HSDPA/HSUPA Handbook

by Borko Furht and Syed Ahson (eds)

Chapter 2: Receiver Designs and Multi-User Extensions to MIMO HSDPA

Shakti Prasad Shenoy, Irfan Ghauri, and Dirk T.M. Slock

2.1 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. *Making the communication link resilient/robust to channel fades.* Diversity techniques have long been considered 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 are statistically independent of 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 [19].

2. *Increasing 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 substreams 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. Because this requires no extra spectral resources, the total data rate (bits per second, b/s) transmitted over the communication link is increased.

3. *Increasing 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 is increased by combining the signals at each receive antenna after application of suitable weights (receive beamforming).

4. *Improving spectral efficiency.* By reusing the multiple access resources [for instance, spreading codes in Code Division Multiple Access (CDMA)] over the spatial dimension, MIMO systems can increase spectral efficiency (b/s/Hz) of the communication system.

However, not all these features can be provided simultaneously. For instance, there exists a trade-off 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 [5,19]. While there has been sustained academic interest in MIMO over the past decade, as witnessed by the huge number 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 can 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 desirable due to space, battery, and cost constraints of mobile terminals. As a result, standardization bodies have to date concentrated more on the subclass of MIMO systems [Multiple-input, single-output (MISO)/single-input, multiple-output (SIMO)], whereby some kind of antenna diversity at the base station (BS) 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 for 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 to support multiple antennas at the receiver. Foreseeing these developments 2 × 2 MIMO has been standardized [1]. In fact, the world's first HSPA+ (High Speed Packet Access) or evolved HSPA network with support for 2 × 2 MIMO was launched in early 2009 [10]. Along with enabling technologies such as adaptive modulation and coding (AMC), fast hybrid automatic repeat request (HARQ), and user feedback based scheduling, MIMO in HSDPA (HS Downlink PA) can lead to peak data rates of 42 Mbps in downlink (DL). However, in the present form, MIMO in HSDPA can support only single-user (SU) scenarios in DL. While shifting from a single user to a multi-user (MU) 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 the frequency non-selective (OFDM) case, where optimal joint-stream maximum a-posteriori (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 maximum likelihood (ML) receivers for multipath channels. For MIMO WCDMA (wideband code division multiple access) transmission in frequency-selective channels, where the multipath mixes up signals 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 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. 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 modeled as a random *i.i.d.* sequence and instead treat it as a deterministic, known sequence. This results in enhanced performance at the cost of increased complexity at the receiver.

This 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 make the customary assumption that the scrambler can be considered a random *i.i.d.* sequence. We also address here the issue of optimal choice of precoding matrix for the receiver. Because 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 do away with the assumption that the scrambler is random and treat it as deterministic. We 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 extend the current single-user MIMO scenarios in HSDPA to the multiuser case. These extensions require minimal changes to existing standards. When multiple UEs must be simultaneously serviced in the downlink, we suggest practical multi-user scheduling strategies that can be employed at the BS so as to maximize the downlink capacity. Finally we wrap up with some concluding remarks.

2.1.1 MIMO in HSDPA

3GPP (3rd Generation Partnership Project) has introduced a variant of Per-Antenna Rate Control (PARC), namely, dual stream transmit antenna array (D-TxAA) for transmit adaptive array 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 (signal-to-interference-plus-noise ratio)-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 have been incorporated by 3GPP as standard in the form of TxAA (Figure 2.1) and its dual stream counterpart D-TxAA (Figure 2.2) for MIMO HSDPA. HSDPA supports a closed loop transmit diversity technique called (TxAA). In the 2 transmit-1 receive (2 × 1) antenna configuration of TxAA, the UE feeds back optimum beamforming weights that the BS uses while transmitting data to UE. D-TxAA is the extension of TxAA when UEs are configured in MIMO mode. 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 must use when transmitting data to the UE. Let us now look at the beamforming/precoding aspect in more detail.

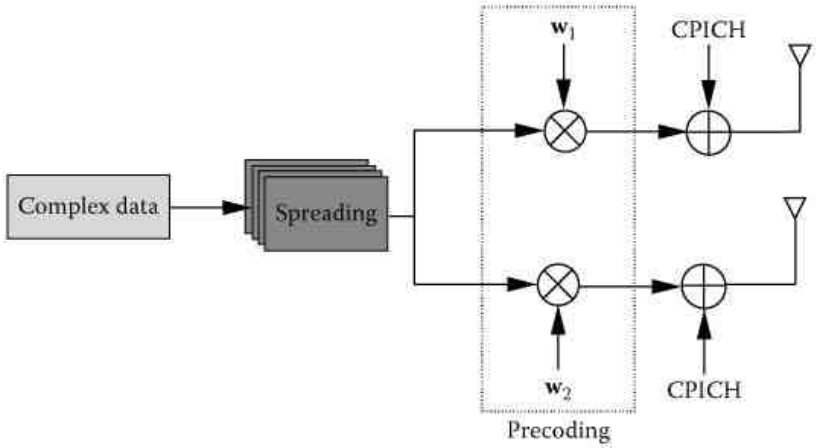

Figure 2.1: Simplified block diagram of transmitter processing for TxAA.

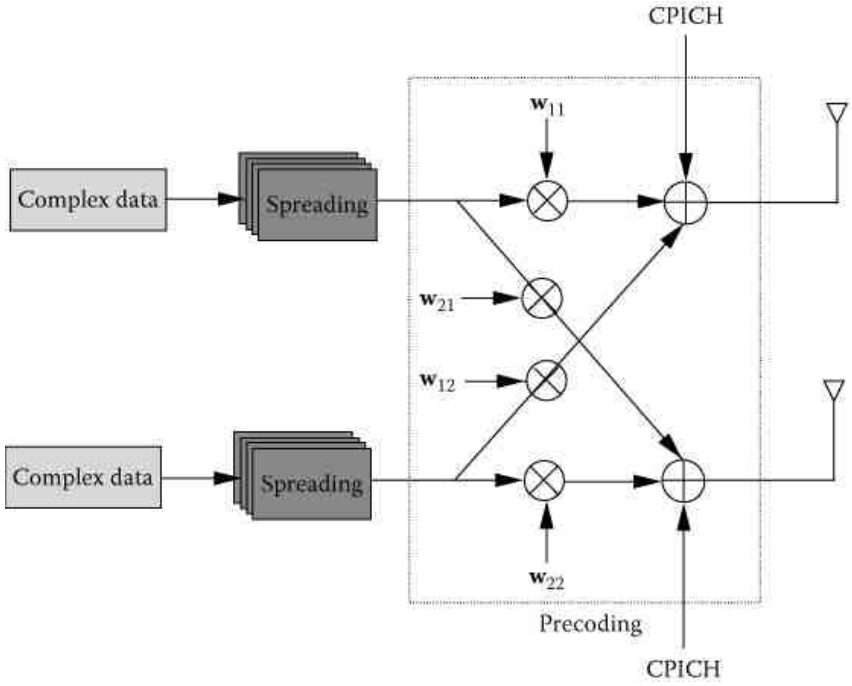

Figure 2.2: Simplified block diagram of transmitter processing for D-TxAA.

2.1.2 Precoding and CQI Feedback

In HSDPA, the UE is required to submit regular CQI and precoding control indicator (PCI) reports to the BS. The CQI 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 predefined packet error rate (PER). The mapping strategy has been subject to significant simulation study (see, e.g., [12]) and the SINR → CQI ↔ PER ↔ 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 (transmission time interval) over which the UE computes the CQI, PCI is computed using the Common Pilot Channel [CPICH(s)] transmitted from both transmit antennas to decide the beam-forming vector. When applied by the BS, this beam forming vector maximizes the aggregate transport block size that the UE can 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 four beamforming weights that control the antenna phase at the 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. Because the pre-coding 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 $w \in \{\frac{1+\sqrt{-1}}{2}, \frac{1-\sqrt{-1}}{2}, \frac{-1+\sqrt{-1}}{2}, \frac{-1-\sqrt{-1}}{2}\}$. One choice of beamforming weight vector, let us call it **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 **H**. However, for frequency-selective channels with a delay spread *L*, there are *L* such MIMO channel taps. In general, it is not possible to choose **w** to match all channel taps; precoding gain in such conditions is in practice very low.

When the UE is configured in MIMO mode and requests two transport blocks to be transmitted, a precoding matrix must 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. 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 four precoding matrices, two of them are related to the remaining as follows. Let $w_1 = \beta$; then by design, $w_3 = \beta$, $w_4 = -w_2$, and

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

Therefore,

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

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

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

2.2 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 proposed here are based on combining chip-level and symbol-level processing for enhanced performance. For each of these receivers, we derive the per-stream SINR expressions. We use the SINR to compute the sum-capacity, which can be interpreted as the 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 the achievable sum-rate of the MIMO channel through its influence on the SINR of streams at the RX output. For D-TxAA with unitary precoding, there exists an optimal choice of the precoding matrix that maximizes 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 and the extent of its impact depends on the MIMO receiver.

