrt{2}$, i.e., the overload region has very low probability (≤ 0.05) of containing any input sample. The process of quantization is as follows. Before transmission, we choose representants between $-1.96/\sqrt{2}$ and $1.96/\sqrt{2}$ to construct an initial set of codebooks that are stored at both transmitter and receiver. The receivers perform a search to find for the components (real and imaginary parts) of the KL coefficients obtained in each time slot the element in the codebook that is closest. Then, the corresponding codebook index is fed back to the transmitter. Finally, the transmitter simply looks at its codebook and builds the precoder parameters from the selected codeword [23]. Therefore, $h_{\mathrm{KL},k}$ lies somewhere in the respective cell, i.e.,

$$h_{\mathrm{KL},k}(q) = h_{\mathrm{Q},k}(q) + n_{\mathrm{Q},k}(q) \tag{16}$$

where $h_{\mathrm{Q},k}$ is the representant and where $n_{\mathrm{Q},k}(q)$ is assumed uniformly distributed over the cell, for simplicity reasons. Even though the input is not uniform if the number of levels for uniform quantization N is large and, therefore, the quantizer step Δ is very small, we can assume that the input pdf is very smooth and that $E(|\epsilon|^2) = \frac{\Delta^2}{12}$ [24], where ϵ is the error in the uniform quantizer.

Remember that $h_{\mathrm{KL},k}(q)$ only lies in the subspace spanned by the columns of V_k. Under the assumption that also the quantizer works in this subspace, i.e., $h_{\mathrm{Q},k}(q) = V_k h'_{\mathrm{Q},k}(q)$, we can follow that $n_{\mathrm{Q},k}(q)$ lies also in the subspace spanned by the columns of V_k. Therefore, we get the rank deficient covariance matrix for the quantization error

$$C_{\mathrm{Q},k} = V_k C_{\mathrm{Q}'} V_k^{\mathrm{H}} \tag{17}$$

where $C_{\mathrm{Q}'} = \frac{\Delta^2}{6} I_{N_t \times N_t}$, since the KL coefficients are uncorrelated and consequently, the quantization errors are too, the variance is $2E[|\epsilon|^2]$.

Taking into account the errors due to estimation, truncation of the KL transform and quantization, we have up to now the following model

$$h_k(q) = h_{\mathrm{Q},k}(q) + n_{\mathrm{Q},k}(q) + n_{\mathrm{KL},k}(q) + h''(q). \tag{18}$$

Note that the error covariance matrix $C_{\mathrm{Q},k} + C_{\mathrm{KL},k} + C_{\mathrm{h''},k}$ has full rank.

$$\boxed{\begin{aligned}
&\boldsymbol{h}_k(D) = \boldsymbol{h}_{\mathrm{Q},k}(D) + \boldsymbol{n}_{\mathrm{Q},k}(D) + \boldsymbol{n}_{\mathrm{KL},k}(D) + \boldsymbol{h}''_k(D) \\
&\quad \text{where } \boldsymbol{h}_{\mathrm{Q},k}(D) \text{ is the quantized version of } \boldsymbol{V}_k \boldsymbol{V}_k^{\mathrm{H}} \boldsymbol{h}_{\mathrm{LS},k}(D) \\
&\quad \text{and} \\
&\quad \boldsymbol{C}_{\mathrm{Q},k}(D) = E[\boldsymbol{n}_{\mathrm{Q},k}(D)\boldsymbol{n}_{\mathrm{Q},k}^{\mathrm{H}}(D)] = \tfrac{\Delta^2}{6}\boldsymbol{V}_k \boldsymbol{V}_k^{\mathrm{H}} \\
&\quad \boldsymbol{n}_{\mathrm{KL},k}(D) \sim \mathcal{N}_{\mathbb{C}}\left(\boldsymbol{0},\, \boldsymbol{V}_k \boldsymbol{V}_k^{\mathrm{H}}\left(\sigma_n^2(\boldsymbol{S}^{\mathrm{H}}\boldsymbol{S})^{-1} + 2\left(1 - J_0\left(2\pi\tfrac{f_{\mathrm{D},k}}{f_{\mathrm{slot}}}D\right)\right)\boldsymbol{C}_{h,k}\right)\boldsymbol{V}_k \boldsymbol{V}_k^{\mathrm{H}}\right) \\
&\quad \boldsymbol{h}''_k(D) \sim \mathcal{N}_{\mathbb{C}}\left(\boldsymbol{0},\, (\boldsymbol{I} - \boldsymbol{V}_k \boldsymbol{V}_k^{\mathrm{H}})\boldsymbol{C}_{h,k}(\boldsymbol{I} - \boldsymbol{V}_k \boldsymbol{V}_k^{\mathrm{H}})\right)
\end{aligned}}$$

