3 SYSTEM MODEL AND UNCODED PERFORMANCE EVALUATION

It was shown that factors such as the scattering environment, the user distribution and the fading environment, amongst others, play important roles in determining the performance of space-time mobile systems. In this chapter three smart antenna techniques, namely beamforming, transmit and receive diversity, are considered in more detail to mitigate the effects due to the inherent impairments of the mobile communication channel. These techniques are especially applicable to mobile CDMA communications, where interference is more pronounced due to the fact that users transmit on the same frequency with unique spreading codes.

The aim of a well-designed mobile communication system is to share the common transmission medium in such a manner that

- the average overall amount of transmitted information will be as large as possible,
- the average probability of error experienced by each user will be as low as possible, and
- the average delay will be as low as possible.

In general, not all of these goals can be achieved simultaneously, and the design process involves a trade-off between these objectives. By making use of coding and smart antenna techniques, it is possible to approach these goals. For instance, beamforming can be used to decrease a system's probability of error by reducing CDMA interference. This can be achieved by intelligent combination of the received signals by multiple antenna elements at the base station or mobile. In a mobile communication system with antenna arrays, the fast fading signal component introduces a random phase and amplitude to the received signal on each antenna element, which perturbs the steering vector of the array. In the case of Rayleigh or Nakagami fading, the phase can take on any value between $(0, 2\pi]$, and the DOA of the waves may be impossible to determine from short-duration observations of the received signal. For this reason, in a fading environment it may not be useful to implement the beamformer to create lobes and nulls toward desired and interfering sources. When the fast fading is highly correlated between the elements, it may be considered as a single scalar which multiplies with the steering vector, affecting all elements equally. On the other hand, no receive

diversity gain can be obtained, since receive diversity relies on uncorrelated fading. The correlation between elements decreases with element spacing and changes according to the scattering environment in which the system operates. There is therefore a conflict between the avoidance of grating lobes and the need for receive diversity gain [127].

3.1 BACKGROUND

Thus far in this thesis, it has been shown that a variety of factors influence the performance of cellular systems incorporating smart antenna techniques. These factors have been combined into an easy to use channel model that makes the analytic evaluation of a space-time smart antenna based cellular CDMA systems possible. In this chapter, the channel model developed in Chapter 2 will be used to evaluate the performance of an uncoded cellular CDMA system.

In order to simplify the analysis of space-time processing systems, a basic model of the communication system which identifies inputs, outputs and the channel is required. For a general space-time processing system where multiple antennas are employed at both the transmitter and the receiver, such a signal model is known as a multiple-input/multiple-output (MIMO) model. Clearly this is due to the fact that the desired signal has multiple inputs into the channel (the transmit antennas) as well as multiple outputs (the receive antennas). Furthermore, a MIMO system can be viewed as multiple single-input/single-output (SISO) sub-channels. The MIMO system's channel capacity is then the sum of the individual capacities of these sub-channels. Fading correlation effects as described in Chapter 2, affects the MIMO system capacity by modifying the distributions of the gains characterizing the SISO sub-channels. As the general MIMO case is not frequently used in practice, a number of alternative channel configurations for single user (SU) and multi user (MU) scenarios are considered. These are

- Uplink and downlink (single-input/single-output - SISO)
  - SU-SISO: Single user with single antenna input/output at the base station and single antenna input/output at the mobile.
  - MU-SISO: Multiuser with single antenna input/output at the base station and single antenna input/output at each of the mobile units.
- Downlink (multi-input/single-output - MISO)
  - SU-MISO: Single user with multiple antenna inputs at the base station and single antenna output at the mobile unit.
  - MU-MISO: Multi user with multiple antenna composite inputs at the base station and single antenna output at each mobile.
- Uplink (single-input/multi-output - SIMO)
  - SU-SIMO: Single user with single antenna input at the mobile and multiple antenna outputs at the base station.
  - MU-SIMO: Multi user with single antenna input at each mobile and multiple antenna composite outputs at the base station.

This research looks at ways and means of using multiple antennas at the BSS transmitter (downlink) and the BSS receiver (uplink) to make the recovery of the transmitted data more reliable. Specifically, the use

of coding in conjunction with transmit and receive diversity combining (space-time coding), and receive beamforming arrays (coded space-time) are considered. The BEP performance analysis presented in this chapter is restricted to the performance of the uncoded space-time CDMA system. In the following chapters, these results will be extended to include the performance of the CDMA systems incorporating space-time coding techniques.

Conventional detection in a single path transmission environment is done by matched filtering and sampling of the received signal, followed by a decision device, e.g., a simple polarity check for BPSK or QPSK. The received signal is matched to the spreading code of the desired user. In a single user environment, this is optimal in the sense that the SNR is maximized which in turn corresponds to ML detection. In a multiuser environment, this is, however, not entirely true. The SNR is still maximized, but the detector is not ML due to the presence of MAI.

In a multi-path environment, the decision statistics for each multi-path component are obtained. As has been discussed in Chapter 1 there are then several strategies for receive diversity combining of these decision statistics pertaining to the same bit. Combining can, of course, also be done after individual detection of each multi-path component which does in fact provide potentially better performance. This is, however, not as common as pre-detection or RAKE combining. For this reason the receiver is restricted to a matched filter front-end followed by RAKE multipath diversity combiner.

In general, adaptive receiver structures, for instance minimum mean-square error detectors, may also be considered. With these receivers the focus is directed towards interference cancellation. In the most general terms, interference cancellation detectors have structures where an explicit estimate of the MAI component in the received signal is generated and then subtracted from the received signal in an iterative manner. In this thesis, the application of interference cancellation, and multiuser detectors in general, has not been considered. It is proposed to be addressed in future research.

Figure 3.1 depicts the basic structure of an M_D-branch diversity system, with

$$\begin{aligned} M_D &= 1 \text{ for a SISO system,} \\ M_D &= M_T \text{ for a MISO (transmit diversity) system,} \\ M_D &= M_R \text{ for a SIMO (receive diversity) system,} \\ M_D &= M_T \times M_R \text{ for a MIMO (combined transmit/receive diversity) system.} \end{aligned}$$

With reference to Figure 3.1, let the signal in the nth channel diversity system have power Ω_n, Nakagami fading parameter m_n, and correlation between the nth and $(n+1)$th branch be $\rho_{n(n+1)}$. It is assumed that each channel is frequency non-selective with channel attenuation factors $\{\beta_n\}, n = 1, \cdots, M_D$ having Nakagami-distributed envelope statistics. Making use of (2.12), the fading distribution of each diversity branch may be written as

$$p_{\beta_n}(\beta_n) = \frac{2}{\Gamma(m_n)} \left(\frac{m_n}{\Omega_n}\right)^{m_n} \beta_n^{2m_n-1} e^{-m_n\beta_n^2/\Omega_n}, \tag{3.1}$$

with $\Omega_n = E\{S_n\} = E\{\beta_n^2\}$. The fading *power*, $S_n = \beta_n^2$, therefore has a gamma distribution [147].

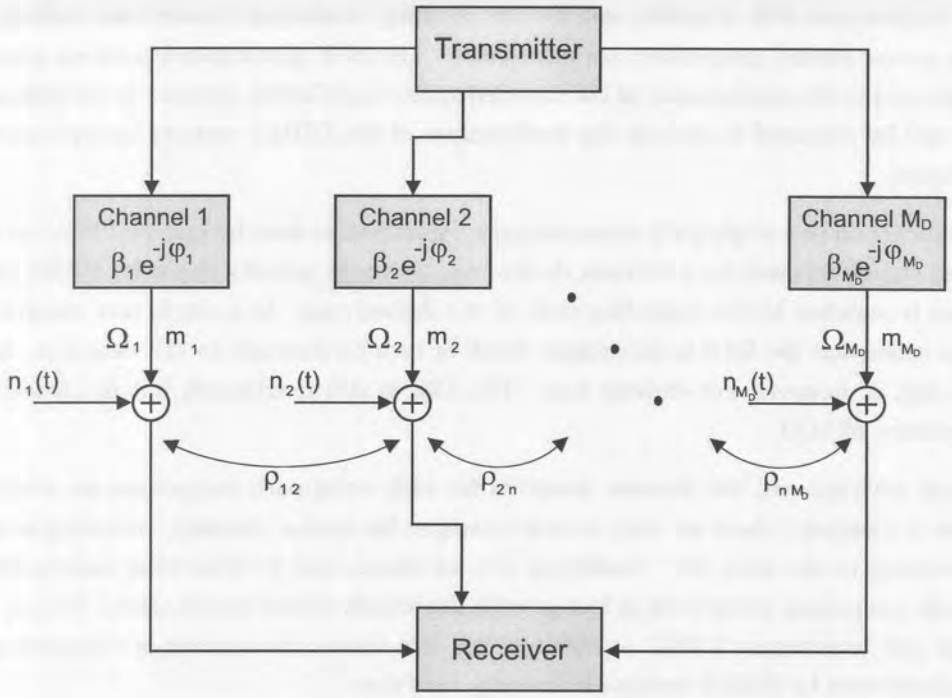

Figure 3.1. Basic structure of a diversity system.

3.2 MULTIPLE TRANSMIT/RECEIVE ANTENNA SYSTEM MODELS

In this section, the different transmit and receive diversity, and beamforming techniques for cellular CDMA are considered.

3.2.1 Transmit Diversity

Two transmit diversity techniques, namely code-division and time-division are proposed for the downlink. As with its receive diversity counterpart, transmit diversity techniques also require a number of uncorrelated diversity branches carrying the same information and a circuit to combine the received signals or to select one of them.

3.2.1.1 Code-Division Transmit Diversity (CDTD). In CDMA it is desirable to transmit orthogonal signals to different users in the downlink and to simultaneously maintain a fixed number of user channels. Clearly both these requirements cannot be met, since the number of available orthogonal channels is fixed. For this reason two main approaches are followed in transmit diversity for cellular CDMA. These are

Orthogonal CDTD (O-CDTD) [63, 148].
For O-CDTD, different complex orthogonal spreading codes are assigned to every antenna element. This maintains the orthogonality between the two output streams, and hence self-interference is eliminated in flat fading. O-CDTD should also be compared with orthogonal transmit diversity (OTD) [149, 150] and Alamouti code transmit diversity (ACTD) [151] proposed for narrowband TDMA.

Non-orthogonal CDTD (NO-CDTD) [54, 55, 152, 153].
For NO-CDTD, the same complex spreading code is assigned to every antenna with an intentional delay

between each antenna element. For this reason NO-CDTD is also known as delayed CDTD, similar to the delayed transmit diversity (DTD) scheme proposed for narrowband TDMA [54, 55, 56]. Typical non-orthogonal spreading sequences used for NO-CDTD are Gold sequences. The advantage of Gold sequences are that they maximize the number of spreading sequences, but compromises orthogonality due to self-interference.