For the spatial multiplexing case in MIMO HSDPA, Figure 2.3 illustrates the equivalent baseband downlink signal model. The received signal vector (chip rate) at the UE can be modeled as:

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

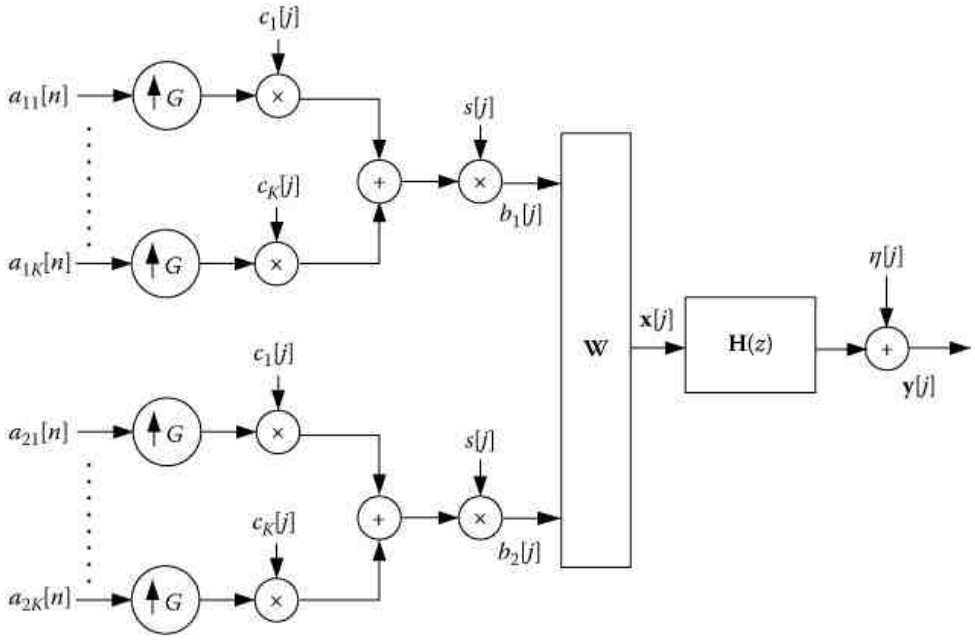

Figure 2.3: MIMO signal model with precoding.

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

(2.3)
$$\mathbf{x}[j] = \underbrace{\mathbf{W}}_{2\times 2}\mathbf{b}[j] = \mathbf{W}\cdot\sum_{k=1}^{K}\underbrace{s[j]c_k[j \mod \mathcal{G}]\mathbf{a}_k[n]}_{\mathbf{b}_k[j]}$$

where *k* is the code index, $n = \lfloor \frac{j}{\mathcal{G}} \rfloor$ is the symbol index, $\mathcal{G}$ is the spreading factor ($\mathcal{G}$ = 16 for HSDPA), **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, $\mathbf{c}_k = [c_k[0] \cdots c_k[\mathcal{G} - 1]]^T$, where $\mathbf{c}_k^T\mathbf{c}_{k'} = \delta_{kk'}$ 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 $\frac{1}{\sqrt{2}}\{\pm 1 \pm \sqrt{-1}\}$.

2.2.1 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 is shown in [9] that for single-input, single-output (SISO) links, linear minimum mean square error (LMMSE) equalizers can be used to restore the orthogonality of these codes prior to the despreading operation, thereby providing superior performance in multipath channels with large delay spreads. In frequency-selective MIMO channels, multipath mixes up signals in space and time, calling for alternative reception strategies. It is shown in [14] that linear chip-level MMSE equalizers not only are able to restore orthogonality of codes, but are also able to efficiently achieve spatial separation in MIMO frequency-selective channels. In the following section, we discuss receivers for MIMO HSDPA based on chip-level equalization.

2.2.1.1 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 the linear (MMSE) FIR estimation of the 2 × 1 chip sequence assuming the channel to have a finite impuse response (FIR). In the spatial multiplexing context, the LMMSE equalization tries to suppress not only all inter-chip

interference (ICI), but also all inter-stream interference (ISI). In Figure 2.3, **b**[*j*] is the input chip vector defined as **b**[*j*] = [$b_1[j]$ $b_2[j]$]^T, where $b_i[j]$ is the *j*th chip of the *i*th input stream. Each chip stream is the sum of *K* spread and scrambled CDMA substreams (one user per CDMA code). Thus, $b_i[j] = \sum_{k=1}^{K} b_{ik}[j]$. The 2 × 2 matrix **H**[*j*] is the *j*th MIMO tap of the FIR channel and **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 2*p* × 1 received signal at the *j*th time instant is given as

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

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

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

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 × 2*pE* LMMSE equalizer **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

(2.6) $$\hat{\mathbf{x}}[j] = \mathbf{F}\ \mathbf{Y}[j] = \mathbf{B}\underbrace{\mathbf{W}\ \mathbf{b}[j]}_{\mathbf{x}[j]} + \underbrace{\overline{\mathbf{B}}\mathcal{W}_{L+E-1}\overline{\mathbf{b}}_{L+E-1}[j] + \mathbf{F}\boldsymbol{\Xi}[j]}_{-\widetilde{\widetilde{\mathbf{x}}}[j]}$$

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

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

which are, respectively, the 2 × 2 matrices that represent the *joint bias* in the equalizer output and the residual ICI. The $\overline{\alpha}^{(ij)}$ are the same as $\alpha^{(ij)}$, with the $\alpha_d^{(ij)}$ term replaced by 0, and *d* is the equalization delay (in chips) associated with **F**.

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

(2.7) $$\hat{\mathbf{x}}[j] = \mathbf{B}\mathbf{x}[j] - \widetilde{\widetilde{\mathbf{x}}}[j]$$

In Equation (2.7), an estimate of the chip sequence **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 Figure 2.4). 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 error covariance matrix corresponding to the error term is denoted by $\mathcal{R}_{\widetilde{\widetilde{x}}\widetilde{\widetilde{x}}}$, from which the MSE can be obtained as follows:

(2.8) $$\mathcal{R}_{\widetilde{\widetilde{x}}\widetilde{\widetilde{x}}} = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}$$

(2.9)
$$r_{11} = \sigma_b^2 \left(\|\overline{\boldsymbol{\alpha}}^{(11)}\|^2 + \|\overline{\boldsymbol{\alpha}}^{(12)}\|^2 \right) + \mathbf{f}_1 \mathcal{R}_{\eta\eta} \mathbf{f}_1^H$$

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

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

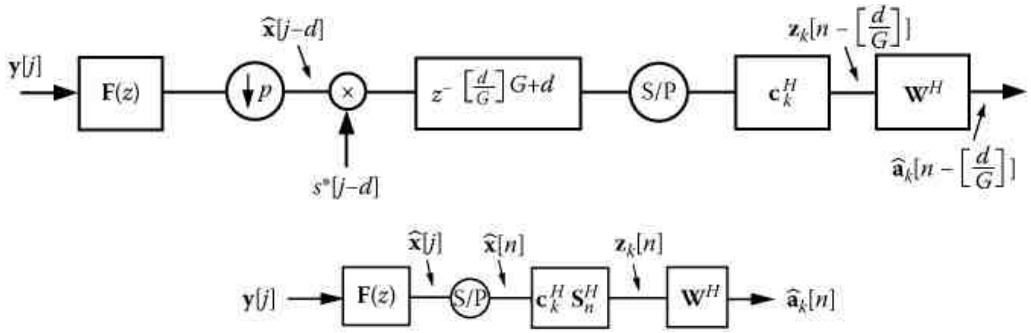

Figure 2.4: LMMSE equalizer and correlator. The second figure is a simplified representation used as chip-equalizer/correlator front-end stage for other receiver structures.

After despreading (for the *k*th code), the 2 × 1 signal at the symbol level is written as

(2.10)
$$\mathbf{z}_k[n] = \mathbf{W}\mathbf{a}_k[n] - \widetilde{\mathbf{z}}_k[n] = \mathbf{B}\mathbf{W}\mathbf{a}_k[n] - \widetilde{\widetilde{\mathbf{z}}}_k[n]$$

Because the LMMSE joint bias is accounted for in **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 the scrambler as a random sequence and taking expectation over the scrambler as well as the 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 $\mathcal{R}_{\widetilde{\widetilde{x}}\widetilde{\widetilde{x}}}$ with the 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

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

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

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

The SINR for the *i*th stream at the output of the output of the LMMSE chip-equalizer/correlator is therefore

(2.11)
$$SINR_i = \frac{\sigma_a^2}{\left(\mathbf{W}^H \mathbf{B}^{-1} \mathcal{R}_{\widetilde{\widetilde{z}}\widetilde{\widetilde{z}}} \mathbf{B}^{-H} \mathbf{W} \right)_{ii}} - 1$$

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 *i*th data stream can now be expressed as

(2.12)
$$\mathcal{C}_i = \log(1 + SINR_i)$$

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

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

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

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 $\mathcal{R}_{\tilde{z}\tilde{z}}$. In other words, the $\mathbf{W}_{opt}$ attempts to maximize the SINR difference between the two streams. [1]

2.2.1.2 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, and re-channelizes the detected data so that the contribution of the detected stream can be subtracted from the received signal (Figure 2.5). The second stream can now be detected using a new FIR LMMSE chip-level receiver obtained as

(2.14) $$\mathbf{H}_{sic} = \overline{\mathbf{W}}\,\mathbf{H}$$

where $\overline{\mathbf{W}}$ is a diagonal matrix with $\mathbf{w}_2$ on its diagonal, and $\mathbf{H}_{sic}$ is the equivalent channel seen by the stream detected last due to cancellation effected at the receiver

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

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 TxAA 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 (because spatial processing, by nature, is simultaneous). Such treatment increases the complexity but may be well worth the effort in terms of SINR gains. Before moving on to combined chip-level and symbol-level processing, we would like to draw the reader's attention to the fact that the receiver structure discussed here can be classified as a symbol-level DFE (decision feedback equalizer). It is true that the interference cancellation is performed at chip level after re-spreading. However, the fact that feedback is based on decisions taken at the symbol level makes this equivalent to cancellation at the symbol level.

