Uplink-Downlink Duality with Regard to Constraints Imposed in Practice

Christian Siegl, Robert F.H. Fischer

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

Abstract

The uplink-downlink duality is reviewed with regard to constraints imposed on the transmission schemes in practice. In particular, the effects of non-Gaussian signaling, the use of implementable precoding schemes, and the availability of only imperfect channel state information at the central base-station are analyzed and discussed. The sources for violating the uplink-downlink duality in practical schemes are identified, and it is shown under which setting still the same (but not optimum) performance of both transmission directions can be achieved. Besides treating signal-to-interference plus noise ratios, as usually done, the duality is studied with respect to bit error rate, a parameter more relevant in practical applications.

I. Introduction

One of the main results developed in information theory over the last years is the *uplink-downlink duality*, e.g., [17], [14], [19], [11]. Basically it states that the performance of the users in an uplink scenario (many user transmitting to one central unit) is potentially exactly the same as in the corresponding downlink scenario (the central unit transmits to the users) when the same sum power is used in the system. This fundamental correspondence holds for linear transmission schemes, as well as successive decoding/encoding. In each case, Gaussian signaling is assumed, and *signal-to-interference plus noise ratios (SINR)* or (equivalently) capacities are considered.

However, in practice, the assumptions made for deriving the uplink-downlink duality are not or not exactly fulfilled. In particular, in most transmission schemes conventional QAM signaling with equiprobable signal points is used, consequently the transmit signals are far away from Gaussian. Moreover, the concept of *Costa precoding* [3], expected in the downlink as the counterpart to *decision-feedback equalization (DFE)* in the uplink, is merely theoretical and in practice implementable precoding schemes, namely *Tomlinson-Harashima precoding (THP)* [6], are employed. Finally, the assumption of perfect *channel state information (CSI)* is far away from reality. Hence, the study of

the uplink-downlink duality for imperfect channel knowledge (done in [4] but still assuming Gaussian signals and Costa precoding) is vital for practical applications.

In this paper we assess the uplink-downlink duality with regards to the following constraints imposed in practice: (i) non-Gaussian signaling, (ii) use of THP instead of Costa precoding, and (iii) imperfect channel knowledge. In these situations, it no longer suffices to study SINR or capacities—as a second, and in most applications more relevant measure of performance, the *bit error rate (BER)* (uncoded transmission) is treated. We identify sources for violating the uplink-downlink duality in practical schemes, and show under which setting it still holds but with some degradation compared to the theoretical optimum.

The paper is organized as follows: In Section II the system model (including imperfect CSI) is introduced. The uplink-downlink duality for linear transmission schemes (linear equalization/pre-equalization) is studied in Section III and that for successive schemes (DFE/THP) is treated in Section IV. In each case, the main results are first derived analytically and then enlightened by presenting and discussing numerical examples. Section V concludes the paper.

II. System Model

We study two communication scenarios: First, transmission from, say N_{U}, non-cooperating users, distributed over some service area, with a central receiver, equipped with N_{C} receive antennas. Assuming flat fading between each pair of antenna, the channel matrix $\boldsymbol{H}_{\mathrm{u}}$ (given in the equivalent complex baseband) of this uplink (UL) scenario or *multiple-access channel (MAC)* is of dimension $N_{\mathrm{C}} \times N_{\mathrm{U}}$.

Given reciprocity, the downlink (DL) scenario, also called *broadcast channel (BC)*, i.e., transmission from the central unit to the individual users, is characterized by the $N_{\mathrm{U}} \times N_{\mathrm{C}}$ channel matrix[1] $\boldsymbol{H}_{\mathrm{d}} = \boldsymbol{H}_{\mathrm{u}}^{\mathsf{H}}$.

Information to be transmitted is carried in the amplitude coefficients $a_{\bullet,k}$, taken form an M-ary QAM constellation $\mathcal{A} = \left\{ a_{\mathrm{I}} + \mathrm{j}a_{\mathrm{Q}} \mid a_{\mathrm{I}}, a_{\mathrm{Q}} \in \{\pm 1/2, \ldots, \pm(\sqrt{M}-1)/2\} \right\}$ with variance σ_a^2. The coefficients are collected in the vectors $\boldsymbol{a}_{\mathrm{u}} = [a_{\mathrm{u},1}, \ldots, a_{\mathrm{u},N_{\mathrm{U}}}]^{\mathsf{T}}$ and $\boldsymbol{a}_{\mathrm{d}} = [a_{\mathrm{d},1}, \ldots, a_{\mathrm{d},N_{\mathrm{U}}}]^{\mathsf{T}}$. Distribution of power over the users in the uplink is accomplished by scaling the data symbols $a_{\mathrm{u},k}$ appropriately; the positive real scaling factors are combined into the diagonal power allocation matrix $\boldsymbol{\Lambda}_{\mathrm{u}} = \mathbf{diag}(\lambda_{\mathrm{u},k})$, with $\mathrm{trace}(\boldsymbol{\Lambda}_{\mathrm{u}}^2) = N_{\mathrm{U}}$ in order to guarantee fixed sum power. In the same way, scaling in the downlink is done by the diagonal matrix $\boldsymbol{\Lambda}_{\mathrm{d}} = \mathbf{diag}(\lambda_{\mathrm{d},k})$, which has to guarantee

[1]Notation: lower case bold letters denote vectors; upper case bold letters matrices; $\boldsymbol{A} = [a_{ij}]$: matrix with elements a_{ij}; $\boldsymbol{A}^{\mathsf{T}}$: transpose, $\boldsymbol{A}^{\mathsf{H}}$: Hermitian (i.e., conjugate) transpose; $\boldsymbol{I}$: Identity matrix; $\mathbf{diag}(\cdot)$: diagonal matrix with indicated elements; $\mathbf{tril}(\cdot)$: lower triangular part of matrix. $\bullet \in \{\text{“u”}; \text{“d”}\}$.

that after further preprocessing, sum power is the same as in the uplink. By scaling the data symbols (and further preprocessing in the downlink), the vector of channel symbols $\boldsymbol{x}_\bullet$ is obtained.

At the receive antennas additive Gaussian noise, collected in the noise vector $\boldsymbol{n}_\bullet$, is present. We assumed the noise samples to be i.i.d. with variance σ_n^2, in both uplink or downlink scenario. In summary, the receive vector hence reads

$$\boldsymbol{y}_\bullet = \boldsymbol{H}_\bullet \boldsymbol{x}_\bullet + \boldsymbol{n}_\bullet \, . \tag{1}$$

One practical aspect we discuss in this paper is the situation where only imperfect *channel state information (CSI)* is available at the central unit. No CSI (except suited scaling) is assumed at the individual, scattered users. In case of imperfect CSI, the channel matrix available, $\hat{\boldsymbol{H}}_\bullet$, differs from the actual channel matrix $\boldsymbol{H}_\bullet$ by the estimation error $\boldsymbol{\Delta}$, i.e.,

$$\hat{\boldsymbol{H}}_\bullet = \boldsymbol{H}_\bullet + \boldsymbol{\Delta}_\bullet \, . \tag{2}$$

Noteworthy, the channel matrix is preferably estimated during uplink transmission and then used for both equalization of the uplink and pre-equalization of the downlink. Then, estimate and estimation error for up- and downlink are hermitian version of each other, as the actual channel matrices are.