Figure 3.2(a) illustrates the CDTD structure for a single user in the downlink. Both O-CDTD and NO-CDTD have the receive diversity property of soft-failure[1].

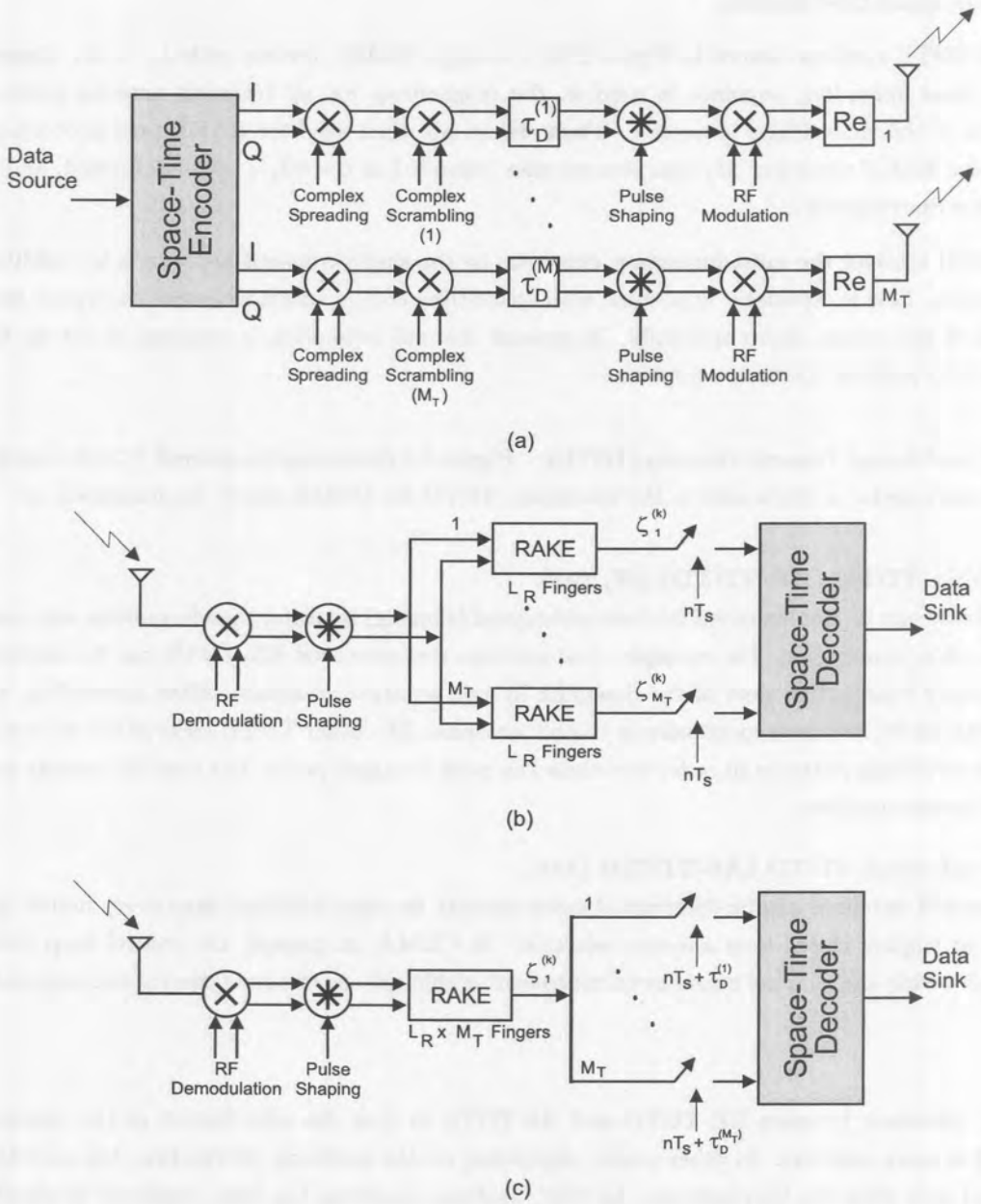

Figure 3.2. Block diagram of single user CDTD system. (a) Transmitter for O-CDTD and NO-CDTD, (b) Receiver for O-CDTD, and (c) Receiver for NO-CDTD.

The main components are the spaced-time encoder, complex spreader, modulator and transmit antennas. Data modulation per antenna is QPSK, where the spaced-time encoded downlink data is mapped to the I and Q branches. The data is spread by a combination of complex (orthogonal and non-orthogonal) variable spreading and complex scrambling codes. In both O-CDTD and NO-CDTD the symbol rate on each transmit antenna M_T is reduced by a factor $1/M_T$ to ensure that the data bits are evenly distributed to each transmit antenna element.

The general receiver structures for O-CDTD and NO-CDTD are shown in Figure 3.2(b) and 3.2(c), respectively. With reference to Figure 3.2(b), a total of M_T RAKE receivers, each with L_R fingers, are employed for O-CDTD. Each of the M_T RAKE receivers is trained on the spreading sequence associated with the corresponding transmit antenna. The M_T complex outputs are then sampled at the symbol rate, T_s, and passed to the space-time decoder.

In the NO-CDTD receiver, shown in Figure 3.2(c), a single RAKE receiver with $L_R \times M_T$ fingers is used. Since the same spreading sequence is used at the transmitter for all transmit antenna paths, *a priori* information of the time delays is needed. These delays are then used for RAKE post-processing. At the output of the RAKE combiner M_T complex samples (sampled at $t = nT_s + \tau_D^m$) are formed, and processed by the space-time decoder.

At the CDTD receiver the most important extension to the single transmit antenna is the addition of M_T RAKE fingers. This is especially important with space-time coding which attempts to exploit the degrees-of-freedom of the system more optimally. In general channel estimation is required to set up the RAKE receiver and to perform diversity reception.

3.2.1.2 Time-Division Transmit Diversity (TDTD). Figure 3.3 illustrates the general TDTD transmitter and receiver structures for a single user in the downlink. TDTD for CDMA can be implemented as

Round-robin TDTD (RR-TDTD) [57, 148].
This scheme can be implemented by time-orthogonal (sharing) by using pseudo-random antenna hopping (round-robin) sequencing. For example, dual antenna time-switched RR-TDTD can be implemented by transmitting consecutive slots of the downlink by two separate antennas. After scrambling, the spread time slots can be switched consecutively to each antenna. The other TDTD users of the system may have different switching patterns in order to reduce the peak transmit power and peak to average power ratio in each power amplifier.

Antenna selection TDTD (AS-TDTD) [148].
The transmit antennas can be determined more optimal, by using feedback from every mobile to the base station to employ closed-loop antenna selection. In CDMA, in general, the control loop delay can be kept well within the channel coherence time to enable efficient use of power control and antenna selection loops.

The main difference between RR-TDTD and AS-TDTD is that the distribution of the encoded bits in AS-TDTD is more selective. In other words, depending on the feedback information, the encoded bits are transmitted only from the best antenna. In [148], feedback signaling has been employed to simultaneously perform fast closed loop power control for downlink diversity.

The drawbacks associated with TDTD are that

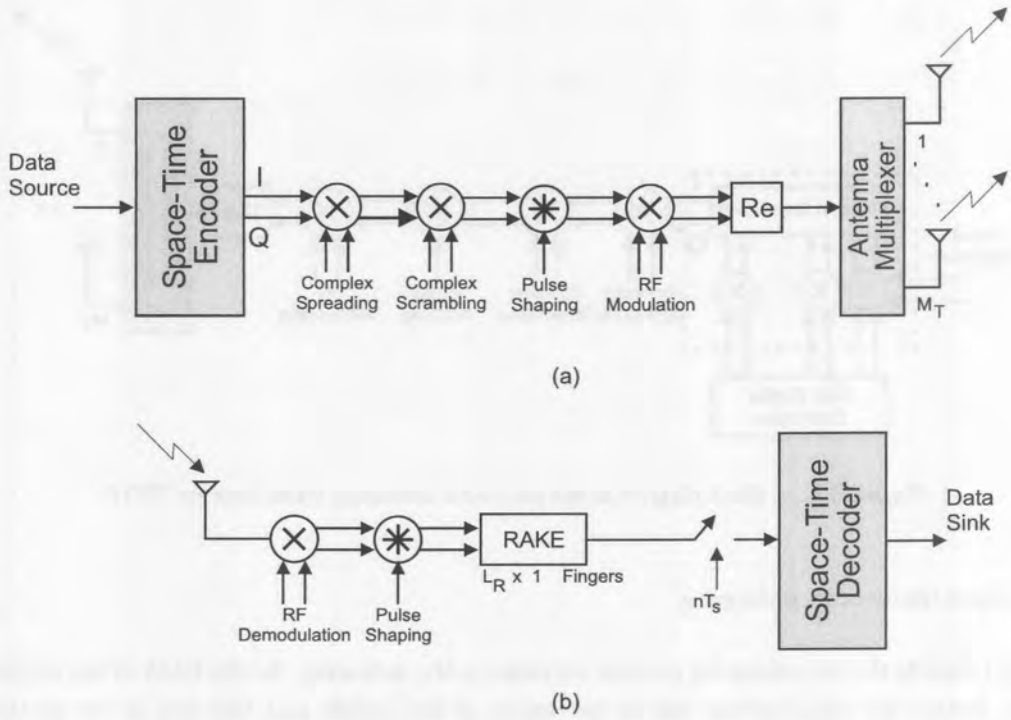

Figure 3.3. Block diagram of single user TDTD system. (a) Transmitter, and (b) Receiver.

- TDTD does not provide soft-failure, and

- the channel has to be re-estimated on a slot-by-slot basis.

3.2.1.3 CDTD and TDTD with Pre-RAKE Combining. As a interesting extension to both the CDTD and TDTD diversity schemes presented in the foregoing, a pre-RAKE configuration may be considered. Figure 3.4 illustrates a general pre-RAKE combining TDTD transmitter. This transmitter is based on the CDMA pre-RAKE combining strategy by Jeong *et al.* [154]. The principle of operation is that the transmitted pre-RAKE signal is a time-reversed replica of the channel impulse response. In this way space and path diversity is possible at the mobile receiver without any conventional receive diversity techniques. The transmitter is based on two diversity principles, pre-RAKE (realized by the tapped delay line) and space diversity (realized by the multiple transmit antennas). A similar extension to CDTD is possible.

3.2.2 Receive Diversity

As has been argued in Chapter 1, space diversity reception in the uplink is one of the effective and, hence, widely applied techniques for mitigating the effects of multipath fading. The classical approach is to use multiple antennae at the receiver and perform combining (i.e., MRC or EGC) or selection (i.e., SC) and switching in order to improve the received signal quality. The receive diversity system employing M_R antennas is shown in Figure 3.5(a) (transmitter) and 3.5(b) (receiver), respectively.

