Draft 1

March 31, 2009

Contents

Chapter 1

Multiuser Extensions to MIMO HSDPA

1.1 Introduction

Any wireless communication system that leverages the use of multiple antennas both at the transmitter and the receiver qualifies as a multiple-input-multiple-output (MIMO) wireless system. Multiple antennas at the transmitter and receiver add an additional spatial dimension to the communication channel. By taking advantage of this fact and by exploiting the spatial properties of the MIMO channel, it is possible to provide the following features to the communication system.

1. Make the communication link resilient/robust to channel fades: Diversity techniques have for long been considered as effective means to combat channel fading. In simple terms, diversity is achieved by combining multiple copies of the same transmit signal. If the fading characteristics of each copy is statistically independent from the rest,the combined signal is more robust to channel fading. In the context of MIMO systems using the concept of *spatial diversity*, it is possible to show that the probability of losing the signal due to deep fades reduces exponentially with the number of decorrelated transmit-receive antenna pairs (spatial links) between the transmitter and receiver.

2. Increase the link capacity: Instead of using the multiple spatial channels to provide diversity,

it is possible to use these channels for multiplexing in the spatial domain. A high-data rate stream is first split into multiple sub-streams of lower data rates. Subject to certain channel conditions [6], $\min N_{tx}, N_{rx}$ streams can be transmitted over the MIMO channel. Here N_{tx}, N_{rx} refer to the number of antennas at the transmitter and receiver respectively. Since this requires no extra spectral resources, the total data rate (bits/s) transmitted over the communication link is increased.

3. Increased coverage area: Transmit beamforming is a technique in which signals transmitted from multiple antennas are multiplied by a complex weighting factor (different for each antenna) such that the transmitted signal power is concentrated in certain spatial directions (or spatial signatures). The resultant signal can now travel over a larger distance in that direction thus increasing the coverage area of the base-station. A similar type of processing can be employed at the receiver whereby the received signal power and be increased by combining the signals at each receive antenna after application of suitable weights (receive beamforming).

4. Improved spectral efficiency: By reusing the multiple access resources (for instance, spreading codes in CDMA) over the spatial dimension, MIMO systems can increase spectral efficiency (bits/s/Hz) of the communication system.

However, not all of these features can be provided simultaneously. For instance, there exists a tradeoff between the coverage range and the link quality in any MIMO system [15]. Similarly using multiple transmit antennas for spatial multiplexing reduces the available spatial degrees of freedom for spatial reuse.

   MIMO systems first attracted attention due mainly to the tremendous increase in channel capacity that is promised by MIMO [5] [19]. While there has been sustained academic interest in MIMO over the decade as evidenced by the huge amount of research publications in this topic, true MIMO systems are only recently being standardized. This has been mainly due to the increased system complexity of MIMO systems. While MIMO an potentially provide huge gains at no extra cost in terms of spectral resources, these gains can only be realized at the cost of increased system and hardware complexity. Moreover, until recently, multiple antennas at the user equipment (UE) were not considered to be desirable due to space, battery and cost constraints of mobile terminals. As a result, standardization bodies have till date concentrated more on the sub-class of MIMO sys-

tems (MISO/SIMO) whereby some kind of antenna diversity at the base station is used to exploit transmit and/or receive diversity in the interest of enhancing link quality or increasing the total system capacity. With the emergence of Internet-centric applications and an increased demand of high-data-rate applications in cellular systems this trend is changing very quickly. The present generation of smart phones and Internet enabled devices have both the form factor as well as the computational powers that can support multiple antennas at the receiver. Foreseeing these developments 2×2 MIMO has been standardized in [1] standards. In fact, the worlds first HSPA+ or evolved HSPA network with support for 2×2 MIMO was launched in early 2009 [10]. Along with enabling technologies like adaptive modulation and coding (AMC), fast hybrid-automatic repeat request (H-ARQ) and user feedback based scheduling, MIMO in HSDPA can lead to peak data rates of 42Mbps in downlink. However, in the present form, MIMO in HSDPA can support only single user (SU) scenarios in DL. While shifting from single user to multiuser paradigm mandates a whole new level of increased system complexity [7], the associated gains are significant. For instance, MU-MIMO opens up the possibility of code-multiplexing which can lead to increased system capacity.

As mentioned earlier, the gains promised by MIMO can only be realized at the cost of increased computational complexity. MIMO has largely been discussed in the context of frequency non-selective (OFDM) case, where optimal joint-stream (MAP) detection can be employed. Spatio-temporal receivers based on ordered successive interference cancellation (OSIC) in frequency-selective environments were considered in [13] while [17] proposed a class of ML receivers for multipath channels. For MIMO WCDMA transmission in frequency selective channels, where the multipath mixes signals up in space and time, proposals for receiver (RX) solutions include chip-level equalization and despreading followed by joint detection of the data streams at symbol level [14]. 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 [21]. In this chapter we explore a class of receivers based on the two-step processing strategy. We will consider receivers for HSDPA downlink that use MIMO-linear MMSE chip equalizer at the chip level followed by further processing at the symbol level after despreading. In order to fully harness the potential of this approach, we will find it necessary to do away with the customary assumption that the scrambler used at the transmitter can be considered to be random and instead treat it as deterministic. This results in enhanced performance at a cost of increased complexity at the

receiver.

The chapter is divided into three sections. First we provide a brief review of MIMO support in HSDPA. Then, in the first section we discuss various receiver designs for HSDPA when the UE is configured in MIMO mode. Throughout all analysis in this section we will make the customary assumption that the scrambler can be considered to be a random i.i.d sequence. We will also address here the issue of optimal choice of precoding matrix for the receiver. Since the receiver is required to choose the precoding matrix that maximizes its aggregate transport block size, we derive analytical expressions for the choice of the optimum precoding matrix that maximizes the sum-capacity of the receiver when it is based on MMSE designs. In the second section we will do away with the assumption that the scrambler is random and treat it as deterministic. We will introduce the time-varying model of the resulting symbol-level spatial channel and show that the deterministic point of view leads to a set of reduced dimension linear receivers and interference cancelers with increased achievable capacity. In the third section we will extend the current single-user MIMO scenarios in HSDPA to the multi-user case. These extensions require minimal changes to existing standards. When multiple UEs have to be simultaneously serviced in the downlink we will suggest practical multi-user scheduling strategies that can be employed at the BS so as to maximize the downlink capacity. Finally we will wrap up with some concluding remarks.

1.1.1   MIMO in HSDPA

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.

Both transmit diversity and spatial multiplexing has been incorporated by 3GPP as standard in the form of TxAA and its dual stream counterpart D-TxAA for MIMO HSDPA. HSDPA supports a closed loop transmit diversity technique called transmit adaptive array (TxAA). In the 2 transmit-1 receive (2×1) antenna configuration of TxAA, the UE and feeds back optimum beamforming weights that the BS uses while transmitting data to UE. D-TxAA is an extension of TxAA when UEs are configured in MIMO model Here two separately encoded, interleaved and spread transport blocks are transmitted in parallel. In this case, the UE decides the precoding matrix that the BS has to use when transmitting data to the UE. Let us now look at beamforming/ precoding aspect in more detail.

1.1.2 Precoding and CQI Feedback

In HSDPA, the UE is required to submit regular channel quality information (CQI) and precoding control information (PCI) reports to the BS. The CQI can be mapped to a particular modulation and coding scheme (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 predefined-defined Packet-Error Rate (PER). The mapping strategy has been subject to significant simulation study (see e.g., [12]) and SINR $\to$ CQI $\leftrightarrow$ PER $\leftrightarrow$ throughput relationship has been agreed to, appearing as CQI to MCS tables in the 3GPP standard document [1]. In addition to this, for each TTI over which the UE computes the CQI, PCI is computed using the CPICH(s) transmitted from both transmit. antennas to decide the beamforming vector that when applied by the BS maximizes the aggregate transport block size that the UE may support given the present channel conditions. To this end, when UE is not configured in MIMO mode, or when it requests transmission of a single transport block, the UE is required to choose one of 4 beamforming weights that control the antenna phase at BS. This then constitutes its PCI feedback. The UE indicates the number of transport blocks to be transmitted to it as part the CQI report. The BS fixes the phase of its primary (reference) antenna and alters the phase of the secondary antenna accordingly. Since the precoding weight applied to the reference antenna is a constant ($1/\sqrt{2}$), the feedback consists of the weight for antenna-2 and is one of the following weights $w \in \left\{ \frac{1+\sqrt{-1}}{2}, \frac{1-\sqrt{-1}}{2}, \frac{-1+\sqrt{-1}}{2}, \frac{-1-\sqrt{-1}}{2} \right\}$. One choice of beamforming weight

vector, let us call it $\mathbf{w}$, might be one that maximizes the received signal power (or equivalently the receive SNR). For frequency-flat channels, this corresponds to the beamforming vector that is "closest" to the maximum right singular eigenvector of the 2×2 channel matrix $\mathbf{H}$. However, for frequency selective channel with a delay spread L, there are L such MIMO channel taps. In general it is not possible to chose $\mathbf{w}$ to match all channel taps, precoding gains in such conditions is in practice very low.

When the UE is configured in MIMO mode and requests 2 transport blocks to be transmitted, a precoding matrix has to be used in place of a single beamforming weight vector. 2×2 unitary precoding based on receiver feedback is applied alongside spatial multiplexing at the base station in HSDPA [1] in D-TxAA. In order to keep feedback overhead low, both columns of the precoding matrix have exactly the same structure as the beamforming weight vector in TxAA. Moreover, the second column of this matrix is a unique function of the first. This severely restricts possible gains due to precoding. In fact, of the 4 precoding matrices, 2 matrices are related to the remaining as follows. Let $w_1 = \beta$, then by design $w_3 = \beta$ and $w_4 = -w_2$ and

$$w_2 \in \left\{ \frac{1+\sqrt{-1}}{2}, \frac{1-\sqrt{-1}}{2}, \frac{-1+\sqrt{-1}}{2}, \frac{-1-\sqrt{-1}}{2} \right\} \rightarrow \in \{\gamma, \theta, -\theta, -\gamma\} \qquad (1.1)$$

Therefore,

$$\mathbf{W} = \begin{bmatrix} w_1 & w_3 \\ w_2 & w_4 \end{bmatrix},$$

$$\mathbf{W}_1 = \begin{bmatrix} \beta & \beta \\ \gamma & -\gamma \end{bmatrix}, \mathbf{W}_2 = \begin{bmatrix} \beta & \beta \\ \theta & -\theta \end{bmatrix},$$

The other two matrices are formed by interchanging the first and second columns of $\mathbf{W}_1$ and $\mathbf{W}_2$. Since the 2 transmitted streams interfere with each other and thereby influence CQI as well as PCI choice, the precoding matrix has to be computed after joint equalization of both streams.

1.2 Receiver Designs for MIMO HSDPA: Part I

In this section, we analyze performance of a variety of receiver designs for unitary precoded D-TxAA MIMO in HSDPA. The receiver structures we propose here are based on combining chip-level and symbol level processing for enhanced performance. For each of these receivers we also derive the per-stream SINR expressions. We will use the SINR to compute the sum-capacity which can be interpreted as upper bound for achievable rates. This will form the basis for comparing the performance of the proposed receivers. The precoding matrix in D-TxAA will influence achievable sum-rate of the MIMO channel through its influence on the Signal-to-Interference-plus-Noise Ratio (SINR) of streams at the receiver (RX) output. Therefore, 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. We will show that precoding choice depends upon the MIMO receiver and the extent of its impact depends on the MIMO receiver.