Assuming an MMSE channel estimator, e.g., based on pilot symbols [15], the entries δ_{kj} of $\boldsymbol{\Delta}$ are i.i.d. complex Gaussian random variables with estimation error variance σ_δ^2. Moreover, in case of MMSE estimator, the coefficients of estimated channel matrix $\hat{\boldsymbol{H}}_\bullet$ and that of the estimation error $\boldsymbol{\Delta}$ are independent, $\mathrm{E}\{\delta_{kj}\hat{h}_{kj}\} = 0$, $\forall k, j$.

III. Uplink-Downlink Duality for MMSE Linear (Pre-)Equalization

In [17], the uplink-downlink duality was derived for linear processing/pre-processing in a very general setting when assuming Gaussian signaling. In this setting it suffices to treat *signal-to-interference plus noise ratios (SINR)*, which, moreover, are directly related to capacities [19], [17], [11]. Here, we specialize the discussion to the important case of linear equalization/pre-equalization optimized according to the *minimum mean-squared error (MMSE) criterion* but assuming QAM signaling with a finite number of signal points. Moreover, the effect of non-perfect CSI is assessed.

A. Perfect CSI, Gaussian Signaling

The block diagrams of up- and downlink when employing (MMSE) linear (pre-)equalization are shown in Fig. 1.

1) Uplink: The MMSE linear equalization matrix $\boldsymbol{F}_{\mathrm{u}}$ is obtained when minimizing the variance of the error between equalizer output signal $\boldsymbol{r}_{\mathrm{u}}$ and channel input $\boldsymbol{x}_{\mathrm{u}}$. Given the channel matrix $\boldsymbol{H}_{\mathrm{u}}$ and the power allocation matrix $\boldsymbol{\Lambda}_{\mathrm{u}}$ it calculates to [16]

$$\boldsymbol{F}_{\mathrm{u}} = \left(\boldsymbol{H}_{\mathrm{u}}^{\mathsf{H}}\boldsymbol{H}_{\mathrm{u}} + \zeta\boldsymbol{\Lambda}_{\mathrm{u}}^{-2}\right)^{-1}\boldsymbol{H}_{\mathrm{u}}^{\mathsf{H}}\,, \tag{3}$$

with $\zeta \stackrel{\mathrm{def}}{=} \sigma_n^2/\sigma_a^2$. It is convenient to define the overall transmission matrix, the cascade consisting of channel and feedforward processing, i.e.,

$$\boldsymbol{C}_{\mathrm{u}} = \boldsymbol{F}_{\mathrm{u}}\boldsymbol{H}_{\mathrm{u}} = \left(\boldsymbol{H}_{\mathrm{u}}^{\mathsf{H}}\boldsymbol{H}_{\mathrm{u}} + \zeta\boldsymbol{\Lambda}_{\mathrm{u}}^{-2}\right)^{-1}\boldsymbol{H}_{\mathrm{u}}^{\mathsf{H}}\boldsymbol{H}_{\mathrm{u}}\,, \tag{4}$$

which has real positive diagonal elements $c_{\mathrm{u},kk}$. From the right hand side it is immediate that the signals are scaled by factors smaller than one (the elements $c_{\mathrm{u},kk}$), i.e., a bias is present in the MMSE solution [1], [8]. To compensate for this bias and, at the same time, for the power scaling at the transmitter, a diagonal scaling matrix $\tilde{\boldsymbol{\Lambda}}_{\mathrm{u}}^{-1}$ with

$$\tilde{\boldsymbol{\Lambda}}_{\mathrm{u}} = \mathbf{diag}(\tilde{\lambda}_{\mathrm{u},k}) \qquad \text{with} \quad \tilde{\lambda}_{\mathrm{u},k} = c_{\mathrm{u},kk}\lambda_{\mathrm{u},k} \tag{5}$$

is employed in front of the decision device/decoder which then operates on the initial signal constellation $\mathcal{A}$.

Using the unbiased MMSE linear receiver, the SINR calculates to [17]

$$\mathrm{SINR}_{\mathrm{u},k} = \frac{c_{\mathrm{u},kk}^2\lambda_{\mathrm{u},k}^2}{\zeta||\boldsymbol{f}_{\mathrm{u},k}||^2 + \sum_{\substack{j=1\\ j\neq k}}^{N_{\mathrm{U}}}|c_{\mathrm{u},kj}|^2\lambda_{\mathrm{u},j}^2} \tag{6}$$

$$= \frac{c_{\mathrm{u},kk}^2\lambda_{\mathrm{u},k}^2}{c_{\mathrm{u},kk}\lambda_{\mathrm{u},k}^2 - c_{\mathrm{u},kk}^2\lambda_{\mathrm{u},k}^2} = \frac{c_{\mathrm{u},kk}}{1-c_{\mathrm{u},kk}}\,, \tag{7}$$

where $\boldsymbol{f}_{\mathrm{u},k}^{\mathsf{T}}$ is the k^{th} row vector of $\boldsymbol{F}_{\mathrm{u}}$. The last expression is due to a fundamental principle of the MMSE solution (cf., e.g., [8]) and follows from applying the matrix inversion lemma [10] to (3) which, after multiplying with $\boldsymbol{\Lambda}_{\mathrm{u}}^{-1}$ results in $\boldsymbol{\Lambda}_{\mathrm{u}}^{-1}\boldsymbol{F}_{\mathrm{u}} = \boldsymbol{\Lambda}_{\mathrm{u}}\boldsymbol{H}_{\mathrm{u}}^{\mathsf{H}}(\boldsymbol{H}_{\mathrm{u}}\boldsymbol{\Lambda}_{\mathrm{u}}^2\boldsymbol{H}_{\mathrm{u}}^{\mathsf{H}} + \zeta\boldsymbol{I})^{-1}$. Multiplying with the term in brackets, taking the hermitian, and multiplying by $\boldsymbol{F}_{\mathrm{u}}$ finally gives

$$\boldsymbol{C}_{\mathrm{u}}\boldsymbol{\Lambda}_{\mathrm{u}}^2 = \boldsymbol{C}_{\mathrm{u}}\boldsymbol{\Lambda}_{\mathrm{u}}^2\boldsymbol{C}_{\mathrm{u}}^{\mathsf{H}} + \zeta\boldsymbol{F}_{\mathrm{u}}\boldsymbol{F}_{\mathrm{u}}^{\mathsf{H}}\,. \tag{8}$$

In particular, for the main diagonal elements, the equation reads

$$c_{\mathrm{u},kk}\lambda_{\mathrm{u},k}^2 = \sum_{j=1}^{N_{\mathrm{U}}}|c_{\mathrm{u},kj}|^2\lambda_{\mathrm{u},j}^2 + \zeta||\boldsymbol{f}_{\mathrm{u},k}||^2\,, \tag{9}$$

with which the denomination in (6) can be simplified as done above.