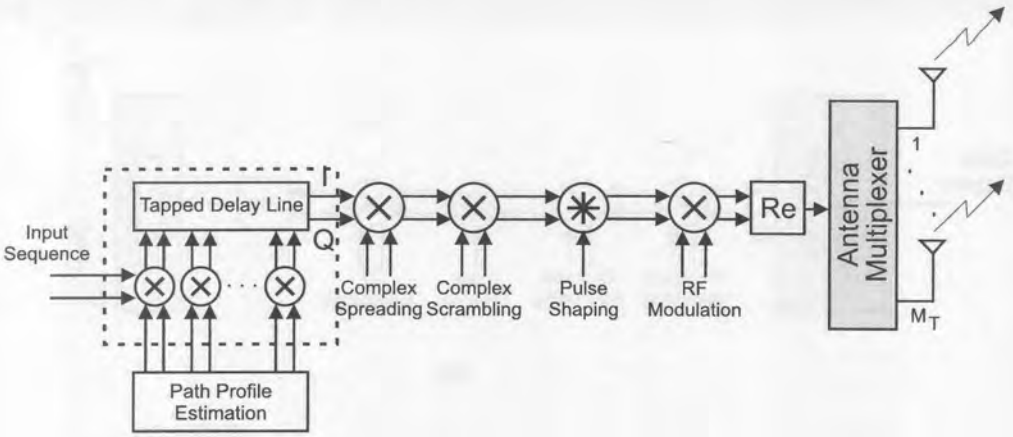

Figure 3.4. Block diagram of the pre-RAKE combining transmitter for TDTD.

3.2.3 Transmit/Receive Beamforming

Figure 3.5(c) depicts the beamforming receiver employing M_D antennas. As the DOA of the received signals at the base station are time-varying due to movement of the mobile and also due to the scattering environment, the co-channel interference, multipath components and even Doppler frequency are time varying functions. By using a beamformer, it is possible to separate signals co-located in frequency, but separated in the spatial domain and to track these time varying signals. In a CDMA system, this specifically results in reducing the interference from unwanted signals by optimizing the array pattern through the adjustment of the weights of the array according to some criteria, or cost function.

3.3 PERFORMANCE EVALUATION

The BEP performance derivation for uncoded space-time beamforming and transmit/receive diversity are considered in this section.

3.3.1 Receive Beamforming Performance

Making extensive use of the analysis carried out by Lötter [126], the BEP performance of a uniform linear array (ULA) beamformer, with M_B antenna elements is considered. Consider Figure 3.5(a), where the output of the transmitter of user k can be written as

$$s^{(k)}(t) = \sqrt{2P}a^{(k)}(t)\, b^{(k)}(t)\, \cos(\omega_c t + \phi^{(k)}), \tag{3.2}$$

where P denotes the average transmitted signal power, $b^{(k)}(t)$ denotes binary data with symbol period T_s seconds and values taken from the set $\{\pm 1\}$, $a^{(k)}(t)$ denotes a random binary spreading sequence with chip period T_c seconds and length $N = T_s/T_c$ [2] with value taken from the set $\{\pm 1\}$. N is the spreading sequence length. In addition, standard BPSK [3] modulation is used with carrier frequency ω_c rad/s and unknown carrier phase $\phi^{(k)}$, a random variable uniformly distributed over $[0, 2\pi)$. The transmitted signal propagates over a radio channel modeled as a Nakagami fading, time invariant, discrete multipath channel with equivalent lowpass response

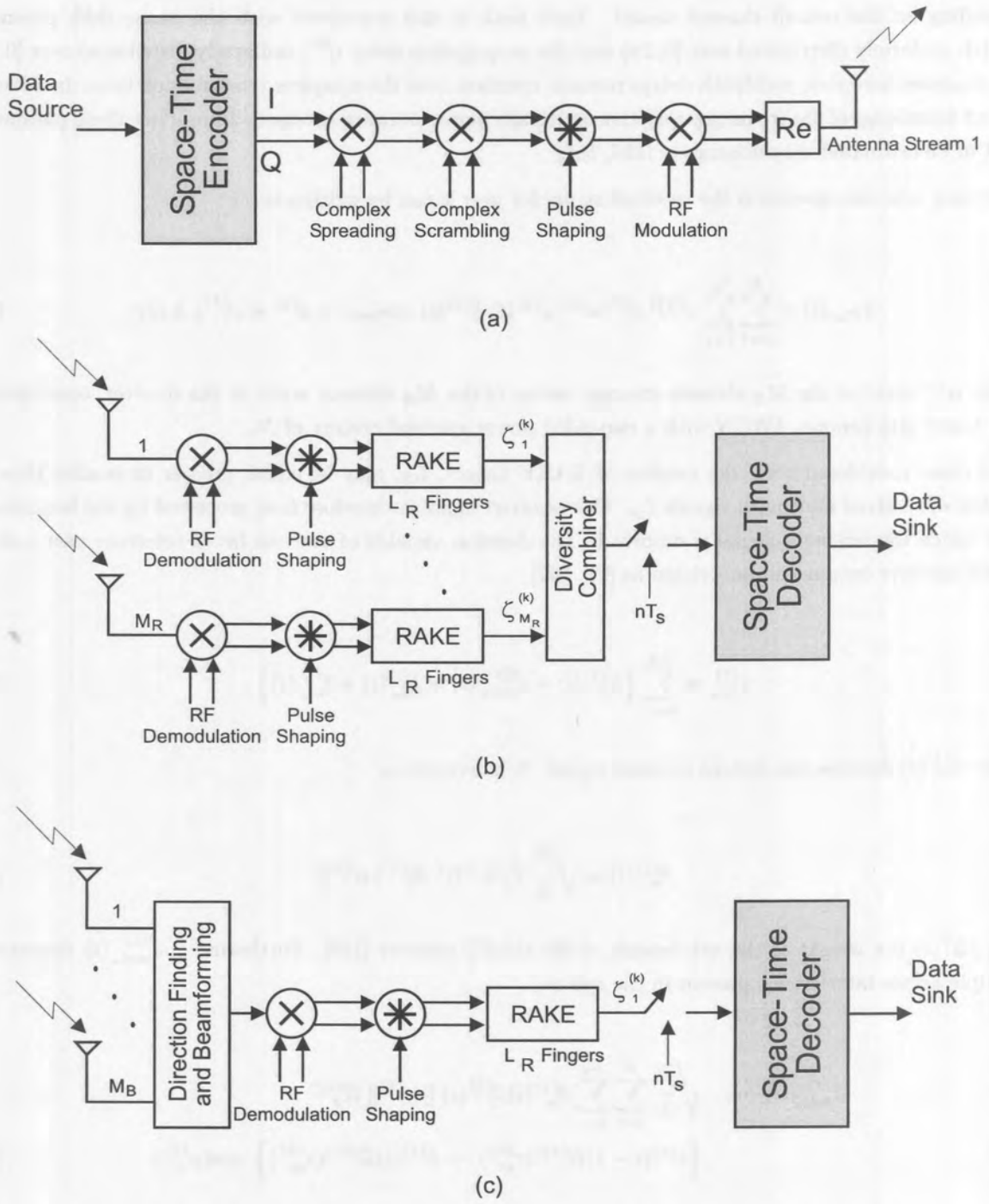

Figure 3.5. Block diagram of receive diversity and beamforming system. (a) Transmitter, (b) Diversity receiver, and (c) Beamforming receiver.

$$h^{(k)}(\tau) = \sum_{l=1}^{L_p} \beta_l^{(k)} \exp\left(j\varphi_l^{(k)}\right) \delta[\tau - \tau_l^{(k)}]. \tag{3.3}$$

Each path is characterized by the random variables $\beta_l^{(k)}(i)$ denoting the strength of path l from user k at symbol interval i. Each of these can be modeled as either Rayleigh distributed or Nakagami-m distributed,

depending on the overall channel model. Each path is also associated with the phase shift parameter $\varphi_l^{(k)}(i)$, uniformly distributed over $[0, 2\pi)$ and the propagation delay $\tau_l^{(k)}$, uniformly distributed over $[0, T_s)$. It is assumed here that multipath delays remains constant over the complete transmission time. In addition, perfect knowledge of the frequency selective multipath parameters are assumed. In practice these parameters need to be estimated very accurately [155, 156].

Assuming coherent operation the received signal for user k can be written as

$$r_{M_B}(t) = \sum_{k=1}^{K} \sum_{l=1}^{L_p} \sqrt{2P}\, \beta_l^{(k)} \mathbf{w}^{(k)}\, a^{(k)}(t)\, b^{(k)}(t)\, \cos(\omega_c t + \phi^{(k)} + \varphi_l^{(k)}) + \eta(t), \tag{3.4}$$

where $\mathbf{w}^{(k)}$ denotes the M_B element steering vector of the M_B element array at the receiver, optimized for user k and $\eta(t)$ denotes AWGN with a two sided power spectral density of N_0.

In all cases considered here, the number of RAKE fingers, L_R, may be equal, greater or smaller than the number of received multipath signals L_p. The received signal is therefore first processed by the beamformer after which the reference signal is despread. The decision variable of the i-th bit of reference user j at the RAKE receiver output can be written as [91, 157]

$$\zeta_{M_B}^{(j)} = \sum_{n=1}^{L_R} \left\{ S_n^{(j)}(i) + I_{\mathrm{mai}_n}^{(j)}(i) + I_{\mathrm{si}_n}^{(j)}(i) + I_{\mathrm{ni}_n}^{(j)}(i) \right\}, \tag{3.5}$$

where $S_n^{(j)}(i)$ denotes the desired received signal. It is written as

$$S_n^{(j)}(i) = \sqrt{\frac{P}{2}}\, T_s\, b^{(j)}(i)\, \beta_n^{(j)^2} \|\mathbf{w}^{(j)}\| \tag{3.6}$$

with $\beta_n^{(j)}(i)$ the weight of the nth branch of the RAKE receiver [129]. Furthermore, $I_{\mathrm{mai}_n}^{(j)}(i)$ denotes the multiple access interference present in the cell or,

$$\begin{aligned} I_{\mathrm{mai}_n}^{(j)}(i) \;=\; & \sqrt{\frac{P}{2}} \sum_{k=2}^{K} \sum_{l=1}^{L_p} \beta_n^{(j)}(i) \beta_l^{(k)}(i)\, \|\mathbf{w}^{(k)}\|\, \mathcal{R}_x^{(kj)} \cdot \\ & \left\{ b^{(k)}(i-1) R^{(kj)}(\tau_{nl}^{(k)}) \;+\; b^{(k)}(i) \hat{R}^{(kj)}(\tau_{nl}^{(k)}) \right\} \cos(\varphi_{nl}^{(k)}). \end{aligned} \tag{3.7}$$