For the spatial multiplexing case in MIMO HSDPA, Fig. 1.1 illustrates the equivalent baseband downlink signal model. The received signal vector (chip-rate) at the UE can be modeled as
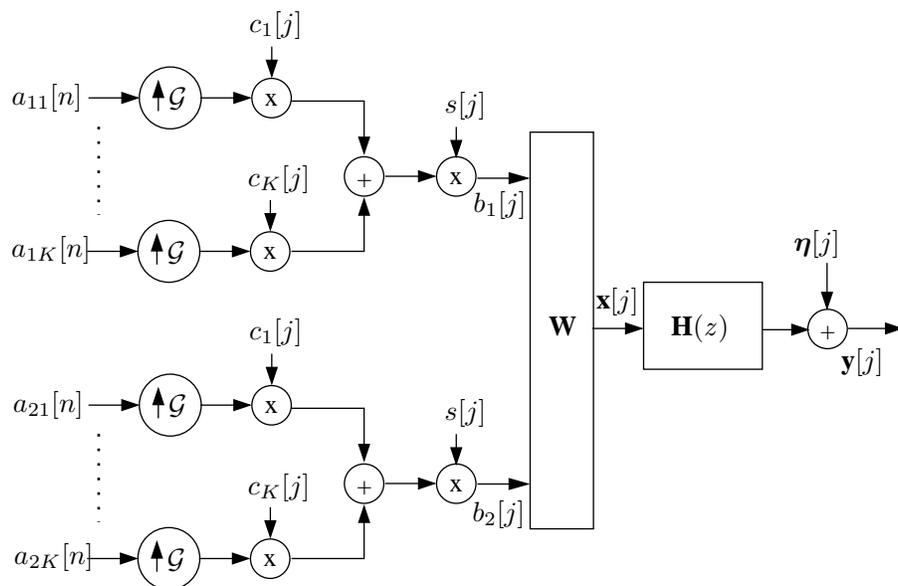
Figure 1.1: MIMO signal model with precoding.

$$\underbrace{\mathbf{y}[j]}_{2p\times 1} = \underbrace{\mathbf{H}(z)}_{2p\times 2}\underbrace{\mathbf{x}[j]}_{2\times 1} + \underbrace{\boldsymbol{\eta}[j]}_{2\times 1}. \tag{1.2}$$

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

$$\mathbf{x}[j] = \underbrace{\mathbf{W}}_{2\times 2}\mathbf{b}[j] = \mathbf{W}\cdot\sum_{k=1}^{K}\underbrace{s[j]c_k[j \mod \mathcal{G}]\mathbf{a}_k[n]}_{\mathbf{b}_k[j]} \tag{1.3}$$

k is the code index, $n = \left\lfloor\frac{j}{\mathcal{G}}\right\rfloor$ is the symbol index, $\mathcal{G}$ is the spreading factor ($\mathcal{G} = 16$ for HSDPA), $\mathbf{W} = [\mathbf{w}_1\ \mathbf{w}_2]$ is the 2×2 precoding matrix with $\mathbf{w}_1 = [\frac{1}{\sqrt{2}}\ w]^T$ and $\mathbf{w}_2 = [\frac{1}{\sqrt{2}}\ -w]^T$. The symbol vector $\mathbf{a}_k[n] = [a_{1k}[n]\ a_{2k}[n]]^T$ represents two independent symbol streams, $\mathbf{c}_k = [c_k[0]\ \dots\ c_k[\mathcal{G} - 1]]^T$, where $\mathbf{c}_k^T\cdot\mathbf{c}_k' = \delta_{kk'}$ 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 $\frac{1}{\sqrt{2}}\{\pm 1 \pm \sqrt{-1}\}$.

1.2.1   Chip-Level Equalization

It is well known that orthogonal codes used in WCDMA downlink experience a loss of orthogonality when the transmission is over multipath channels with significant time-dispersion. Such multipath channels induce inter-code-interference in the classical correlator based receivers rendering them suboptimal in such scenarios. It was shown in [9] that for single-input-single-output (SISO) links, linear minimum mean square (LMMSE) equalizers can be used to restore the orthogonality of these codes prior to despreading operation thereby providing superior performance in multipath channels with large delay spreads. In frequency selective MIMO channels, multipath mixes signals up in space and time calling for alternative reception strategies. It was shown in [14] that linear chip-level MMSE equalizers are not only able to restore orthogonality of codes but are also able to efficiently achieve spatial separation in MIMO frequency selective channels. In the following, we will discuss receivers for MIMO HSDPA based on chip level equalization.

LMMSE Chip Equalizer-Correlator

The classical MMSE chip equalizer-correlator receiver is an SINR maximizing chip equalizer followed by code correlation and soft symbol estimate generation at the output of the correlator.

Consider linear (MMSE) FIR estimation of the 2×1 chip sequence. 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). In Fig. 1.1, $\mathbf{b}[j]$ is the input chip vector defined as $\mathbf{b}[j] = [b_1[j] \ b_2[j]]^T$, 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 $\mathbf{H}[j]$ is the jth MIMO element of the FIR channel and $\mathbf{W}$ is the precoding matrix. Denoting by L, the maximum delay spread of the frequency-selective channel (in chips) and assuming an arbitrary oversampling factor p at the receiver, the $2p \times 1$ received signal at the jth time instant is given as

$$\mathbf{y}[j] = \sum_{l=0}^{L-1} \mathbf{H}[l]\mathbf{W}\mathbf{b}[j-l] + \boldsymbol{\eta}[j] = \mathbf{H}\mathcal{W}_L\mathbf{b}_L[j] + \boldsymbol{\eta}[j], \tag{1.4}$$

where $\mathbf{H} = [\mathbf{H}_1 \ \mathbf{H}_2]$, with $\mathbf{H}_i$ being the $2p \times L$ FIR channel from the ith transmit antenna to the 2 RX antennas. $\mathcal{W}_L = \mathbf{W} \otimes \mathbf{I}_L$ and $\mathbf{b}_L[j] = [\mathbf{b}_{1,L}^T[j] \ \mathbf{b}_{2,L}^T[j]]^T$ where $\mathbf{b}_{i,L}[j] = [b_i[j-L+1] \ \ldots \ b_i[j]]^T$ is chip sequence vector of the ith stream. Stacking E successive samples of the received signal $\mathbf{y}[j]$, we can express the received signal as

$$\mathbf{Y}[j] = \mathcal{T}_E(\mathbf{H})\mathcal{W}_{L+E-1}\mathbf{b}_{L+E-1}[j] + \boldsymbol{\Xi}[j], \tag{1.5}$$

where $\mathcal{T}_E(\mathbf{H}) = [\mathcal{T}_E(\mathbf{H}_1) \ \mathcal{T}_E(\mathbf{H}_2)]$ and $\mathcal{T}_E(\mathbf{H}_i)$ is a block Toeplitz matrix with $[\mathbf{H}_i \ \mathbf{0}_{2p \times E-1}]$ as the first block row. Let us assume a $2 \times 2pE$ LMMSE equalizer $\mathbf{F} = [\mathbf{f}_1^T \ \mathbf{f}_2^T]^T$. The output of the equalizer is a linear estimate of the chip sequence given by

$$\widehat{\mathbf{x}}[j] = \mathbf{F}\mathbf{Y}[j] = \mathbf{B} \underbrace{\mathbf{W}\mathbf{b}[j]}_{\mathbf{x}[j]} + \underbrace{\overline{\mathbf{B}}\mathcal{W}_{L+E-1}\overline{\mathbf{b}}_{L+E-1}[j] + \mathbf{F}\boldsymbol{\Xi}[j]}_{-\widetilde{\widetilde{\mathbf{x}}}[j]}, \tag{1.6}$$

Defining $\boldsymbol{\alpha}^{(ij)} = \mathbf{f}_i\mathcal{T}_E(\mathbf{H}_j)$, we have

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

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

We can thus write the equalizer output as the sum of an arbitrarily scaled desired term and an error term.

$$\widehat{\mathbf{x}}[j] = \mathbf{B}\mathbf{x}[j] - \widetilde{\widehat{\mathbf{x}}}[j], \tag{1.7}$$

In (1.7), an estimate of the chip sequence $\mathbf{b}[j]$ can be obtained after a further stage of processing where the precoding is undone to separate streams. The latter represented by $\mathbf{W}^H$ is a linear operation and can be carried out before or after despreading (the latter case is shown in fig. 1.2). d represents the equalization delay in chips. The joint-bias can also be interpreted as a spatial
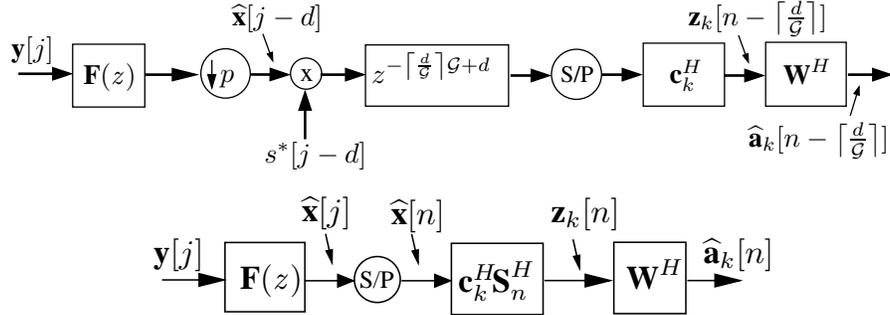
Figure 1.2: LMMSE equalizer and correlator. The second figure is a simplified representation used as chip-equalizer /correlator front-end stage for other receiver structures.

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 $\mathbf{B}$ 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 error covariance matrix corresponding to the error term is denoted by $\mathcal{R}_{\widetilde{\widehat{x}}\widetilde{\widehat{x}}}$ from which the MMSE can be obtained as below.

$$\mathcal{R}_{\widetilde{\widehat{x}}\widetilde{\widehat{x}}} = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}, \tag{1.8}$$

$$r_{11} = \sigma_b^2 \left(\|\overline{\boldsymbol{\alpha}}^{(11)}\|^2 + \|\overline{\boldsymbol{\alpha}}^{(12)}\|^2 \right) + \mathbf{f}_1 \mathcal{R}_{\eta\eta} \mathbf{f}_1^H$$

$$r_{22} = \sigma_b^2 \left(\|\overline{\boldsymbol{\alpha}}^{(21)}\|^2 + \|\overline{\boldsymbol{\alpha}}^{(22)}\|^2 \right) + \mathbf{f}_2 \mathcal{R}_{\eta\eta} \mathbf{f}_2^H \qquad (1.9)$$

$$r_{12} = r_{21}^* = \sigma_b^2 \left(\overline{\boldsymbol{\alpha}}^{(11)} \cdot \overline{\boldsymbol{\alpha}}^{(21)H} + \overline{\boldsymbol{\alpha}}^{(12)} \cdot \overline{\boldsymbol{\alpha}}^{(22)H} \right) + \mathbf{f}_1 \mathcal{R}_{\eta\eta} \mathbf{f}_2^H$$

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

$$\mathbf{z}_k[n] = \mathbf{W} \mathbf{a}_k[n] - \widetilde{\mathbf{z}}_k[n] = \mathbf{B} \mathbf{W} \mathbf{a}_k[n] - \widetilde{\widetilde{\mathbf{z}}}_k[n], \qquad (1.10)$$

Since the LMMSE joint bias is accounted for in $\mathbf{B}$, the quantity $\widetilde{\widetilde{\mathbf{z}}}_k[n]$ contains no desired symbol contribution. Note that in this receiver structure we assume $\mathbf{W}^H \mathbf{z}_k[n]$ to be the decision statistic. Considering scrambler as a random sequence and taking expectation over the scrambler as well as input data symbol sequence, one can show that the covariance matrix of the estimation error $\mathcal{R}_{\widetilde{\widetilde{z}}\widetilde{\widetilde{z}}}$ is similar to the chip-equalizer output error covariance matrix $\mathcal{R}_{\widetilde{x}\widetilde{x}}$ with scaling of the interference quantities by the number of users (codes). The elements of $\mathcal{R}_{\widetilde{\widetilde{z}}\widetilde{\widetilde{z}}}$ are given by

$$r_{11} = \sigma_a^2 \frac{K}{\mathcal{G}} \left(\|\overline{\boldsymbol{\alpha}}^{(11)}\|^2 + \|\overline{\boldsymbol{\alpha}}^{(12)}\|^2 \right) + \mathbf{f}_1 \mathcal{R}_{\eta\eta} \mathbf{f}_1^H$$

$$r_{22} = \sigma_a^2 \frac{K}{\mathcal{G}} \left(\|\overline{\boldsymbol{\alpha}}^{(21)}\|^2 + \|\overline{\boldsymbol{\alpha}}^{(22)}\|^2 \right) + \mathbf{f}_2 \mathcal{R}_{\eta\eta} \mathbf{f}_2^H$$

$$r_{12} = r_{21}^* = \sigma_a^2 \frac{K}{\mathcal{G}} \left(\overline{\boldsymbol{\alpha}}^{(11)} \cdot \overline{\boldsymbol{\alpha}}^{(21)H} + \overline{\boldsymbol{\alpha}}^{(12)} \cdot \overline{\boldsymbol{\alpha}}^{(22)H} \right) + \mathbf{f}_1 \mathcal{R}_{\eta\eta} \mathbf{f}_2^H$$

The SINR for the ith stream at the output of the output of the LMMSE chip equalizer/correlator is therefore

$$SINR_i = \frac{\sigma_a^2}{\left(\mathbf{W}^H \mathbf{B}^{-1} \mathcal{R}_{\widetilde{\widetilde{z}}\widetilde{\widetilde{z}}} \mathbf{B}^{-H} \mathbf{W} \right)_{ii}} - 1, \qquad (1.11)$$

Once MIMO joint bias is properly taken into account, 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. We address this issue later in this chapter.