Table 2. The errors model.

4.4. Statistical model for feedback delay errors (Type D)

The transmission over the feedback channel introduces a delay of D slots. This delay can equivalently be modeled as follows. The estimator gets outdated training data, i.e., the observation of the estimator is delayed by D slots. Then, the respective feedback channel has no delay. The LS estimate for delayed training data reads as

$$\boldsymbol{h}_{\mathrm{LS},k}(D) = \boldsymbol{h}_k(0) + \boldsymbol{n}_{\mathrm{LS},k}(D) \tag{19}$$

where $\boldsymbol{n}_{\mathrm{LS},k}(D)$ has the same statistical properties as described above in (10). Clearly,

$$\begin{aligned}
\boldsymbol{h}_{\mathrm{LS},k}(D) &= \boldsymbol{h}_k(D) + \boldsymbol{h}_k(0) - \boldsymbol{h}_k(D) + \boldsymbol{n}_{\mathrm{LS},k}(D) \\
&= \boldsymbol{h}_k(D) + \boldsymbol{n}'_{\mathrm{LS},k}(D)
\end{aligned} \tag{20}$$

being $\boldsymbol{n}'_{\mathrm{LS},k}(D) = \boldsymbol{h}_k(0) - \boldsymbol{h}_k(D) + \boldsymbol{n}_{\mathrm{LS},k}(D)$. With the properties of $\boldsymbol{h}_k(q)$ and $\boldsymbol{h}_{\mathrm{w},k}(q)$ showed in Section 3, and taking into account that

$$\begin{aligned}
\boldsymbol{C}_{h_{\mathrm{w}},k}(D) &= E[\boldsymbol{h}_{\mathrm{w},k}(q)\boldsymbol{h}_{\mathrm{w},k}^{\mathrm{H}}(q-D)] \\
&= J_0\left(2\pi\frac{f_{\mathrm{D},k}}{f_{\mathrm{slot}}}D\right)\boldsymbol{I}_{N_{\mathrm{t}}\times N_{\mathrm{t}}}
\end{aligned} \tag{21}$$

where J_0 denotes the zero–th order Bessel function of the first kind, $f_{\mathrm{D},k}$ is the *maximum* Doppler frequency, and f_{slot} the slot rate, we obtain

$$\begin{aligned}
E[\boldsymbol{e}_D \boldsymbol{e}_D^{\mathrm{H}}] &= 2\boldsymbol{C}_{h,k} - E[\boldsymbol{h}_k(0)\boldsymbol{h}_k^{\mathrm{H}}(D)] - E[\boldsymbol{h}_k(D)\boldsymbol{h}_k^{\mathrm{H}}(0)] \\
&= 2\left(1 - J_0\left(2\pi\frac{f_{\mathrm{D},k}}{f_{\mathrm{slot}}}D\right)\right)\boldsymbol{C}_{h,k}.
\end{aligned} \tag{22}$$

Here $\boldsymbol{e}_D = \boldsymbol{h}_k(0) - \boldsymbol{h}_k(D)$.

Hence, the new LS error has the property

$$\boldsymbol{n}'_{\mathrm{LS},k}(D) \sim \mathcal{N}_{\mathbb{C}}\left(\boldsymbol{0}, \boldsymbol{C}'\right) \tag{23}$$

with $\boldsymbol{C}' = \sigma_n^2(\boldsymbol{S}^{\mathrm{H}}\boldsymbol{S})^{-1} + 2\left(1 - J_0\left(2\pi\frac{f_{\mathrm{D},k}}{f_{\mathrm{slot}}}D\right)\right)\boldsymbol{C}_{h,k}$.

