Optimal Precoding and MMSE Receiver Designs for MIMO WCDMA

Shakti Prasad Shenoy*†, Irfan Ghauri*, Dirk T.M. Slock†

*Infineon Technologies France SAS, GAIA, 2600 Route des Crêtes, 06560 Sophia Antipolis Cedex, France
Email: shakti.shenoy@eurecom.fr, irfan.ghauri@infineon.com

†Mobile Communications Department, Institut Eurécom, 2229 Route des Crêtes, 06904 Sophia Antipolis Cedex, France
Email: dirk.slock@eurecom.fr

Abstract—2×2 **unitary precoding based on receiver feedback is applied alongside spatial multiplexing at the base station in HSDPA (D-TxAA) when the mobile terminal supports MIMO transmissios [1]. This precoding will influence achievable sum-rate of the MIMO channel if it influences the Signal-to-Interference-plus-Noise Ratio (SINR) of streams at the receiver (RX) output. We propose a set of MIMO HSDPA receivers, all based upon a LMMSE chip-level matrix filter (equalizer) front end, and introduce the notion of *joint bias* for the MIMO chip equalizer. Statistical properties of the spatial model thus obtained are exploited to analyze the performance of proposed MIMO receivers. It is shown that precoding choice depends upon the MIMO receiver and the extent of its impact depends on the MIMO RX.**

I. INTRODUCTION

3GPP has introduced a variant of Per-Antenna Rate-Control (PARC), namely D-TxAA for Dual-stream Transmit Diversity for Multi-Input Multi-Output (MIMO) transmissions [1] in UMTS WCDMA. Code reuse is made across the two streams and the scrambling sequence is also common to both transmit (TX) streams. All (15) spreading codes are allocated to the same user in the HSDPA MIMO context. In general, all UEs served by a BS feed an SINR-based (or based on some other appropriate measure) Channel Quality Indicator (CQI) back to the BS. In addition, the UE also computes (and feeds back) the weighting vector(s) that would ideally provide the best instantaneous rate for the next time slot. Together, these feedbacks translate into a specific transport block size and a specific Modulation and Coding Scheme (MCS) for each UE. Based on this information, the BS is capable of maximizing the downlink throughput for each transmission time-interval.

MIMO has largely been discussed in the context of frequency non-selective (OFDM) case, where optimal joint-stream (MAP) detection can be employed. In CDMA, on the other hand, multipath mixes signals up in space and time calling for alternative reception strategies. Proposals for receiver (RX) solutions include chip-level equalization and depreading followed by joint detection of the data streams at symbol level [2]. More generally, a two stage approach is considered where the first stage is the chip-equalizer correlator followed by some kind of joint processing or decision-feedback approach [3]. In practice, the symbol-level spatial channel can now be seen as a per-code spatial mixture.

Alternatively, more general FIR channel shortening can be considered [3] leading asymptotically (in SNR) introducing MIMO spatial *joint* channel which can indeed be followed up by a CDMA code-correlator and per-code symbol-level multi-stream MAP detector. Other receiver options with varying degrees of complexity could also be considered, e.g., (a) symbol-level MMSE, which is the true linear MMSE receiver for the symbol sequence but is time-varying due to the aperiodic scrambler and (b) non-linear (turbo)-iterated serial and/or parallel (SIC/PIC) interference canceler for all user codes. In general, all attempts to simplify processing fall well short of optimal time-varying symbol-level processing.

For D-TxAA with unitary precoding, there exists an optimal choice of the precoding matrix that would maximize the sum rate across the two streams. In principle, the receiver can evaluate the SINR corresponding to all precoding choices and request the application of the SINR-maximizing weights for the next TX frame. The receiver further signals a CQI for each stream that can be mapped to a particular MCS. The data packet size associated with a particular MCS can then be mapped to obtain the supported throughput for each stream for a certain pre-defined Packet-Error Rate (PER). The mapping strategy has been subject to significant simulation study (see e.g., [4]) and SINR → CQI ↔ PER ↔ throughput relationship has been agreed to, appearing as CQI to MCS tables in the 3GPP standard document [1].

In this paper, we analyze performance of a variety of simple receivers for unitary precoded D-TxAA MIMO in the HSDPA context. We propose several receiver structures and derive SINR expressions per stream for each of them. We then compare their performance in terms of their sum-rate capacity which can be interpreted as upper bound for achievable rates.

II. MIMO SIGNAL MODEL

For the spatial multiplexing case in MIMO HDSPA, Fig. 1 illustrates the equivalent baseband downlink signal model. The received
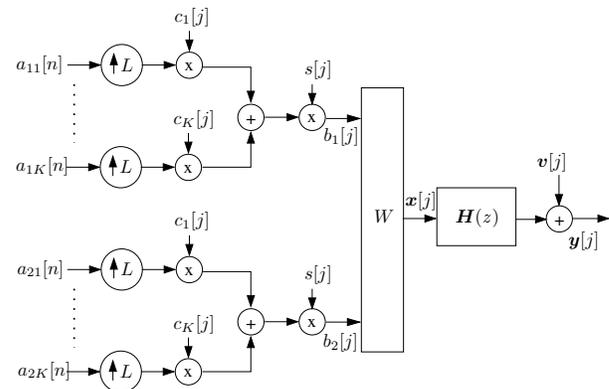
Fig. 1. MIMO signal model with precoding.