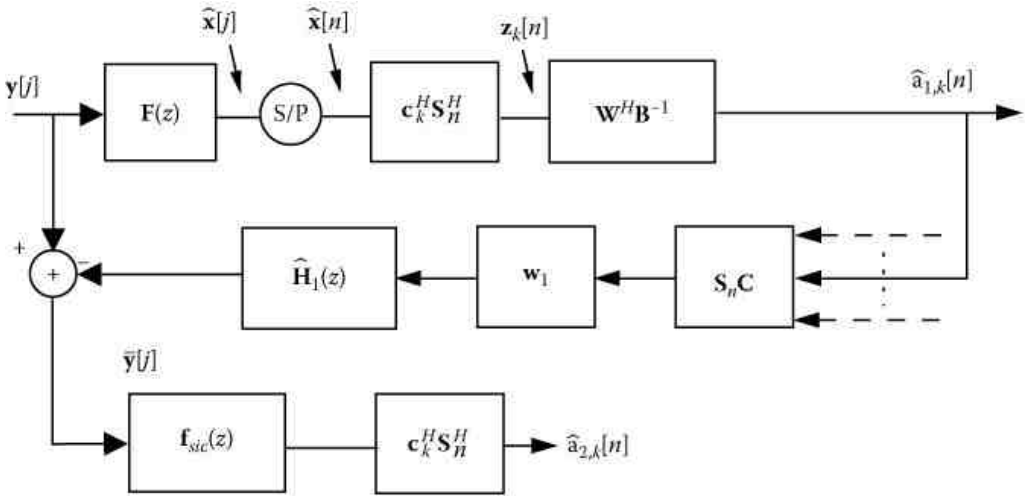

Figure 2.5: LMMSE chip equalizer/correlator and chip-level SIC for stream 2.

2.2.2 Combined Chip-Level and Symbol-Level Equalization

We start by motivating the need for combined chip-level and symbol-level processing and then move on to some receiver designs based on this approach.

Optimal linear receivers for WCDMA are symbol-level (deterministic), time-varying, multi-user receivers that are known to be prohibitively complex. One class of such receivers is based on symbol-level multi-user detection (MUD), where linear or nonlinear 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. Nonlinear MUD methods focus on estimating, reconstructing, and subtracting signals of interfering codes and are, 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 the loss of orthogonality due to the multipath channel, and this problem is effectively solved by attempting to restore orthogonality through an 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 the LMMSE chip equalizer in this context is not perfect and therefore mandates additional processing that can be performed at the 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 a channel *sparsifier* or indeed a more specific spatio-temporal → spatial channel-shortener (e.g., 2*N* × 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. To be precise, the chip-level processing stage will always consist of the MIMO LMMSE chip equalizer, which will be followed by the correlator. We then consider various symbol-level processing stages that can be employed at the receiver.

2.2.2.1 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 Figure 2.6. As discussed in the previous subsection, "Chip-Level Equalization," the output of the correlator is $\mathbf{z}_k[n]$, given by Equation (2.10). $\mathcal{F}_{sp}$ denotes the spatial MMSE at the output of which we have a linear estimate of the symbol vector as

(2.16) $$\widehat{\mathbf{a}}_k[n] = \mathbf{a}_k[n] - \widetilde{\mathbf{a}}_k[n]$$

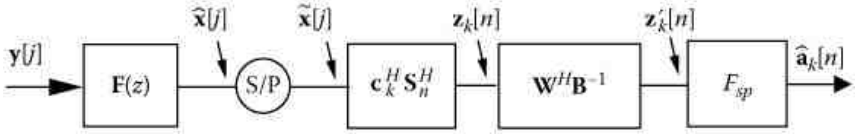

Figure 2.6: Chip LMMSE equalizer and correlator followed by symbol-level (spatial) MMSE.

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

(2.17) $$\mathcal{R}_{\widetilde{a}\widetilde{a}} = \mathcal{R}_{aa} - \mathcal{R}_{az'}\mathcal{R}_{z'z'}^{-1}\mathcal{R}_{z'a}$$

(2.18) $$= \sigma_a^2\mathbf{I} - \sigma_a^4\mathbf{W}^H\left(\sigma_a^2\mathbf{I} + \mathbf{B}^{-1}\mathcal{R}_{\widetilde{z}\widetilde{z}}\mathbf{B}^{-H}\right)^{-1}\mathbf{W}$$

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

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

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

(2.20) $$\mathbf{B} = \mathbf{I} - \mathcal{R}_{\widetilde{z}\widetilde{z}}\mathcal{R}_{zz}^{-1}$$

as with the LMMSE chip-level equalizer/correlator receiver, this translates to a sum-capacity expression similar to the one derived in the previous subsection:

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

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

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

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.

2.2.2.2 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 Figure 2.7. Denote the output of the correlator as $\mathbf{u}_k[n]$, written as

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

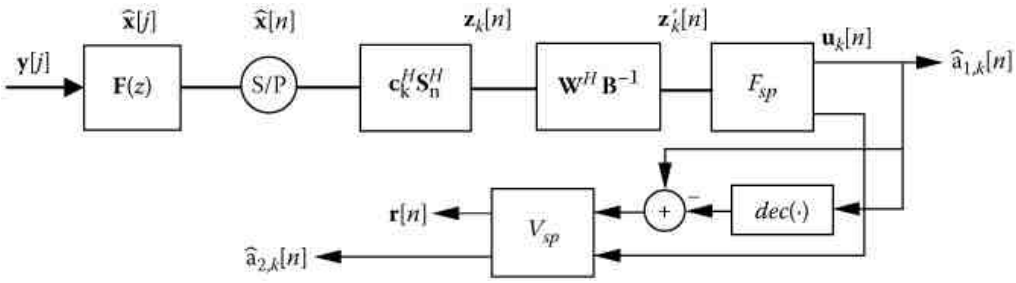

Figure 2.7: Chip LMMSE equalizer/correlator followed by spatial MMSE and symbol-level SIC for stream 2.

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

(2.23) $$\mathcal{R}_{\tilde{\tilde{u}}\tilde{\tilde{u}}} = \mathcal{F}_{sp}\mathbf{W}^H\mathbf{B}^{-1}\mathcal{R}_{\tilde{\tilde{z}}\tilde{\tilde{z}}}\mathbf{B}^{-H}\mathbf{W}\mathcal{F}_{sp}^H$$

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

(2.24) $$\mathcal{R}_{\tilde{r}\tilde{r}} = \mathcal{V}_{sp}\mathcal{R}_{\tilde{\tilde{u}}\tilde{\tilde{u}}}\mathcal{V}_{sp}^H$$

which is minimized if $\mathcal{R}_{\tilde{r}\tilde{r}}$ = **D**, which is a diagonal matrix, and the problem boils down to the estimation of the error term in stream 2 from stream 1. Toward this end, consider LDU factorization of $\mathcal{R}_{\tilde{\tilde{u}}\tilde{\tilde{u}}}$ = **LDL**^{H}. Then, $\mathcal{V}_{sp}$ = **L**$^{-1}$ minimizes

Equation (2.24). Denoting elements of $\mathcal{R}_{\tilde{i}\tilde{i}}$ as r_{ij}, the elements of **D** are given as $\sigma^2_{\tilde{r}_1} = r_{11}$ and

(2.25) $$\begin{aligned}\sigma^2_{\tilde{r}_2} &= r_{22} - r_{21} r_{11}^{-1} r_{12} \\ &= \det(\mathcal{R}_{\tilde{i}\tilde{i}}) \\ &= \det(\mathcal{F}_{sp}) \det\left(\mathbf{B}^{-1} \mathcal{R}_{\tilde{z}\tilde{z}} \mathbf{B}^{-H}\right) \det\left(\mathcal{F}_{sp}^{H}\right),\end{aligned}$$

Thus, MMSE for stream 1 is $\sigma^2_{\tilde{r}_1}$ and that of stream 2 is $\sigma^2_{\tilde{r}_2}$. As depicted in Figure 2.7, 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 natched filter bound).