Therefore, at the end, we find the model of the errors described in Table 2.

$$\boxed{\begin{aligned}
&\boldsymbol{\Phi}^{-1} = \left(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}} + \tfrac{1}{\gamma}\mathbf{I}\right)^{-1} \\
&\boldsymbol{P} = \boldsymbol{I}_K,\, \boldsymbol{D} = \boldsymbol{0}_K \\
&\textbf{for } i = K,\ldots,1 \\
&\quad q = \underset{q'=1,\ldots,i}{\arg\min}\, \sigma_{v_{q'}}^2\, \boldsymbol{\Phi}^{-1}(q',q') \\
&\quad \boldsymbol{P}_i = \boldsymbol{I}_K \text{ whose } i\text{-th and } q\text{-th rows are exchanged} \\
&\quad \boldsymbol{P} = \boldsymbol{P}_i \boldsymbol{P} \\
&\quad \boldsymbol{\Phi}^{-1} = \boldsymbol{P}_i \boldsymbol{\Phi}^{-1} \boldsymbol{P}_i^{\mathrm{T}} \\
&\quad \boldsymbol{D}(i,i) = \boldsymbol{\Phi}^{-1}(i,i) \\
&\quad \boldsymbol{\Phi}^{-1}(1:i,i) = \boldsymbol{\Phi}^{-1}(1:i,i)/\boldsymbol{D}(i,i) \\
&\quad \boldsymbol{\Phi}^{-1}(1:i-1,1:i-1) = \boldsymbol{\Phi}^{-1}(1:i-1,1:i-1) \\
&\qquad -\boldsymbol{\Phi}^{-1}(1:i-1,i)\boldsymbol{\Phi}^{-1}(1:i-1,i)^{\mathrm{H}}\boldsymbol{D}(i,i) \\
&\boldsymbol{L}^{\mathrm{H}} = \text{upper triangular part of } \boldsymbol{\Phi}^{-1} \\
&\boldsymbol{B}^{\mathrm{H}} = \boldsymbol{L}^{-1},\, \boldsymbol{F}^{\mathrm{H}} = \boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{L}^{\mathrm{H}}\boldsymbol{D} \\
&\chi = \|\boldsymbol{F}^{\mathrm{H}}(:,1)\|_2^2 + \sigma_v^2\|\boldsymbol{F}^{\mathrm{H}}(:,2:K)\|_2^2 \\
&\beta = \sqrt{E_{\mathrm{tr}}/\chi} \\
&\boldsymbol{F}^{\mathrm{H}} = \beta\boldsymbol{F}^{\mathrm{H}}
\end{aligned}}$$

Table 3. Calculation of robust THP filter with optimum ordering.

5. ROBUST THP DESIGN

It is very well–known that THP performance degrades strongly due to imperfect CSI at the transmitter. For a multi–user scenario we have non–cooperative receivers and, therefore, linear equalization cannot be performed at the receiver. Thus, it will be necessary to stochastically model the errors (as was made in Section 4), designing an average robust THP with it, studied for a similar application in [14].

Our model for the channel matrix is given by

$$\boldsymbol{H}(q) = \hat{\boldsymbol{H}}(q) + \boldsymbol{\Theta}_{\mathrm{T}}(q) \tag{24}$$

where $\hat{\boldsymbol{H}}(q)$ is the quantized version of the channel matrix and $\boldsymbol{\Theta}_{\mathrm{T}}(q)$ is the error matrix. We follow the robust design of [14], i.e., we solve a robust optimization according to our error model of the feedback channel. For the efficient THP implementation described in [10] and shown in Table 3, we start from

$$\boldsymbol{\Phi}^{-1} = \left(\hat{\boldsymbol{H}}(q)\hat{\boldsymbol{H}}^{\mathrm{H}}(q) + \boldsymbol{C}_{\boldsymbol{\Theta}_{\mathrm{T}}} + \frac{1}{\gamma}\boldsymbol{I}\right)^{-1} \tag{25}$$