signal vector (chip-rate) at the UE can be modeled as

$$\underbrace{\boldsymbol{y}[j]}_{2m \times 1} = \underbrace{\boldsymbol{H}(z)}_{2m \times 2} \underbrace{\boldsymbol{x}[j]}_{2 \times 1} + \underbrace{\boldsymbol{v}[j]}_{2 \times 1}. \tag{1}$$

In this model, j is the chip index, $\boldsymbol{H}(z)$ is the frequency selective MIMO channel the output of which is sampled m times per chip and $\boldsymbol{v}[j]$ represents the vector of noise samples that are zero-mean circular Gaussian random variables. The sequence $\boldsymbol{x}[j]$ introduced into the channel is itself a linear combination (D-TxAA see [1]) of the two steams and is expressed as

$$\boldsymbol{x}[j] = \underbrace{\boldsymbol{W}}_{2 \times 2} \boldsymbol{b}[j] = \boldsymbol{W} \cdot \sum_{k=1}^{K} \underbrace{s[j]c_k[j \mod L]\boldsymbol{a}_k[n]}_{\boldsymbol{b}_k[j]} \tag{2}$$

k is the code index, p is the index of the symbol on code k given by $n = \lfloor \frac{j}{L} \rfloor$, L is the spreading factor ($L = 16$ for HSDPA), $\boldsymbol{W} = [\boldsymbol{w}_1 \, \boldsymbol{w}_2]$ is the 2×2 precoding matrix with $\boldsymbol{w}_1 = [\frac{1}{\sqrt{2}} \, w]^T$ and $\boldsymbol{w}_2 = [\frac{1}{\sqrt{2}} \, -w]^T$. The symbol vector $\boldsymbol{a}_k[n] = [a_{1k}[n] \, a_{2k}[n]]^T$ represents two independent symbol streams, $\boldsymbol{c}_k = [c_k[0] \, \ldots \, c_k[L-1]]^T$, where $\boldsymbol{c}_i^T \cdot \boldsymbol{c}_j = \delta_{ij}$ are unit-norm spreading codes common to the two streams, and $s[j]$ the common scrambling sequence element at chip time j, which is zero-mean *i.i.d.* with elements from $\frac{1}{\sqrt{2}}\{\pm 1 \pm j\}$.

III. MIMO HSDPA RECEIVER STRUCTURES

A. Receiver 1: MMSE Chip Equalizer-Correlator

In the spatial multiplexing context, the LMMSE equalization tries not only to suppress all Inter-Chip Interference (ICI) but also all Inter-Stream Interference (ISI). The 2×2 linear FIR MMSE chip-level equalizer is $\boldsymbol{F} = \boldsymbol{R_{xy}} \boldsymbol{R_{yy}^{-1}}$ (see fig. 2). We can write the equalizer output as the sum of an arbitrarily scaled desired term and an error term

$$\widehat{\boldsymbol{x}}[j] = \boldsymbol{x}[j] - \widetilde{\boldsymbol{x}}[j]. \qquad (3)$$

The error $\widetilde{\boldsymbol{x}}[j]$ is a zero-mean complex normal random variable. The error covariance matrix is denoted by $\boldsymbol{R}_{\widetilde{x}\widetilde{x}}$.

In (3), an estimate of the chip sequence can be obtained after a further stage of processing where the precoding is undone to separate streams. The latter represented by $\boldsymbol{W}^H$ is a linear operation and can be carried out before or after despreading (the latter case is shown in fig. 2). d represents the equalization delay in chips.
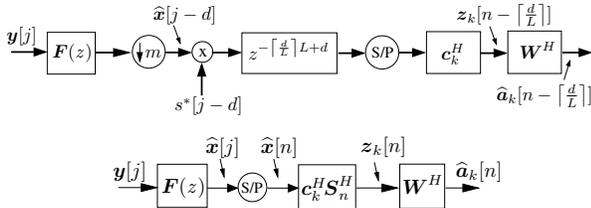
Fig. 2. LMMSE equalizer and correlator. The second figure is a simplified representation used as chip-equalizer /correlator front-end stage for other receiver structures.

After despreading (for the kth code) the 2×1 signal at the symbol level is written as

$$\boldsymbol{z}_k[n] = \boldsymbol{W} \boldsymbol{a}_k[n] - \widetilde{\boldsymbol{z}}_k[n] = \boldsymbol{B} \boldsymbol{W} \boldsymbol{a}_k[n] - \widetilde{\widetilde{\boldsymbol{z}}}_k[n]. \qquad (4)$$

In this expression, $\boldsymbol{B}$ is the LMMSE *joint bias* at the output of the chip-equalizer/correlator (see appendix A), and by consequence the quantity $\widetilde{\widetilde{\boldsymbol{z}}}_k[n]$ therefore contains no desired symbol contribution.

Note that in this RX structure we assume $\boldsymbol{W}^H \boldsymbol{z}_k[n]$ to be the decision statistic.

1) Estimation of $\boldsymbol{R}_{\widetilde{z}\widetilde{z}}$: Under the assumption of a FIR signal model, the estimation error covariance matrices $\boldsymbol{R}_{\widetilde{x}\widetilde{x}}$ (chip-level) and $\boldsymbol{R}_{\widetilde{z}\widetilde{z}}$ (symbol-level) are derived in appendix A.