$I_{\mathrm{si}_n}^{(j)}(i)$ denotes the self-interference present in the cell,

$$\begin{aligned} I_{\mathrm{si}_n}^{(j)}(i) \;=\; & \sqrt{\frac{P}{2}} \sum_{l=1}^{L_p} \beta_n^{(j)}(i) \beta_l^{(j)}(i)\, \|\mathbf{w}^{(j)}\|\, \mathcal{R}_x^{(jj)} \cdot \\ & \left\{ b^{(j)}(i-1) R^{(jj)}(\tau_{nl}^{(j)}) \;+\; b^{(j)}(i) \hat{R}^{(jj)}(\tau_{nl}^{(j)}) \right\} \cos(\varphi_{nl}^{(j)}). \end{aligned} \tag{3.8}$$

$I_{\mathrm{ni}_n}^{(j)}(i)$ denotes the AWGN interference, with variance $\sigma^2 = N_o T/2$. In (3.7) to (3.8), $b^{(j)}(i)$ denotes the information bit to be detected, and $b^{(j)}(i-1)$ the preceding bit. In addition, $\tau_{nl}^{(k)} = \tau_l^{(k)} - \tau_n^{(j)}$, $\varphi_{nl}^{(k)} = \varphi_l^{(k)} - \varphi_n^{(j)}$,

$$R^{(kj)}(\tau) = \int_0^{\tau} a^{(k)}(t-\tau)\, a^{(j)}(t)\, dt$$
$$\hat{R}^{(kj)}(\tau) = \int_{\tau}^{T} a^{(k)}(t-\tau)\, a^{(j)}(t)\, dt, \tag{3.9}$$

and, the spatial correlation parameter

$$\mathcal{R}_x^{(kj)} = \frac{Re[\mathbf{w}^{(j)^H}\mathbf{w}^{(k)}]}{\|\mathbf{w}^{(j)}\|\,\|\mathbf{w}^{(k)}\|}. \tag{3.10}$$

In (3.10), $(\cdot)^H$ denotes the Hermitian transpose and $\mathbf{w}^{(k)}$ the array steering vector optimizing the response of the antenna array for user j, or

$$\mathbf{w}^{(j)} = (w_1, w_2, \cdots, w_{M_B})^T. \tag{3.11}$$

From (3.10) it should be clear that $\mathcal{R}_x^{(kk)} = 1$ for all k, and independent of n, the specific diversity branch (or multipath signal). This means that the array response has been optimized in such a way that the antenna radiation pattern is a maximum in the direction of each of the L_p multipath components. Clearly this assumption assumes that the DOA of each multipath component has been perfectly estimated[4]. Furthermore, the antenna array elements are assumed to be sufficiently closely spaced to ensure that the signals received at each antenna element is highly correlated. More specifically, it is assumed that the correlation between the signals received at each element of the antenna array is greater than 0.8. If this correlation factor is lower, the antenna pattern synthesized by the adaptive antenna array will exhibit grating lobes [127]. The high correlation levels would enable digital beamforming techniques to be used to implement a spatial filter.

3.3.1.1 Correlation Influence on Beam Pattern. It is well known that lack of correlation influences the beamforming capabilities of an antenna array, but increases the receive diversity gain.

Figures 3.6 and 3.7 show the beam pattern for a $M_B = 12$ element uniform linear array (ULA) with antenna element spacing of $\lambda/2$. For each correlation value, 1000 snapshots were simulated and the corresponding radiation patterns generated.

Shown in these figures are the maximum beam pattern, the minimum beam pattern and the average beam pattern. A correlation of $\rho_{ij} = 0.8$ already degrades the beamforming characteristics and the beam pattern tends towards an omni-directional antenna pattern. However, the main beam is still at 0° with the side lobe levels altered. With the correlation between the antenna elements equal to $\rho_{ij} = 0.3$, it is clear that the antenna array cannot be used as a beamformer – the antenna pattern tends toward an omni-directional pattern, with very little gain relative to the side lobes in the desired direction of 0°.

3.3.1.2 BEP Performance. To arrive at an expression for the BEP of a beamforming system with a RAKE receiver, the signal power, U_S^2, and the total interference power, σ_T^2 need to be calculated. With these variables known, the received SNR is

$$\text{SNR} = \frac{U_S^2}{\sigma_T^2}. \tag{3.12}$$

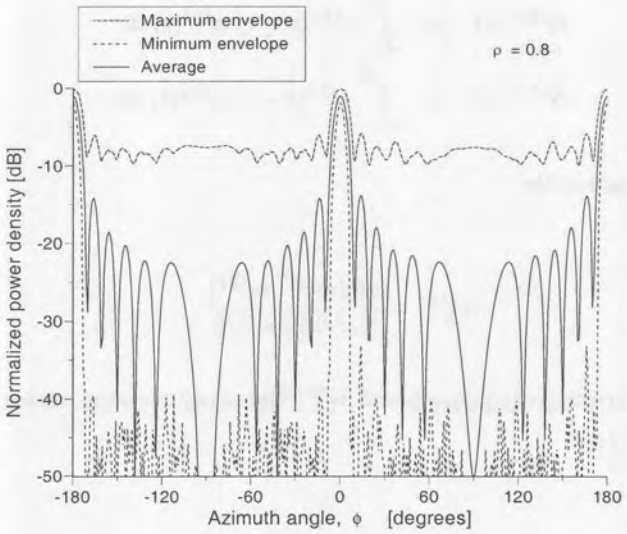

Figure 3.6. ULA beam pattern with envelope correlation of $\rho_{ij} = 0.8$.

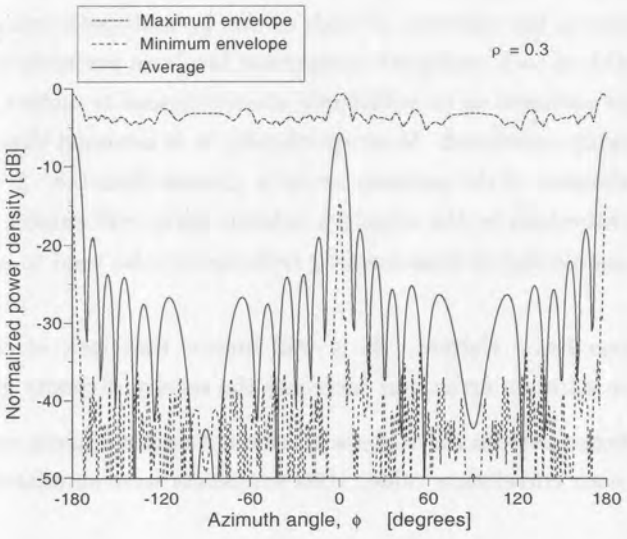

Figure 3.7. ULA beam pattern with envelope correlation of $\rho_{ij} = 0.3$.

From (3.11), the ULA beamformer for the reference user (j), will calculate [126]

$$\|\mathbf{w}^{(j)}\|^2 = (\mathbf{w}^{(j)})^H \mathbf{w}^{(j)} = \sum_{i=1}^{M_B} w_i^* w_i = M_B. \tag{3.13}$$

To simplify the analysis, it is assumed that the sum of all interference terms of (3.5) are Gaussian distributed [70, 85, 91, 115]. This assumption has been shown to be accurate, even for a small number of users K when the BEP is 10^{-3} or greater. Therefore, expanding on the results of [91, 115] to include beamforming, the variance of each interference term (equations (3.6), 3.7) and (3.8)) on the n^{th} RAKE tap, conditioned on the fading parameter, $\beta_n^{(j)}$, can be written as

$$\left(\sigma_{\mathrm{si}_n}^{(j)}\right)^2 = \frac{E_b T_s}{4N} \sum_{\substack{l=1 \\ l \neq n}}^{L} \left(\beta_n^{(j)} \|\mathbf{w}^{(j)}\| E\left\{\mathcal{R}^{(jj)}\right\}\right)^2 \Omega_l^{(j)}, \tag{3.14}$$

$$\left(\sigma_{\mathrm{mai}_n}^{(j)}\right)^2 = \frac{E_b T_s}{6N} \sum_{\substack{k=1 \\ k \neq j}}^{K} \sum_{l=1}^{L} \left(\beta_n^{(j)} \|\mathbf{w}^{(k)}\| E\left\{\mathcal{R}^{(jk)}\right\}\right)^2 \Omega_l^{(k)}, \tag{3.15}$$

$$\sigma^2 = \frac{T_s N_0}{4} \cdot \left(\beta_n^{(j)}\right)^2, \tag{3.16}$$

where $\Omega_l^{(k)}$ denotes the average signal power of path l received from user k. This in turn yields a total interference term of

$$\sigma_T^2 = \sum_{n=1}^{L_R} \left(\left(\sigma_{si_n}^{(j)}\right)^2 + \left(\sigma_{mai_n}^{(j)}\right)^2 + \sigma^2 \right), \tag{3.17}$$

where L_R denotes the number of branches in the RAKE receiver. Furthermore, the desired signal output of the RAKE combining receiver can be written as

$$U_S = \sqrt{\frac{E_b T_s}{2}} \sum_{n=1}^{L_R} \|\mathbf{w}^{(j)}\| (\beta_n^{(j)})^2. \tag{3.18}$$

In (3.18), it is assumed that the RAKE receiver will recover the strongest (i.e. largest average received signal power) L_R multipath components. For convenience, and without loss of generality, the strongest multipath components are assumed to be the L_R components that arrive at the receiver first.

The variance of the fading parameters of each interfering user, $E\left\{(\beta_l^{(k)})^2\right\}$ is equal to the average signal power received from that user, $\Omega_l^{(k)}$. This variance is not a function of the antenna array or steering vector, as the fading process is caused by physical scattering processes that occur at the mobile. The effect of the array is contained in the spatial correlation parameter (given by (3.10)), which will be a minimum if the array has a null in the direction of a specific interfering multipath signal arriving at the base station.

For coherent demodulation, the BEP conditioned on the instantaneous SNR, S, can be expressed as [126, 158]

$$P_{e|S} = Q\left(\sqrt{\Gamma_0 \cdot s}\right), \tag{3.19}$$

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} dt$ is the Q-function [147]. The output SNR (as defined in (3.12)), can be written in the form required by (3.19) as

$$S = \sum_{n=1}^{L_R} \left(\beta_n^{(j)}\right)^2, \tag{3.20}$$

and

$$\Gamma_0 = \left(\frac{1}{2N} \sum_{\substack{l=1 \\ l \neq n}}^{L} \left(E\left\{ \mathcal{R}^{(jj)} \right\} \right)^2 \Omega_l^{(j)} + \frac{1}{3N} \sum_{\substack{k=1 \\ k \neq j}}^{K} \sum_{l=1}^{L} \left(E\left\{ \mathcal{R}^{(jk)} \right\} \right)^2 \Omega_l^{(k)} + \frac{N_0}{2E_b \|\mathbf{w}^{(j)}\|^2} \right)^{-1}, \tag{3.21}$$

assuming that $\|\mathbf{w}^{(k)}\|^2$ is equal for all k [126].