with γ denoting the *signal to noise ratio* (SNR) that is defined as the ratio between the total transmitted energy, E_{tr}, and the

noise spectral density, and where

$$C_{\Theta_{\mathrm{T}}} = E[\Theta_{\mathrm{T}}(q)\Theta_{\mathrm{T}}^H(q)] = \sum_{k=1}^{K} C_{\mathrm{T},k}^* \qquad (26)$$

denoting $(\bullet)^*$ the conjugation and where

$$\begin{aligned}
C_{\mathrm{T},k} &= C_{\mathrm{Q},k} + V_k V_k^{\mathrm{H}}(\sigma_n^2(S^{\mathrm{H}}S)^{-1} \\
&+ 2(1 - J_0(2\pi\frac{f_{\mathrm{D},k}}{f_{\mathrm{slot}}}D))C_{h,k})V_k V_k^{\mathrm{H}} \\
&+ (I - V_k V_k^{\mathrm{H}})C_{h,k}(I - V_k V_k^{\mathrm{H}}). \qquad (27)
\end{aligned}$$

Thus, by considering the statistical properties derived for the errors in the previous section in the algorithm used to obtain the filters, we can compensate in advance the impairments of the channel state information at the transmitter side to avoid an increase of the complexity of the receiver terminals.

5.1. MMSE receiver equalizer

Additionally, when we design the gain factors according to the *Minimum Mean Square Error* (MMSE) equalizer [25, 26] instead of using the modulo gain $1/\beta$ at the receiver side, we can improve the overall system performance with a small increase in the receivers complexity.

The received signal corresponding to the k–th user is y_k. The output w_k of the MMSE gain factor is obtained as,

$$w_k = (p^* R^{-1})y_k \qquad (28)$$

where R is the variance of y_k and p is the cross correlation between the k-th user's received signal y_k and the desired value $d_k = e_k^{\mathrm{T}} P^{\mathrm{T}} B^{\mathrm{H}} v$ for user k. These desired symbols are obtained from the linear representation [10] of the nonlinear modulo operator at the transmitter in Fig. 2. According to (1), the values of R and p are

$$\begin{aligned}
R = E\left[|y_k|^2\right] &= h_k^{\mathrm{T}} F^{\mathrm{H}} \Gamma F h_k^* + \sigma_n^2 \\
p = E\left[y_k d_k^*\right] &= E\left[(h_k^{\mathrm{T}} F^{\mathrm{H}} v + n_k)v^{\mathrm{H}} B P e_k\right] \\
&= h_k^{\mathrm{T}} F^{\mathrm{H}} \Gamma B P e_k. \qquad (29)
\end{aligned}$$

where Γ is a diagonal matrix corresponding to the covariance matrix of the signal v and whose entries are equal to $\tau^2/6$ except the first one that is 1.

Clearly, p cannot be computed from a channel estimate, because the receiver has no knowledge about the precoder F^{H} and the permutation matrix P. However, the transmitter can linearly precode the symbols during the training phase such that symbols for the k-th receiver are weighted prior to transmission with $F^{\mathrm{H}} \Gamma B P e_k$. Consequently, the received signal corresponding to the transmitted training symbols contains the term $F^{\mathrm{H}} \Gamma B P e_k s_{\mathrm{training},k}$ and the needed cross correlation p can be estimated.
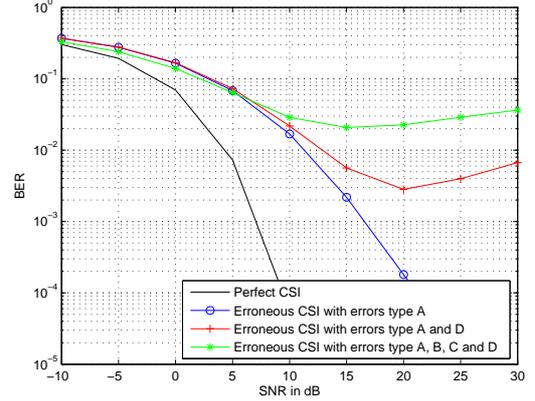
Fig. 3. BER performance vs. SNR for non-robust THP and v=10 km/h.