2) Downlink: From the uplink-downlink duality [17], as the channels are hermitian versions of each other, the matrices for signal processing in the downlink are the hermitian versions of that present in the uplink, i.e., $\boldsymbol{F}_{\mathrm{d}} = \boldsymbol{F}_{\mathrm{u}}^{\mathsf{H}}$ and in turn $\boldsymbol{C}_{\mathrm{d}} = \boldsymbol{C}_{\mathrm{u}}^{\mathsf{H}}$. Then, the SINR calculates to [17]

$$\mathrm{SINR}_{\mathrm{d},k} = \frac{c_{\mathrm{d},kk}^2 \lambda_{\mathrm{d},k}^2}{\zeta + \sum_{\substack{j=1 \\ j \neq k}}^{N_{\mathrm{U}}} |c_{\mathrm{d},kj}|^2 \, \lambda_{\mathrm{d},j}^2} \; . \tag{10}$$

However, in order to achieve the same SINRs for all users, $\mathrm{SINR}_{\mathrm{u},k} = \mathrm{SINR}_{\mathrm{d},k}$, $\forall k$, an adapted power allocation $\boldsymbol{\Lambda}_{\mathrm{d}}$ has to be used. Following the derivations in [17], for the present situation we have

$$\boldsymbol{\Lambda}_{\mathrm{d}}^2 = \mathbf{diag}(\zeta \boldsymbol{\Theta}_{\mathrm{d}}^{-1} \mathbf{1}) \; , \tag{11}$$

where $\boldsymbol{\Theta}_{\mathrm{d}}$ is a matrix with elements $\Theta_{\mathrm{d},kk} = c_{\mathrm{d},kk} - |c_{\mathrm{d},kk}|^2$, $k = 1, \ldots, N_{\mathrm{U}}$, and $\Theta_{\mathrm{d},kj} = -|c_{\mathrm{d},kj}|^2$, $k, j = 1, \ldots, N_{\mathrm{U}}$, $k \neq j$. This choice, moreover, guarantees that the same total downlink transmit power (after feedforward filtering) as in the uplink is present.

B. Perfect CSI, QAM Signaling

If noise and residual interference are Gaussian, treatment of the SINR is sufficient. On the one hand, assuming also Gaussian transmit signals, there is a one-to-one relation between SINR and capacities. On the other hand, for (uncoded) M-ary QAM transmission, the well-known relation between SINR and BER via the standard Q function holds for Gray labeling of the signal points (cf. [13])

$$\mathrm{BER}_k = \frac{N_{\min}}{\log_2(M)} \mathrm{Q}\left(\sqrt{\frac{3}{M-1} \mathrm{SINR}_k} \right) \; . \tag{12}$$

Here, $N_{\min}$ is the average number of minimum neighbor signal points, and $\mathrm{Q}(x)$ is the standard complementary Gaussian integral function.

However, in practice, where all users employ QAM signaling, residual interference and hence the total noise $e_k \stackrel{\mathrm{def}}{=} z_k - a_k$ (see Fig. 1) is not Gaussian. Error rate is then dependent on the actual probabilty density function (pdf) of the error e_k, not only on its power. Additionally, dependent on the interference situation, the pdfs of the e_k's will differ. Especially, in up- and downlink the contribution of (filtered) noise and residual interference to the total noise power may be extremely different.

Hence, the same SINR will not lead to the same error rates of the users and the same SINR in up- and downlink does not guarantee same performance in both directions. Only for high signal-to-noise ratios (as residual interference tend to zero) or huge number of interferer (due to the central limiting theorem), SINR is meaningful. Table I shows that in case of linear processing and perfect CSI the only source for violating the uplink-downlink duality is the non-Gaussianity of total noise.

In order to visualize the violation of the uplink-downlink duality in practical schemes numerical simulations are performed. For that a 3×3 system with channel matrix

$$\boldsymbol{H}_{\mathrm{u}} = \begin{bmatrix} -0.43 - 1.15\mathrm{j} & 0.13 + 1.19\mathrm{j} & 0.02 + 1.73\mathrm{j} \\ -1.67 + 1.19\mathrm{j} & 0.29 - 0.04\mathrm{j} & 0.20 + 0.01\mathrm{j} \\ 0.11 - 0.02\mathrm{j} & -0.50 + 1.32\mathrm{j} & -0.24 - 0.28\mathrm{j} \end{bmatrix}, \tag{13}$$

randomly drawn with i.i.d. Gaussian coefficients, is considered. The properties discussed for this particular channel are representative for the general case (except for pathological cases) of channel matrices with i.i.d. coefficients.

On the left hand side of Fig. 2 achievable SINR pairs for user 1 and 3 are plotted; on the right hand side the corresponding pairs of BER are displayed; the top row is valid for linear schemes with perfect CSI. The solid gray lines give the theoretical results (6) and (12), where the power of user 2 is fixed to σ_a^2 and the power of users 1 and 3 is chosen to $p\sigma_a^2$ and $(1-p)\sigma_a^2$, respectively, and p ranges from 0 to 1. Two different channel SNRs $\bar{E}_{\mathrm{b,Tx}}/N_0 \stackrel{\mathrm{def}}{=} \sigma_a^2/(\log_2(M)\sigma_n^2) \mathrel{\widehat{=}} 0$ dB and 3 dB are assumed. For five different power allocations numerical simulations using $(M = 4)$-ary QAM with Gray labeling for each user are preformed and the SINR measured. The crosses are valid for the uplink, whereas the circles correspond to the downlink.

It is clearly visible that the measured SINRs are in perfect agreement with the theoretical prediction and that in up- and downlink the same SINR points are achieved. This, however, changes slightly when BER (uncoded transmission) is considered. Here, the solid gray lines give the result derived via (12). Especially for very low error rates the non-Gaussian residual interference lead to different results in up- and downlink.

C. Imperfect CSI

We now turn to the case where no perfect CSI is available at the central unit—a situation highly relevant in practical systems. Dependent on the degree of channel knowledge two strategies for adjusting the feedforward matrices can be distinguished.

First, a channel estimator produces the estimate $\hat{\boldsymbol{H}}_{\mathrm{u}}$ and no further knowledge (about their reliability) is available. In this "naive approach" $\hat{\boldsymbol{H}}_{\mathrm{u}}$ is immediately used in (3) as for perfect CSI (cf. [4]). Because of channel errors, the uplink $\mathrm{SINR}_{\mathrm{u},k}$ is then not given by (6). However, straightforward calculations[2] show that $\mathrm{SINR}_{\mathrm{u},k}$ can be calculated by (6) if $\zeta = \sigma_n^2/\sigma_a^2$ is replaced by $\hat{\zeta} = (\sigma_n^2 + N_{\mathrm{U}}\sigma_a^2\sigma_\delta^2)/\sigma_a^2$.

[2]If, in the uplink, $\boldsymbol{F}_{\mathrm{u}}$ is fit to $\hat{\boldsymbol{H}}_{\mathrm{u}}$ instead of the actual channel $\boldsymbol{H}_{\mathrm{u}}$, in addition the the filtered noise $\boldsymbol{F}_{\mathrm{u}}\boldsymbol{n}_{\mathrm{u}}$ the error $-\boldsymbol{F}_{\mathrm{u}}\boldsymbol{\Delta}_{\mathrm{u}}\boldsymbol{a}_{\mathrm{u}}$ is present at the decision device. Since the elements of $\boldsymbol{\Delta}$ are assumed to be i.i.d. with variance σ_δ^2, the additional error variance is $||\boldsymbol{f}_k||^2 N_{\mathrm{U}}\sigma_\delta^2\sigma_a^2$; the total noise variance is hence $(\sigma_n^2 + N_{\mathrm{U}}\sigma_a^2\sigma_\delta^2)||\boldsymbol{f}_k||^2$. Similar argumentation is valid for the downlink.