2) Output SINR: From analysis of $\boldsymbol{R}_{\widetilde{z}\widetilde{z}}$, it can be shown that the SINR for the ith stream at the output of the output of the LMMSE chip equalizer/correlator is given by

$$SINR_i = \frac{\sigma_a^2}{\left(\boldsymbol{W}^H \boldsymbol{B}^{-1} \boldsymbol{R}_{\widetilde{z}\widetilde{z}} \boldsymbol{B}^{-H} \boldsymbol{W}\right)_{ii}} - 1. \qquad (5)$$

Taking expectation over the time-varying (random) scrambling sequence as is customary, the bias term can be considered to be constant at the equalizer/correlator output.

Once MIMO joint bias is properly taken into account (see appendix A), the expression for the LMMSE chip equalizer output SINR is exact. The situation is different at the symbol-level where the bias,

in practice, varies over time. However, this issue is beyond the scope of this paper and will be discussed elsewhere.

The per-code capacity of the ith data stream therefore corresponds to

$$\begin{aligned} C_i &= \log(1 + SINR_i) \\ C_i &= \log\left(\frac{\sigma_a^2}{MMSE_i}\right) \end{aligned} \qquad (6)$$

Our objective is to choose the precoding matrix $\boldsymbol{W}$ to maximize the sum-capacity of two streams. This boils down to the following optimization problem:

$$\boldsymbol{W}_{opt} = \arg\max_{\boldsymbol{W}} \left[\log\left(\frac{\sigma_a^4}{MMSE_1 \cdot MMSE_2}\right)\right]. \qquad (7)$$

The optimum precoding matrix can be seen to minimize the product of MMSEs of the streams. By exploiting the structure of the matrices in the unitary codebook specified in the HSDPA standard [1], the optimum precoding matrix $\boldsymbol{W}_{opt}$ maximizes $\Re(|wr_{12}|)$, where r_{12} is defined in appendix A as the off-diagonal term of the error covariance matrix $\boldsymbol{R}_{\widetilde{z}\widetilde{z}}$. In other words, the $\boldsymbol{W}_{opt}$ attempts to maximize the SINR difference between the two streams[1].

B. Receiver 2: MMSE Chip Equalizer-Symbol Level LMMSE

In an alternative receiver structure, the output of the chip-equalizer is fed into a symbol level (spatial) LMMSE filter after the descrambler/correlator block. This is shown in Fig. 3. As discussed in III-A, the output of the correlator is $\boldsymbol{z}_k[n]$ given by (4). $\mathcal{F}_{sp}$ denotes the
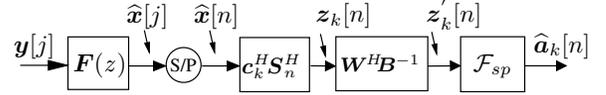
Fig. 3. Chip LMMSE equalizer and correlator followed by symbol-level (spatial) MMSE.

spatial MMSE at the output of which we have a linear estimate of the symbol vector as

$$\widehat{\boldsymbol{a}}_k[n] = \boldsymbol{a}_k[n] - \widetilde{\boldsymbol{a}}_k[n]. \qquad (8)$$

The error covariance matrix for the LMMSE estimate of $\boldsymbol{a}_k[n]$ is given by

$$\begin{aligned} \boldsymbol{R}_{\widetilde{a}\widetilde{a}} &= \boldsymbol{R}_{aa} - \boldsymbol{R}_{az'} \boldsymbol{R}_{z'z'}^{-1} \boldsymbol{R}_{z'a} & (9) \\ &= \sigma_a^2 \boldsymbol{I} - \sigma_a^4 \boldsymbol{W}^H \left(\sigma_a^2 \boldsymbol{I} + \boldsymbol{B}^{-1} \boldsymbol{R}_{\widetilde{z}\widetilde{z}} \boldsymbol{B}^{-H}\right)^{-1} \boldsymbol{W}. & (10) \end{aligned}$$

Expressing the above relation in terms of the correlator output covariances, $\boldsymbol{B} \boldsymbol{R}_{\widetilde{z}\widetilde{z}} \boldsymbol{B}^{-H}$ and using some algebra leads to the expression

$$R_{\widetilde{a}\widetilde{a}} = \sigma_a^2 \boldsymbol{I} - \sigma_a^4 \boldsymbol{W}^H \left(\sigma_a^2 \boldsymbol{I} + \left(\boldsymbol{R}_{\widetilde{z}\widetilde{z}}^{-1} - \boldsymbol{R}_{zz}^{-1}\right)^{-1}\right)^{-1} \boldsymbol{W}. \qquad (11)$$

Like the LMMSE chip level equalizer/correlator RX, this translates to a sum-capacity expression similar to the one derived in the previous section.

$$C_1 + C_2 = \log\left(\frac{\sigma_a^4}{\det(\text{diag}(\boldsymbol{R}_{\widetilde{a}\widetilde{a}}))}\right) \qquad (12)$$

The throughput maximizing precoding matrix can therefore be shown to be the one with element w that maximizes

$$\Re\left(\left|w \left[\left(\sigma_a^2 \boldsymbol{I} + \left(\boldsymbol{R}_{\widetilde{z}\widetilde{z}}^{-1} - \boldsymbol{R}_{zz}^{-1}\right)^{-1}\right)^{-1}\right]_{12}\right|\right)$$

One may remark that spatial MMSE processing after the equalizer/correlator stage should lead to further suppression of residual

---

[1]to its best abilities given the limited resolution of $\boldsymbol{W}$.

interference and lends itself to low-complexity per-code implementation.