An interesting observation is that the SINR expression for stream 2 in the symbol-level SIC case is independent of the precoding **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], because 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 have 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, thus reducing chances of error propagation during the interference cancellation stage and hence directly impacting the detection performance of stream 2.

It should, however, be noted that any low-complexity, symbol-level processing is hardly comparable to the 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 the noise-plus-interference correlation between streams to improve the SINR of the symbol detected last, the latter benefits from stripping the spatiotemporal interference of the entire detected stream, where for the stream detected last, all streams can hence-forth be considered nonexistent (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 any remaining interference. For SIC, the stream detected last is known to attain the matched filter bound (MFB).

2.2.2.3 Spatial ML Receiver

Figure 2.8 shows another possible receiver structure, where the chip-equalizer correlator front end is followed, as before, by the spatial MMSE stage. The resulting spatial mixture

(2.26) $$\mathbf{u}_k[n] = \mathcal{F}_{sp} \mathbf{z}'_k[n] = \mathbf{a}_k[n] - \tilde{\tilde{\mathbf{u}}}_k[n]$$

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

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

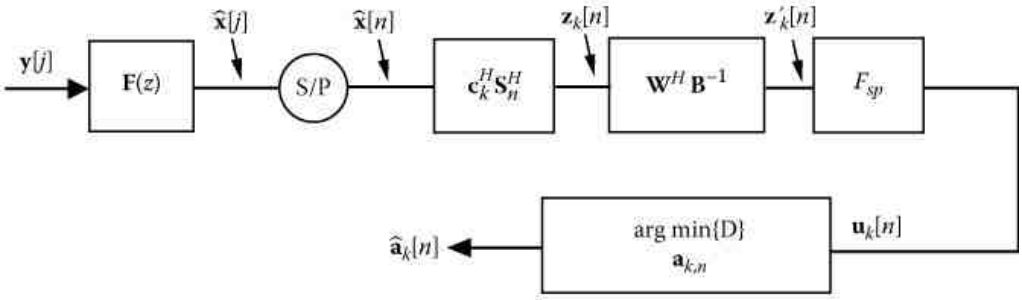

Figure 2.8: Chip LMMSE equalizer/correlator followed by spatial MMSE and joint detection.

This metric can be solved for $\mathbf{a}_k[n]$. It is shown in [21] that joint detection outperforms SIC. However, the SIC structure in [21] addresses an 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 *i*th stream corresponds to the MFB of spatial channel resulting from the cascade of $\mathcal{F}_{sp}$ and **B**. The MFB can be interpreted as the SNR of the *i*th stream when it is detected, assuming that the symbols of the other stream(s) are known. $\mathcal{R}_{\tilde{i}\tilde{i}}$ is the noise variance.

2.2.3 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 $2p \times 2$ FIR MIMO 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 cumulative distribution function (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 the sum-rate) for the LMMSE/correlator design. The difference between them is 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 either of the two matrices, but one can also choose which stream of 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. Figure 2.9 shows the 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 very small complexity, we are able to see some gain in the achievable rates of the receiver.

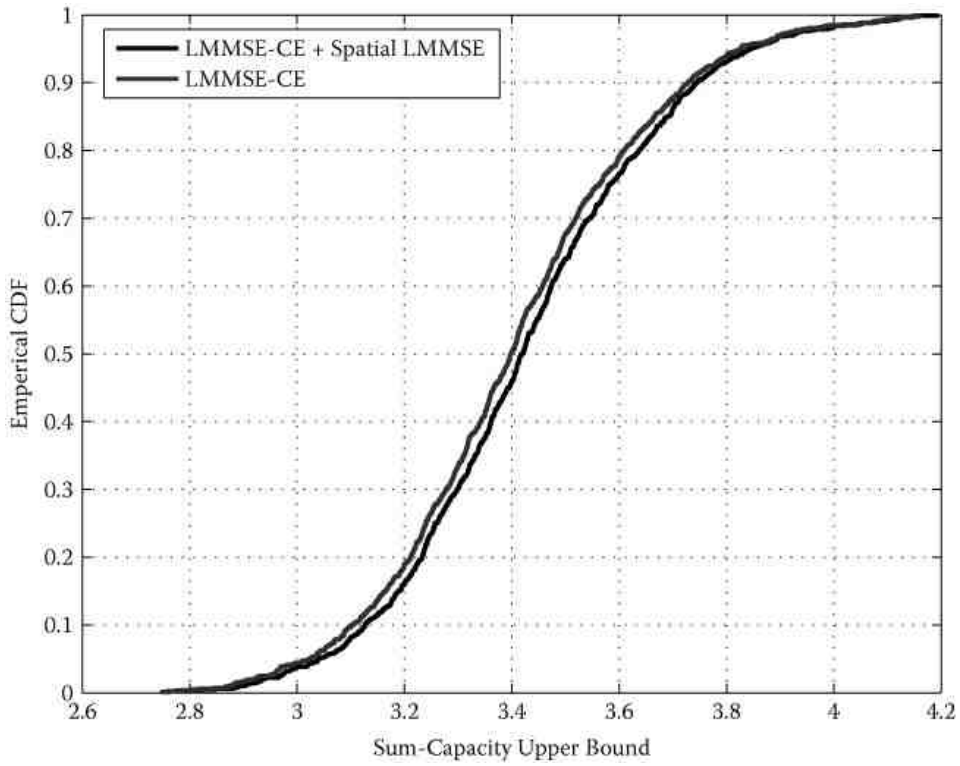

Figure 2.9: Performance of LMMSE chip equalizer/correlator receiver and LMMSE chip equalizer and spatial MMSE receiver.

In Figure 2.10 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 nonlinear equalization stage for stream 2. That the gain is not considerable is due to the fact that stream 1 does not benefit from nonlinear equalization. Because 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 obtain much better performance. The chip-level SIC in Figure 2.11, as can be expected, outperforms all other receivers at the cost of a significant processing delay and architectural complexity at the receiver.

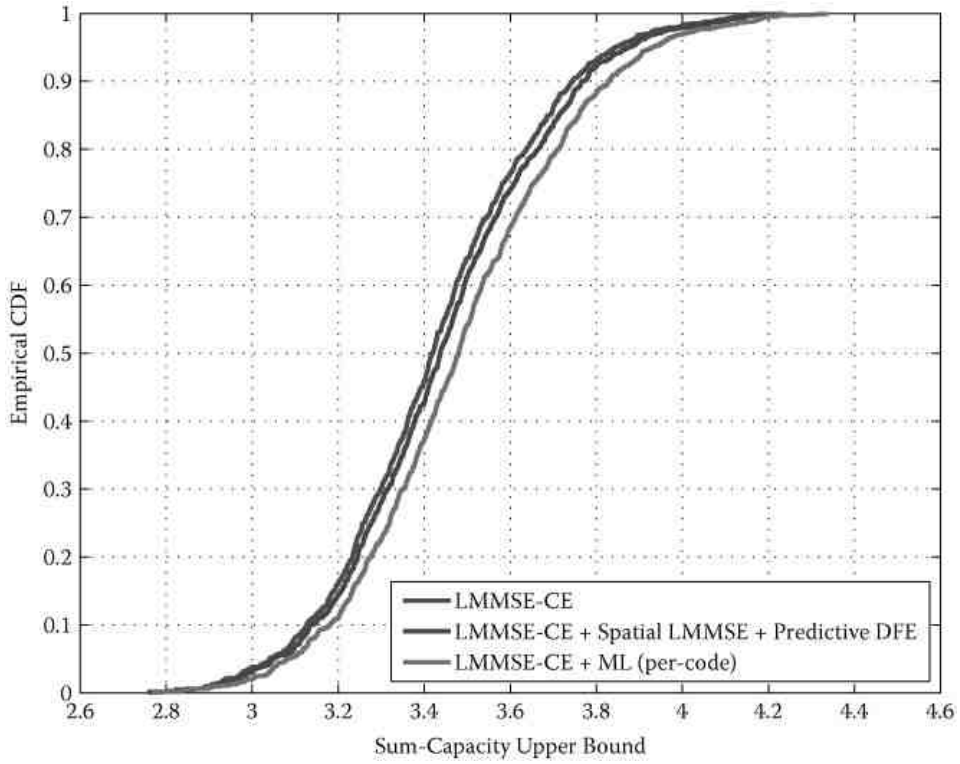

Figure 2.10: Comparison of sum-capacity upper bounds for different receiver structures.