Calculating the downlink power distribution by (11), a mismatch is present, too, and a non-optimum power allocation is used. Moreover, because of channel errors, the downlink $\mathrm{SINR}_{\mathrm{d},k}$ is given by (10) if ζ is again replaced by $\hat{\zeta}$. Using (3), (6), (7), (10), and (11) a relationship between the actual SINR_k and that if $\hat{\boldsymbol{H}}_{\mathrm{u}}$ would be the true channel matrix can be derived. Defining $c_{kk} = c_{\mathrm{u},kk} = c_{\mathrm{d},kk}$, we have

$$\alpha_{\mathrm{u},k}\mathrm{SINR}_{\mathrm{u},k} = \frac{c_{kk}}{1-c_{kk}} = \alpha_{\mathrm{d},k}\mathrm{SINR}_{\mathrm{d},k} \ . \tag{14}$$

Due to the chosen power allocation, the factors $\alpha_{\bullet,k}$ calculate to

$$\alpha_{\mathrm{u},k} = \frac{N_{\mathrm{U}}\sigma_\delta^2 ||\boldsymbol{f}_{\mathrm{u},k}||^2 \lambda_{\mathrm{u},k}^{-2}}{c_{kk}\left(1-c_{kk}\right)} + 1 \ , \tag{15}$$

$$\alpha_{\mathrm{d},k} = \frac{N_{\mathrm{U}}\sigma_\delta^2 \lambda_{\mathrm{d},k}^{-2}}{c_{kk}\left(1-c_{kk}\right)} + 1 \ . \tag{16}$$

In both situations, $\alpha_{\bullet,k}$ are larger than 1, i.e., as expected, the actual SINR is smaller than for the case when $\hat{\boldsymbol{H}}_{\mathrm{u}}$ is the true channel matrix. The loss increases as (i) the number N_{U} of users increases, (ii) the channel error variance σ_δ^2 increases, or (iii) the channel SNR increases, as for high SNRs (or close to orthogonal channels) c_{kk} tends to one and $\alpha_{\bullet,k}$ would be unbounded.

Moreover, the SINRs in up- and downlink would be the same if $\alpha_{\mathrm{u},k} = \alpha_{\mathrm{d},k}$, which would hold for $||\boldsymbol{f}_{\mathrm{u},k}||^2 \lambda_{\mathrm{u},k}^{-2} = \lambda_{\mathrm{d},k}^{-2}$. A study of (7), (10), and (11) reveals that this equality could exactly be achieved if in (11) ζ is replaced by $\hat{\zeta}$ as already done in the SINR calculation. However, this would require an additional degree of CSI, namely the error variance σ_δ^2.

Hence, in summary, in case of imperfect CSI and naive filter calculation the SINRs of up- and downlink are decreased compared to perfect channel knowledge and differ. This difference is due to the mismatch in power allocation and is small, if $N_{\mathrm{U}}\sigma_a^2\sigma_\delta^2$ is small compared to σ_n^2. Finally, a second source for violating the uplink-downlink duality is, as in case of perfect CSI, the non-Gaussianity of the transmit signals which again effects the BER.

If at least some knowledge about the reliability of the channel estimate $\hat{\boldsymbol{H}}_{\mathrm{u}}$ is available (the variance σ_δ^2) at the central unit, this fact can be utilized in the calculation of the matrices. Taking the channel error as additional noise into account, i.e., replacing σ_n^2 by $\sigma_n^2 + N_{\mathrm{U}}\sigma_a^2\sigma_\delta^2$, the uplink receive matrix should be calculated using (3) with $\hat{\zeta}$ instead of ζ [18], [12], [4]. In literature, this strategy of utilizing partial CSI, is sometimes called "conditional mean estimator", we simply call it MMSE solution.

Here, replacing ζ at all occurrences by $\hat{\zeta}$, (6) and (10) give the actual SINRs of the users. Moreover, using this value of ζ for the downlink power allocation (11), up- and downlink SINR are exactly equal. Hence, imperfect channel knowledge effects up- and downlink in the same way; both directions experience the same degradation. Finally, a comparison of the SINRs for "naive" and adapted MMSE filter calculation reveals that utilizing partial CSI is always rewarding. Both, the length of the row

of $\boldsymbol{F}_{\mathrm{u}}$ and the coefficients $c_{\mathrm{u},kj}$ are (slightly) smaller if ζ is increased, which, studying (6) and (10) give larger SINRs.

As above for perfect CSI, the uplink-downlink duality in practical schemes is now visualized for imperfect channel knowledge (second and third row of Fig. 2). All parameters are chosen the same as before and the channel estimation variance is firstly fixed to $\sigma_\delta^2 = 0.01$. Compared to perfect channel knowledge (outermost gray line), some degradation is observable. For the MMSE approach, the SINR in up- and downlink are identical; for the naive approach and with respect to BER slight differences between up- and downlink are visible. However, for the present linear schemes, up- and downlink performance are still in very good agreement. Finally, from the last row, the variation of SINR/BER over the channel estimation variance $\sigma_\delta^2 = 0.01,\ 0.03,\ 0.1$ can be assessed for the MMSE variant (for $10\log_{10}(\bar{E}_{\mathrm{b,Tx}}/N_0) = 3$ dB). As expected, only for poor channel estimates, naive and MMSE approach show different results. In each situation, up- and downlink perform almost the same.

IV. Uplink-Downlink Duality for MMSE DFE and Precoding

Besides linear equalization, *decision-feedback equalization (DFE)*, Fig. 3 (top), is a very popular strategy for separating the users. In wireless applications, DFE is often called *successive interference cancellation* [16]; in case of MIMO (or block) transmission known as *generalized DFE (GDFE)* [2]. The counterpart to DFE at the receiver side is (nonlinear) precoding at the transmitter. In [17] the duality of DFE to the theoretical concept of *Costa precoding* is derived. Both schemes perform successive decoding and encoding, respectively. Here, we discuss the connection of DFE to *Tomlinson-Harashima precoding (THP)* (for a detail description of THP see [6]), bottom of Fig. 3, which is a low-complexity realization of Costa precoding.

A. Perfect CSI, Gaussian Signaling

1) Uplink: Using DFE, equalization is done successively, taking already detected users into account by subtracting out their part to the receive signal. The receiver consists of two parts: a linear feedforward processing with matrix $\boldsymbol{F}_{\mathrm{u}}$ for achieving spatial causality, and a non-linear feedback loop with (lower triangular, unit main diagonal) feedback matrix $\boldsymbol{B}_{\mathrm{u}}$, cf. Fig. 3. In DFE, the decision order can be optimized; a permutation matrix $\boldsymbol{P}_{\mathrm{u}}$ (a single one in each row and column) reestablishes the original order.

Optimizing the DFE according to the MMSE criterion and given a particular permutation $\boldsymbol{P}_{\mathrm{u}}$ (which may by optimized according to any criterion, e.g., the BLAST ordering [9]), the feedback matrix is obtained by calculating a Cholesky factorization [6]

$$\boldsymbol{P}_{\mathrm{u}}^{\mathsf{T}} \left(\boldsymbol{H}_{\mathrm{u}}^{\mathsf{H}} \boldsymbol{H}_{\mathrm{u}} + \zeta \boldsymbol{\Lambda}_{\mathrm{u}}^{-2} \right) \boldsymbol{P}_{\mathrm{u}} = \boldsymbol{B}_{\mathrm{u}}^{\mathsf{H}} \boldsymbol{\Sigma}_{\mathrm{u}} \boldsymbol{B}_{\mathrm{u}} \tag{17}$$

with real diagonal matrix $\mathbf{\Sigma}_{\mathrm{u}} = \mathbf{diag}(\varsigma_{\mathrm{u},1}, \ldots, \varsigma_{\mathrm{u},N_{\mathrm{U}}})$. The corresponding feedforward matrix is given by

$$\boldsymbol{F}_{\mathrm{u}} = \mathbf{\Sigma}_{\mathrm{u}}^{-1} \boldsymbol{B}_{\mathrm{u}}^{-\mathsf{H}} \mathbf{\Lambda}_{\mathrm{u}} \boldsymbol{H}_{\mathrm{u}}^{\mathsf{H}} \; . \tag{18}$$

As above, cascade of channel and feedforward matrix is denoted as $\boldsymbol{C}_{\mathrm{u}}$. However, the actual values of the elements of $\boldsymbol{C}_{\mathrm{u}}$ (in particular the main diagonal elements) differ from that when using linear equalization.