6. SIMULATIONS

In this section, we present the results of several computer simulations that we carried out to validate the proposed system. The figures illustrate how each type of error degrades the system more and more (see Fig. 3) and how robust THP can be employed to improve the performance and compensate the channel effects when the mismatch is caused by the different error sources that have been presented in this paper.

The results are the mean of 5000 channel realizations and 100 symbols were transmitted per channel realization. The input bits are QPSK modulated.

Figures show the performance for $N_{\mathrm{t}} = K = 8$. We consider an f_{slot} of 1500 Hz at a center frequency of 2 GHz. We also consider errors due to the feedback delay, being D equal to 1 for all users. The figures show the estimated BER curves for a Doppler frequency normalized to the slot period of 0.0123 (v= 10 km/h), i.e., relatively fast fading.

Fig. 4 plots the loss in performance when KL compression is applied. It can be seen, however, that $r = 2$ coefficients can be enough to ensure a suitable system performance with the enormous advantage of reducing the overhead of the feedback channel, especially for a high number of antennas at the transmitter. Note that a compression ratio given by $\frac{N_t \times L}{r \times m}$, where L and m is the number of bits employed for, respectively, coding the real and imaginary part of the channel coefficients $h_{k,\ell}$ (1) and representing the codebook index, can be reached employing only that truncation of the KL coefficients and the described quantization process prior to transmission. For example, in these simulation results, we use a codebook of 1024 ($m = 10$) entries and if $L = 32$ bits are used to encode the real and imaginary value of each channel coefficient, we obtain in a very simple way a compression ratio of 12.8. With a reasonable codebook size, we are obtaining good performance. Obviously, larger sizes improves the performance at the cost of decreasing the compression for the CSI sent
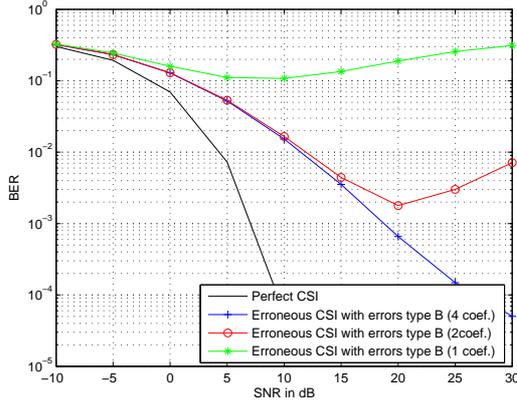
Fig. 4. BER performance vs. SNR when only truncation of KL coefficients is applied for non-robust THP and v=10 km/h.

Fig. 5. BER performance vs. SNR for different user speeds.

through the feedback channel and greatly extending the storage capability necessary at the receivers [23].

The precoding performance for different user speeds is plotted in the Fig. 5. You can see the curves for a speed of 10, 30 and 60 km/h (normalized Doppler frequency of 0.0123, 0.0370 and 0.0741, respectively). Obviously, the loss in performance for faster fading is greater than for slower fading, although with robust THP we can always get better performance with not much fed back CSI, as you can see in the figure.

Thus, Fig. 6 shows the estimated BER curves when we have a macro-cell environment which determines an offset of 2 degrees according to the model described in [20] and in Section 3. You can see how the non-robust THP curves go up for high SNR due to the sensitivity of our THP scheme to imperfect CSI. The figure plots the improvement in performance with respect to standard THP when the proposed robust THP scheme is applied. As you can see in the figure, the loss due to imperfect CSI can be reduced if we also apply the MMSE equalizer described in Section 5, so that the channel effects are compensated between both the transmitter and receiver side. Finally, if we have a macro-cell environment which determines an offset of 35 degrees, you can see in the Fig. 7 how the performance is worse than for 2 degrees, given that the environment obstructs even more the signal transmission.