The spatial channel sees a non-negligible contribution from the kth code (desired code), as seen in section IV, RX 2 does mprove on RX 1 but its performance is still limited by the temporal (inter-chip) interference that is still sufficiently strong at the correlator output.

C. Receiver 3: MMSE Chip Equalizer - Predictive DFE

A noise-predicive decision feedback equalizer (DFE) [5] uses past noise estimates to predict the current noise sample. This is readily applied to our spatial-multiplexing problem where once one stream is detected, spatial correlation of noise (spatial interference) can be exploited to improve estimation of the stream detected last (second in this case). With some abuse of terminology this can be branded Successive Interference Cancellation (SIC).

The SIC receiver is shown in Fig. 4. Denote the output of the correlator as $u_k[n]$, written as

$$\boldsymbol{u}_k[n] = \boldsymbol{W}^H \boldsymbol{B}^{-1} \mathcal{F}_{sp} \boldsymbol{z}'_{k,n} = \boldsymbol{a}_k[n] - \underbrace{\mathcal{F}_{sp} \boldsymbol{W}^H \boldsymbol{B}^{-1} \widetilde{\widetilde{\boldsymbol{z}}}_k[n]}_{\widetilde{\widetilde{\boldsymbol{u}}}_k[n]} \quad (13)$$

The covariance matrix $\boldsymbol{R}_{\widetilde{\widetilde{\boldsymbol{u}}}\widetilde{\widetilde{\boldsymbol{u}}}}$, the diagonal bias matrix $\boldsymbol{B}$ and $\boldsymbol{R}_{\widetilde{\widetilde{\boldsymbol{z}}}\widetilde{\widetilde{\boldsymbol{z}}}}$, the covariance matrix of $\widetilde{\widetilde{\boldsymbol{z}}}$ can be related as

$$\boldsymbol{R}_{\widetilde{\widetilde{\boldsymbol{u}}}\widetilde{\widetilde{\boldsymbol{u}}}} = \mathcal{F}_{sp} \boldsymbol{W}^H \boldsymbol{B}^{-1} \boldsymbol{R}_{\widetilde{\widetilde{\boldsymbol{z}}}\widetilde{\widetilde{\boldsymbol{z}}}} \boldsymbol{B}^{-H} \boldsymbol{W} \mathcal{F}_{sp}^H \quad (14)$$

Assume a 2×2 lower triangular filter $\mathcal{G}_{sp}$ with unit diagonal and the remaining element g_{21} such that $\widetilde{\boldsymbol{r}}[n] = \mathcal{G}_{sp} \widetilde{\widetilde{\boldsymbol{u}}}_k[n]$. Then the new error covariance matrix is given as

$$\boldsymbol{R}_{\widetilde{\boldsymbol{r}}\widetilde{\boldsymbol{r}}} = \mathcal{G}_{sp} \boldsymbol{R}_{\widetilde{\widetilde{\boldsymbol{u}}}\widetilde{\widetilde{\boldsymbol{u}}}} \mathcal{G}_{sp}^H. \quad (15)$$

which is minimized if $\boldsymbol{R}_{\widetilde{\boldsymbol{r}}\widetilde{\boldsymbol{r}}} = \boldsymbol{D}$, i.e., a diagonal matrix and the problem boils down to the estimation of the error term in stream 2 from stream 1. Towards this end, consider LDU factorization of $\boldsymbol{R}_{\widetilde{\widetilde{\boldsymbol{u}}}\widetilde{\widetilde{\boldsymbol{u}}}} = \boldsymbol{L}\boldsymbol{D}\boldsymbol{L}^H$. Then, $\mathcal{G}_{sp} = \boldsymbol{L}^{-1}$ minimizes (15). Denoting elements of $\boldsymbol{R}_{\widetilde{\widetilde{\boldsymbol{u}}}\widetilde{\widetilde{\boldsymbol{u}}}}$ as r_{ij} as in (25), the elements of $\boldsymbol{D}$ are given as $\sigma_{\widetilde{r}_1}^2 = r_{11}$ and

$$\begin{aligned} \sigma_{\widetilde{r}_2}^2 &= r_{22} - r_{21} r_{11}^{-1} r_{12} \\ &= \det(\boldsymbol{R}_{\widetilde{\widetilde{\boldsymbol{u}}}\widetilde{\widetilde{\boldsymbol{u}}}}) \\ &= \det(\mathcal{F}_{sp}) \det(\boldsymbol{B}^{-1} \boldsymbol{R}_{\widetilde{\widetilde{\boldsymbol{z}}}\widetilde{\widetilde{\boldsymbol{z}}}} \boldsymbol{B}^{-H}) \det(\mathcal{F}_{sp}^H). \end{aligned} \quad (16)$$