Using the MMSE approach, a bias is present in DFE, too [1]. To compensate for this bias and for the power scaling at the transmitter, a diagonal scaling matrix $\tilde{\mathbf{\Lambda}}_{\mathrm{u}}^{-1}$ with $\tilde{\lambda}_{\mathrm{u},k} = c_{\mathrm{u},kk} \lambda_{\mathrm{u},P(k)}$ is placed in front of the decision device. Here, $P(k) = l$ gives the number l of the user which is decided at step k, i.e., $P(k) = l$, if the element in the l^{th} row and k^{th} column of $\boldsymbol{P}_{\mathrm{u}}$ is one. Since this device operates on the initial signal constellation $\mathcal{A}$ but the users are permuted, scaling by $\bar{\mathbf{\Lambda}}_{\mathrm{u}} = \boldsymbol{P}_{\mathrm{u}}^{\mathsf{T}} \mathbf{\Lambda}_{\mathrm{u}} \boldsymbol{P}_{\mathrm{u}} = \mathbf{diag}(\lambda_{\mathrm{u},P(k)})$ has also to be done in the feedback part.

Using unbiased MMSE DFE, the SINR calculates to [17]

$$\begin{aligned} \mathrm{SINR}_{\mathrm{u},P(k)} &= \frac{c_{\mathrm{u},kk}^2 \lambda_{\mathrm{u},P(k)}^2}{\zeta ||\boldsymbol{f}_{\mathrm{u},k}||^2 + \sum_{j=k+1}^{N_{\mathrm{U}}} |c_{\mathrm{u},kj}|^2 \lambda_{\mathrm{u},P(j)}^2} & (19) \\ &= \frac{c_{\mathrm{u},kk}}{1 - c_{\mathrm{u},kk}} \; . & (20) \end{aligned}$$

The last expression can be derived similarly as for linear schemes.

2) Downlink: Basically, downlink transmission can be derived from the uplink by flipping the entire structure and employing the hermitian versions of all matrices, $\boldsymbol{F}_{\mathrm{d}} = \boldsymbol{F}_{\mathrm{u}}^{\mathsf{H}}$, $\boldsymbol{B}_{\mathrm{d}} = \boldsymbol{B}_{\mathrm{u}}^{\mathsf{H}}$, $\boldsymbol{P}_{\mathrm{d}} = \boldsymbol{P}_{\mathrm{u}}^{\mathsf{H}}$, and $\boldsymbol{C}_{\mathrm{d}} = \boldsymbol{C}_{\mathrm{u}}^{\mathsf{H}}$. Moreover, following [17], compared to DFE the processing order has to be reversed, i.e., encoding in THP starts with the user detected last in DFE.

Using the theoretical optimum precoding scheme ("Costa precoding"), the SINR in the downlink calculates to [17]

$$\mathrm{SINR}_{\mathrm{d},P(k)} = \frac{c_{\mathrm{d},kk}^2 \lambda_{\mathrm{d},P(k)}^2}{\zeta + \sum_{j=1}^{k-1} |c_{\mathrm{d},kj}|^2 \lambda_{\mathrm{d},P(j)}^2} \; . \tag{21}$$

In order to achieve the same SNR for successive detection/encoding

$$\mathbf{\Lambda}_{\mathrm{d}}^2 = \mathbf{diag}(\zeta \, \mathbf{tril}^{-1}(\mathbf{\Theta}_{\mathrm{d}}) \mathbf{1}) \; , \tag{22}$$

where $\mathbf{tril}(\cdot)$ returns only the lower triangular part (including main diagonal) of a matrix, and $\mathbf{\Theta}_{\mathrm{d}}$ is defined as above in the linear case.

B. Perfect CSI, QAM Signaling

In practice, non-Gaussian signaling is used and SINR does no longer immediately relate to BER. Moreover, the theoretical concept of Costa precoding cannot be implemented in the downlink but its scalar approximation THP is employed.

In THP, assuming square M-ary QAM signal sets, a (usually) scalar modulo operation reduces real and quadrature component the transmit symbols before scaling and feedforward processing individually to the interval $\mathcal{R} = [-\sqrt{M}/2,\ +\sqrt{M}/2)$. The symbols of the user encoded first are still drawn from $\mathcal{A}$, but the symbols of all following users tend to be continuous uniformly distributed over $\mathcal{R}$ (cf. [6, Fig. E.4]). In turn, for the users treated last, the variance of the transmit symbols increases by the factor

$$\gamma_{\mathrm{p}}^2 \stackrel{\mathrm{def}}{=} \frac{M}{M-1} \tag{23}$$

the *precoding loss*. Using, as often proposed (e.g., [5]) a *dither sequence* $\boldsymbol{d}_{\mathrm{d}}$, whose elements are drawn uniformly distributed over $\mathcal{R}$, all users experience this loss in SINR. At the transmitter $\boldsymbol{d}_{\mathrm{d}}$ is added right in front of the modulo device, and at the receiver it is subtracted right in front of the slicer.

Comparing the transmission schemes based on equal transmit power, the precoding loss equivalently corresponds to an increase in noise power. Hence not the inverse SNR ζ is active but $\zeta_{\mathrm{p}} \stackrel{\mathrm{def}}{=} \zeta\gamma_{\mathrm{p}}^2$. The SINRs of the users can then be calculated using (21) but replacing ζ by ζ_{p} leading to somewhat smaller values than for the optimum Costa precoding. The same loss would only appear if dither is also used in DFE, cf. [7]. Hence, in practice, where a loss occurs only for precoding schemes, uplink-downlink duality is violated, cf. Table I. However, the increase in average transmit power should already be taken into account when calculating the required matrices. Hence, for THP and DFE with dither ζ should be replaced by ζ_{p} in (17) and in the power allocation (22).

Next, since THP uses modulo congruent signal points and hence modulo congruent decisions, the number $N_{\min}$ of nearest neighbor signal points is increased from $4(\sqrt{M}-1)/\sqrt{M}$ to 4 for square QAM constellations. Since in DFE without dither this effect does not occur, this is a further source for violating the uplink-downlink duality (cf. Tab. I).

Noteworthy, in THP and DFE using dither, the generated transmit symbols are statistically independent of the data symbols, e.g., [5]. This enables a slightly different MMSE optimization; in particular no bias problem occurs. In turn, bias compensation should not be done. In Fig. 3 the scaling matrices $\tilde{\boldsymbol{\Lambda}}_{\bullet}$ and $\bar{\boldsymbol{\Lambda}}_{\bullet}$ have then to be replaced by $\boldsymbol{\Lambda}_{\bullet}$ and $\bar{\tilde{\boldsymbol{\Lambda}}}_{\bullet}$ (permuted version of $\tilde{\boldsymbol{\Lambda}}_{\bullet}$), respectively. In the downlink transmission, this version of MMSE solution without the need for bias compensation is sometimes called *inflated lattice precoding* [5], [7]. Using DFE and THP in this way, up- and downlink have the same power loss and increase in nearest neighbors; they perform the same.

Finally, in DFE error propagation occurs. Decisions error of users decoded first are fed back and increase the error rate and lower the actual SINR of users detected later on. Since at the transmitter the symbols are correctly known, this effect does not apply to THP.