The corresponding per-code capacity of the ith data stream can now be expressed as

$$\mathcal{C}_i = \log(1 + SINR_i)$$

$$\mathcal{C}_i = \log \left(\frac{\sigma_a^2}{MMSE_i} \right) \qquad (1.12)$$

Our objective is to choose the precoding matrix $\mathbf{W}$ to maximize the sum-capacity of two streams.

This boils down to the following optimization problem:

$$\mathbf{W}_{opt} = \arg \max_{\mathbf{W}} \left[\log \left(\frac{\sigma_a^4}{MMSE_1 \cdot MMSE_2} \right) \right], \tag{1.13}$$

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 $\mathbf{W}_{opt}$ maximizes $\Re(|wr_{12}|)$, where r_{12} is the top-right off-diagonal term of the error covariance matrix $\mathcal{R}_{\widetilde{\widetilde{z}}\widetilde{\widetilde{z}}}$. In other words, the $\mathbf{W}_{opt}$ attempts to maximize the SINR difference between the two streams [1].

Chip Level Successive Interference Cancellation

Consider now a chip-level Successive Interference Cancellation (SIC) receiver that detects data symbols from one stream, say stream 1 and re-spreads, re-scrambles, re-channelizes the detected data so that the contribution of the detected 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

$$\mathbf{H}_{sic} = \text{diag}(\mathbf{w}_2)\mathbf{H}, \tag{1.14}$$

where, $\mathbf{H}_{sic}$ is the equivalent channel seen by the stream detected last due to the cancellation effected at the receiver

$$\overline{\mathbf{Y}}_2[j] = \mathcal{T}(\mathbf{H}_{sic})\mathbf{b}_{2,L+E-1}[j] + \mathbf{\Xi}[j], \tag{1.15}$$

and $\mathbf{b}_{2,L+E-1}[j] = [b_2[j - L - E + 2] \cdots b_2[j]]^T$ is the chip sequence vector of the second stream. This case, assuming perfect cancellation of stream 1, is analogous to single stream Tx-AA communications and the SINR achieved for stream 2 is much improved. The SINR expressions for this SIC receiver are straightforward. The SINR expression for the first stream remains the same as that of the chip-equalizer correlator receiver and the expression for SINR for the second stream is similar to that for the MISO LMMSE chip-level equalizer/correlator case. One further consideration in this receiver is that if stream 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

---

[1]to its best abilities given the limited resolution of $\mathbf{W}$.
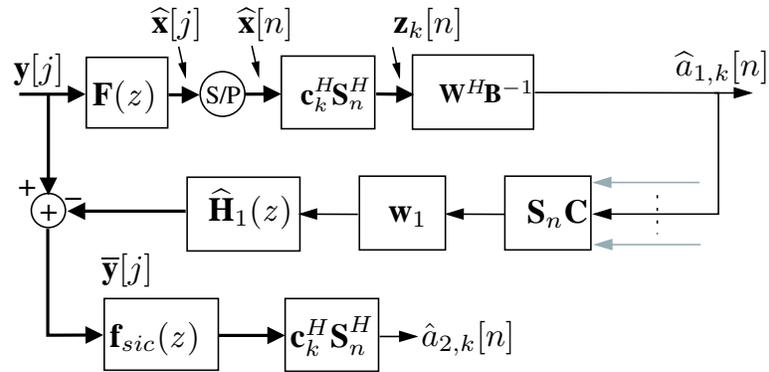
Figure 1.3: Chip LMMSE equalizer/correlator and chip-level SIC for stream 2.

is simultaneous). Such treatment increases complexity but may be well worth the effort in terms of SINR gains. We shall discuss this later in the chapter in the context of joint chip-level and symbol level processing, that such processing increases the quality of the estimates of stream 1 before feedback. The receiver structure discussed here is in fact similar to the one proposed in [3] where it is called chip-level DFE but is in fact symbol-level DFE since the decisions are on the symbols, not on chips. Even though the feedback interference cancellation is performed at chip level after re-spreading, but that is equivalent to canceling at symbol level.

1.2.2 Combined Chip-Level and Symbol-Level Equalization

Optimal linear receivers for WCDMA are symbol level (deterministic) time-varying multiuser receivers that are known to be prohibitively complex. One class of such receivers is based on symbol-level multiuser detection (MUD) where linear or non-linear transformations can be applied to the output of the channel matched filter (RAKE). Linear methods in this category are decorrelating and MMSE MUD both requiring inverses of large time-varying code cross-correlation matrices across symbols thus leading to impractical computational complexities. Non-linear MUD methods focus on estimating, reconstructing and subtracting signals of interfering codes and in general called interference canceling (IC) receivers. A less complex alternative is *dimensionality reducing* linear chip equalization followed by further linear or nonlinear interference canceling or joint detection stages to improve symbol estimates [2]. The basis of these receivers is that interference arises from loss of orthogonality due to the multipath channel and this problem is effectively

solved by attempting to restore orthogonality through a SINR maximizing LMMSE equalizer. In MIMO WCDMA, in addition to inter-code interference due to loss of orthogonality, the MIMO channel also introduces inter-stream interference. The spatial separation effected by LMMSE chip equalizer in this context is not perfect and therefore mandates additional processing that can be performed at chip or symbol level. This type of processing can be intuitively treated as a dimensionality reduction stage in MUD. It may take for example, the form of a general chip-level filter carrying out functions of channel *sparsifier* or indeed a more specific spatio-temporal $\rightarrow$ spatial channel-shortener (e.g., $2N \times 2$ to 2×2 in MIMO HSDPA) [21]. This stage precedes either per-code joint detection of data streams at symbol level [14] or can be followed-up by one of the several possible decision-feedback approaches [21] and [3]. In general, for MIMO, if the scrambler is treated as i.i.d. random, the resulting symbol-rate spatial channel can now be seen as a per-code spatial mixture and is constant. To this mixture simplified (per-code) processing can now be applied. In this section we investigate such class of MIMO HSDPA receivers based on LMMSE chip-level MIMO equalizer and a further stage of symbol level processing.

LMMSE Chip Equalizer- Symbol level LMMSE

Consider a receiver structure where 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. 1.4. As discussed in 1.2.1, the output of the correlator is $\mathbf{z}_k[n]$ given by (1.10). $\mathcal{F}_{sp}$ denotes the spatial
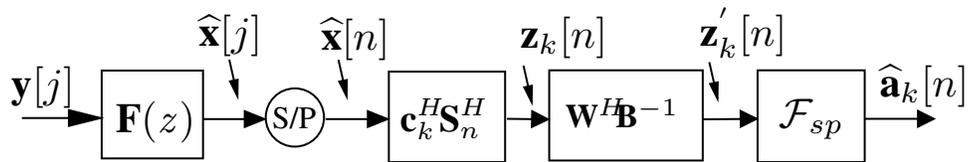
Figure 1.4: Chip LMMSE equalizer and correlator followed by symbol-level (spatial) MMSE.

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

$$\widehat{\mathbf{a}}_k[n] = \mathbf{a}_k[n] - \widetilde{\mathbf{a}}_k[n], \tag{1.16}$$

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

$$
\begin{aligned}
\mathcal{R}_{\widetilde{a}\widetilde{a}} &= \mathcal{R}_{aa} - \mathcal{R}_{az'}\mathcal{R}_{z'z'}^{-1}\mathcal{R}_{z'a} && (1.17) \\
&= \sigma_a^2\mathbf{I} - \sigma_a^4\mathbf{W}^H\left(\sigma_a^2\mathbf{I} + \mathbf{B}^{-1}\mathcal{R}_{\widetilde{z}\widetilde{z}}\mathbf{B}^{-H}\right)^{-1}\mathbf{W}, && (1.18)
\end{aligned}
$$

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

$$
\mathcal{R}_{\widetilde{a}\widetilde{a}} = \sigma_a^2\mathbf{I} - \sigma_a^4\mathbf{W}^H\left(\sigma_a^2\mathbf{I} + \left(\mathcal{R}_{\widetilde{z}\widetilde{z}}^{-1} - \mathcal{R}_{zz}^{-1}\right)^{-1}\right)^{-1}\mathbf{W}, \tag{1.19}
$$

$\mathcal{R}_{\widetilde{z}\widetilde{z}}$ is in the above expression is related to the joint-bias $\mathbf{B}$ through

$$
\mathbf{B} = \mathbf{I} - \mathcal{R}_{\widetilde{z}\widetilde{z}}\mathcal{R}_{zz}^{-1}, \tag{1.20}
$$

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

$$
\mathcal{C}_1 + \mathcal{C}_2 = \log\left(\frac{\sigma_a^4}{\det(\mathrm{diag}(\mathcal{R}_{\widetilde{a}\widetilde{a}}))}\right) \tag{1.21}
$$

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

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

One may remark that spatial MMSE processing after the equalizer/correlator stage should lead to further suppression of residual interference and lends itself to low-complexity per-code implementation. The spatial channel sees a non-negligible contribution from the kth code (desired code), therefore this receiver does improve on the MMSE chip-equalizer correlator receiver, but its performance is limited by temporal (inter-chip) interference that is still sufficiently strong at the correlator output.

LMMSE Chip Equalizer - Predictive DFE

A noise-predictive decision feedback equalizer (DFE) [4] 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. 1.5. Denote the output of the correlator as $\mathbf{u}_k[n]$, written as

$$\mathbf{u}_k[n] = \mathbf{W}^H \mathbf{B}^{-1} \mathcal{F}_{sp} \mathbf{z}'_{k,n} = \mathbf{a}_k[n] - \underbrace{\mathcal{F}_{sp} \mathbf{W}^H \mathbf{B}^{-1} \widetilde{\mathbf{z}}_k[n]}_{\widetilde{\widetilde{\mathbf{u}}}_k[n]} \tag{1.22}$$

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

$$\mathcal{R}_{\widetilde{\widetilde{u}}\widetilde{\widetilde{u}}} = \mathcal{F}_{sp} \mathbf{W}^H \mathbf{B}^{-1} \mathcal{R}_{\widetilde{\widetilde{z}}\widetilde{\widetilde{z}}} \mathbf{B}^{-H} \mathbf{W} \mathcal{F}_{sp}^H \tag{1.23}$$

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

$$\mathcal{R}_{\widetilde{r}\widetilde{r}} = \mathcal{V}_{sp} \mathcal{R}_{\widetilde{\widetilde{u}}\widetilde{\widetilde{u}}} \mathcal{V}_{sp}^H, \tag{1.24}$$

which is minimized if $\mathcal{R}_{\widetilde{r}\widetilde{r}} = \mathbf{D}$, 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 $\mathcal{R}_{\widetilde{\widetilde{u}}\widetilde{\widetilde{u}}} = \mathbf{LDL}^H$. Then, $\mathcal{V}_{sp} = \mathbf{L}^{-1}$ minimizes (1.24). Denoting elements of $\mathcal{R}_{\widetilde{\widetilde{u}}\widetilde{\widetilde{u}}}$ as r_{ij}, the elements of $\mathbf{D}$ are given as $\sigma_{\widetilde{r}_1}^2 = r_{11}$ and

$$\begin{aligned} \sigma_{\widetilde{r}_2}^2 &= r_{22} - r_{21} r_{11}^{-1} r_{12} \\ &= \det(\mathcal{R}_{\widetilde{\widetilde{u}}\widetilde{\widetilde{u}}}) \\ &= \det(\mathcal{F}_{sp}) \det(\mathbf{B}^{-1} \mathcal{R}_{\widetilde{\widetilde{z}}\widetilde{\widetilde{z}}} \mathbf{B}^{-H}) \det(\mathcal{F}_{sp}^H), \end{aligned} \tag{1.25}$$

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. 1.5, this can be interpreted as stream 1 achieving the same performance as for the chip-level LMMSE/correlator

- spatial MMSE, while stream 2 benefits from stripping (and thus achieves the spatial MFB). The
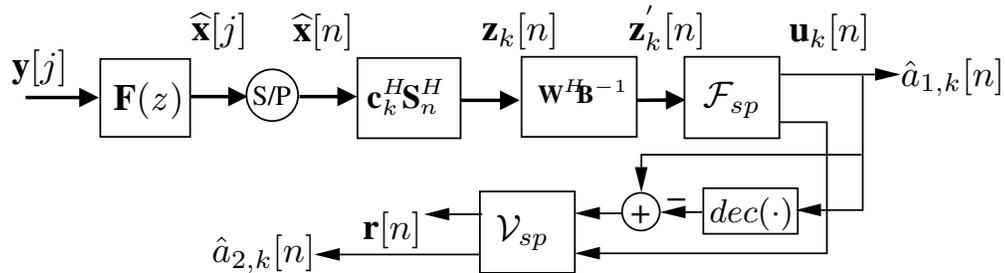
Figure 1.5: Chip LMMSE equalizer/correlator followed by spatial MMSE and symbol-level SIC for stream 2.

rates are therefore expressed as

$$\mathcal{C}_i = \log(1 + SINR_i) \tag{1.26}$$

An interesting observation is that the SINR expression for stream 2 in the symbol-level SIC case is independent of the precoding $\mathbf{W}$ applied. In this receiver, stream 1 should exhibit better performance than in the case of the chip-equalizer/correlator receiver. An alternative receiver structure proposed in [21] 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, the receiver discussed above is a better alternative to [21], 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.