To obtain the average BER, (3.19) must be averaged over the pdf of S. As discussed in Chapter 2, the distribution of S should accommodate different values of the fading parameter m for the different received paths. If it is assumed that the fading amplitude, $\beta_n^{(j)}$, is Nakagami distributed, the *power*, $\left(\beta_n^{(j)}\right)^2$, of the received fading amplitude will be gamma distributed. From (3.20) it is clear that the pdf of the sum of L_R gamma distributed random variables is required to obtain the average error rate. In Appendix A a general pdf for the sum of an arbitrary number of correlated gamma distributed random variables are derived, and repeated here in terms of the characteristic function

$$p_S(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi_S(t) e^{-its} dt, \tag{3.22}$$

where $\Phi_S(t)$ is the characteristic function defined in Appendix A. The BEP for a beamforming system with a RAKE receiver can now be written as

$$P_e = \int_{-\infty}^{\infty} P_{e|S} p_S(s) ds, \tag{3.23}$$

which can be solved using numerical methods [126].

3.3.2 *Transmit/Receive Diversity Performance*

In this section, the BEP performance of a CDTD transmit and MRC receive diversity system with arbitrary correlated fading on each of the transmit/receive diversity branches will be determined.

In beamforming applications it is assumed that the beamforming antenna array receives L_p multipath components, with the received signal envelope correlation at each of the array elements equal to one. The same signal is therefore received by all elements of the beamformer. In the case of diversity, the situation changes since the elements of the diversity array are separated by a larger physical distance. This means that the L_p uncorrelated multipath signals arriving at the first element in the diversity array is no longer the exact same set of multipath components received by the other elements in the array. In fact, when some of the elements of the diversity array are separated by a large distance (typically 20λ or more), the L_p uncorrelated multipath signals received at one element in the diversity array are completely different to the L_p multipath signals received at any other element in the diversity array.

Moving on from the conceptual discussion above, the decision variable of a M_D-branch diversity system can be written as

$$\zeta_{M_D}^{(j)} = \sum_{m=1}^{M_D} \zeta_{m_D}^{(j)}, \tag{3.24}$$

where each

$$\zeta_{m_D}^{(j)} = S_n^{(j)} + I_{\mathrm{si}_n}^{(j)} + I_{\mathrm{mai}_n}^{(j)} + I_{\mathrm{ni}_n}^{(j)}, \tag{3.25}$$

and represents the output of the RAKE receiver on a specific diversity branch. In 3.25, S_n, I_{mai_n} and I_{si_n} are defined in (3.6), (3.7) and (3.8), respectively, with the weight vector, $\mathbf{w}$ unity. As before, $I_{\mathrm{ni}_n}^{(j)}$ is the AWGN sample.

3.3.2.1 BEP Performance. In order to determine the BEP of a CDTD or MRC diversity system, it is necessary to determine the pdf of the SNR at the output of the combiner at the receiver. The SNR pdf is a function of

- The characteristics of the composite signal transmitted/received at each diversity branch (the effective value of the Nakagami fading parameter m). It has been shown [115, 158] that coherently combined, uncorrelated multipath components result in a composite signal envelope with effective Nakagami fading parameter, m_{eff}, given by

$$m_{eff} = \sum_{l=1}^{L_R} m_l. \tag{3.26}$$

  For instance, if an antenna element in a diversity array receives three multipath echoes, each with a Nakagami fading parameter $m_l = 1$ and coherently combines these signals using a RAKE combiner, the composite fading signal will have an effective Nakagami parameter $m_{eff} = 3$.

- The relation (correlation) between the signals received at each branch. The correlation between the composite fading envelopes received at each diversity branch is a function of the antenna height and also of the scattering environment as described in Chapter 2.

From [158] and (3.19) it follows readily that the conditional BEP of a M_D-branch diversity system is also

$$P_{e|S} = Q\left(\sqrt{\Gamma_0 \cdot s}\right), \tag{3.27}$$

where the received signal power random variable, S, will have a different pdf in (3.27) than in (3.19). Specifically for receive diversity with MRC, the pdf of S is given in Appendix A, with the transformation of $M_D = M_R$ and $m = m_{eff}$. For Rayleigh fading (Nakagami fading parameter $m = 1$) with equal path strength ($\Omega_l = \Omega$) and equal correlation, a special case for the pdf of S follows as (see Appendix A)

$$\begin{aligned} p_S(s) &= \frac{1}{\Omega^2\Gamma(M_R)}\left(\frac{s}{\Omega^2}\right)^{M_R-1} \\ &\times \frac{\exp\left(-\frac{s}{(1-\rho)\Omega^2}\right) \cdot {}_1F_1\left(1, M_R, \frac{\rho M_R s}{(1-\rho)(1-\rho+\rho M_R)\Omega^2}\right)}{(1-\rho)^{(M_R-1)}(1-\rho+\rho M_R)}, \end{aligned} \tag{3.28}$$

where M_R denotes the number of MRC receive diversity branches.

A similar expression for the pdf of S for a $M_D = M_T$ CDTD transmit diversity scheme with equal powered transmissions, with $m = m_{eff} = 1$, and with constant correlation between the branches, and transmitted over a Rayleigh fading channel, can be written. Specifically, the components of the received power vector for the transmit diversity system can be written as

$$\begin{aligned} p_{\mathbf{S_n}}(\mathbf{s_n}) &= \frac{1}{\Omega^2\Gamma(M_T \cdot L_R)} \left(\frac{\mathbf{s_n}}{\Omega^2}\right)^{M_T \cdot L_R - 1} \\ &\times \frac{\exp\left(-\frac{\mathbf{s_n}}{(1-\rho)\Omega^2}\right) \cdot {}_1F_1\left(1, M_T \cdot L_R, \frac{\rho M_T \cdot L_R \mathbf{s_n}}{(1-\rho)(1-\rho+\rho M_T \cdot L_R)\Omega^2}\right)}{(1-\rho)^{(M_T \cdot L_R - 1)}(1-\rho+\rho M_T \cdot L_R)}. \end{aligned} \tag{3.29}$$

With reference to (3.27), the unknown variables required to determine the BEP performance of a diversity system are the interference term Γ_0 and the correlation matrix given in Appendix A. The interference term has been defined in (3.21) for CDMA with beamforming. However, as has been shown by Lötter [126], the analysis is equally valid for diversity with $M_B = 1$ elements.

3.3.3 Numerical Results

Using (3.23), (3.28), (3.29) and the system parameters outlined in Table 3.1, the BEP performance of a CDMA system using beamforming, together with transmit and receive diversity can be determined numerically under various physical and implementation conditions. The BEP performance of the space-time systems will be presented as a function of system load, $V = N/K$, with operating point taken as $E_b/N_0 = 20$ dB. Figures 3.8 to 3.10 compare the (uncoded) performance of the different space-time processing techniques covered in this chapter.

Parameter	**Simulation value**
Spreading sequence length	$N = 32$
Operating environment	2-Path, equal strength.
User distribution	uniform
Number of multipath signals	$L_p = 2$
Number of users	$K = 1, 2, \ldots, N$
Number of RAKE fingers	$L_R = 2$
Beamforming elements	$M_B = 1, 2, 3;\ \rho = 1, 0.5$
Transmit diversity elements	$M_T = 1, 2, 3\ \rho = 0, 0.5$
Receive diversity elements	$M_R = 1, 2, 3\ \rho = 0, 0.5$

Table 3.1. System parameters for numerical evaluation of BEP performance.

In Figure 3.8, the influence of the beamforming antenna array size on the capacity of a cellular CDMA system is shown assuming a constant correlation model. As would be expected, the BEP performance of the system improves with increasing beamformer size. This is due to the fact that larger beamforming arrays can synthesize narrower beams and thereby reduce the MAI seen by the reference user. As the BER

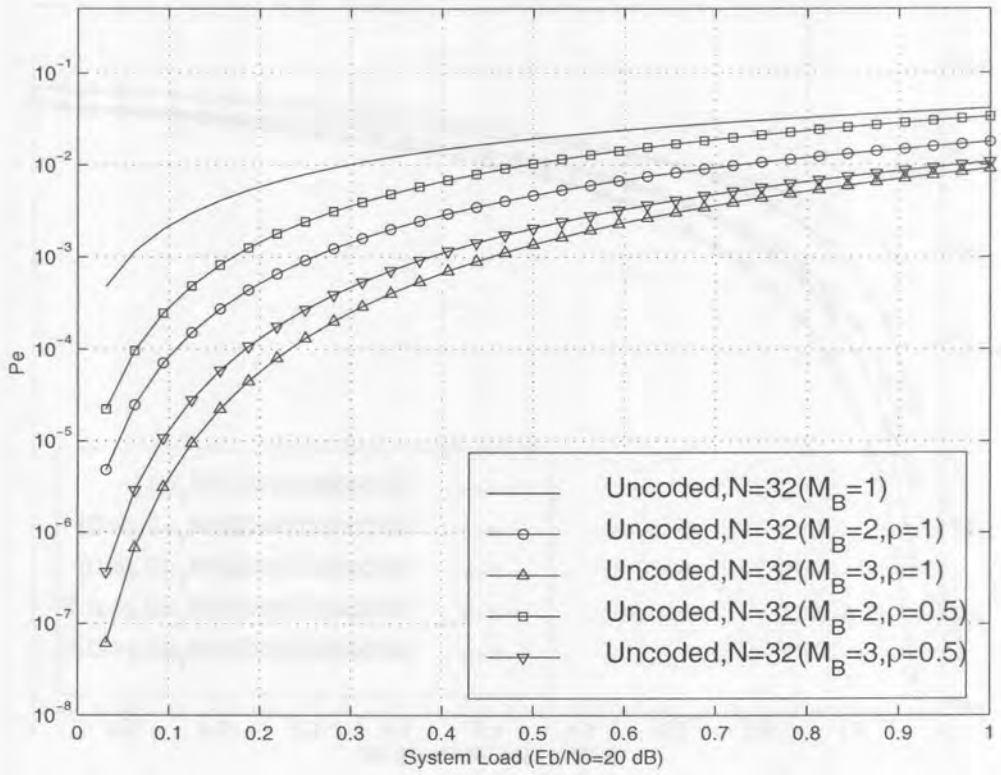

Figure 3.8. Uncoded receive beamforming performance.

probability is reduced by increasing the number of elements in the beamformer, the capacity of the cellular system is also increased.

In the beamforming system the signal processor performs the calculation of the direction of the desired signal, as well as the set of antenna weights required to focus the antenna radiation pattern in the direction of the desired signal. Typically, the radiating elements of an adaptive antenna array are separated by $\lambda/2$ where λ denotes the wavelength of the carrier frequency. Transmission from the target mobiles occurs at the same time instant, and the beam of the base station antenna is formed to maximize the received signal power from these target mobiles, while the received power from other interfering mobiles (inside this particular cell, as well as in adjacent cells) is minimized through the introduction of nulls in the antenna radiation pattern.