7. CONCLUSION

In this paper, we have demonstrated the feasibility of quantization and truncation of KL decomposition for MU–MISO Tomlinson–Harashima precoding in FDD systems. Thanks to these techniques, the feedback channel overhead is strongly reduced. We have also considered the effect of estimating the channel using supervised methods and the delay inherent to the feedback channel, so that we model these errors and design a robust THP scheme which clearly outperforms the
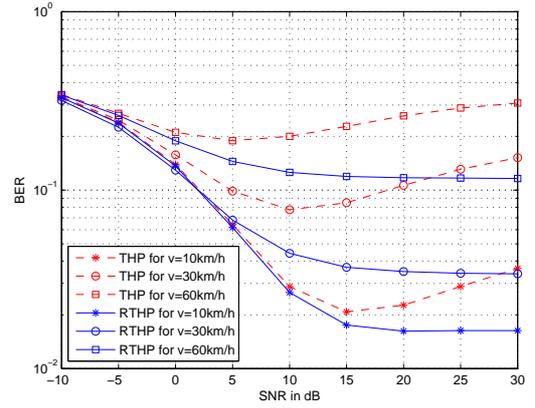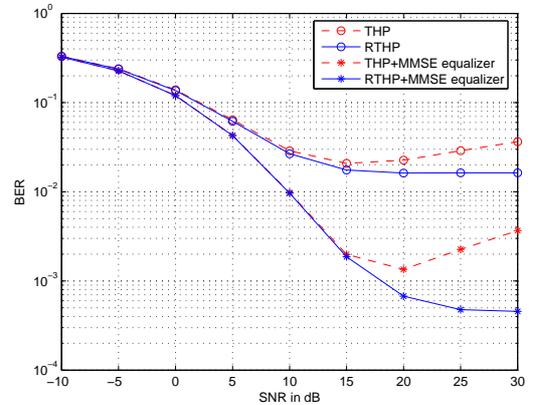
Fig. 6. BER performance vs. SNR for robust and non-robust THP when an offset of 2 deg is considered for the *Spatial Channel Model*.
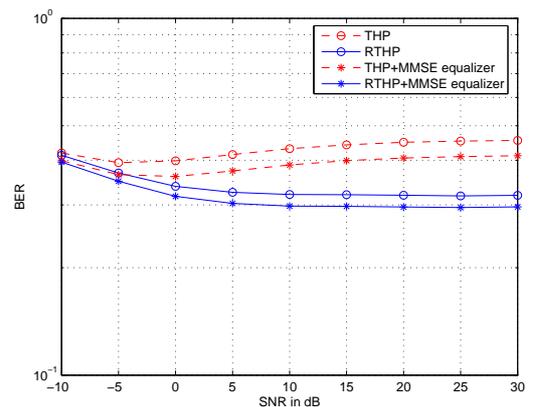
Fig. 7. BER performance vs. SNR for robust and non-robust THP when an offset of 35 deg is considered for the *Spatial Channel Model*.

conventional THP design. Thus, we are capable of adapting the precoder parameters to channel variations with a limited feedback channel.

8. REFERENCES

[1] M. Costa, *Writing on Dirty Paper*. IEEE Transactions on Information Theory, vol. 29, no. 3, pp. 439–441, May 1983.

[2] P. Viswanath and D. N. C. Tse, *Sum Capacity of the Vector Gaussian Broadcast Channel and Uplink-Downlink Duality*. IEEE Transactions on Information Theory, vol. 49, no. 8, pp. 1912–1921, August 2003.

[3] M. Schubert and H. Boche, *Iterative Multiuser Uplink and Downlink Beamforming Under SINR Constraints*. IEEE Transactions on Signal Processing, vol. 53, no. 7, pp. 2324–2334, July 2005.

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

[5] M. Tomlinson. *New Automatic Equalizer Employing Modulo Arithmetic*. Electronic Letters, vol. 7, no. 5/6, pp. 138-139, March 1971.

[6] H. Harashima and H. Miyakawa. *Matched–Transmission Technique for Channels with Intersymbol Interference*. IEEE Journal on Communications, vol. 20, no. 4, pp. 774-780, August 1972.