It should however be noted that any low-complexity symbol level processing 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 spatio-temporal 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).

Spatial ML Receiver

Yet another possible receiver structure is shown in Fig. 1.6 where the chip-equalizer correlator front end is followed up, as before, by the spatial MMSE stage. The resulting spatial mixture

$$\mathbf{u}_k[n] = \mathcal{F}_{sp}\mathbf{z}'_k[n] = \mathbf{a}_k[n] - \widetilde{\widetilde{\mathbf{u}}}_k[n], \tag{1.27}$$

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

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

This metric can be solved for $\mathbf{a}_k[n]$. It was shown in [21] that joint detection outperforms SIC.
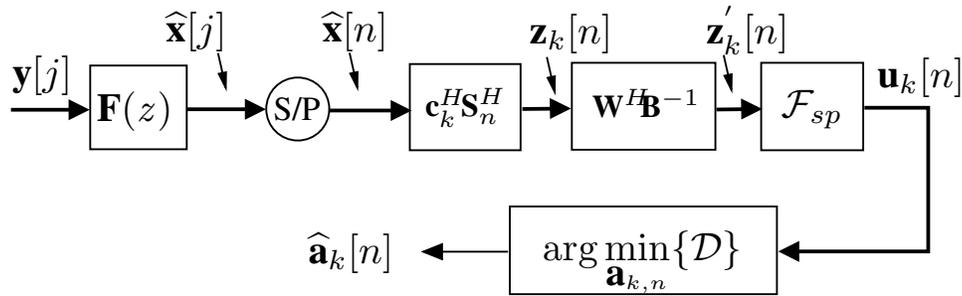
Figure 1.6: Chip LMMSE equalizer/correlator followed by spatial MMSE and joint detection.

However, the SIC structure in [21] 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 $\mathbf{B}$. 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. $\mathcal{R}_{\widetilde{\widetilde{u}}\widetilde{\widetilde{u}}}$ is the noise variance.

1.2.3 Numerical Results

We present here the simulation results and compare the performance of the different receiver structures that were discussed in this section based on their sum-capacity. For a fixed SNR and over several realizations of a frequency selective $2p \times 2$ MIMO FIR channel $\mathbf{H}(z)$, we compute the optimal precoding matrices and use the corresponding SINRs of both streams at the output of the receivers to calculate an upper bound on the sum capacity. The channel coefficients are complex valued zero-mean Gaussian 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
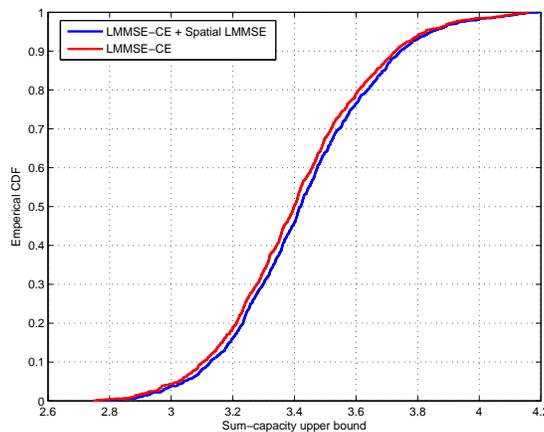
receivers. The structure of the precoding matrices used in HSDPA is such that two out of the four possible precoding matrices give the same SINR (and thus sum-rate) for the LMMSE/correlator design. The difference between them being that one favors stream 1 by bestowing a higher SINR for stream 1, and 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. 1.2.3 shows distribution of sum-capacity at the output of the MMSE chip-equalizer correlator receiver and that of the spatial MMSE receiver. With an additional processing stage of a very small complexity we are able to see some gain in the achievable rates of the receiver.

In Fig. 1.2.3 we compare the performance of LMMSE chip equalizer- correlator receiver with the receiver that performs spatial MMSE as well as predictive-DFE and the per-code ML receiver.
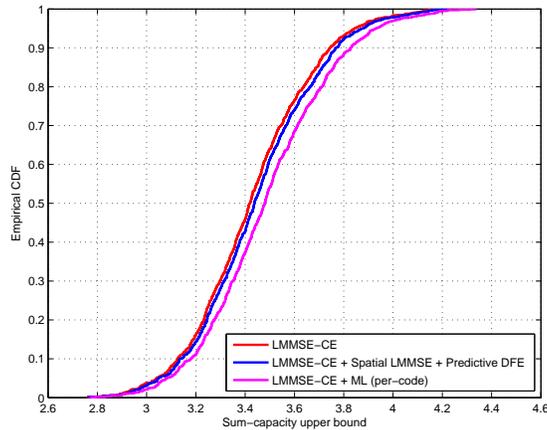
As before, optimal precoding matrices are used at the base-station. The receiver that performs spatial LMMSE and DFE benefits slightly from the additional spatial processing for both streams and a non-linear equalization stage for stream-2. That the gain is not considerable is due to the fact that stream-1 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. By performing spatial ML detection one is able to get much better performance. The chip-level SIC, in Fig. 1.2.3 as can be expected, outperforms all other receivers at the cost of a significant complexity at the receiver.
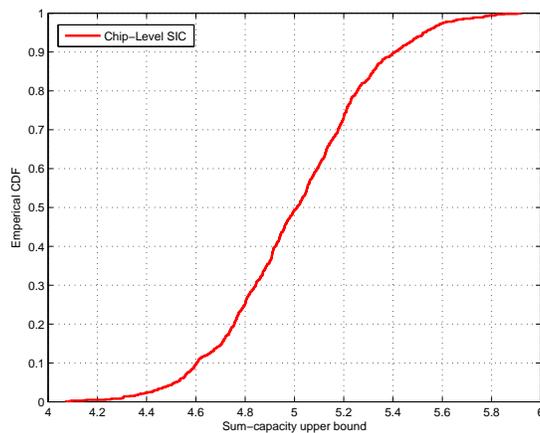
1.3   Receiver Designs for MIMO HSDPA: Part II

Till this point, we discussed various receiver designs that assumed the scrambler to be random i.i.d. Random treatment of scrambler leads to a time-invariant spatial signal model which in turn leads to intuitively pleasing RX solutions. While this assumption simplifies receiver designs for the second stage of the two-step processing employed in the receivers it limits the performance of these receivers. Since the first step in the two-stage approach can be interpreted as a dimensionality reduction step, the limitation on the gain obtained by this design over classical chip-equalization is can be linked to the efficacy of the dimensionality reduction achieved at the output of the chip-equalizer and also by the type of processing at symbol level. In the general MIMO case, the resulting symbol-rate spatial channel can now be seen as only a per-code spatial mixture. When the scrambler is treated as random, this mixture becomes time-invariant and therefore simplified (per-code) processing can be applied. For a processing gain $\mathcal{G}$, assuming N_t to be the number of TX streams, N_r the number of RX antennas, and p to be the oversampling factor *w.r.t.* the chip rate, this can be seen as a dimensionality-reduction from $p \cdot \mathcal{G} \cdot N_r$ to N_t. Given this drastic reduction, it is not surprising to see performance falling well short of optimal time-varying symbol-level processing (linear and non-linear MUD solutions). In the previous section, we chose to tradeoff performance in the interest of reduced complexity symbol level processing in order to point out that despite their shortcomings, their complexity/performance equation encourages use of these solutions. In this section, in an attempt to further increase the performance of our receiver designs, we put forth the idea of deterministic treatment of the scrambler. We then focus on the resulting spatial channel model while treating the scrambler as *deterministic* [8]. Such a treatment demands time-varying processing after the equalizer-correlator stage but offsets some of the performance losses of the dimensionality reduction stage and random scrambler assumption. We show that deterministic treatment of scrambler allows us to retrieve the time-varying contribution of joint-bias which would have otherwise been relegated to noise.

In the interest of simplicity, we will not consider here, the precoding aspect of DL transmission. However, we stress that introduction of precoding does not in any way alter the results obtained in this section. The DL signal model remains exactly the same as before, apart from the absence of linear precoding before transmission and we illustrate it here for convenience.
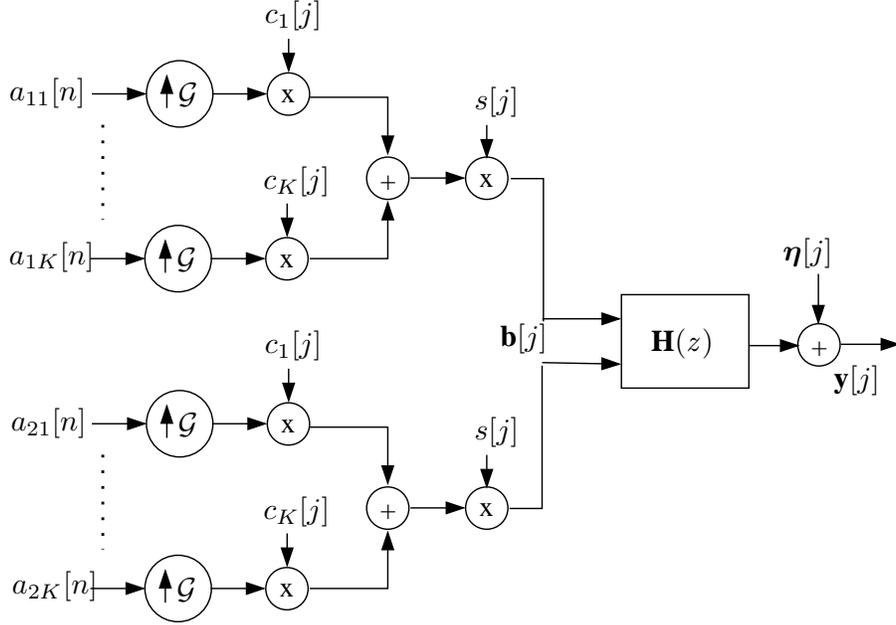
Figure 1.7: MIMO signal model without precoding.

The received signal vector (chip-rate) at the UE is now modeled as

$$\underbrace{\mathbf{y}[j]}_{2p\times 1} = \underbrace{\mathbf{H}(z)}_{2p\times 2}\,\underbrace{\mathbf{b}[j]}_{2\times 1} + \underbrace{\boldsymbol{\eta}[j]}_{2\times 1}, \tag{1.28}$$

where

$$\mathbf{b}[j] = \sum_{k=1}^{K} s[j]c_k[j \mod \mathcal{G}]\mathbf{a}_k[n], \tag{1.29}$$

1.3.1   Chip-Level Equalization

When the scrambler is treated as deterministic, the desired signal contribution at the correlator output is not only concentrated in one tap of the channel-equalizer cascade (as is the case when the scrambler is treated as random) but also contains a scrambler dependent time-varying component (thus not only a mean but also a variance). This leads to a relationship linking the LMMSE chip-equalizer output joint-bias and the time-varying correlator output joint-bias. Here we will first consider the LMMSE chip-equalizer correlator receiver and derive an analytical expression for the bias term and evaluate the SINR including the explicit contribution of this quantity.