Note that in the case of systems using CDMA, the situation changes as follows. In general, it can be assumed that the number of subscribers active in a cell will be larger than the number of elements in an array, that is greater than the freedom levels of the adaptive system. All of these subscribers are transmitting in the same frequency band at the same time, meaning null steering cannot be used to cancel all interfering signals. Therefore, in the case of CDMA, at best beamsteering techniques in conjunction with limited null steering can be used to point the main beam of the antenna array in the direction of a desired user or group of users [112] and to place nulls in the direction of the main interfering signals.

Also shown in Figure 3.8, is the effects of correlation (or lack thereof), indicating the reduced effectiveness of the beamformer to "remove" unwanted users from the system under conditions of reduced correlation.

The uncoded BEP performance of transmit O-CDTD and receive MRC diversity systems under different constant correlation conditions are shown in Figure 3.9 and 3.10, respectively.

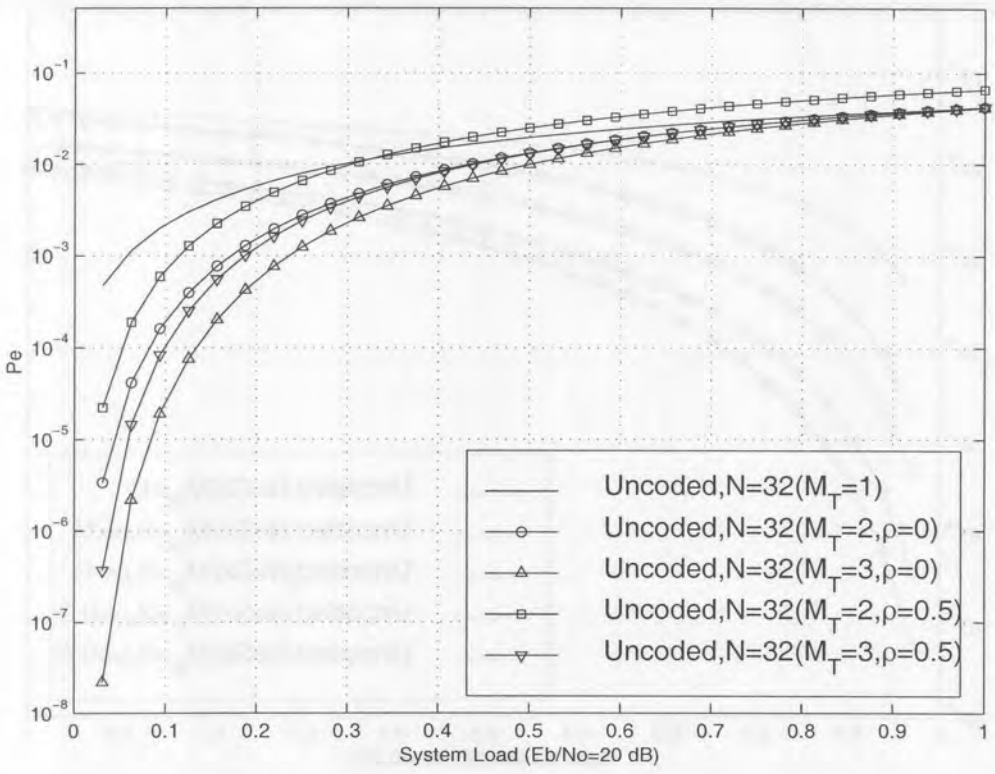

Figure 3.9. Uncoded O-CDTD transmit diversity performance.

In both cases it is clear that the correlation between the transmitted/received signals on the diversity branches has a significant impact on the BEP performance of the system. This is evident from the figures where the influence of the number of diversity branches, as well as the correlation between the signals at the various branches is clearly shown. The results indicate that variations in the correlation between received signal envelopes as a function of the user's position or other spatial parameters must be taken into account when determining the capacity of a diversity system.

It is also important to note that even when employing orthogonal spreading sequences (as in O-CDTD), the downlink will not be perfectly orthogonal due to multipath propagation. In [159, 160], the downlink orthogonality factor has been calculated for different environments. This factor, expressed as a percentage and shown in Table 3.2, is the fraction of the total output power that will be experienced as intra-cell interference. An orthogonality factor of zero corresponds to a perfectly orthogonal downlink, while a factor of one is a completely non-orthogonal downlink. As shown in Table 3.2, 40% of the power transmitted from the reference cell will act as intra-cell interference in a vehicular environment.

Propagation model	Orthogonality factor (%)
Indoor office	10 %
Outdoor to indoor and pedestrian	6 %
Vehicular	40 %

Table 3.2. Orthogonality factor for the different WCDMA channel environments.

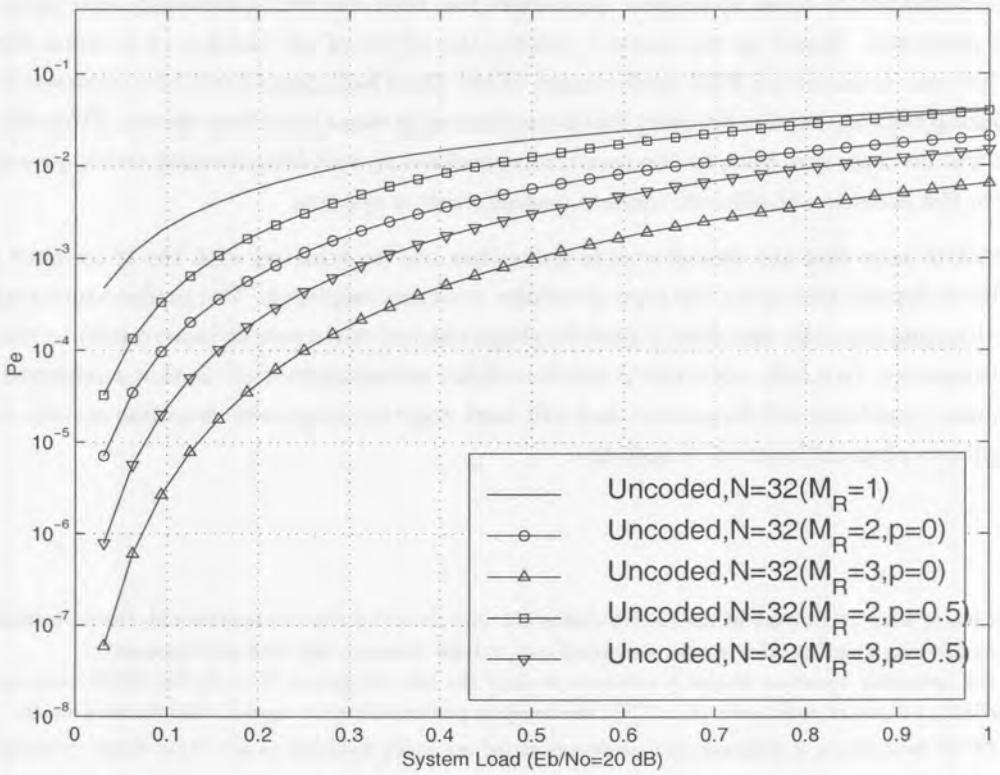

Figure 3.10. Uncoded MRC receive diversity performance.

Orthogonality, therefore, removes 60% of the interference or, stated differently, an orthogonality factor of 0.4 is obtained (40% of orthogonality remains). Under these conditions, the effectiveness of O-CDTD transmit diversity system is reduced, and it is expected that NO-CDTD should provide improved performance.

Also, from Figures 3.9 and 3.10 it is clear that the performance of the transmit diversity is inferior to that of receive diversity employing the same number of diversity branches. This is also attributed to the loss of orthogonality in the downlink.

Receive beamforming provided the best performance of all the techniques considered. Since the antenna beam is focused on a specific user, the antenna gain (or SNR) in the direction of the user is increased, and the transmissions received from interfering users are suppressed by the introduction of nulls in antenna pattern.

From the results presented in Figure 3.8 and 3.10 it can be seen that the performance of perfect beamforming ($\rho = 1$) and perfect receive diversity ($\rho = 0$) is similar as has been shown in [126].

3.4 SUMMARY

The reasoning behind the use of space-time processing techniques is the optimization of the cellular spectral efficiency of the network. This is realized by implementing more than one antenna element to optimally transmit or receive signals by using both temporal and spatial signal processing techniques in the transceiver. This chapter considered in detail the main smart antenna techniques, namely transmit and receive diversity, and beamforming.

The BEP performance of these space-time processors has been derived analytically and some numerical results were presented. Based on the analytic results, the effects of the number of antenna elements and correlation between branches on BEP performance of the three techniques have been addressed. Also, the close relationship between receive diversity and beamforming systems have been shown. From the presented results it should be clear that both the transmit/receive diversity and beamforming techniques are of great importance in the provision of efficient wireless communication systems.

It is important to note that the overall system gains that can be achieved with the space-time techniques described above, depend heavily on the type of cellular structure employed. The performance may therefore differ when changing from the described 2-path Rayleigh channel environment (macro-cell) to micro- or pico cellular environments. In a fully operational mobile cellular environment, such as that envisioned in UMTS, all these cellular structures will be present and will work together to provide seamless service, irrespective of the subscriber's physical location or mobility.

Notes

1. Soft-failure states that, should one of the receive chains fail, and the other chain is operational, the performance loss is of the order of the diversity gain. In other words, the signal may still be detected, but with inferior quality.
2. Although the spreading sequence length is assumed to span the the bit period ($T = T_b$ for BPSK) and symbol period ($T = T_s$ for QPSK), i.e., short codes with $N = T/T_c$, the analysis presented here is equally valid for long codes.
3. Although BPSK modulation is assumed, the results presented are easily extended to any other linear modulation scheme.
4. The estimation of the DOA of multipath signals fall outside the scope of this thesis.

4 CHANNEL CODING FOR CDMA

Following the derivation of the uncoded system performance in Chapter 3, based on the presented channel model of Chapter 2, this chapter shifts the focus to channel coding techniques and their performance. Specifically, the use and performance of classical convolutional, turbo codes and trellis codes for cellular CDMA will be addressed.

4.1 FEC BACKGROUND

4.1.1 Block Codes

The approach to error correction coding taken by modern digital communication systems started in the late 1940's with the ground breaking work of Shannon [161], Hamming [162], and Golay [163]. The next main class of linear block codes to be discovered were the Reed-Muller (RM) codes in 1954 [164]. The latter codes provided a significant improvement on the Hamming and Golay codes because they allowed more flexibility in the size of the code word and the number of correctable errors per code word [165, 166]. Following the discovery of RM codes came the discovery of cyclic codes [164, 165].