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

[8] M. Schubert and S. Shi. *MMSE Transmit Optimization with Interference Pre-Compensation*. In *Proc. VTC 2005 Spring*, vol. 2, pp. 845–849, May 2005.

[9] M. Joham and W. Utschick. *Ordered Spatial Tomlinson Harashima Precoding*. In Smart Antennas — State-of-the-Art, ser. EURASIP Book Series on Signal Processing and Communications, T. Kaiser, A. Bourdoux, H. Boche, J. R. Fonollosa, J. B. Andersen, and W. Utschick, Eds., EURASIP, Hindawi Publishing Corporation, 2006, vol. 3, ch. III. Transmitter, pp. 401–422.

[10] K. Kusume, M. Joham, W. Utschick, and G. Bauch. *Efficient Tomlinson–Harashima Precoding for Spatial Multiplexing on Flat MIMO Channel*. In Proc. ICC 2005, vol. 3, pp. 2021-2025, May 2005.

[11] R. El Assir, F. A. Dietrich, M. Joham, and W. Utschick. *Min-Max MSE Precoding for Broadcast Channels based on Statistical Channel State Information*. In Proc. SPAWC 2006, July 2006.

[12] B. Chun. *A Downlink Beamforming in Consideration of Common Pilot and Phase Mismatch*. In European Conference on Circuit Theory and Design, September 2005.

[13] A. P. Liavas. *Tomlinson-Harashima Precoding With Partial Channel Knowledge*. IEEE Transactions on Communications, vol. 53, no. 1, pp. 5–9, January 2005.

[14] F. A. Dietrich, and W. Utschick. *Robust Tomlinson-Harashima Precoding*. In Proc. of the 16th IEEE Symposium on Personal, Indoor and Mobile Radio Communications, vol. 1, pp. 136–140, Berlin, Germany, 2005.

[15] D. J. Love, R. W. Heath Jr., W. Santipach, and M. L. Honig. *What is the value of the limited feedback for MIMO channels?*. IEEE Communications Magazine, vol. 42, pp. 54-59, October 2004.

[16] M. Joham, D. A. Schmidt, J. Brehmer, and W. Utschick. *Finite-Length MMSE Tomlinson-Harashima Precoding for Frequency Selective Vector Channels*. Accepted for publication in IEEE Transactions on Signal Processing, 2006.

[17] W. Jakes. *Microwave Mobile Communications*, New York: Wiley, 1974.

[18] T. Croft, P. Dent, and G. E. Bottomley. *Jakes fading model revisited*. IEEE Electronics Letter, vol. 29, 1993.

[19] T. Zemen. *Time–Variant Channel Estimation Using Discrete Prolate Spheroidal Sequences*. IEEE Trans. on Signal Processing, vol. 53, No. 9, September 2005.

[20] *3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Spatial channel model for Multiple Input Multiple Output (MIMO) simulations (Release 6)*, 2003.

[21] J. Balakrishnan, M. Rupp, and H. Viswanathan. *Optimal channel tracking for multiple antenna systems*. In Proc. of the conference on multiaccess, mobility and teletraffic for wireless communications, December 2000.

[22] F.A. Dietrich, and W. Utschick. *Pilot–Assisted Channel Estimation Based on Second-Order Statistics*. IEEE Trans. Signal Processing, vol. 53(3), pp. 1178-1193, March 2005.

[23] P. M. Castro and L. Castedo. *Adaptive Vector Quantization for precoding using blind channel prediction in frequency selective MIMO mobile channels*. In Proc. ITG/IEEE Workshop on Smart Antennas 2005, April 2005.

[24] A. Gersho and R. Gray. *Vector Quantization and Signal Compression*. Kluwer Academic Publishers, 1992.

[25] E. A. Lee and D. G. Messerschmitt. *Digital Communication*. Second Edition. Kluwer Academic Publishers, 1994.

[26] J. G. Proakis and M. Salehi. *Communication systems engineering*. Prentice Hall, 1994.