MMSE Chip Equalizer-Correlator Revisited

As usual, we start by deriving the expression for the output energy of this receiver. Without loss of generality, we consider linear MMSE estimation of the 2×1 MIMO symbol sequence, $\mathbf{a}_k[n]$, of the kth code among K codes (each stream has K codes). In any case, this corresponds to the 2×2 MIMO case in HSDPA. Refer to fig. 1.8, for a vectorized TX signal model where $\mathbf{b}[n]$ is the $2\mathcal{G} \times 1$ chip vector defined as $\mathbf{b}[n] = [\mathbf{b}_0^T[n] \cdots \mathbf{b}_{\mathcal{G}-1}^T[n]]^T$, where $\mathbf{b}_m[n]$ is the mth multi-code (K codes) MIMO (2×1) chip corresponding to the nth MIMO symbol vector, $\mathbf{a}[n]$ of size $2K \times 1$. Then, at the symbol level, the input output relationship can be compactly represented using the q
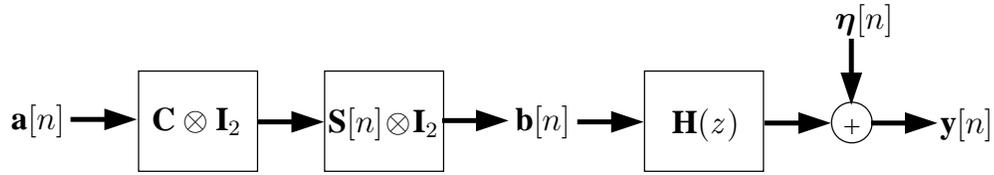
Figure 1.8: MIMO TX signal model.

operator as

$$\mathbf{y}[n] = \mathbf{H}(q)\mathbf{a}[n] + \boldsymbol{\eta}[n], \tag{1.30}$$

where

$$\mathbf{H}(q) = \sum_{m=0}^{\lceil L/\mathcal{G}\rceil-1} \mathbf{H}(m)q^{-m}, \tag{1.31}$$

This expression compactly represents the convolution equation since $q^{-m}\mathbf{a}[n] = \mathbf{a}[n - m]$. In order to disambiguate the above expression from the z-transform operation we use q^{-1} to denote the unit delay operator. Assuming an oversampling factor of p, the symbol level channel $\mathbf{H}(z) = \sum_m z^{-m}\mathbf{H}[m]$ consists of $p\mathcal{G} \times \mathcal{G}$ matrix taps. If the delay spread is L chips, there are $\lceil L/\mathcal{G}\rceil$ *pseudo-circulant* matrices that fully represent the channel. These matrices are defined as

$$\mathbf{H}[m] = \begin{bmatrix} \mathbf{h}[m\mathcal{G}] & \mathbf{h}[m\mathcal{G}+1] & \dots & \mathbf{h}[(m+1)\mathcal{G}-1] \\ \mathbf{h}[m\mathcal{G}-1] & & & \vdots \\ \vdots & & \ddots & \\ \mathbf{h}[(m-1)\mathcal{G}+1] & \dots & \dots & \mathbf{h}[m\mathcal{G}] \end{bmatrix}$$

with $\mathbf{h}[.]$ being the $2p \times 2$ chip-level MIMO channel coefficients. The LMMSE equalizer $\mathbf{F}(z)$ in Fig. 1.9 can be represented in a similar fashion and visualized to be composed of $\mathbf{f}[.]$ which would be the $2 \times 2p$ equalizer coefficients defined at the chip-level. The channel equalizer cascade is then given by

$$
\begin{aligned}
\mathbf{G}(z) &= \mathbf{F}(z)\mathbf{H}(z), &(1.32)\\
&= \sum_{\kappa=0}^{N-1} \mathbf{F}[\kappa]z^{-\kappa} \sum_{m=0}^{M-1} \mathbf{H}[m]z^{-m}, &(1.33)\\
&= \sum_{\nu=0}^{N+M-2} \mathbf{G}[\nu]z^{-\nu}, &(1.34)
\end{aligned}
$$

where, assuming the chip-equalizer length to be E chips, we have $M = \lceil L/\mathcal{G} \rceil$ and $N = \lceil E/\mathcal{G} \rceil$. The channel-equalizer cascade at symbol level can therefore be defined similarly to be composed of 2×2 chip-level matrix-coefficients $\mathbf{g}[k] = \sum_{l=0}^{L-1} \mathbf{f}[k-l]\mathbf{h}[l]$

$$
\mathbf{y}[n] \longrightarrow \boxed{\mathbf{F}(z)} \longrightarrow \boxed{\mathbf{S}^*[n] \otimes \mathbf{I}_2} \longrightarrow \boxed{\mathbf{c}_k^H \otimes \mathbf{I}_2} \longrightarrow \widehat{\mathbf{a}}[n]
$$
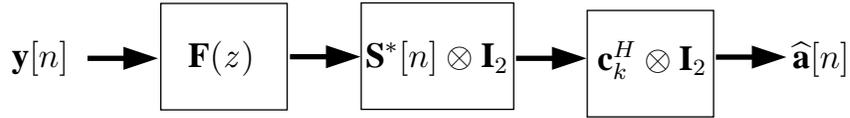
Figure 1.9: MIMO RX model.

Let the equalizer delay be d, and define the corresponding channel-equalizer cascade matrix at d

$$
\mathbf{G}[0] = \mathbf{F}(q)\mathbf{H}(q)|_{[0]} = \begin{bmatrix} \mathbf{g}[0] & \mathbf{g}[1] & \cdots & \mathbf{g}[\mathcal{G}-1] \\ \mathbf{g}[-1] & & & \vdots \\ \vdots & & \ddots & \\ \mathbf{g}[-\mathcal{G}+1] & \cdots & \cdots & \mathbf{g}[0] \end{bmatrix}
$$

as the $2\mathcal{G} \times 2\mathcal{G}$ zeroth MIMO matrix-tap of the channel equalizer cascade. $\overline{\mathbf{G}}(z) = \sum_{m \neq 0} z^{-m}\mathbf{G}[m]$ represents the MIMO inter-symbol interference (ISI). We can henceforth write

$$
\widehat{\mathbf{a}}_k[n] = \left(\mathbf{c}_k^H \otimes I_2\right)\left(\mathbf{S}^*[n] \otimes \mathbf{I}_2\right)\left\{\mathbf{G}(q)\left(\mathbf{S}[n] \otimes \mathbf{I}_2\right)\left(\mathbf{C} \otimes \mathbf{I}_2\right)\mathbf{a}[n] + \mathbf{F}(q)\boldsymbol{\eta}[n]\right\}, \qquad (1.35)
$$

Defining

$$\mathbf{B}_{n,k}(z) = (\mathbf{c}_k^H \otimes \mathbf{I}_2)(\mathbf{S}^*[n] \otimes \mathbf{I}_2)\mathbf{G}(z)(\mathbf{S}[n] \otimes \mathbf{I}_2)(\mathbf{C} \otimes \mathbf{I}_2)$$

as the symbol-rate channel at time instant n (also a $\overline{\mathbf{B}}_{n,k}(z)$ corresponding to $\overline{\mathbf{G}}(z)$) , we can write the correlator output as

$$\mathbf{z}_k[n] = \underbrace{\mathbf{B}_{n,k}[0]\mathbf{a}_k[n]}_{kth\ code} + \underbrace{\mathbf{B}'_{n,k}[0]\overline{\mathbf{a}}[n]}_{other\ codes} + \underbrace{\sum_m \mathbf{B}_{n,k}[m]\mathbf{a}[n+m]}_{all\ codes\ other\ symbols} + \underbrace{\mathbf{F}(z)\boldsymbol{\eta}[n]}_{noise}, \qquad (1.36)$$

In this expression, $\mathbf{B}_{n,k}[0]$ is the desired user channel at symbol-time n (time-varying channel), which one can split into a time invariant part $\mathbf{E}_n\mathbf{B}_{n,k}[0] = \mathbf{B}[0] = \mathbf{B} \cdot I_{\mathcal{G}}$ (assuming the scrambler to be white), and a time-varying part (if scrambler treated as deterministic). When the scrambler is treated as white, we refer to the 2×2 channel as *spatial channel* or even as *joint MIMO bias* and denote it as $\mathbf{g}_0 = \mathbf{B}$. As discussed in [8], treating the scrambler as white has the effect of capturing the mean signal energy (corresponding to the $\mathbf{g}[0]$ contribution) at the output of the per code MIMO channel while consigning the variance (off-diagonal part in $\mathbf{G}[0]$) definitively and irrecoverably to the interference term.

It may be noticed that each element of $\mathbf{G}[m]$ is a 2×2 MIMO matrix coefficient. The former can therefore be split into four $\mathcal{G} \times \mathcal{G}$ SISO submatices $\mathbf{G}_{r\kappa}[m]$, for $r, \kappa \in \{1, 2\}$. A corresponding $\mathcal{G} \times \mathcal{G}$ matrix coefficient $\overline{\mathbf{G}}_{r\kappa}[m] = \mathbf{G}_{r\kappa}[m] - \mathbf{g}_{r\kappa}[m] \cdot I_{\mathcal{G}}$ is also defined and so is $\mathbf{g}_{r\kappa}[m]$, the $r\kappa^{th}$ element of the spatial channel $\mathbf{g}[m]$.

Taking expectation over the scrambler, we can express the output energy of the receiver as

$$\mathcal{R}_{zz} = \mathcal{R}_{des} + \underbrace{\mathcal{R}_{MUI} + \sum_m \mathcal{R}_{m,ISI} + \mathbf{F}\mathcal{R}_{\eta\eta}\mathbf{F}^H}_{\mathcal{R}_{\widetilde{z}\widetilde{z}}}, \qquad (1.37)$$

where,

$$
\boldsymbol{\mathcal{R}}_{des} = \begin{bmatrix} |g_{11}[0]|^2 + |g_{12}[0]|^2 & \displaystyle\sum_{\kappa=1}^{2} g_{1\kappa}[0]g_{2\kappa}^*[0] \\ \displaystyle\sum_{\kappa=1}^{2} g_{2\kappa}[0]g_{1\kappa}^*[0] & |g_{21}[0]|^2 + |g_{22}[0]|^2 \end{bmatrix} + \\
\frac{1}{\mathcal{G}^2} \cdot \begin{bmatrix} \displaystyle\sum_{\kappa=1}^{2} tr\{\overline{\mathbf{G}}_{1\kappa}[0]\overline{\mathbf{G}}_{1\kappa}^{H}[0]\} & \displaystyle\sum_{\kappa=1}^{2} tr\{\overline{\mathbf{G}}_{1\kappa}[0]\overline{\mathbf{G}}_{2\kappa}^{H}[0]\} \\ \displaystyle\sum_{\kappa=1}^{2} tr\{\overline{\mathbf{G}}_{2\kappa}[0]\overline{\mathbf{G}}_{1\kappa}^{H}[0]\} & \displaystyle\sum_{\kappa=1}^{2} tr\{\overline{\mathbf{G}}_{2\kappa}[0]\overline{\mathbf{G}}_{2\kappa}^{H}[0]\} \end{bmatrix} ,
$$

$$
\boldsymbol{R}_{MUI} = \frac{K-1}{\mathcal{G}^2} \cdot \begin{bmatrix} \displaystyle\sum_{\kappa=1}^{2} tr\{\overline{\mathbf{G}}_{1\kappa}[0]\overline{\mathbf{G}}_{1\kappa}^{H}[0]\} & \displaystyle\sum_{\kappa=1}^{2} tr\{\overline{\mathbf{G}}_{1\kappa}[0]\overline{\mathbf{G}}_{2\kappa}^{H}[0]\} \\ \displaystyle\sum_{\kappa=1}^{2} tr\{\overline{\mathbf{G}}_{2\kappa}[0]\overline{\mathbf{G}}_{1\kappa}^{H}[0]\} & \displaystyle\sum_{\kappa=1}^{2} tr\{\overline{\mathbf{G}}_{2\kappa}[0]\overline{\mathbf{G}}_{2\kappa}^{H}[0]\} \end{bmatrix} ,
$$

and, the ISI contribution from the mth symbol can be expressed as

$$
\boldsymbol{\mathcal{R}}_{m,ISI} = \frac{K}{\mathcal{G}^2} \cdot \begin{bmatrix} \displaystyle\sum_{\kappa=1}^{2} tr\{\overline{\mathbf{G}}_{1\kappa}[m]\overline{\mathbf{G}}_{1\kappa}^{H}[m]\} & \displaystyle\sum_{\kappa=1}^{2} tr\{\overline{\mathbf{G}}_{1\kappa}[m]\overline{\mathbf{G}}_{2\kappa}^{H}[m]\} \\ \displaystyle\sum_{\kappa=1}^{2} tr\{\overline{\mathbf{G}}_{2\kappa}[m]\overline{\mathbf{G}}_{1\kappa}^{H}[m]\} & \displaystyle\sum_{\kappa=1}^{2} tr\{\overline{\mathbf{G}}_{2\kappa}[m]\overline{\mathbf{G}}_{2\kappa}^{H}[m]\} \end{bmatrix} ,
$$

In these relations, the $\boldsymbol{\mathcal{R}}_{des}$ is composed of two contributions shown above as the sum of two 2×2 matrices. When the scrambler is treated as random the term scaled by $1/\mathcal{G}^2$ is the quantity that ceases being a part of the signal energy contribution and is associated instead with the interference for reasons explained earlier.