An important subclass of the cyclic codes was discovered simultaneously by Hocquenghem in 1959 and by the team of Bose and Ray-Chaudhuri in 1960 [15, 164], known as BCH codes. BCH codes were extended to the non-binary case ($q > 2$) by Reed and Solomon in 1960 [167]. Reed Solomon (RS) codes constituted a major improvement since their non-binary nature allows for protection against bursts of errors.

Despite the success of block codes, there are several fundamental drawbacks to their use. Firstly, due to the frame (block) oriented nature of block codes, the entire code word must be received before decoding can be completed. This can introduce an intolerable latency into the system, particularly for large block lengths. A second drawback is that block codes require precise frame synchronization. A third drawback is that algebraic-based decoders for block codes usually employ hard-decision decoding (HDD), rather than the unquantized, or "soft", outputs of the demodulator. It is actually possible to perform soft-decision decoding of block codes, although until recently soft-decision decoding has been regarded as too complex. Recent

work in the area of errors-and-erasures decoding for RS codes [165] and trellis-based soft-decision decoding algorithms for other classes of block codes, swung the interest and approach towards soft-decision decoding (SDD) [168].

4.1.2 *Convolutional Codes*

The drawbacks of block codes can be avoided by taking a different approach to coding, namely that of convolutional coding, which was first introduced in 1955 by Elias [164]. Rather than segmenting data into distinct blocks, convolutional encoders add redundancy to a continuous stream of input data by using a linear shift register. Each set of n output bits is a linear combination of the current set of k input bits and the m bits stored in the shift register. The total number of bits that each output depends on is called the constraint length, denoted by K_{cc}[1]. Just as the data is continuously encoded, it can also be continuously decoded with only nominal latency. Furthermore, the decoding algorithms based on the Viterbi and maximum *a posteriori* (MAP) algorithms, can make full use of soft-decision information from the demodulator.

A key weakness of convolutional codes is that they are very susceptible to burst errors. This weakness can be alleviated by using an interleaver, which scrambles the order of the code bits prior to transmission. By scrambling the code bits' order at the transmitter and then reversing the process at the receiver, burst error patterns can be broken up so that they appear independent to the decoder. All of the 2G and 3G digital cellular standards use some form of block interleaving.

It should be noted that in many ways convolutional codes have properties that are complimentary to those of RS codes. While convolutional codes are susceptible to burst errors, RS codes handle burst errors quite well. However, convolutional codes with soft-decision decoding generally outperform RS codes of similar complexity at low SNRs [165]. In severely power limited channels, an interesting and efficient system design can be obtained by using the concatenation of a RS "outer" code and a convolutional "inner" code [169].

4.1.3 *Concatenated Codes and Iterative Decoding*

Since conventional block- and convolutional codes are highly structured, encoders and decoders with reasonable implementation complexity are possible. However, the very same structure that facilitates practical implementation, results in significantly inferior performance gains relative to the random coding bounds predicted by Shannon.

With this in mind, perhaps the most exiting and potentially important development in coding theory in recent years has been the introduction of parallel concatenated convolutional codes by Berrou *et al.* [170]. The term "turbo code" was adopted to describe this new class of code. The introduction of turbo coding has opened a whole new way of looking at the problem of constructing good codes with low complexity decoding. Although turbo codes possess random-like properties, they still contain enough structure to admit practical encoding and decoding algorithms. As a consequence, the performance of turbo codes comes much closer to the Shannon bound than conventional block and convolutional codes.

4.1.3.1 Turbo Encoding.

Turbo codes are iteratively decoded parallel concatenated convolutional codes (referred to here as PCCC) which consist of two convolutional encoders, one of which encodes the information bits directly, while the other encodes the information bits following interleaving. The key to solving the decoding complexity of PCCC schemes is the existence of a sub-optimal decoding algorithm which achieves performance very close to that of a maximum likelihood decoder. This algorithm iteratively decodes each code separately using soft-input/soft-output algorithms such as MAP or SOVA [166, 171].

In [172] the serial dual of turbo codes were introduced. These iteratively decoded serially concatenated convolutional codes (SCCC) are constructed from the same constituent codes and interleaver elements as PCCC, but are concatenated in a serial rather than a parallel fashion. Again, an iterative decoding algorithm is used which achieves near-optimum results. SCCC achieve comparable performance to PCCC, and in some cases can offer superior performance [172, 173, 174, 175].

The turbo or PCCC encoder is composed of two or more recursive systematic convolutional (RSC) encoders, which are in general identical. The constituent encoders receive the "same" data, the only difference is that the stream to each encoder is permuted by an interleaver, with the result that turbo codes appear random. Because the interleaver must have a fixed structure and generally works on data in a block-wise manner, turbo codes are by nature block codes.

Recall that the minimum distance of a linear block code is a good first order estimate of the code's performance. For linear block codes, the minimum distance is the smallest non-zero Hamming weight of all valid code words. The combination of interleaving and RSC encoding ensures that most code words produced by a turbo coder have a high Hamming weight. Because of its infinite impulse response properties, the output of an RSC encoder generally has a high Hamming weight. There are, however, some input sequences which cause an RSC encoder to produce low weight outputs. Because of the interleaver, the two RSC encoders do not receive their inputs in the same order. Thus, if one encoder receives an input that causes a low weight output, then it is improbable that the other encoder also receives an input that produces a low weight output. Unfortunately, since there will always be a few input messages that cause both RSC encoders to produce low weight outputs, the minimum distance of a turbo code may, in general, not be particularly high. But the *multiplicity* of low weight code words in well designed turbo codes is low. It is because of the relatively small number of low weight code words that turbo codes can perform well at low SNR [176, 166].

However, the performance of turbo codes at higher SNRs becomes limited by the relatively small minimum distance of the code. While the goal of traditional code design is to increase the minimum distance of the code, the objective of turbo code design is to reduce the multiplicity of low weight code words.

4.1.3.2 Turbo Decoding. The problem of estimating the states of a Markov process in the presence of noise has two well known trellis-based solutions — the Viterbi algorithm [177] and the (symbol-by-symbol) MAP algorithm [178, 179]. The two algorithms differ in their optimality criterion. The Viterbi algorithm finds the most probable transmitted sequence, while the MAP algorithm, on the other hand, attempts to find the most likely transmitted symbol, given the received sequence [166, 180, 181].

One drawback of both PCCC and SCCC is decoder complexity. For some applications, such as hand-held mobile handsets, it may be desirable to tradeoff some of the high coding gain of SCCC for lower decoder complexity. Several types of algorithms can be used within a turbo decoder to perform soft/input soft-output decoding of the constituent codes. An excellent overview of the trellis-based soft-input soft-output decoding algorithms has been presented in [166]. Shown in Figure 4.1 are the trellis based algorithms. These algorithms can be partitioned into two main classes, depending on whether they were derived from the Viterbi algorithm or from the MAP algorithm. Generally speaking, the algorithms at the bottom of the diagram are more computationally complex and perform better than the algorithms at the top of the diagram. Also, the MAP-based algorithms are generally more computational intensive and perform better than the algorithms based on the Viterbi algorithm.

The soft output Viterbi algorithm (SOVA) is an extension of the classic Viterbi algorithm that provides the reliability of the bit estimates [180, 182, 183, 184]. In addition, the improved SOVA algorithm, utilizing a multiplicative correction factor to improve the reliability estimates, may also be considered.

The MAP algorithm calculates the *a posteriori* probabilities directly. However, the algorithm suffers from a high computational complexity and numerical sensitivity. The Max-Log-MAP and Log-MAP algorithms perform the MAP algorithm in the log domain, which significantly reduces complexity and numerical sensitivity [171].

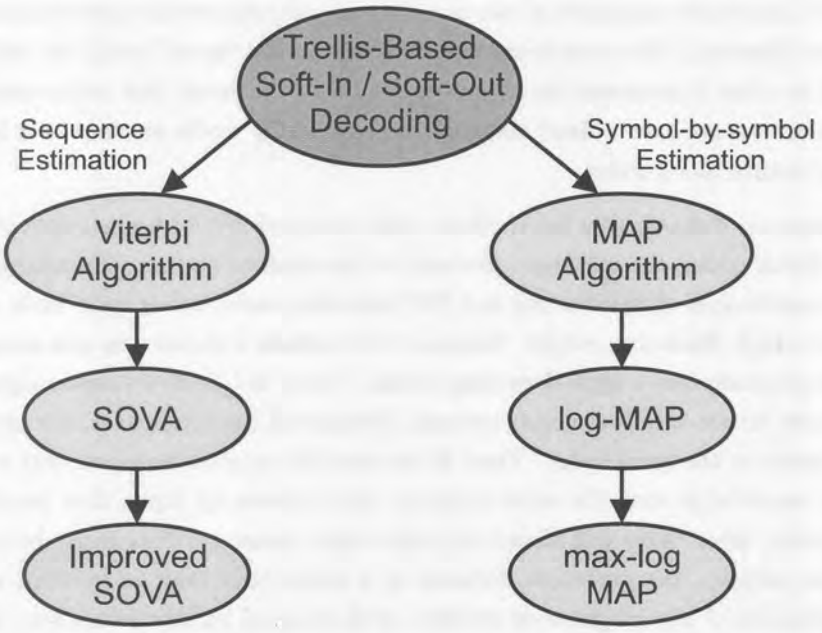

Figure 4.1. Trellis-based soft-input soft-output algorithms.

4.1.4 *Turbo Codes Performance*

Figure 4.2 depicts the turbo code design space [185]. The design space can be grouped into service dependent and implementation dependent components. The service dependent components influence typically the quality of service and the data rate. The implementation dependent components influence the maximum decoding delay, the implementation complexity, system flexibility, modularity and integratability.

Below, a short description of the more important blocks of Figure 4.2 is given.