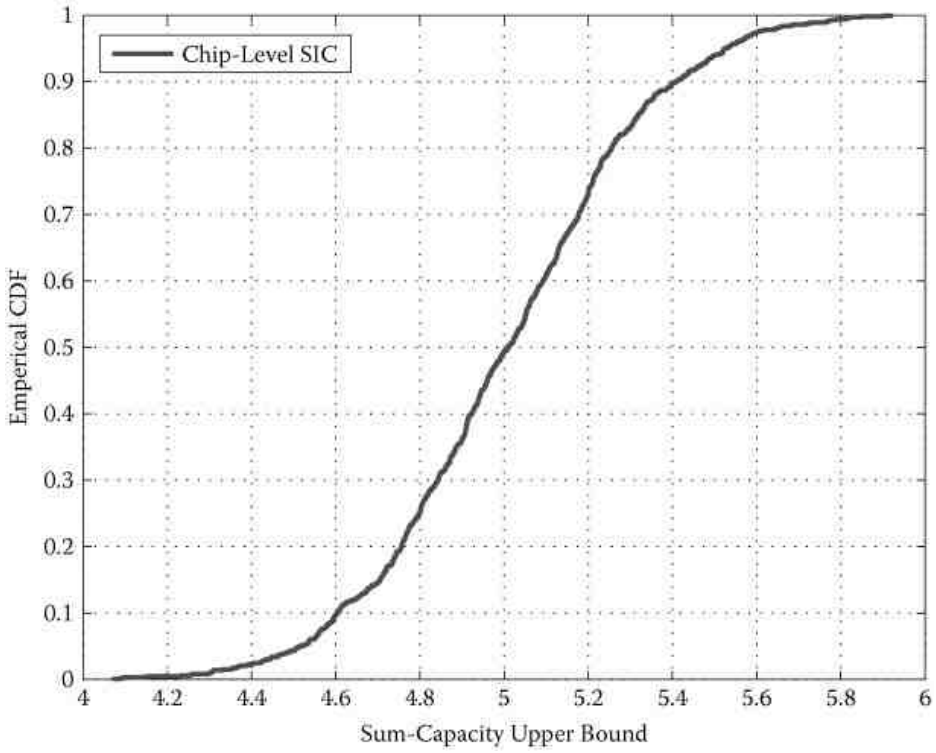

Figure 2.11: Upper bound for sum-capacity for the chip-level SIC receiver.

[1]To its best abilities given the limited resolution of **W**.

2.3 Receiver Designs for MIMO HSDPA: Part II

Thus far, we have discussed various receiver designs that have assumed that the scrambler was random *i.i.d.* Modeling the scrambler as random *i.i.d* 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. Because 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 can be linked to the efficacy of the dimensionality reduction achieved at the output of the chip equalizer and also the type of processing at the 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 that N_t is the number of TX streams, N_r the number of RX antennas, and p 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 nonlinear MUD solutions). In the previous section, we chose to trade off 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 and focus on the resulting spatial channel model [8]. Such a treatment mandates 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 here that deterministic treatment of the 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 do not consider here the precoding aspect of DL transmission. However, we do stress that introduction of pre-coding 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 2.12).

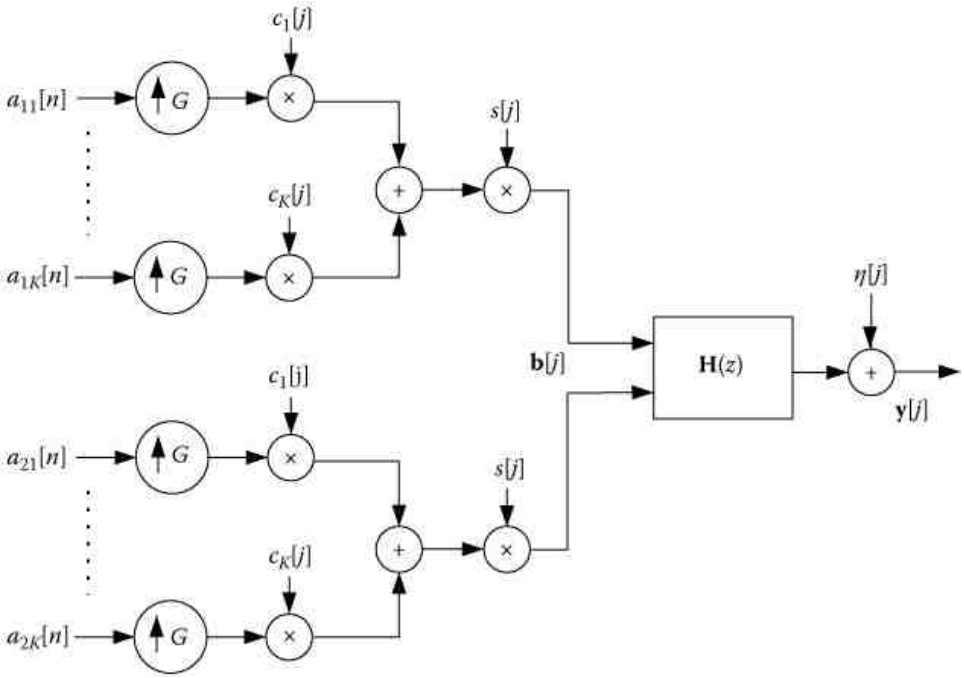

Figure 2.12: MIMO signal model without precoding.

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

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

where

(2.28) $$\mathbf{b}[j] = \sum_{k=1}^{K} s[j]c_k[j \mod \mathcal{G}]\mathbf{a}_k[n]$$

2.3.1 Chip-Level Equalization

When the scrambler is treated as deterministic, the desired signal contribution at the correlator output not only is 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 first consider the LMMSE chip-equalizer correlator receiver and then derive an analytical expression for the bias term and evaluate the SINR, including the explicit contribution of this quantity.

2.3.1.1 MMSE Chip-Equalizer Correlator Revisited

As before, 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 Figure 2.13 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 $2K$ × 1 and is given by $\mathbf{a}[n] = [\mathbf{a}_1^T[n] \cdots \mathbf{a}_K^T[n]]^T$. Here, $\mathbf{a}_k[n]$ represents the symbol on kth code. Figure 2.13 depicts the vectorized transmit signal model where $\mathbf{C}$ represents the $\mathcal{G} \times \mathcal{G}$ spreading matrix, and the diagonal matrix $\mathbf{S}[n]$ of the same dimension represents multiplication of the scrambling sequence for the nth symbol instant. Then, at the symbol level, the input-output relationship at the symbol can be compactly represented using the q operator as

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

where

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

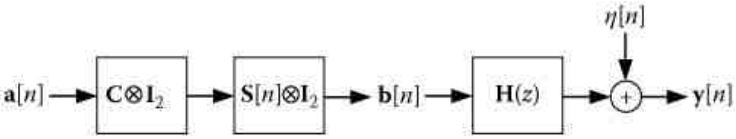

Figure 2.13: MIMO TX signal model.

This expression represents the convolution equation in a compact fashion because $q^{-m}\mathbf{a}[n] = \mathbf{a}[n - m]$. 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 $\mathbf{H}(z) = \Sigma_m z^{-m}\mathbf{H}[m]$ consists of $p\mathcal{G} \times \mathcal{G}$ matrix taps. If the delay spread is L chips, then there are $\lceil L/\mathcal{G} \rceil$ *pseudo-circulant* matrices that fully represent the channel. These matrices are defined as

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

with $\mathbf{h}[.]$ being the $2p$×2 chip-level MIMO channel coefficients. The LMMSE equalizer $\mathbf{F}(z)$ in Figure 2.14 can be represented in a similar fashion and visualized to be composed of $\mathbf{f}[.]$, which would be the 2 × $2p$ equalizer coefficients defined at the chip level. The channel-equalizer cascade is then given by

(2.31) $$\mathbf{G}(z) = \mathbf{F}(z)\mathbf{H}(z)$$

(2.32) $$= \sum_{\kappa=0}^{N-1} \mathbf{F}[\kappa] z^{-\kappa} \sum_{m=0}^{M-1} \mathbf{H}[m] z^{-m}$$

(2.33) $$= \sum_{v=0}^{N+M-2} \mathbf{G}[v] z^{-v}$$

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

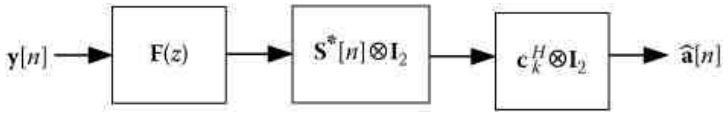

Figure 2.14: MIMO RX model.

Let the equalizer delay be d. Define the corresponding channel-equalizer cascade matrix at d as

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

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

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

Defining

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

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

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

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 is treated as deterministic). When the scrambler is treated as white, we refer to the 2 × 2 channel as a 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 submatrices $\mathbf{G}_{r\kappa}[m]$, for r, $\kappa \in \{1, 2\}$. A corresponding $\mathcal{G} \times \mathcal{G}$ matrix coefficient $\overline{\mathbf{G}}_{rk}[m] = \mathbf{G}_{rk}[m] - \mathbf{g}_{rk}[m] \cdot I_G$

is also defined and so is $\mathbf{g}_{r\kappa}[m]$, the $r\kappa$th element of the spatial channel $\mathbf{g}[m]$.