At the output of the despreader for the kth code, one can therefore express the signal as

$$
\mathbf{z}_k[n] = \mathbf{B}_{n,k}[0]\mathbf{a}_k[n] - \widetilde{\widetilde{\mathbf{z}}}_k[n], \tag{1.38}
$$

where the time varying MIMO joint-bias $\mathbf{B}_{n,k}[0]$ is no longer constant and varies for each symbol. The per-user SINR of stream i is given by (1.39).

$$SINR_{k,i} = \frac{\sigma_{a_k}^2 \left(|g_{ii}[0]|^2 + \frac{1}{\mathcal{G}^2}\mathrm{tr}\left\{ \overline{\mathbf{G}}_{ii}[0]\overline{\mathbf{G}}_{ii}^H[0] \right\} \right)}{\sigma_{a_k}^2 \left(\frac{(K-1)}{\mathcal{G}^2} \sum_{\kappa=1}^{2} \mathrm{tr}\left\{ \overline{\mathbf{G}}_{i\kappa}[0]\overline{\mathbf{G}}_{i\kappa}^H[0] \right\} + \frac{K}{\mathcal{G}^2} \sum_{m} \sum_{\kappa=1}^{2} \mathrm{tr}\left\{ \overline{\mathbf{G}}_{i\kappa}[m]\overline{\mathbf{G}}_{i\kappa}^H[m] \right\} \right) + \sigma_\eta^2\|\mathbf{f}_i\|^2},$$

$$(1.39)$$

Chip Level SIC Revisited

We now consider the effect of treatment of scrambler as deterministic on the chip-level SIC. The SIC structure in [21] qualifies as a symbol-level SIC, feeding back symbol decisions to the output of equalizer-correlator. This assumes a time-invariant symbol-level channel $\mathbf{B}$ resulting from treating the scrambler as random. The suboptimalities introduced in all earlier stages, i.e., dimensionality reduction through chip-equalization and reducing the symbol-level channel to its mean value (random scrambler) take their toll; this SIC does not provide significant gains over chip equalizer/correlator solution.

On the other end of the SIC spectrum is the chip-level SIC, discussed here, that is an entirely different solution that considers all stages of the channel-equalizer correlator stage as deterministic and re-creates all components (ISI and MUI) of the stream detected first before subtracting it from the input signal. Subsequently, the second stream can be dealt with in much improved conditions where interference from the first stream is (ideally) entirely suppressed.

1.3.2 Combined Chip-Level and Symbol-Level Equalization

As discussed in the previous section, treating the scrambler as deterministic mandates computation of joint-bias for each symbol since it can no longer be considered as a constant. Here we discuss briefly the effect of deterministic treatment of scrambler on further symbol level processing stages.

LMMSE Chip Equalizer Symbol Level MMSE Revisited

We consider again the effect of deterministic treatment of scrambler for symbol-level spatial MMSE receiver. In order to claim the quantity $\frac{1}{\mathcal{G}^2}\text{tr}\left\{\overline{\mathbf{G}}_{rr}[0]\overline{\mathbf{G}}_{rr}^H[0]\right\}$ in (1.39) as part of signal energy, it suffices to put in place time-varying processing at the correlator output, where the nth symbol vector on the kth code, $z_k[n]$ is given by (1.36). As a result of time-varying symbol level *joint-bias*, the 2×2 MMSE equalizer will now have to be computed for each symbol. This will indeed provide higher gains than the spatial MMSE receiver above which treats the time varying signal contribution as noise.

Spatial ML Receiver Revisited

In treating the scrambler as random the spatial channel (**B**), the ML metrics will deal with a time-invariant channel. A continuous processing matched filter bound can therefore be defined per stream. The ith stream MFB is therefore proportional to the energy in the corresponding SIMO channel. On the contrary, if a deterministic scrambler is assumed, time-variation in the channel must be accounted for in ML metrics. Strictly speaking, the MFB is only defined per symbol as the SINR of the nth symbol considering all other symbols to be known (correctly detected). We can nevertheless argue that deterministic treatment of the scrambler leads to reduced interference variance $\mathcal{R}_{\widetilde{\widetilde{z}}\widetilde{z}}$ and increased recoverable signal power that will lead to performance improvement for the ML solution.

1.3.3   Numerical Results

As before we compare the performance of the different receiver structures based on their sum-capacity. We simulate here a single-user situation where 15 codes are assigned to the same user. Furthermore, we assume code-reuse across antennas. The length of FIR MIMO equalizers of is comparable to channel delay spread in chips. In Fig. 1.3.3 we plot the capacity bounds for two cases. In the first instance, we treat the scrambler as random. The symbol energy for code k is therefore given by the symbol variance for the code scaled by an arbitrary time-invariant
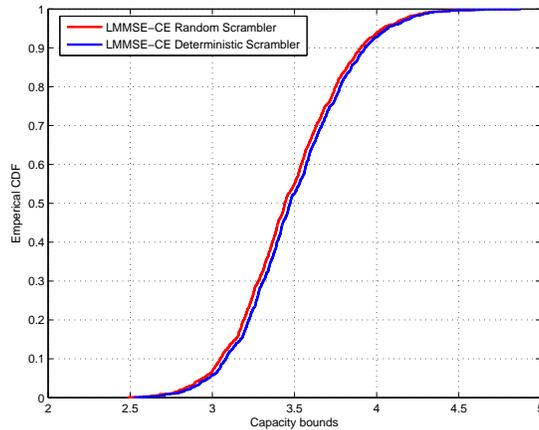
scale factor. In the second case, we treat the scrambler as known (deterministic treatment). In this case, firstly, the signal power now is time-varying at symbol rate. This time varying signal power can be seen as the sum of a "mean" power contribution equal to the signal power when the scrambler is assumed to be random, and time-varying contribution due to deterministic treatment of the scrambler. Note that the SINR distribution for the deterministic treatment of the scrambler in Fig. 1.3.3 represents the *average* gains and not the true gain. The actual gain will be higher than that seen in Fig. 1.3.3.
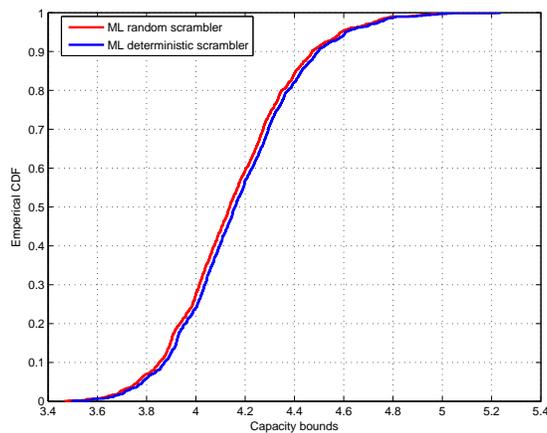
1.4   Multi-user Extensions to HSDPA MIMO

In this section, we shift our focus to extending MIMO in HSDPA to support multiple users in the downlink (MU-MIMO). In its present form, the standard only supports 2×2 SU-MIMO in the downlink (DL) in the form of D-TxAA. It is possible for the BS to employ SDMA and service multiple UEs in DL instead. In this case, the limitation of 2 transmit antennas implies that a maximum of 2 spatially separated users can be simultaneously served by the BS with the same code. In general, MU extensions for closed loop transmit diversity schemes (both TxAA and D-TxAA) introduce multi-user interference in downlink since there exists the possibility of different users feeding back different beamforming vectors in TxAA or different precoding matrices in D-TxAA.

There is a large amount of literature available for multiuser MIMO communication in the general case. It has been studied previously in [18] and more recently in [7] where multiuser transmission techniques are classified into linear and non-linear transmission algorithms. Non-linear algorithms involving multiuser signal designs that avoid interference generation to other users based on dirty paper coding techniques remain currently impractical due to the requirement of perfect channel state information at the transmitter (CSIT). They also suffer from all the drawbacks associated with outdated CSIT due to scheduling delays at the base station and/or rapidly changing downlink channels. Linear processing of transmitted signals like multiuser beamforming remain by far the most practical solution for multiuser transmission. Theoretical research in multiuser communications tends to consider frequency-flat channels. In reality most mobile communication channels are frequency selective. There exists some literature on multiuser extension of HSDPA. In [11] the authors propose code reuse in D-TxAA based on a multi-user beamforming (MUB) scheme which schedules users with orthogonal weight vectors to separate them in space. They however limit their analysis to flat-channels. In [20], the authors consider MU-TxAA for frequency selective channels and propose the so-called "interference-aware" receiver which in addition to requiring multiple antennas at the receiver also assumes knowledge of beamforming weight vectors of all the users at the receiver. On the other hand, in this section, we look at the problem of maximizing system capacity in the frequency selective MISO/MIMO downlink channels assuming the receivers select weights that maximize receive SINR (and thus increase their individual data rates). In the HSDPA

context, the BS is equipped with 2 transmit antennas i.e. $N_{tx} = 2$. In our treatment, we do not assume any explicit knowledge of beamforming weight vectors of other users, for single stream transmission we consider single antenna UE and study different beamforming strategies that can be adopted by the BS and for dual stream transmission we consider UE with two antennas and compare the performance of SDMA against spatial multiplexing to a single user by extending D-TxAA to a MU configuration where at most N_{tx} users can be synchronously served by the BS. Each transmit stream is assigned to a different user. This rules out simultaneously serving any two users that feed back the same beamforming weight vector. Users that request linearly independent weight vectors can however be served simultaneously.

1.4.1 Multiuser TxAA

We consider a 2-transmit, 1-receive antenna configuration for TxAA. For the rest of the section, whenever we refer to a MU-TxAA system, we consider U seperare UEs each having a single receive antenna. The number of codes assigned to each user is denoted by $K_1, K_2, \ldots, K_U$ and $K = \sum_{u=1}^{U} K_u$. Then, for TxAA, from Fig. 1.10 the transmit and beamformed chip sequence is given by

$$\mathbf{x}[j] = \sum_{u=1}^{U} \mathbf{w}_u \cdot s_n[j \mod \mathcal{G}] \sum_{k \in K_u} c_k[j \mod \mathcal{G}] a_{u,k}[\lfloor \frac{j}{\mathcal{G}} \rfloor n], \tag{1.40}$$

where j is the chip index, n is the symbol index, u is the user index, k is the code index, $\mathcal{G}$ is the spreading gain, s_n denotes the scrambler for the nth symbol, c_k denotes the kth spreading code, $\mathbf{w}_u = [w_{u,1} w_{u,2}]^T$ is the weight vector corresponding to u^{th} user and finally $a_{u,k}$ is the uth user's symbol on code index k given that $k \in K_u$. The transmitted signal propagates through a multipath channel which we denote here by $\mathcal{H}_u^0, \mathcal{H}_u^1, \ldots, \mathcal{H}_u^{L-1}$. For an oversampling factor of p at the receiver, each $\mathcal{H}_u^l$ matrix is a $p \times 2$ matrix corresponding to the lth tap of the uth user's multipath channel. For simplicity we assume that all UEs see a channel with a maximum delay spread of L chips and employ an equalizer of length E (in chips). The chip-rate received signal at each UE is given by

$$\mathbf{y}_u = \mathbf{H}_u \mathbf{x} + \boldsymbol{\eta} \tag{1.41}$$
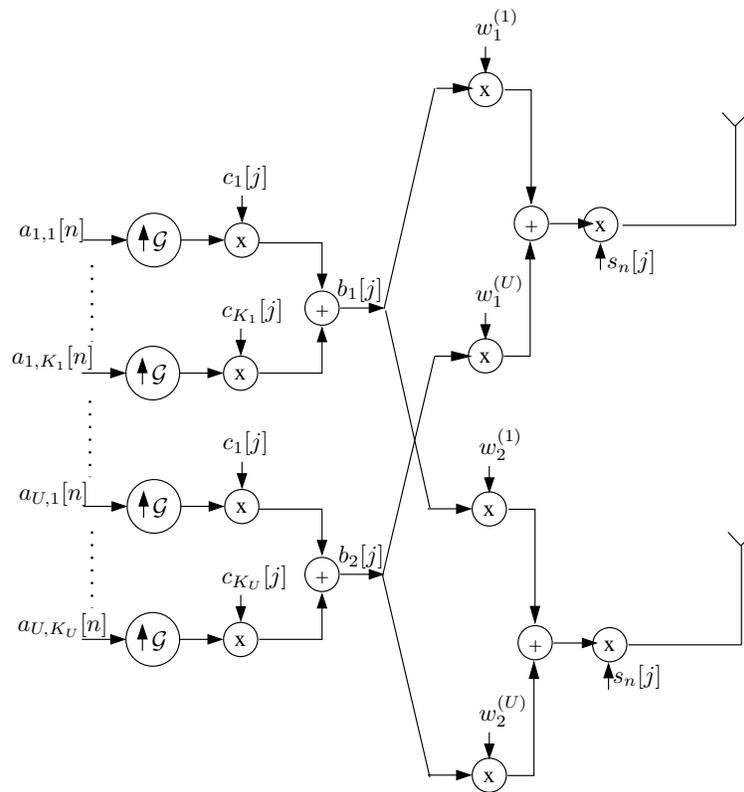
Figure 1.10: Multiuser TxAA transmit signal model.