Thus MMSE for stream 1 is $\sigma_{\widetilde{r}_1}^2$ and that of stream 2 is $\sigma_{\widetilde{r}_2}^2$. As depicted in fig. 4, this can be interpreted as stream 1 achieving the same performance as for the chip-level LMMSE/correlator - spatial MMSE (RX 2 above), while stream 2 benefits from stripping (and thus achieves the spatial MFB). The rates are therefore expressed as
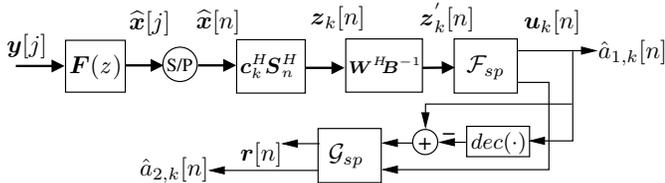
Fig. 4. Chip LMMSE equalizer/correlator followed by spatial MMSE and symbol-level SIC for stream 2.

$$C_i = \log(1 + SINR_i) \quad (17)$$

Another interesting observation is that the SINR expression for stream 2 in the symbol-level SIC case is independent of the precoding $\boldsymbol{W}$ applied.

1) Discussion: In this receiver, stream 1 should exhibit better performance than in the case of RX 1. An alternative receiver structure proposed in [3] is also possible where stream 1 processing is just limited to the chip equalizer-correlator cascade and stream 2 is subjected to symbol-level SIC as above. However, RX 3 is a better alternative to [3], since in this case, stream 1 should get an additional boost in SINR due to the spatial MMSE processing. This should not only amplify stream 1 rate, but also has the desirable effect of improving stream 1 detection. This improved reliability, although not relevant in this discussion where we assume ideal suppression of stream 1 is all-important in practical implementations, reducing chances of error-propagation during the interference cancellation stage and hence directly impacting detection performance of stream 2.

D. Receiver 4: Chip Level SIC

Indeed a better known SIC receiver detects data symbols from one stream, say stream 1 and respreads, rescrambles, rechannelizes detected data, the contribution of that stream can be subtracted from the received signal. The second stream can now be detected using a new FIR LMMSE chip-level receiver obtained as

$$\boldsymbol{f}_{sic} = \sigma_b^2 \left(\frac{1}{\sqrt{2}} \widetilde{\boldsymbol{h}}_1^H - w \widetilde{\boldsymbol{h}}_2^H \right) \boldsymbol{R}_{\overline{y}\overline{y}}^{-1}, \quad (18)$$

where,

$$\overline{\boldsymbol{y}}[n] = \mathcal{T}(\overline{H})\boldsymbol{S}[n]\boldsymbol{C}\boldsymbol{A}_2[n] + \boldsymbol{V}[n], \quad (19)$$

and $\mathcal{T}(\overline{H}) = \frac{1}{\sqrt{2}}\mathcal{T}(H_1) - w\mathcal{T}(H_2)$.

This case, assuming perfect cancellation of stream 1, is analogous to single stream communications and the SINR achieved for stream 2 is much improved. The SINR expressions for this SIC receiver are straightforward and similar to the ones for the MISO LMMSE chip-level equalizer/correlator case (see appendix A). It must be noted that there is significant structural differences between the two SIC receivers that also translate to behavior differences - one such consideration is the possibility of chip-level SIC to cancel intercell interference. One further consideration in RX 4 is that if stream
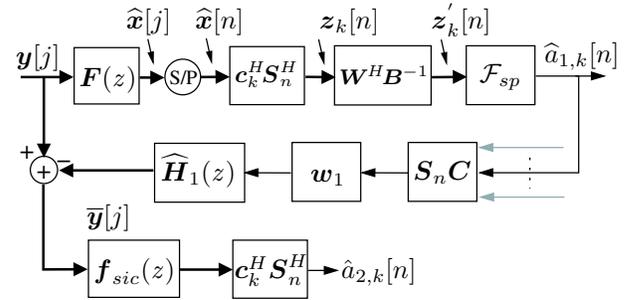
Fig. 5. Chip LMMSE equalizer/correlator followed by spatial MMSE and chip-level SIC for stream 2.

1 symbol estimates are obtained at the output of a spatial MMSE, this would also imply spatial processing for stream 2 (since spatial processing by nature is simultaneous). Such treatment increases complexity but may be well worth the effort in terms of SINR gains and as discussed for RX 3 above, the quality of the estimates of stream 1 before feedback.

1) Different Types of SIC Receivers: The noise-predictive DFE is hardly comparable to chip-level SIC receiver in any other way except that symbols on streams are detected in the order of decreasing SINR. While the former exploits noise plus interference correlation between streams to improve SINR of symbol detected last, the latter benefits from stripping of spatiotemporal interference of the entire detected stream, where for stream detected last, all streams can henceforth be considered non-existent (assuming perfect cancellation). Not only do streams see different levels of interference, a new chip-equalizer can be calculated at each stage that benefits from

a larger noise-subspace to cancel remaining interference. For SIC, stream detected last is known to attain the Matched-Filter Bound (MFB).

A more general feedback has also been proposed in [6] where where it was called chip-level DFE but in fact it is symbol-level DFE (the decisions are on the symbols, not on chips). Even though the feedback interference cancellation is performed at chip level after respreading, but that is equivalent to canceling at symbol level and the equivalence of that solution with RX 4 is not straightforward. In general, many DFE/SIC designs are possible.