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

(2.37) $$\mathcal{R}_{zz} = \mathcal{R}_{des} + \underbrace{\mathcal{R}_{MUI} + \sum_{m} \mathcal{R}_{m,ISI} + \mathbf{F}\mathcal{R}_{\eta\eta}\mathbf{F}^{H}}_{\mathcal{R}_{\tilde{z}\tilde{z}}}$$

where

$$\mathcal{R}_{des} = \begin{bmatrix} |g_{11}[0]|^2 + |g_{12}[0]|^2 & \sum_{\kappa=1}^{2} g_{1\kappa}[0]g_{2\kappa}^{*}[0] \\ \sum_{\kappa=1}^{2} g_{2\kappa}[0]g_{1\kappa}^{*}[0] & |g_{21}[0]|^2 + |g_{22}[0]|^2 \end{bmatrix} +$$

$$\frac{1}{\mathcal{G}^2} \cdot \begin{bmatrix} \sum_{\kappa=1}^{2} tr\{\overline{\mathbf{G}}_{1\kappa}[0]\overline{\mathbf{G}}_{1\kappa}^{H}[0]\} & \sum_{\kappa=1}^{2} tr\{\overline{\mathbf{G}}_{1\kappa}[0]\overline{\mathbf{G}}_{2\kappa}^{H}[0]\} \\ \sum_{\kappa=1}^{2} tr\{\overline{\mathbf{G}}_{2\kappa}[0]\overline{\mathbf{G}}_{1\kappa}^{H}[0]\} & \sum_{\kappa=1}^{2} tr\{\overline{\mathbf{G}}_{2\kappa}[0]\overline{\mathbf{G}}_{2\kappa}^{H}[0]\} \end{bmatrix}$$

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

where the superscript $*$ represents complex conjugation. The ISI contribution from the mth symbol can be expressed as

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

In these relations, $\mathcal{R}_{des}$ 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 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

(2.38) $$\mathbf{z}_k[n] = \mathbf{B}_{n,k}[0]\mathbf{a}_k[n] - \tilde{\tilde{\mathbf{z}}}_k[n]$$

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 Equation (2.39):

(2.39) $SINR_{k,i}$

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

2.3.1.2 Chip-Level SIC Revisited

We now consider the effect of treatment of the 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 **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 time-varying, symbol-level channel to its mean value) take their toll; this SIC does not provide significant gains over the chip-equalizer correlator solution.

On the other end of the SIC spectrum is the chip-level SIC, discussed here, which is an entirely different solution that considers all stages of the channel-equalizer correlator stage as deterministic and recreates 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 under much improved conditions where interference from the first stream is (ideally) entirely suppressed.

2.3.2 Combined Chip-Level and Symbol-Level Equalization

We will now briefly discuss the effect of deterministic treatment of scrambler on further linear or non-linear symbol level processing stages when the receiver design is based on combined chip and symbol level equalization. For the spatial MMSE receiver, in order to claim the quantity $\frac{1}{\mathcal{G}^2} tr\{\overline{\mathbf{G}}_{rr}[0]\overline{\mathbf{G}}_{rr}^H[0]\}$ in Equation (2.39) as part of signal energy, it suffices to put in place time-varying processing at the correlator output, where the *n*th symbol vector on the *k*th code, $z_k[n]$ is given by Equation (2.36). 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. [In the case of the spatial-ML receiver, treating the scrambler as a random sequence, the spatial channel **B** in the ML metric is time-invariant.] A continuous processing matched filter bound can therefore be defined per stream. The *i*th 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 *n*th symbol considering all other symbols are known (correctly detected). We can, nevertheless, argue that deterministic treatment of the scrambler leads to reduced interference variance $\mathcal{R}_{\tilde{z}\tilde{z}}$ and increased recoverable signal power that will lead to performance improvement for the ML solution.

2.3.3 Numerical Results

As before, we compare the performance of 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 is comparable to the channel delay spread in chips. Figure 2.15 plots 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 a known sequence. In this case, 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 Figure 2.16 represents the *average* gains and not the true gain. The actual gain will be higher than that seen in Figure 2.16.

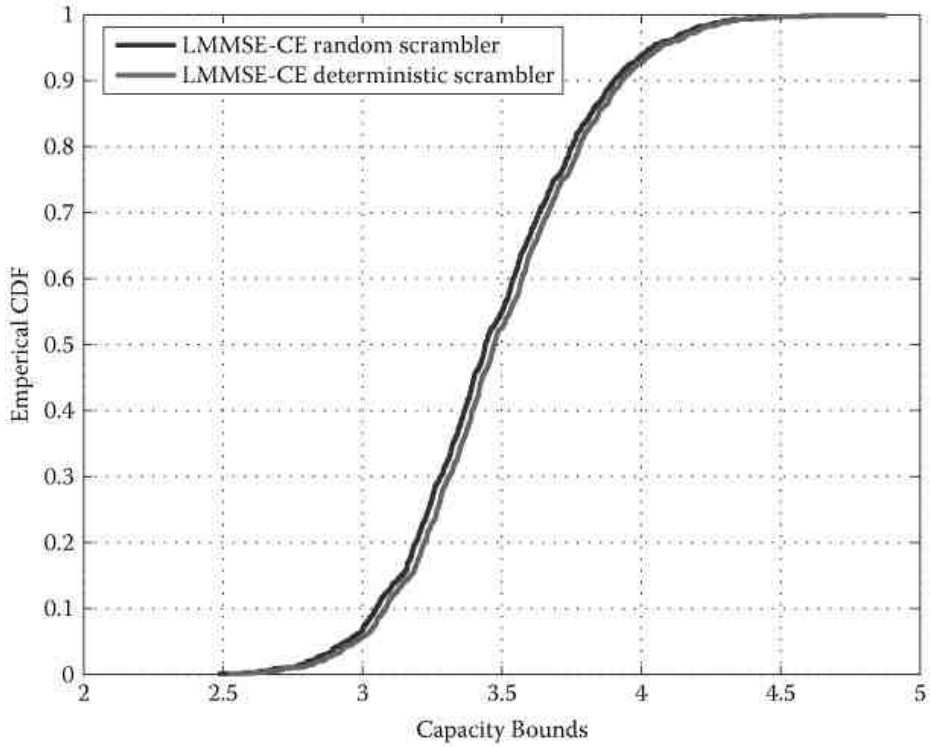

Figure 2.15: Performance of LMMSE chip-equalizer correlator with random and deterministic scrambler.

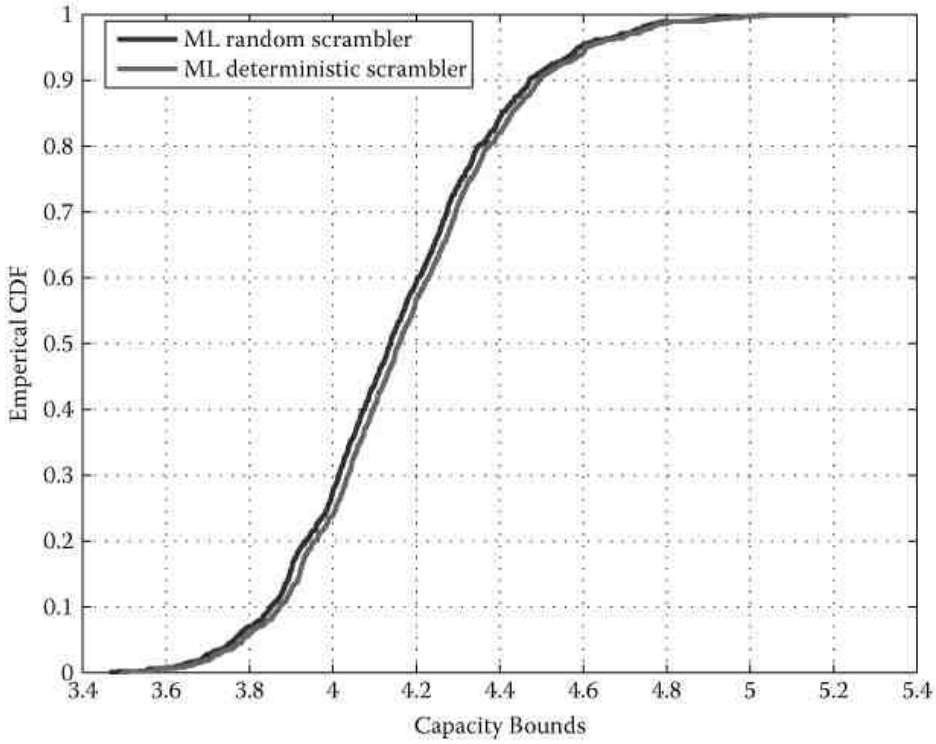

Figure 2.16: Sum-capacity at the output of spatial-ML receiver with deterministic and random scrambler.

2.4 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 spatial division multiple access (SDMA) and service multiple UEs in DL instead. In this case, the limitation of

two transmit antennas implies that a maximum of two 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 because 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 multi-user MIMO communication in the general case. It was studied previously in [18] and more recently in [7], where multi-user transmission techniques were classified into linear and nonlinear transmission algorithms. Nonlinear algorithms involving multi-user 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, such as multi-user beamforming, remain by far the most practical solution for multi-user transmission. Theoretical research in multi-user communications tends to consider frequency-flat channels: In reality, most mobile communication channels are frequency selective. There exists some literature on multi-user extension of HSDPA. In [11] the authors propose code reuse in D-TxAA based on a multi-user beamforming (MUB) scheme that 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 two transmit antennas, that is, $N_{tx} = 2$. 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 an 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 who request linearly independent weight vectors can, however, be served simultaneously.