where $\mathbf{H}_u$ is the channel convolution matrix for the uth user given by

$$
\mathbf{H}_u =
\begin{bmatrix}
\mathcal{H}_u^0 & \mathcal{H}_u^1 & \cdots & \mathcal{H}_u^{L-1} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathcal{H}_u^0 & \cdots & \cdots & \mathcal{H}_u^{L-1} & \vdots \\
\mathbf{0} & \mathbf{0} & \ddots & \ddots & \ddots & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \ddots & \mathcal{H}_u^0 & \ddots & \mathcal{H}_u^{L-1}
\end{bmatrix},
\tag{1.42}
$$

$\mathbf{x}$ is the transmit chip-vector formed by stacking $L + E - 1$ vectors and can be expressed as

$$
\mathbf{x} = [\mathbf{x}^T[j], \mathbf{x}^T[j-1], \ldots, \mathbf{x}^T[j-L-E+2]],
\tag{1.43}
$$

and $\boldsymbol{\eta}$ is zero mean, circularly symmetric, Gaussian distributed, additive white noise of variance σ_η^2. In addition, we also define the $p \times 1$ vector $\mathbf{r}_{u,v}^l = \mathcal{H}_u^l \mathbf{w}_v$, $v \in 1, 2, \ldots, U$ and use this to define the lth beamformed channel tap of user u, due to beamforming weight of another synchronous DL user v. We denote this by $\mathbf{R}_{u,v}$ and express this as

$$
\mathbf{R}_{u,v} =
\begin{bmatrix}
\mathbf{r}_{u,v}^0 & \mathbf{r}_{u,v}^1 & \cdots & \mathbf{r}_{u,v}^{L-1} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{r}_{u,v}^0 & \cdots & \cdots & \mathbf{r}_{u,v}^{L-1} & \vdots \\
\mathbf{0} & \mathbf{0} & \ddots & \ddots & \ddots & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \ddots & \mathbf{r}_{u,v}^0 & \ddots & \mathbf{r}_{u,v}^{L-1}
\end{bmatrix},
\tag{1.44}
$$

Beamforming Strategies at Transmitter

Consider the case where the base station serves U simultaneous users in the downlink. We assume standard MMSE chip equalizer-correlator receivers. Let $\mathbf{f}_u$ represent the MMSE filter of length E applied at user u, then the equivalent channel-equalizer cascade at the output of the chip equalizer for user u is given by

$$
\boldsymbol{\alpha}^{(u)} = \mathbf{f}_u \mathbf{R}_{u,u} + \mathbf{f}_u \sum_{v \neq u}^{U} \mathbf{R}_{u,v},
\tag{1.45}
$$

which can be represented by

$$
\boldsymbol{\alpha}^{(u)} = \boldsymbol{\alpha}_{u,u} + \sum_{v \neq u}^{U} \boldsymbol{\alpha}_{u,v},
\tag{1.46}
$$

where $\boldsymbol{\alpha}_{u,u}$, is the channel-equalizer cascade for codes assigned to user u and $\boldsymbol{\alpha}_{u,v}$ is the channel-equalizer cascade for codes assigned to user v at user u. $\boldsymbol{\alpha}_{u,u}$ can in turn be split into the desired equalizer response and the residual inter-chip-interference and represented as

$$\boldsymbol{\alpha}_{u,u} = \boldsymbol{\alpha}_{u,u}^{d} + \overline{\boldsymbol{\alpha}}_{u,u} \tag{1.47}$$

$$\boldsymbol{\alpha}_{u,u}^{d} = \left[\; \overbrace{0 \ldots 0}^{d-1} \; \alpha_{u,u}^{d} \; \overbrace{0 \ldots 0}^{L+E-2-d} \; \right] \tag{1.48}$$

where d is the equalizer delay. The LMMSE equalizer is considered to be followed by a stacking operation allowing despreading and symbol decision.

Simple multiuser beamforming

To understand the effect of multiple-users with distinct beamforming weights in DL, it is insightful to derive the per-code SINR at the receiver for the case where multiple users are served in the downlink with different beamforming weights. When the BS employs different user-defined beamforming weights in downlink for MU transmission, at each receiver, codes assigned to different users propagate through U distinct *beamformed channels* even though the physical channel through which they propagate is the same. Without explicit knowledge of all beamforming weights used in the downlink, which is the so called interference aware [20] receiver, the receiver will not be able to effectively mitigate the effect of MUI. Since each user is aware only of beamforming weights that will be applied for codes assigned to itself and not of other users, the equalizer at each user is only matched to the beamformed channel seen by the codes assigned to this user. In computing the ideal beamforming weights for itself, a UE has to make some hypothesis on the beamforming weight vectors of other users in DL and choose the weight vector that maximizes the SINR corresponding to that hypothesis. For the general case where there exist U different users, defining K_u as the index set containing code indices of the uth user, the SINR per-code $SINR_{k \in K_u}$

that is seen by the code assigned to the user is given by

$$\frac{\sigma_k^2 |\alpha_{u,u}^d|^2}{\frac{1}{\mathcal{G}} \sum_{k \in K_u} \sigma_k^2 \|\overline{\boldsymbol{\alpha}}_{u,u}\|^2 + \sum_{v \neq u} \frac{1}{\mathcal{G}} \sum_{k \in K_v} \sigma_k^2 \|\boldsymbol{\alpha}_{u,v}\|^2 + \sigma_\eta^2 \mathbf{f}_u \mathbf{f}_u^H} \tag{1.49}$$

Where σ_k^2 denotes the chip variance of the kth code. In a simple extension of beamforming with multiple users with different beamforming weight vectors, each UE makes the assumption that all users in DL have the same beamforming weight vectors and computes the ideal beamforming weight vector under this assumption. The BS however makes no attempt to group users with same beamforming weights. As a result, it is expected that the downlink capacity drops significantly

Weight optimization by average interference criterion

Alternatively UE can anticipate that in reality, any of the four weights may be chosen by the other users in DL. Assuming that other users choose one of four beamforming weights with equal likelihood, it is reasonable to choose that beamforming weight which has the maximum SINR when averaged over all four hypothesis for the other users weights. Each UE computes therefore computes the ideal beamforming weight by plugging into (1.49), all possible combinations of weight vectors and feeds back the weight vector with the best average SINR over all the hypothesis for all the other users in DL. The idea is that while the true SINR at the receiver may still not be the same as expected SINR, the resulting SINR is higher than obtained by assuming the same beamforming weight is requested by all users scheduled in DL. Thus this beamforming vector must will perform better on average and increase the average data rate per user when compared to the simple multiuser beamforming case.

Cooperative beamforming

If the BS were to have the knowledge of the SINR seen by a particular user for all possible combinations of weight vectors applied at the base station, then, the BS can choose the optimal combination of weights that maximizes the downlink capacity. We call this cooperative beamforming because, in this case, all the users compute all possible SINRs corresponding to the weight vectors in the codebook. From (1.49) we see that for a given weight-vector, the SINR is highest when all other users also have the same beamforming weight-vector. Each user therefore feeds back as

many SINRs as the codebook size. Thus it is a form of cooperation between the users and BS to maximize system capacity. In practice, this involves considerable amount of receiver processing and also a lot of feedback to the BS. Nonetheless, the gains in such a case is worth investigating.

Scheduled beamforming

The practical and indeed the best solution to this problem with least complexity is for the BS to schedule in the DL, only those users that request the same beamforming weights. Each user assumes that same weights are applied to all codes in DL and computes the weight vector that maximizes the per code SINR. For this case, the user can then restore the orthogonality of all codes with the MMSE chip equalizer-correlator receiver. The per-code SINR for the uth user is then given by

$$\frac{\sigma_k^2 |\alpha_{u,u}^d|^2}{\frac{K}{\mathcal{G}}\sigma_k^2 \|\overline{\boldsymbol{\alpha}}_{u,u}\|^2 + \sigma_\eta^2 \mathbf{f}_u \mathbf{f}_u^H} \tag{1.50}$$

The combination of scheduling at BS and the choice of weight vector that maximizes the individual SINR at the receiver results in maximization of DL capacity.

1.4.2   Multiuser D-TxAA

For MU-D-TxAA system, we consider 2 separate UEs with N_{rx} receive antennas each. In a MU-D-TxAA system, the BS transmits 2 transport blocks for as many users scheduled in DL. All codes of a single stream are assigned to one user and re-used across the two streams. From Fig. 1.11, we see that the transmit signal vector in downlink can be modeled as

$$\mathbf{x}[j] = \underbrace{\mathbf{W}}_{2\times 2}\mathbf{b}[j] = \mathbf{W} \cdot \sum_{k=1}^{K} s[j]c_k[j \mod \mathcal{G}]\mathbf{a}_k[n] \tag{1.51}$$

$\mathbf{W} = [\mathbf{w}_1\, \mathbf{w}_2]$ is the 2×2 unitary precoding matrix. The columns of $\mathbf{W}$ are made up of the beamforming weight vectors corresponding to the two downlink users. The symbol vector $\mathbf{a}_k[n] = [a_{1k}[n]\ a_{2k}[n]]^T$ represents two independent symbol streams belonging to two different users. The spreading codes are common to the two streams and so is the scrambling sequence $s[j]$.
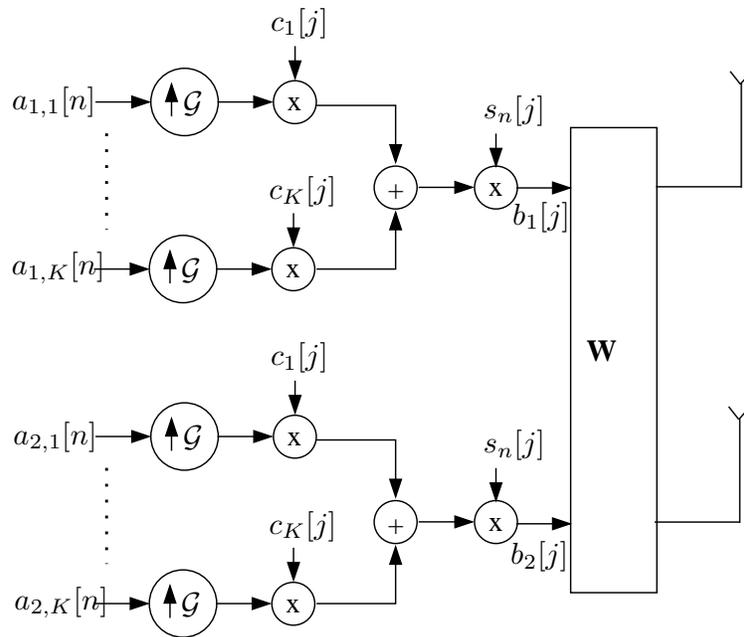
Figure 1.11: Multiuser D-TxAA transmit signal model.