E. Receiver 5: Spatial ML Receiver

Another possible receiver structure is shown in Fig. 6 where the chip-equalizer correlator front end is followed up, as before, by the spatial MMSE stage. The resulting spatial mixture

$$\boldsymbol{u}_k[n] = \mathcal{F}_{sp}\boldsymbol{z}_k^{'}[n] = \boldsymbol{a}_k[n] - \widetilde{\widetilde{\boldsymbol{u}}}_k[n], \tag{20}$$

is later processed for joint detection (code-wise ML detection) of the two symbol streams. The ML metric is given as follows.

$$\mathcal{D} = \{\boldsymbol{u}_k[n] - \boldsymbol{a}_k[n]\}^H \boldsymbol{R}_{\widetilde{u}\widetilde{u}}^{-1} \{\boldsymbol{u}_k[n] - \boldsymbol{a}_k[n]\}.$$

This metric can be solved for $\boldsymbol{a}_k[n]$. It was shown in [3] that joint
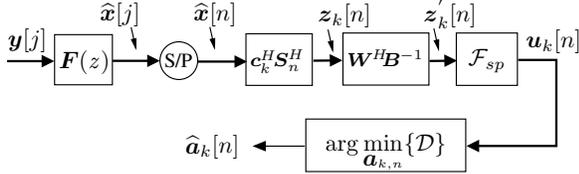
Fig. 6. Chip LMMSE equalizer/correlator followed by spatial MMSE and joint detection.

detection outperforms SIC. It must be however be noted that the SIC structure in [3] addresses a SIC applied directly at the output of the chip equalizer-correlator output. Thus stream 1 gets the same SINR as the chip-equalizer while in our case, stream 1 would also reap the benefits of spatial MMSE processing. For joint detection, the SINR for the ith stream corresponds to the MFB of spatial channel resulting from the cascade of $\mathcal{F}_{sp}$ and $\boldsymbol{B}$. The MFB can be interpreted as the SNR of ith stream when it is detected assuming that symbols of the other stream(s) are known. $\boldsymbol{R}_{\widetilde{u}\widetilde{u}}$ is the noise variance.

IV. SIMULATION RESULTS

We present here the simulation results and compare the performance of the different receiver structures based on their sumcapacity. For a fixed SNR and over several realizations of a frequency selective $2m \times 2$ MIMO FIR channel $\boldsymbol{H}(z)$, we compute the optimal precoding matrices and use the corresponding SINRs of both streams at the output of the receivers to calculate an upperbound on the sum capacity. The channel coefficients are complex valued zero-mean Gaussain of length 20 chips. We assume FIR MIMO equalizers of length comparable to the channel.

The sum-capacity CDF is thus used as a performance measure for all receivers. The structure of the precoding matrices used in HSDPA is such that two out of the four possible precoding matrices give the same SINR (and thus sum-rate) for the LMMSE/correlator design. The difference between them being that one favors stream 1 by bestowing a higher SINR for stream 1, and the other matrix does just the reverse. This means that one can not only achieve the same sum-rate by choosing any of the two matrices, but one can also choose which stream among the two, contributes a larger fraction of the sum. Without loss of generality, in all our simulations, we choose the matrix that maximizes the SINR of stream 1.

Fig. 7 shows distribution of sum-capacity at the output of the MMSE chip-equalizer correlater receiver and that of the spatial MMSE receiver. With an additional processing stage of a very small
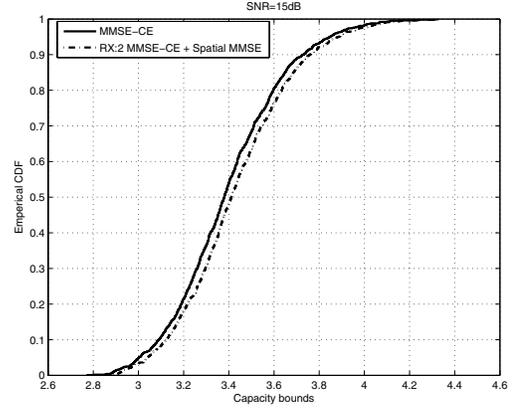
Fig. 7. Upper bounds on the sum-capacity at the output of RX 1 and RX 2 .
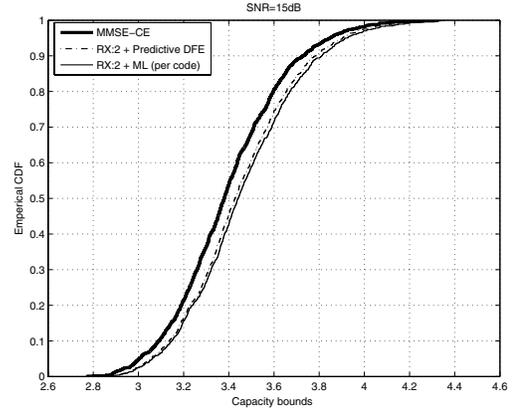
Fig. 8. Sum-capacity at the output of RX 1, RX 3, and RX 5.

complexity we are able to see some gain in the achievable rates of the receiver.