As in the case of linear schemes, the uplink-downlink duality in practical situations between DFE and THP is assessed by numerical examples. The parameters are the same as in Figure 2, additionally, in DFE the decision order is 1, 2, and user 3 is detected last (i.e., $\boldsymbol{P}_{\mathrm{u}} = \boldsymbol{I}$, which is the optimum order according to the BLAST algorithm for the present channel); THP encodes user 3 first and user 1 last. The plots in the first and second row of Fig. 4 are valid for perfect CSI whereas third and fourth row correspond to imperfect CSI.

Comparing the SINRs in uplink (DFE) and downlink (THP), some degradation of the third user in DFE due to error propagation in the successive decision process is apparent. Assuming perfect feedback symbols ("genie-aided" DFE), this loss vanishes and both transmission directions achieve approximately the same SINR. With respect to BER, significant differences (much more as for the linear schemes) are visible. Again, DFE suffers from error propagation, leading to a flattening of the BER of the user detected last. Moreover, the non-Gaussian interference in THP—its pdf differ for each power distribution—lead to a non-convex $(\mathrm{BER}_1, \mathrm{BER}_3)$ curve. Ignoring error propagation, the BER performance of DFE is exactly predicted by (12). It should be noted, that in comparison with Figure 2 the higher SINRs and lower BERs of user 3 (and user 2) for the successive schemes over the linear ones can be observed. User 1 (detected first/encoded last) does not gain in neither case.

In the second row, dither is used in THP and DFE—the filters are then adjusted according to the biased MMSE solution, cf. [7]. Due to the increased transmit power (factor γ_{p}^2), higher SINRs are achieved for all schemes. THP and genie-aided DFE now match perfectly, even with respect to error rates. However, due to the periodic extension of the signal constellation (which is worst for the present case of 4-ary QAM) higher error rates occur. Hence, dither is only a theoretical concept but it is a means to perfectly match up- and downlink.

For comparison, the performance when using different constellation sizes is compiled in Figure 5. The results for 4-ary QAM at $3\ \mathrm{dB}$ are repeated and the results for 16-ary QAM at $7\ \mathrm{dB}$ and 64-ary QAM at $11.5\ \mathrm{dB}$ are shown—the SNRs are chosen such that DFE exhibits (almost) the same error rates for all schemes. Going to higher modulations, the differences in error rate between DFE (uplink) and THP (downlink) get smaller. This is due to the smaller precoding loss in THP and because interference tends to become closer to Gaussian.

C. Imperfect CSI

As a further source for violating the uplink-downlink duality, the impact of imperfect CSI is addressed. For the successive schemes, exactly the same arguments as for the linear ones apply. If the estimation error is not taken into account in the design of the matrices ("naive" approach), the SINR_k is degraded and can be calculated by (19) and (21), respectively, but replacing ζ by $\hat{\zeta}$, or in

case of THP and DFE with dither by $\hat{\zeta}_{\mathrm{p}} \stackrel{\mathrm{def}}{=} \hat{\zeta}\gamma_{\mathrm{p}}^2$.

Moreover, if the power allocation is calculated as for perfect CSI, a mismatch between up- and downlink occurs. Interestingly, the connection between up- and downlink SINR is again given by (14), (15), and (16). Hence, the degradation due to imperfect CSI is (almost) the same for DFE in the uplink and THP in the downlink.

Finally, if at least the error variance σ_δ^2 is known, in the Cholesky factorization (17) ζ should be replaced by $\hat{\zeta}$ (DFE) or $\hat{\zeta}_{\mathrm{p}}$ (THP and DFE with dither), respectively. Utilizing this partial CSI will (slightly) gain in SINR.

The achievable SINR/BER pairs for DFE and THP in case of imperfect CSI are plotted in the last two rows of Fig. 4. No dither is used. Ignoring error propagation, up- and downlink again exhibit the same SINR, but somewhat degraded compared to perfect CSI. For moderate channel error variances, even the naive approach for filter calculation causes almost no loss. With respect to BER, THP suffers from the increased number of nearest neighbors and the non-Gaussian interference. In practice, where error propagation in DFE cannot be avoided, dependent on the power distribution either up- or downlink can perform best.

Comparing the degradation due to imperfect CSI for DFE and THP on the one side and the linear schemes on the other side, similar losses are observable (compare the last rows of Figures 2 and 4). It can be stated that the nonlinear schemes are not more sensitive to channel variations than the linear ones—even in case of imperfect CSI, THP and (genie-aided) DFE outperform linear (pre-)equalization.

V. Conclusion

In this paper we have assessed the uplink-downlink duality, taking constraints imposed in practice into considerations. First, in practical schemes, non-Gaussian signaling is employed, second, implementable precoding schemes—and not the concept of Costa precoding—are used, and, finally, only imperfect channel knowledge is available at the central receiver/transmitter. It is shown that in these situations, the study of SINRs (or capacities) is no longer sufficient as the residual interference is non-Gaussian. In turn, there is no one-to-one relation between SINR and BER, which in this paper is studied in parallel to the SINR.

Sources for violating the uplink-downlink duality in practical schemes are identified, namely error propagation in the DFE feedback loop, the use of modulo congruent signal points (and hence higher number of nearest neighbors) in THP, and increased transmit power do to precoding loss. It is shown that using dither in up- and downlink, identical performance of both transmission direction can be forced. Finally, even in case of imperfect channel knowledge at the central unit, the same SINR

performance of up- and downlink can be guaranteed. However, with respect to error rates, some deviations are observable.

Using channel coding instead of uncoded transmission, similar effects will occur. Non-Gaussian interference and error propagation might be less parasitic for very strong codes but the increased number of nearest neighbors and the precoding loss will be detrimental in each case.

REFERENCES

[1] J.M. Cioffi, G.P. Dudevoir, M.V. Eyuboğlu, and G.D. Forney. MMSE Decision-Feedback Equalizers and Coding — Part I: Equalization Results, Part II: Coding Results. *IEEE Transactions on Communications*, COM-43:2582–2604, October 1995.

[2] J.M. Cioffi and G.D. Forney. Canonical Packet Transmission on the ISI Channel with Gaussian Noise. In *IEEE Global Telecommunications Conference'96*, pages 1405–1410, London, November 1996.

[3] M.H.M. Costa. Writing on Dirty Paper. *IEEE Transactions on Information Theory*, 29(3):439–441, May 1983.

[4] F.A. Dietrich, W. Utschick, and P. Breun. Linear precoding based on a stochastic MSE criterion. In *Proceedings of the 13th European Signal Processing Conference*, Antalya, Türkei, September 2005.

[5] U. Erez, S. Shamai, and R. Zamir. Capacity and Lattice Strategies for Canceling Known Interference. *IEEE Transactions on Information Theory*, 51(11):3820–3833, November 2005.

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

[7] R.F.H. Fischer and C. Siegl. Inflated Lattice Precoding, Bias Compensation, and the Uplink/Downlink Duality: The Connection. *IEEE Communications Letters*, 11(2):185–187, February 2007.

[8] G.D. Forney. On the role of MMSE estimation in approaching the information- theoretic limits of linear Gaussian channels: Shannon meets Wiener. In *41st Annual Allerton Conf. on Comm., Control, and Computing*, October 2003.

[9] G.D. Golden, G.J. Foschini, R.A. Valenzuela, and P.W. Wolniansky. Detection Algorithm and Initial Laboratory Results using V-BLAST Space-Time Communication Architecture. *Electronics Letters*, pages 14–16, January 1999.

[10] G. Golub and C. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, 1996.

[11] N. Jindal, S. Vishwanath, and A.J. Goldsmith. On the Duality of Gaussian Multiple-Access and Broadcast Channels. *IEEE Transactions on Information Theory*, 50(4):768–783, April 2004.