2.4.1 Multi-User TxAA

We consider a 2-transmit, 1-receive antenna configuration for TxAA. For the remainder of this section, whenever we refer to a MU-TxAA system, we consider U separate 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 Figure 2.17, the transmit and beamformed chip sequence is given by

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

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 uth user, and finally, $a_{u,k}[n]$ 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

(2.41) $$\mathbf{y}_u = \mathbf{H}_u \mathbf{x} + \boldsymbol{\eta}$$

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

(2.42)
$$\mathbf{H}_u = \begin{bmatrix} \mathcal{H}_u^0 & \mathcal{H}_u^1 & \cdots & \mathcal{H}_u^{L-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathcal{H}_u^0 & \cdots & \cdots & \mathcal{H}_u^{L-1} & \vdots \\ \mathbf{0} & \mathbf{0} & \ddots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \ddots & \mathcal{H}_u^0 & \ddots & \mathcal{H}_u^{L-1} \end{bmatrix}$$

x is the transmit chip-vector formed by stacking *L* + *E* - 1 vectors and can be expressed as

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

and η is zero mean, circularly symmetric, Gaussian distributed, additive white noise of variance σ_η^2. In addition, we also define the *p*×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 *l*th beamformed channel tap of user *u*, due to the beamforming weight of another synchronous DL user *v*. We denote this by $\mathbf{R}_{u,v}$ and express this as

(2.44)
$$\mathbf{R}_{u,v} = \begin{bmatrix} \mathbf{r}_{u,v}^0 & \mathbf{r}_{u,v}^1 & \cdots & \mathbf{r}_{u,v}^{L-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{r}_{u,v}^0 & \cdots & \cdots & \mathbf{r}_{u,v}^{L-1} & \vdots \\ \mathbf{0} & \mathbf{0} & \ddots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \ddots & \mathbf{r}_{u,v}^0 & \ddots & \mathbf{r}_{u,v}^{L-1} \end{bmatrix}$$

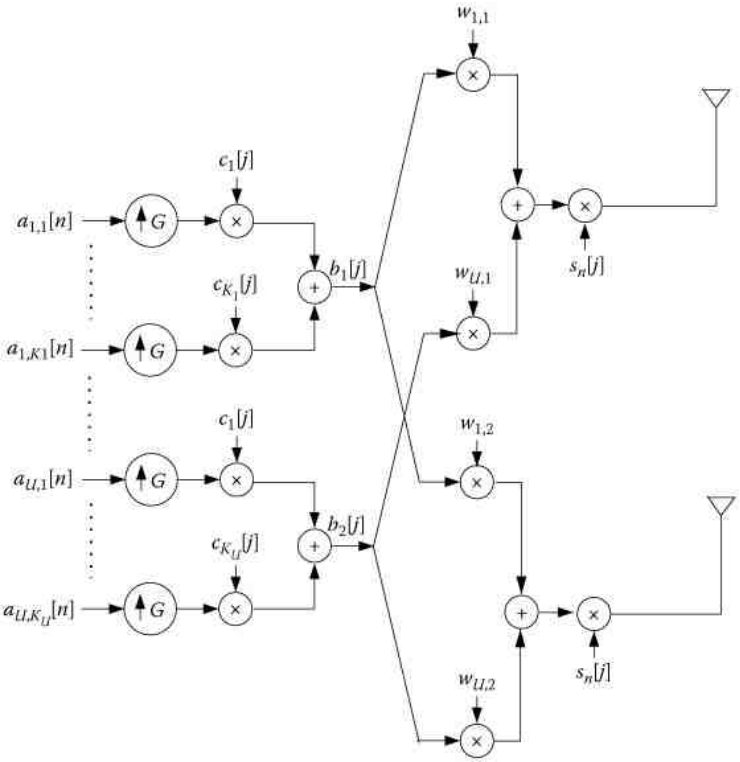

Figure 2.17: Multi-user TxAA transmit signal model.

2.4.1.1 Beamforming Strategies at the 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

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

which can be represented by

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

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

(2.47) $$\boldsymbol{\alpha}_{u,u} = \boldsymbol{\alpha}^{d}_{u,u} + \overline{\boldsymbol{\alpha}}_{u,u}$$

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

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

2.4.1.1.1 Simple Multi-User 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 beam-forming weights in the 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. Because 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 *u*th user, the SINR per-code $SINR_{k \in K_u}$ that is seen by the code assigned to the user is given by

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

where σ_k^2 denotes the chip variance of the *k*th 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 compute the ideal beamforming weight vector under this assumption. The BS, however, makes no attempt to group users with the same beamforming weights. As a result, it is expected that the downlink capacity drops significantly.

2.4.1.1.2 Weight Optimization by Average Interference Criterion

Alternatively, a UE can anticipate that, in reality, any of the four weights may be chosen by the other users in the DL. Assuming that other users choose one of four beamforming weights with equal likelihood, it is reasonable to choose the beamforming weight which has the maximum SINR when averaged over all four hypotheses for the other users' weights. Each UE therefore computes the ideal beamforming weight by plugging into Equation (2.49) all possible combinations of weight vectors and feeds back the weight vector with the best average SINR over all the hypotheses 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 the expected SINR, the resulting

SINR is higher than obtained by assuming that the same beamforming weight is requested by all users scheduled in the DL. Thus, this beamforming vector must perform better on average, and increase the average data rate per user when compared to the simple multi-user beamforming case.

2.4.1.1.3 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 could 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 Equation 2.49 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 the BS to maximize system capacity. In practice, this involves a considerable amount of receiver processing and also a lot of feedback to the BS. Nonetheless, the gains in such a case are worth investigating.

2.4.1.1.4 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 *u*th user is then given by

(2.50)
$$\frac{\sigma_k^2 |\alpha_{u,u}^d|^2}{\frac{K}{G}\sigma_k^2 \|\overline{\boldsymbol{\alpha}}_{u,u}\|^2 + \sigma_\eta^2 \mathbf{f}_u \mathbf{f}_u^H}$$

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.

2.4.2 Multi-User D-TxAA

For the MU-D-TxAA system, we consider two separate UEs with N_{rx} receive antennas each. In an MU-D-TxAA system, the BS transmits two transport blocks for as many users scheduled in the DL. All codes of a single stream are assigned to one user and reused across the two streams. From Figure 2.18 we see that the transmit signal vector in downlink can be modeled as

(2.51)
$$\mathbf{x}[j] = \underbrace{\mathbf{W}}_{2\times 2}\mathbf{b}[j] = \mathbf{W} \cdot \sum_{k=1}^{K} s[j]c_k[j \mod G]\mathbf{a}_k[n]$$

$\mathbf{W} = [\mathbf{w}_1\ \mathbf{w}_2]$ is the 2 × 2 unitary precoding matrix. The columns of **W** are composed 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 $s[j]$.

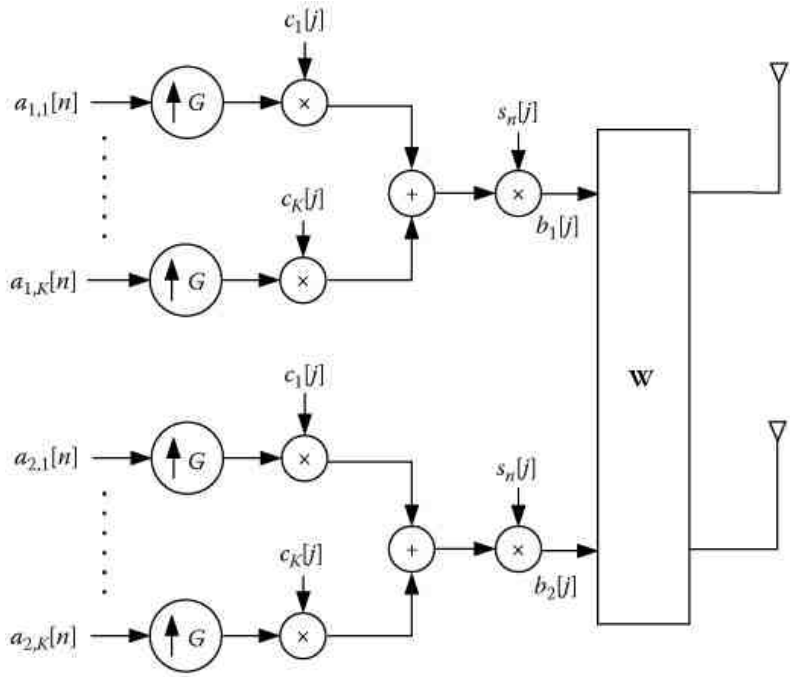

Figure 2.18: Multi-user D-TxAA transmit signal model.