In Fig. 8 we compare the performance of RX 1 with RX 3 and RX 5. As before, optimal precoding matrices are used at the base-station. RX 3 benefits slightly from the additional spatial processing for both streams and a non-linear equalization stage for stream-2. That the gain is not considerable is due to the fact that stream-1 does not benefit from non-linear equalization. Since the performance measure is the sum-capacity of both streams, the performance of this receiver is limited by the performance of stream-1.RX 5 on the otherhand performs better than RX 3 thanks to spatial ML detection performed on a per-code basis.

In Fig. 10 we draw attention to the fact that one should exercise caution while choosing the metric for ML detector in order to compute the correct MFB. The correct metric takes into account the correlation in noise at input of the detector.The chip-level SIC, in Fig. 9 as can be expected, outperforms all other receivers at the cost of a significant complexity at the receiver.

V. CONCLUSIONS

In this contribution, we derived analytical expressions for the choice of the precoding matrix when the precoding matrices are unitary and the receivers are based on MMSE designs. We also compared five distinct receiver structures for D-TxAA MIMO HSDPA all based on the LMMSE chip-level equalizer/correlator as the first processing stage and presented performance comparison of these receivers. The MIMO precoding scheme for HSDPA is such that one can favour any one of the two streams. Two versions of SIC receivers were shown and the fundamental differences between chip-level decision-feedback (SIC) and symbol-level SIC were pointed out. Chip-level SIC receiver indeed performs far better than all per-code symbol-level receivers while in the class of latter, joint (ML) detection outperforms
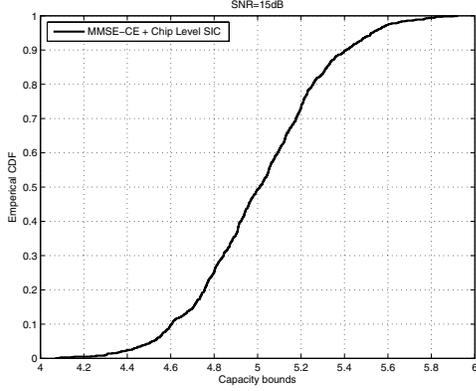
Fig. 9. Sum-capacity at the output of RX 4.
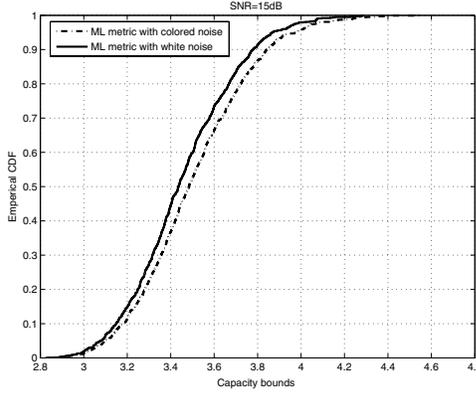
Fig. 10. MFB for RX 5

all others if the metric for ML properly takes into account the spatial correlation among the two streams.

APPENDIX

A. Estimation of Error Covariance ($R_{\widetilde{x}\widetilde{x}}$) at LMMSE Chip-Equalizer Output

We first consider linear (MMSE) FIR estimation of the 2×1 chip sequence. Referring back to fig. 1, $b[j]$ is the input chip vector defined as $b[j] = [b_1[j]\, b_2[j]]^T$, where $b_i[j]$ is the jth chip of the ith input stream. Each chip stream is the sum of K spread and scrambled CDMA sub-streams (1 user per CDMA code). Thus $b_i[j] = \sum_{k=1}^{K} b_{ik}[j]$. The 2×2 matrix $H[j]$ is the jth MIMO element of the FIR channel and W is the precoding matrix.

Let us assume an arbitrary oversampling factor m. Then, the $2m \times 1$ received signal at the jth time instant is given as

$$y[j] = \sum_{l=0}^{N-1} H[l]Wb[j-l] + v[j] = H\mathcal{W}_N b_N[j] + v[j], \quad (21)$$

where $H = [H_1\, H_2]$, with H_i being the $2m \times N$ is the FIR channel from the ith transmit antenna to the 2 RX antennas. $\mathcal{W}_N = W \otimes I_N$ and $b_N[j] = [b_{1,N}^T[j]\, b_{2,N}^T[j]]^T$ where $b_{i,N}[j] = [b_i[j - N + 1] \ldots b_i[j]]^T$ is chip sequence vector of the ith stream.

Stacking E successive samples of the received signal $y[j]$, we can express the received signal as

$$Y[j] = \mathcal{T}_E(H)\mathcal{W}_{N+E-1}b_{N+E-1}[j] + V[j], \quad (22)$$

where $\mathcal{T}_E(H) = [\mathcal{T}_E(H_1)\, \mathcal{T}_E(H_2)]$ and $\mathcal{T}_E(H_i)$ is a block Toeplitz matrix with $[H_i\, 0_{2m \times E-1}]$ as the first block row. Let us assume a $2 \times 2mE$ LMMSE equalizer $F = [f_1^T\, f_2^T]^T$. The output

of the equalizer is a linear estimate of the chip sequence given by

$$\widehat{x}[j] = FY[j] = \underbrace{B\, \underbrace{W b[j]}_{x[j]}} + \underbrace{\overline{B}\mathcal{W}_{N+E-1}\overline{b}_{N+E-1}[j] + FV[j]}_{-\widetilde{\widehat{x}}[j]}. \quad (23)$$