Spatial Multiplexing Vs SDMA

In the spatial multiplexing context, there is only a single user in downlink and the precoding matrix corresponds to the weight vectors applied to the two separate streams transmitted to the same user. For such a case, we can write the equalizer output as the sum of an arbitrarily scaled desired term and an error term

$$\widehat{\mathbf{x}}[j] = \mathbf{x}[j] - \widetilde{\mathbf{x}}[j]. \tag{1.52}$$

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

In (1.52), 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 $\mathbf{W}^H$ is a linear operation and can be carried out before or after despreading.

Under the assumption of a FIR signal model, the estimation error covariance matrices $\mathcal{R}_{\widetilde{x}\widetilde{x}}$ (chip-level) and $\mathcal{R}_{\widetilde{z}\widetilde{z}}$ (symbol-level) are derived in [16]. It can be shown that the SINR for the qth

stream at the output of the output of the LMMSE chip equalizer/correlator is given by [16]

$$SINR_q = \frac{\sigma_a^2}{\left(\mathbf{W}^H \boldsymbol{\mathcal{R}}_{\tilde{z}\tilde{z}} \mathbf{W}\right)_{qq}} - 1. \tag{1.53}$$

where σ_a^2 corresponds to the symbol variance.

In the SDMA context, the BS transmits a single stream for each of the two downlink users. The BS applies the precoding matrix $\mathbf{W}$ whose columns correspond to the weight vectors fed back by the two users. It is obvious that two users who feedback the same weight-vector cannot be scheduled simultaneously for transmission in the downlink. At the receiver, each UE receives both the streams but processes only the stream assigned to itself. In HSDPA, 2×2 unitary precoding is used, this implies that the two columns of the precoding matrix are orthogonal. Moreover, knowledge of a single column automatically fixes the other column of $\mathbf{W}$. Thus, the BS does not have to explicitly inform one UE of the weight vector applied for the other UE. The SINR for the stream assigned to the user in question is therefore the same as in (1.53)

1.4.3   Numerical Results

In this section, we present Monte-Carlo simulation results and performance comparison of different beamforming strategies proposed in the paper. We consider a multipath channel with a maximum delay spread L of 10-chips with uniform power in all channel taps. At any given time BS simultaneously serves 2 users. The beamforming weights are calculated to maximize the per-code SINR at the output of the equalizer correlator combination. Simulations were carried out for a fixed SNR at each receive antenna while keeping the total transmit power is normalized to 1. The cumulative distribution function of the sum-capacity upper-bound in DL is then used as a performance metric to compare different strategies. Depending on the number of independent transport blocks at the transmitter the other simulation parameters are given as below

TxAA

Each UE is assumed to have single receive antenna. Normally, each UE feeds back only its preferred weight vector index, only in case of cooperative-operative beamforming, it feeds back SINR values to the BS. For the sake of simplicity we assume that each UE is allocated 7 of the 15 codes in the DL all with the same power.

D-TxAA

Each independent transport block is assumed to be allocated to a different user. Thus all codes of a stream are allocated to one user. For SDMA with single antenna receivers, we assume users with orthogonal weights are scheduled together. For SDMA with 2-antenna receivers, users with different beamforming weight vectors are assumed to be scheduled together. In the spatial multiplexing case, a 2×2 MIMO system is assumed with all codes and both streams transmitted to a single user. Fig.1.12 compares the sum-capacity in the DL for the case of TxAA. The DL ca-
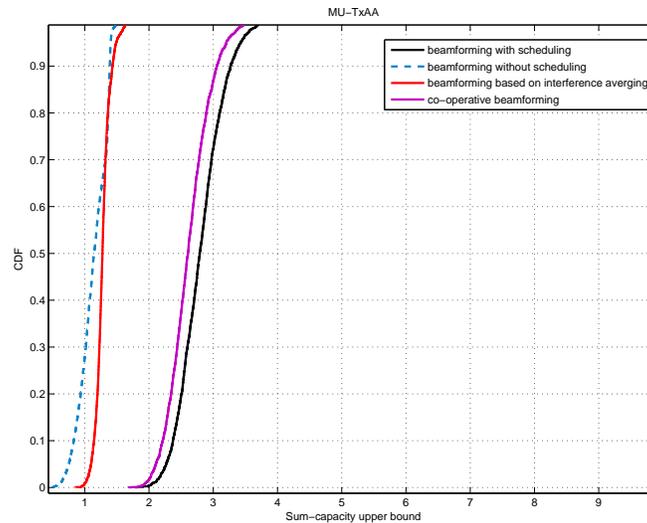
Figure 1.12: Performance of different beamforming schemes for MU-TxAA.

pacity is worst for the case of beamforming without scheduling. This is because of the inability of the receivers to effectively restore orthogonality for all codes and hence effectively mitigate MUI since they do not know the actual beamforming weight of the other user. When the beamforming weight is optimized by the average interference criterion, the weights are not just chosen based on
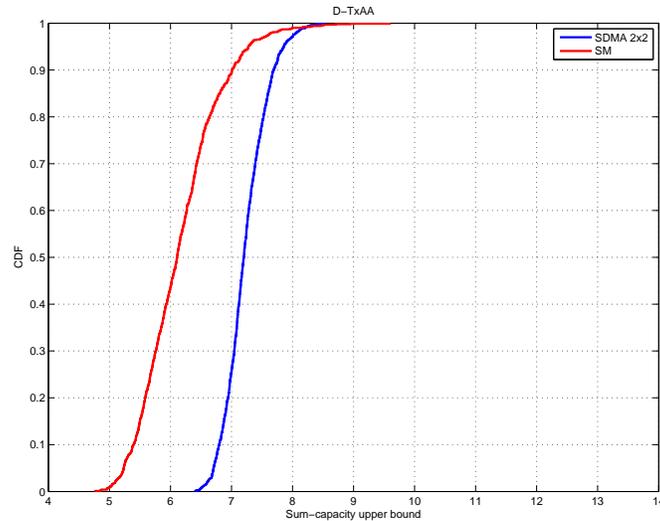
Figure 1.13: DL sum-capacity for MU-D-TxAA.

the channel seen by each user, but also based on the capability of these weights to reduce the average multi user interference due to different beamforming weights of the other user.the downlink capacity is thus better than that in the case of simple multiuser beamforming. At the cost of an increase in complexity and feedback, cooperative-operative beamforming performs better than that of the earlier schemes, even so, it is does not do better than the scheduled beamforming because the UEs need not necessarily be assigned the weight vector that maximizes their individual SINR. Scheduled beamforming thus outperforms all the other schemes since in this case each user is able to effectively mitigate MUI due to the same beamformed channel seen by all codes in downlink. It should be noted that for the case where the total number of users in DL far exceed the number of users actually scheduled in the DL, the performance of cooperative-operative beamforming is expected to improve. In Fig.1.13, we compare the performance of D-TxAA in spatial multiplexing mode with that of the multiuser (SDMA) mode. Simulation results show that the DL sum-capacity is greater for the case of SDMA with single stream transmission to both users.

1.5   Concluding Remarks

In this chapter we first discussed advanced receiver designs for MIMO HSDPA based on the concept of combined chip-level and symbol level processing. In particular, the chip-level processing

stage was the SINR maximizing LMMSE chip-equalizer which in addition to restoring the orthogonality of the codes also achieves spatial separation to a certain degree. Further processing stages at symbol level was introduced to enhance the performance of the receivers. When MIMO HSDPA receivers are based on MMSE designs, we showed that there exists an optimal choice of precoding matrix to be employed at the transmitter that maximizes the sum-capacity of these receivers and derived analytical expressions for the choice of the optimal precoding matrix. The receiver designs that we discussed in the first part of the chapter operated on the assumption that the scrambler used at the transmitter can be treated as random. Since random treatment of scrambler contributes to degradation of receiver performance, in the second section, we looked at receiver designs that treat the scrambler as deterministic. We saw that such receivers can then resort to time-varying symbol level processing after the equalizer-correlator stage in order to re-gain the time varying signal contribution which would otherwise be treated as noise. This leads to additional gains in SINR which ultimately effects achievable capacity of the receivers. Simulation results show that indeed these receivers outperform conventional receivers that treat the scrambler as random. Finally we discussed multi-user extensions to closed loop transmit diversity techniques that have been standardized in [1] and proposed multi-user beamforming strategies that can be employed at the BS in order to maximize the downlink capacity. Simulation results show that for MIMO HSDPA, downlink capacity is maximized by using the MIMO channel to service multiple single stream users (SDMA) instead of single user spatial multiplexing which is currently supported in the standards.

References

[1] 3GPP. *TS 25.214 Physical layer procedures (FDD) (Release 7)*, May 2007. Version 7.5.0.

[2] Ahmet Bastug and Dirk T M Slock. Downlink WCDMA receivers based on combined chip and symbol level equalization. *European Transactions on Telecommunications*, 10, Feb 2005.

[3] Jinho Choi, Seong-Rag Kim, Yan Wang, and Cheng-Chew Lim. Receivers for chip-level decision feedback equalizer for CDMA downlink channels. *IEEE Transactions on Wireless Communications*, 3(1):300–313, January 2004.

[4] 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*, 43(10), Oct. 1995.

[5] G. J. Foschini and M. J. Gans. On limits of wireless communications in a fading environment when using multiple antennas. *Wirel. Pers. Commun.*, 6(3):311–335, 2009.

[6] G. J. Foschini. Layered space-time architecture for wireless communication in a fading environment when using multielement antennas. Technical report, Bell Labs, Autumn 1996.

[7] David Gesbert, Marios Kountouris, Robert Heath, Chan-Byoung W;Chae, and Thomas Slzer. Shifting the MIMO paradigm. *IEEE Signal Processing Magazine*, 24(5):36–46, September 2007.

[8] Irfan Ghauri, Shakti P. Shenoy, and Dirk T. M. Slock. On LMMSE bias in CDMA SIMO/MIMO receivers. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, Las Vegas, NV, March 2008.

[9] Irfan Ghauri and Dirk T. M. Slock. Linear receivers for the DS-CDMA downlink exploiting orthogonality of spreading sequences. In *Proc. 32nd Asilomar Conf. on Signals, Systems & Computers*, volume 1, pages 650–654, Pacific Grove, CA, November 1998.

[10] Global Mobile Suppliers Association. 3G/WCDMA-HSPA Fact Sheet. Technical report, GSA, March 2009.

[11] V. Haikola, M. Lampinen, and M. Kuusela. Practical multiuser beamforming in WCDMA. In *Proc. IEEE Vehicular Technology Conference*, Fall 2006.

[12] 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.

[13] A. Lozano and C. Papadias. Layered space-time receivers for frequency-selective wireless channels. *Communications, IEEE Transactions on*, 50(1):65–73, Jan 2002.

[14] Laurence Mailaender. Linear MIMO equalization for CDMA downlink signals with code reuse. *IEEE Transactions on Wireless Communications*, 4(5):2423–2434, September 2005.

[15] A. J. Paulraj and C. B. Papadias. Space-time processing for wireless communications. *IEEE Signal Processing Magazine*, pages 49–83, November 1997.

[16] Shakti. Prasad Shenoy, Irfan Ghauri, and Dirk T M Slock. Optimal precoding and MMSE receiver designs for MIMO WCDMA. In *Proc. IEEE 67th Vehicular Technology Conference*, Singapore, May 2008.

[17] D.K.C. So and R.S. Cheng. Detection techniques for v-blast in frequency selective fading channels. volume 1, pages 487–491 vol.1, Mar 2002.

[18] Q. H. Spencer, C. B. Peel, A. L. Swindlehurst, and M. Haardt. An introduction to multi-user MIMO downlink. *IEEE Communications Magazine*, October 2004.

[19] I. Emre Telatar. Capacity of multi-antenna gaussian channels. *European Transactions on Telecommunications*, 10:585–595, 1999.

[20] M. Wrulich, C. Mehlfuhrer, and M. Rupp. Interference Aware MMSE Equalization for MIMO TxAA. In *Proc. 2008 3rd International Symposium on Communications, Control and Signal Processing (ISCCSP2008)*, March 2008. St.Julians, Malta.

[21] Jianzhong (Charlie) Zhang, Balaji Raghothaman, Yan Wang, and Giridhar Mandyam. Receivers and CQI measures for MIMO-CDMA systems in frequency-selective channels. *EURASIP Journal on Applied Signal Processing*, (11):1668–1679, November 2005.