2.4.2.1 Spatial Multiplexing versus SDMA

In the spatial multiplexing context, there is only a single user in the 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:

(2.52) $$\widehat{\mathbf{x}}[j] = \mathbf{x}[j] - \widetilde{\mathbf{x}}[j]$$

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 Equation (2.52), 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 *q*th stream at the output of the output of the LMMSE chip equalizer/correlator is given in [16] as

(2.53) $$SINR_q = \frac{\sigma_a^2}{\left(\mathbf{W}^H \mathcal{R}_{\widetilde{z}\widetilde{z}} \mathbf{W}\right)_{qq}} - 1$$

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 **W** whose columns correspond to the weight vectors fed back by the two users. It is obvious that two users who feedback the same weightvector 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 **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 Equation (2.53).

2.4.3 Numerical Results

In this section we present Monte-Carlo simulation results and performance comparison of different beamforming strategies

proposed in this chapter. 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, the BS simultaneously serves two 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 normalized to 1. The cumulative distribution function (CDF) 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.

2.4.3.1 TxAA

Each UE is assumed to have a single receive antenna. Normally, each UE feeds back only its preferred weight vector index, only in the 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.

2.4.3.2 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 two-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. Figure 2.19 compares the sum-capacity in the DL for the case of TxAA. The DL capacity 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 because 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 chosen based solely on the channel seen by each user, but also 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 multi-user 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 because in this case each user is able to effectively mitigate MUI due to the same beamformed channel seen by all codes in the downlink. It should be noted that for the case where the total number of users in DL far exceeds the number of users actually scheduled in the DL, the performance of cooperative-operative beamforming is expected to improve. In Figure 2.20 we compare the performance of D-TxAA in spatial multiplexing mode with that of the multi-user (SDMA) mode. Simulation results show that the DL sum-capacity is greater for the case of SDMA with single stream transmission to both users.

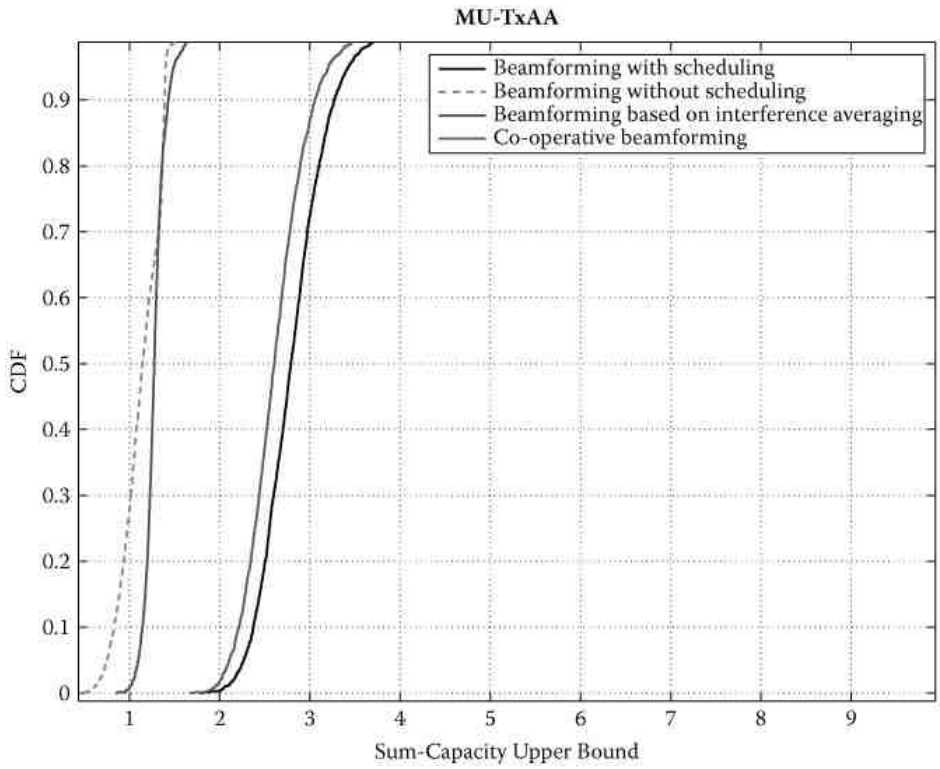

Figure 2.19: Performance of different beamforming schemes for MU-TxAA.

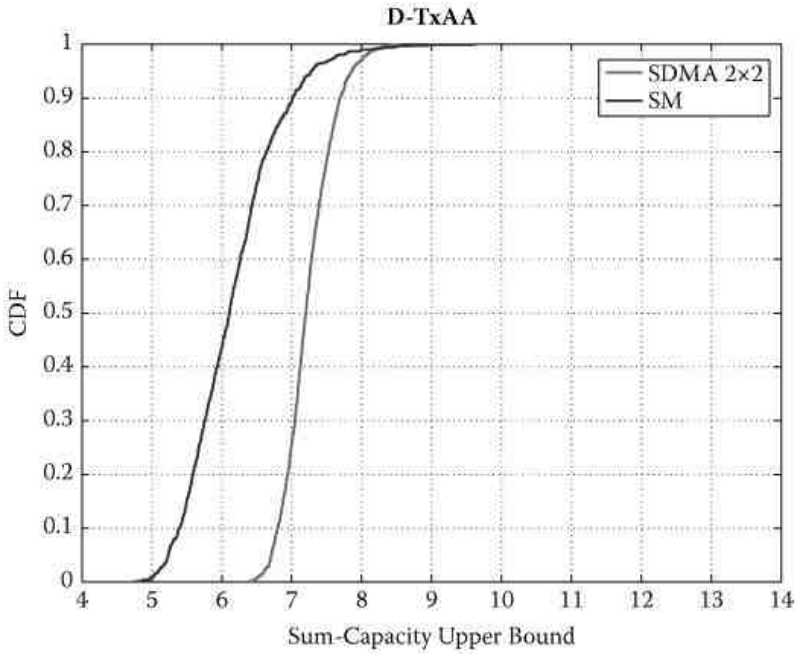

Figure 2.20: DL sum-capacity for MU-D-TxAA.

2.5 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 the symbol level were 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 optimal precoding matrix. The receiver designs discussed in the first part of the chapter operated on the assumption that the scrambler used at the transmitter can be modeled as a random sequence. Because such a treatment contributes to the 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 regain the time-varying signal contribution that would otherwise be treated as noise. This leads to additional gains in SINR, which ultimately affects the 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] A. Baştuğ and D.T.M. Slock. Downlink WCDMA receivers based on combined chip and symbol level equalization. *European Trans. on Telecoms.*, 16: 51–63, 2005.

[3] J. Choi, S.R. Kim, Y. Wang, and C.C. Lim. Receivers for chip-level decision feedback equalizer for CDMA downlink channels. *IEEE Trans. on Wireless Comm.*, 3(1): 300–313, 2004.

[4] J.M. Cioffi, G.P. Dudevoir, M.V. Eyuboglu, and G.D. Forney. MMSE Decision-feedback equalizers and coding. 1. Equalization results. *IEEE Trans. on Comm.*, 43(10): 2582–2594, 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] D. Gesbert, M. Kountouris, R. Heath, C.B.W. Chae, and T. Slzer. Shifting the MIMO paradigm. *IEEE Signal Processing Mag.*, 24(5): 36–46, September 2007.

[8] I. Ghauri, S.P. Shenoy, and D.T.M. Slock. On LMMSE bias in CDMA SIMO/MIMO receivers. In *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, Las Vegas, NV, March 2008.

[9] I. Ghauri and D.T.M. Slock. Linear receivers for the DS-CDMA downlink exploiting orthogonality of spreading sequences. In *Proc. 32nd Asilomar Conf. Signals, Systems & Computers*, 1: 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*, Montreal, Quebec, Canada, 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 Conf.*, Montreal, Quebec, Canada, Fall 2006.

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

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

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

[16] S.P. Shenoy, I. Ghauri, and D.T.M. Slock. Optimal precoding and MMSE receiver designs for MIMO WCDMA. In *Proc. IEEE 67th Vehicular Technol. Conf.*, Singapore, May 2008.

[17] D.K.C. So and R.S. Cheng. Detection techniques for V-BLAST in frequency selective fading channels. *Wireless Commun. Networking Cont.*, 1(17–21): 487–491, March 2002.

[18] Q.H. Spencer, C.B. Peel, A.L. Swindlehurst, and M. Haardt. An introduction to multi-user MIMO downlink. *IEEE Commun. Mag.*, 42(10): 60–67, October 2004.

[19] I.E. Telatar. Capacity of multi-antenna gaussian channels. *Eur. Trans. Telecommun.*, 10: 585–595, 1999.

[20] M. Wrulich, C. Mehlführer, and M. Rupp. Interference aware MMSE equalization for MIMO TxAA. In *Proc. 2008 3rd Int. Symp. Commun., Control and Signal Processing (ISCCSP2008)*, St. Julians, Malta, March 2008.

[21] J. (Charlie) Zhang, B. Raghothaman, Y. Wang, and G. Mandyam. Receivers and CQI measures for MIMO-CDMA systems in frequency-selective channels. *EURASIP J. Appl. Signal Processing*, (11): 1668–1679, November 2005.