Defining $\alpha^{(ij)} = f_i \mathcal{T}_E(H_j)$, we have

$$B = \begin{bmatrix} \alpha_d^{(11)} & \alpha_d^{(12)} \\ \alpha_d^{(21)} & \alpha_d^{(22)} \end{bmatrix} \quad \text{and} \quad \overline{B} = \begin{bmatrix} \overline{\alpha}^{(11)} & \overline{\alpha}^{(12)} \\ \overline{\alpha}^{(21)} & \overline{\alpha}^{(22)} \end{bmatrix},$$

respectively are the 2×2 matrix that represents the *joint bias* in the equalizer output, and the residual inter-chip interference (ICI). The $\overline{\alpha}^{(ii)}$ are the same as $\alpha^{(ii)}$ with the $\alpha_d^{(ii)}$ term replaced by 0, and d is the equalization delay associated with F.

The joint-bias can also be interpreted as a spatial mixture at the chip-equalizer correlator output facilitating formulation of the spatial signal model to be treated henceforth. It must be pointed out that the spatial channel B is so definable assuming the scrambler to be a random sequence. The resulting spatial channel is per-code, while still being the same for all codes.

The MMSE of the LMMSE chip-equalizer is given by $R_{\widetilde{x}\widetilde{x}}$. We can show that

$$R_{\widetilde{\widehat{x}}\widetilde{\widehat{x}}} = R_{\widetilde{x}\widetilde{x}} - R_{\widetilde{x}\widetilde{x}}R_{xx}^{-1}R_{\widetilde{x}\widetilde{x}}, \quad (24)$$

and the error variance can be expressed as

$$R_{\widetilde{\widehat{x}}\widetilde{\widehat{x}}} = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}, \quad (25)$$

from which the MMSE can be obtained. In the above,

$$r_{11} = \sigma_b^2 \left(\|\overline{\alpha}^{(11)}\|^2 + \|\overline{\alpha}^{(12)}\|^2 \right) + f_1 Rvv f_1^H$$

$$r_{22} = \sigma_b^2 \left(\|\overline{\alpha}^{(21)}\|^2 + \|\overline{\alpha}^{(22)}\|^2 \right) + f_2 Rvv f_2^H \quad (26)$$

$$r_{12} = r_{21}^* = \sigma_b^2 \left(\overline{\alpha}^{(11)} \cdot \overline{\alpha}^{(21)H} + \overline{\alpha}^{(12)} \cdot \overline{\alpha}^{(22)H} \right) + f_1 Rvv f_2^H$$

B. Estimation of Error Covariance ($R_{\widetilde{z}\widetilde{z}}$) at Correlator Output

Considering scrambler as a random sequence and taking expectation over the scrambler $s[j]$ as well as input data symbol sequence, one can show that the covariance matrix of the estimation error $R_{\widetilde{z}\widetilde{z}}$ is similar to the chip-equalizer output error covariance matrix $R_{\widetilde{\widehat{x}}\widetilde{\widehat{x}}}$ with scaling of the interference quantities by the number of users (codes). We can show that

$$r_{11} = \sigma_a^2 \frac{K}{L} \left(\|\overline{\alpha}^{(11)}\|^2 + \|\overline{\alpha}^{(12)}\|^2 \right) + f_1 Rvv f_1^H$$

$$r_{22} = \sigma_a^2 \frac{K}{L} \left(\|\overline{\alpha}^{(21)}\|^2 + \|\overline{\alpha}^{(22)}\|^2 \right) + f_2 Rvv f_2^H$$

$$r_{12} = r_{21}^* = \sigma_a^2 \frac{K}{L} \left(\overline{\alpha}^{(11)} \cdot \overline{\alpha}^{(21)H} + \overline{\alpha}^{(12)} \cdot \overline{\alpha}^{(22)H} \right) + f_1 Rvv f_2^H$$

REFERENCES

[1] 3GPP, *TS 25.214 Physical layer procedures (FDD) (Release 7)*, May 2007, version 7.5.0.
[2] L. Mailaender, "Linear MIMO equalization for CDMA downlink signals with code reuse," *IEEE Transactions on Wireless Communications*, vol. 4, no. 5, pp. 2423–2434, September 2005.
[3] J. C. Zhang, B. Raghothaman, Y. Wang, and G. Mandyam, "Receivers and CQI measures for MIMO-CDMA systems in frequency-selective channels," *EURASIP Journal on Applied Signal Processing*, no. 11, pp. 1668–1679, November 2005.
[4] K. Ko, D. Lee, M. Lee, and H. Lee, "Novel sir to channel-quality indicator (CQI) mapping method for HSDPA system," in *Proc. IEEE Vehicular Technology Conference*, Montreal, Canada, September 2006.
[5] J. M. Cioffi, G. P. Dudevoir, M. V. Eyuboglu, and G. D. Forney, "MMSE Decision-Feedback Equalizers and Coding. Part 1: Equalization Results," *IEEE Trans. Communications*, vol. 43, no. 10, Oct. 1995.
[6] J. Choi, S.-R. Kim, Y. Wang, and C.-C. Lim, "Receivers for chip-level decision feedback equalizer for CDMA downlink channels," *IEEE Transactions on Wireless Communications*, vol. 3, no. 1, pp. 300–313, January 2004.