[12] D.P. Palomar, M.A. Lagunas, and J.M. Cioffi. Optimum Linear Joint Transmit-Receive Processing for MIMO Channels with QoS Constraints. *IEEE Transactions on Signal Processing*, 52(5):1179–1197, May 2004.

[13] J.G. Proakis. *Digital Communications*. McGraw-Hill, 2001.

[14] M. Schubert and H. Boche. A Unifying Theory for Uplink and Downlink Multi-User Beamforming. In *Proc. IEEE Intern. Zurich Seminar (IZS 2002)*, pages 27/1–27/6, Zurich, Switzerland, February 2002.

[15] G. Taricco and E. Biglieri. Space-Time Decoding with Imperfect Channel Estimation. *IEEE Transactions on Wireless Communications*, 4(4):1874–1888, July 2005.

[16] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication*. Cambridge Univ. Press, Cambridge, UK, 2005.

[17] P. Viswanath and D.N.C. Tse. Sum Capacity of the Vector Gaussian Broadcast Channel and Uplink-Downlink Duality. *IEEE Transactions on Information Theory*, 49(8):1912–1921, August 2003.

[18] C. Windpassinger. *Detection and Precoding for Multiple Input Multiple Output Channels*. PhD thesis, Universtät Erlangen-Nürnberg, 2004.

[19] W. Yu and J.M. Cioffi. Sum Capacity of Gaussian Vector Broadcast Channels. *IEEE Transactions on Information Theory*, 50(9):1875–1892, September 2004.

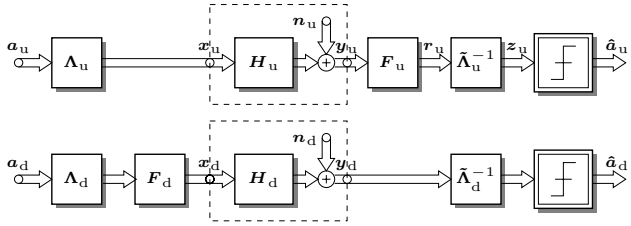

Fig. 1. Linear equalization in the uplink (top) and linear pre-equalization in the downlink (bottom).

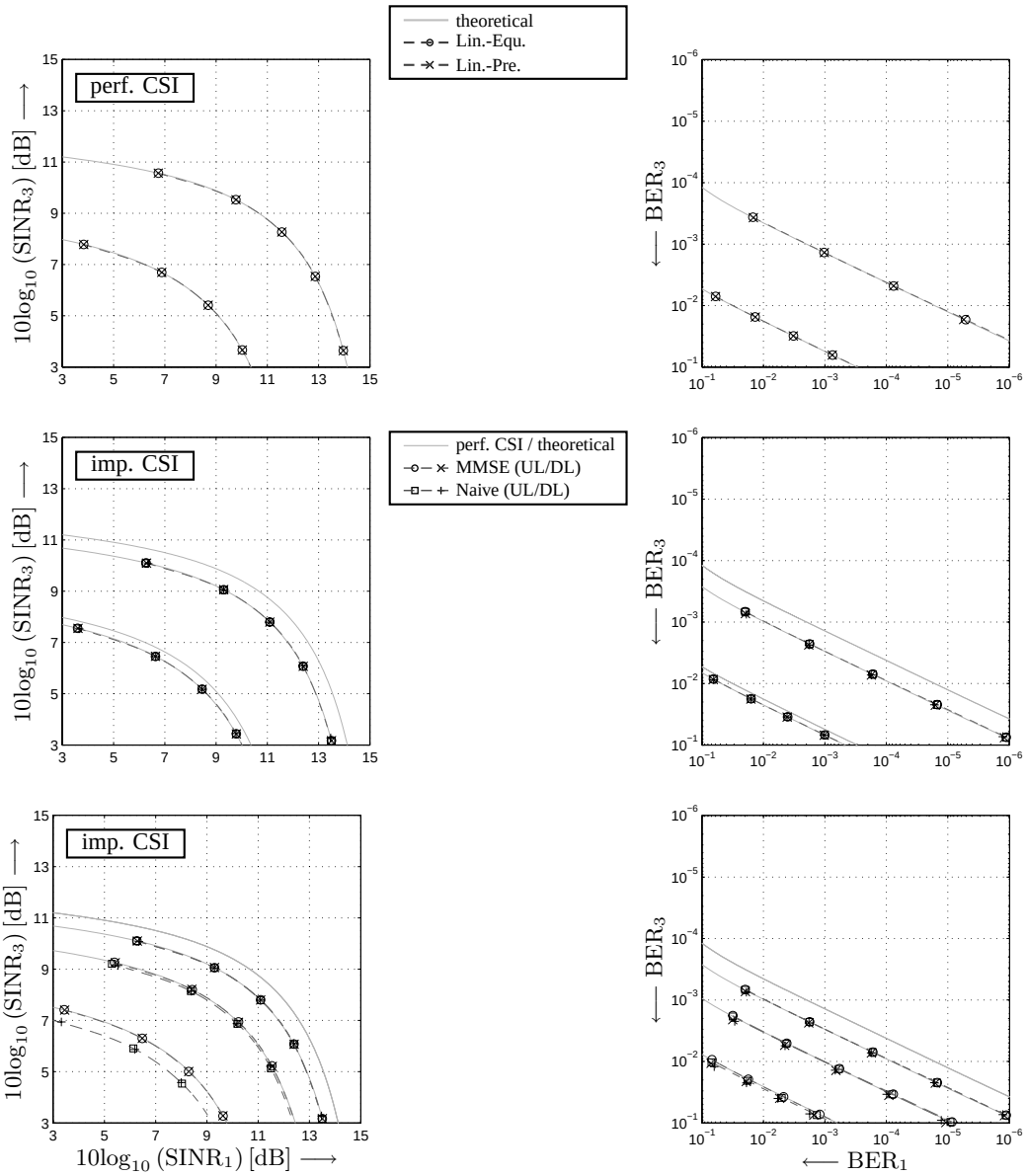

Fig. 2. Comparison of achievable pairs of SINRs (left) and BERs (right) for linear equalization/pre-equalization. Variation of the uplink power distribution p–$(1-p)$ (from left to right): 17%–83%, 33%–67%, 50%–50%, 67%–33%, 83%–17%. Channel according to (13). Solid: theoretical result; Circles (uplink) and crosses (downlink) 4-QAM modulation. Top row: perfect CSI; middle row: imperfect CSI with naive and MMSE filter calculation, $\sigma_\delta^2 = 0.01$ (-20 dB); top and middle rows: $10\log_{10}(\bar{E}_{\mathrm{b,Tx}}/N_0) = 0$ dB, 3 dB; bottom row: imperfect CSI with MMSE filter calculation, $\sigma_\delta^2 = 0.01,\ 0.03,\ 0.1$ ($-20, -15, -10$ dB), $10\log_{10}(\bar{E}_{\mathrm{b,Tx}}/N_0) = 3$ dB.

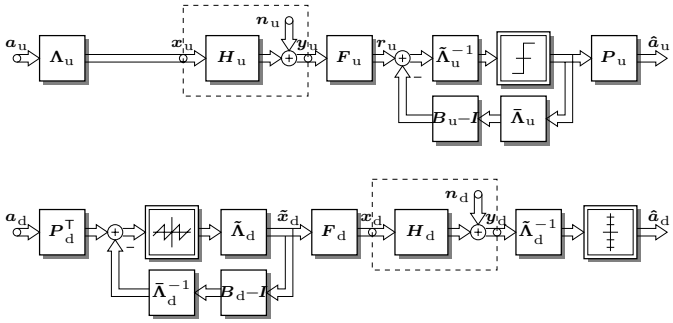

Fig. 3. Decision-feedback equalization in the uplink (top) and Tomlinson-Harashima precoding in the downlink (bottom). The decision device in the THP scheme (without the saturation) takes the modulo congruence of signal points into account.

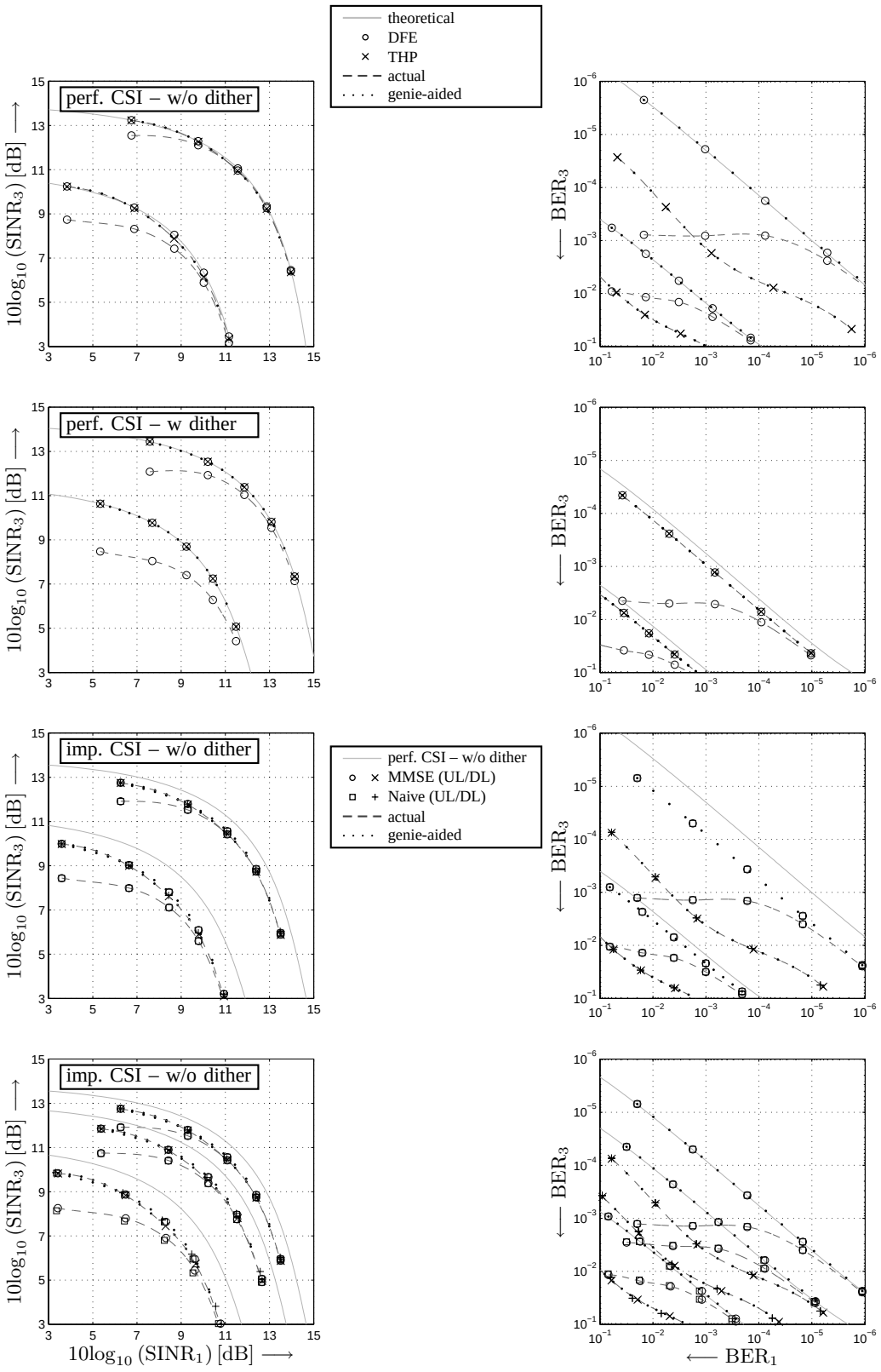

Fig. 4. Comparison of achievable pairs of SINRs (left) and BERs (right) for successive schemes (DFE, THP). Uplink power distribution p–$(1-p)$ (from left to right): 17%–83%, 33%–67%, 50%–50%, 67%–33%, 83%–17%. Channel according to (13). Solid: theoretical result; Circles (uplink) and crosses (downlink) 4-QAM modulation. Top row: perfect CSI; second row: perfect CSI and dither; third row: imperfect CSI with naive and MMSE filter calculation, $\sigma_\delta^2 = 0.01$ (-20 dB); first three rows: $10\log_{10}(\bar{E}_{\mathrm{b,Tx}}/N_0) = 0$ dB, 3 dB; bottom row: imperfect CSI with MMSE filter calculation, $\sigma_\delta^2 = 0.01,\ 0.03,\ 0.1$ ($-20, -15, -10$ dB), $10\log_{10}(\bar{E}_{\mathrm{b,Tx}}/N_0) = 3$ dB.

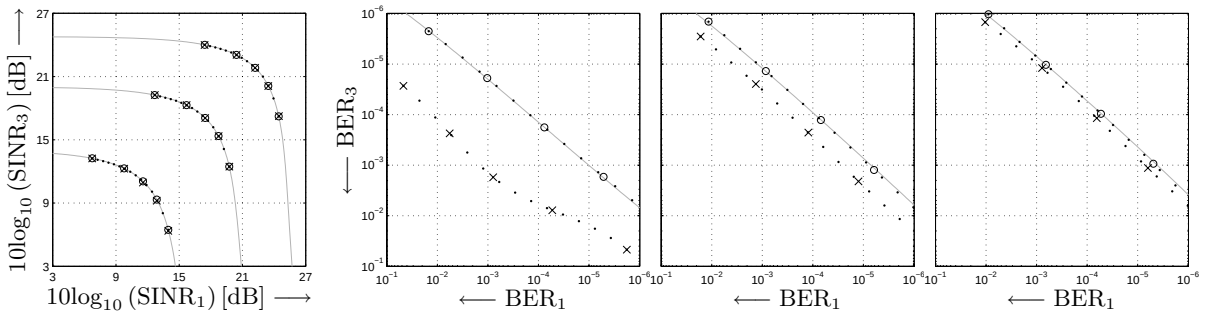

Fig. 5. Comparison of achievable pairs of SINRs (left) and BERs (the three other plots) successive schemes (DFE, THP) and uplink power distributions as above. Solid: theoretical result; Circles (uplink) and crosses (downlink) Left to right: 4-ary QAM at $10\log_{10}(\bar{E}_{\mathrm{b,Tx}}/N_0) = 3$ dB, 16-ary QAM at 7 dB, and 64-ary QAM at 11.5 dB. Perfect CSI.

TABLE I

SOURCES FOR OF VIOLATING THE UPLINK-DOWNLINK DUALITY IN PRACTICAL SCENARIOS.

		non-Gaussian Interference	Variation of SINR		Increase in $N_{\min}$	Error Propagation
			imperfect CSI	TX power increase		
linear	perfect CSI	$\times$	—	—	—	—
	imperfect CSI	$\times$	$\times$	—	—	—
DFE/THP	perfect CSI	$\times$	—	THP, DFE w/ dither	$\rightarrow 4$	DFE
	imperfect CSI	$\times$	$\times$	THP, DFE w/ dither	$\rightarrow 4$	DFE