Turbo Interleaver/Permuter. The interleaver (or permuter) component of the turbo encoder directly defines the service dependent part of the system design space. The weight distribution of the codewords produced by the turbo decoder depends on how the codewords from one of the basic codes are teamed with codewords from the other encoder(s). Stated differently, the performance of the turbo code depends on how effectively the data sequences that produce low encoded weights at the output of one encoder, are matched with permutations of the same data sequence that yield higher encoded weights at the outputs of the others. Two characteristics of the interleaver is of particular importance

- Interleaver size, N_{tc}. This is the most important factor influencing the turbo code performance, and it is well known that performance improves as the interleaver size increases [186]. The gain, in terms of error performance, with increased interleaver size is formally known as the *interleaver gain*. However, as the interleaver size (gain) increases, so does decoding delay, and a balance must be found between

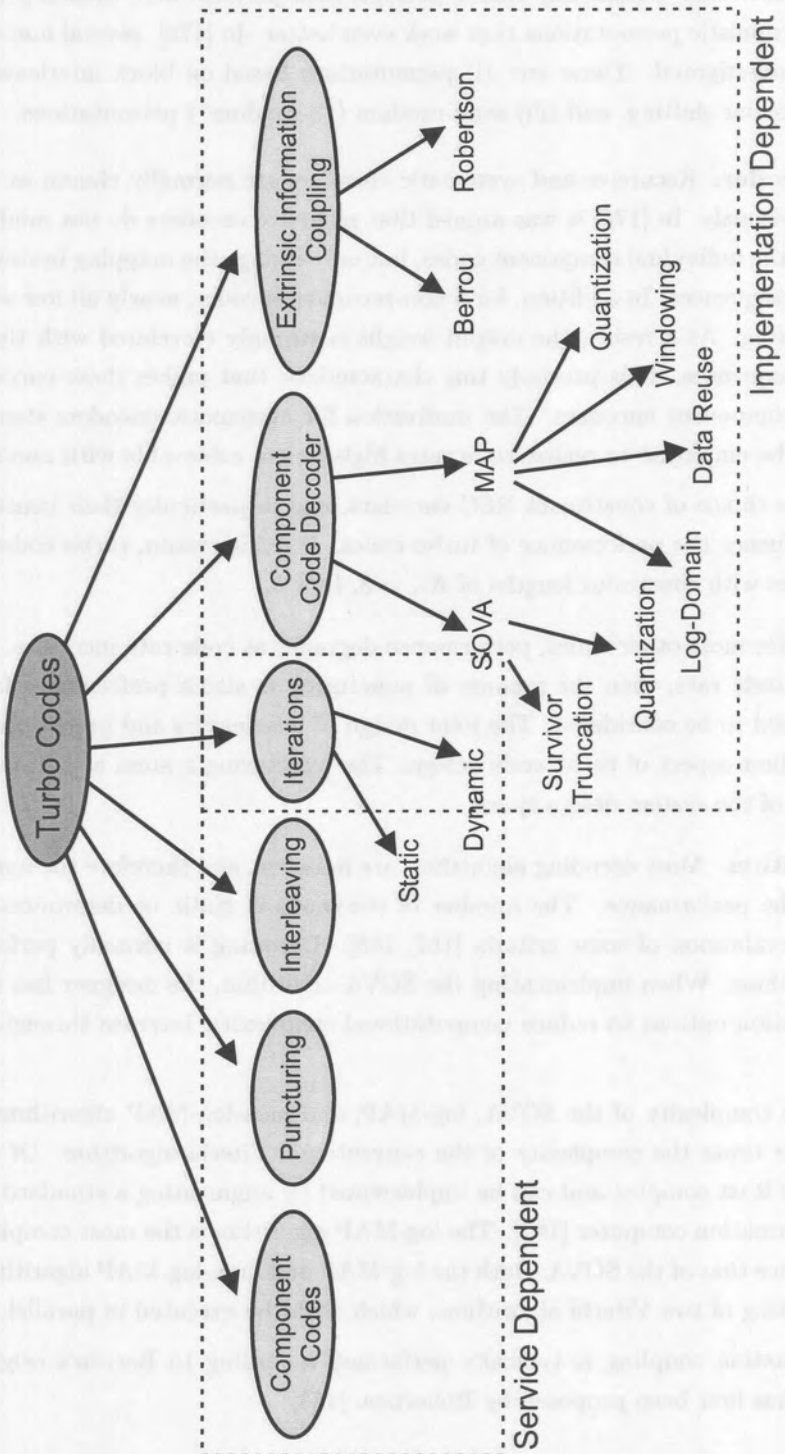

Figure 4.2. Turbo coded system design space [185].

acceptable performance and tolerable latency. At high SNRs, the interleaver design becomes critical [173], where the performance is dominated by the low weight code words. At low SNRs, turbo codes perform well with almost any (randomly permutated) interleaver, provided that the inputs at the RSC encoders are sufficiently uncorrelated.

- Interleaver selection. If randomly chosen permutations perform well, then in principle it is possible to design deterministic permutations that work even better. In [176], several non-random permutations have been investigated. These are: (i) permutations based on block interleavers, (ii) permutations based on circular shifting, and (iii) semi-random ("S-random") permutations.

Constituent Encoder. Recursive and systematic encoders are normally chosen as constituent encoders, as discussed previously. In [176] it was argued that recursive encoders do not modify the output weight distributions of the individual component codes, but only change the mapping between the input data and output encoded sequences. In addition, for a non-recursive encoder, nearly all low weight input sequences are self-terminating. As a result, the output weight is strongly correlated with the input weight for all possible input sequences. It is precisely this characteristic that makes these encoders very undesirable as constituent component encoders. The motivation for systematic encoders stems from the fact that puncturing can be employed to realize code rates higher than achievable with non-systematic encoders.

Surprisingly, the choice of constituent RSC encoders, and in particular their constraint lengths, do not significantly influence the performance of turbo codes. For this reason, turbo codes typically use simple constituent codes with constraint lengths of $K_{tc} = 3, 4$, or 5.

Puncturing. As for most other codes, performance degrades as code rate increases. If puncturing is used to increase the code rate, then the manner of puncturing is also a performance factor and puncturing matrices may need to be considered. The joint design of interleavers and puncturing matrices is perhaps the most important aspect of turbo code design. The puncturing system also directly define the service dependent part of the system design space.

Decoding Algorithm. Most decoding algorithms are iterative, and therefore the number of iterations has an impact on the performance. The number of iterations is static or determined dynamically during decoding after evaluation of some criteria [187, 188]. Decoding is normally performed with the MAP or SOVA algorithms. When implementing the SOVA algorithm, the designer has to choose among several implementation options to reduce computational complexity, increase throughput, or reduce power consumption.

The algorithmic complexity of the SOVA, log-MAP, and max-log-MAP algorithms are similar, ranging from two to four times the complexity of the conventional Viterbi algorithm. Of the three algorithms, the SOVA is the least complex and can be implemented by augmenting a standard Viterbi decoder with a reliability information computer [182]. The log-MAP algorithm is the most complex, with a complexity that is about twice that of the SOVA. Both the log-MAP and max-log-MAP algorithms have an attractive structure consisting of two Viterbi algorithms, which could be executed in parallel.

Extrinsic information coupling is typically performed according to Berrou's original method [170] or directly, which has first been proposed by Robertson [188].

4.1.5 Trellis Coded Modulation

Up until the mid 1970's, coding and modulation were regarded as two separate processes. Ungerboeck changed this thinking in 1976 with the introduction of trellis coded modulation (TCM) [189, 190]. The

main advantage of TCM over classical coding schemes is the fact that trellis coding, and the resulting data-transmission strategy, does not expand the transmission bandwidth. It is both a power- and a bandwidth-efficient modulation scheme.

In the application of trellis coding techniques to cellular CDMA communication, two different approaches have been considered. The first approach is based on classical TCM techniques which combines coding and modulation into a single entity [44]. It has been shown that for a given complexity, chip rate and throughput, these codes provide no advantage over medium to low rate convolutional codes. The second technique uses a different approach and combines coding and spreading, instead of combining coding and modulation. This technique, referred to as trellis coded spreading (TCS), performs better than standard error control techniques for the same complexity and code rate [45].

TCM for CDMA. Boudreau *et al.* [44] considered the use of trellis codes in a DS/CDMA system and compared their performance with that of convolutionally coded DS/CDMA. Their codes were constructed over an M-PSK signal set by taking a standard Ungerboeck type code for M-PSK modulation and then spreading this M-PSK signal with a binary m-sequence, over a large bandwidth. The authors in [44] reported that this approach did not yield a performance advantage over standard convolutional codes when combined in a CDMA system. It was argued that this is mainly due to the fact that a convolutional code can be employed without any bandwidth expansion or decrease in processing gain in a CDMA system. The latter directs one to rather exploit the lower distance properties of lower rate convolutional codes.

TCS for CDMA. A different approach to the idea of trellis coded CDMA was investigated by Woerner and Stark [45]. In this approach the trellis code is constructed over the set of possible signature sequences rather, than over some $2D$ signal constellation. Instead of expanding the number of signal points in the $2D$ constellation, the number of possible spreading sequences used is expanded. A carefully designed trellis then allows only certain combinations of sequences that have a large total minimum distance. Now since the number of sequences has been increased, the actual minimum distances between sequences decrease, but fortunately the trellis code more than compensates for this by increasing the minimum distance of the code above that of the uncoded system.

Coded performance can be calculated by extending the space-time mathematical model derived in Chapter 2 and using well known error control bounding techniques. In the remainder of this chapter, the performance of classical convolutional (including orthogonal extensions), turbo and trellis coded CDMA is considered.

4.2 PERFORMANCE EVALUATION

4.2.1 Convolutional Code (CC) Performance

The bounds presented here are based on block error probability bounds, originally derived by Shannon [161]. Specifically, to determine upper bounds on BEP with convolutional encoding and ML decoding represented by an equivalent (n, k) linear block code, it is useful to recall the state diagram and associated generating function approach. Due to code linearity, it is assumed that the all-zero message has been transmitted, and the upper bound on the word error probability can be written as

$$P_w \leq \sum_{d=d_{free}}^{n} A_d \, P_d(\mathbf{c} \rightarrow \hat{\mathbf{c}}), \tag{4.1}$$

where A_d is the number of codewords with Hamming weight d, obtained from the series expanded transfer function $T(L, I, D)$, and given by (for path length, $L = 1$)

$$\left.\frac{\delta T(I,D)}{\delta I}\right|_{I=1} = \sum_{d=1}^{n} A_d D^d, \tag{4.2}$$

where D is the channel parameter. By setting $I = 1$ after differentiation, the number of bit errors, corresponding to an error event of length d, equals the multiplicity of term D^d, where $D = e^{-R_c E_b/N_o}$.

The *free distance*, denoted by d_{free}, of any code is the minimum Hamming distance between any two distinct code sequences. D is a function of the channel transition probabilities and the message decoding metric only.

The conditional pairwise error probability, $P_d(\mathbf{c} \to \hat{\mathbf{c}})$, is the probability of incorrectly choosing a codeword with weight d, that is, the probability that the incorrectly encoded sequence $\hat{\mathbf{c}}_n = (c_1, c_2, \ldots, c_n)$ is chosen instead of the correctly encoded sequence $\mathbf{c}_n = (c_1, c_2, \ldots, c_n)$.

For the continuous output soft-decision AWGN channel it can be shown that the single user $P_d(\mathbf{c} \to \hat{\mathbf{c}})$ is given by [22, 16, 13]

$$P_d(\mathbf{c} \to \hat{\mathbf{c}}) \;=\; Q\left(\sqrt{\frac{2dR_cE_b}{N_o}}\right). \tag{4.3}$$

From [176], the more general expression for the average weight distribution can be written as

$$A_d = \sum_{i=1}^{k} \binom{n}{i} p(d \mid i), \tag{4.4}$$

where $\binom{k}{i}$ is the number of input words with Hamming weight i and $p(d \mid i)$ is the probability that an input word with Hamming weight i produces a codeword with Hamming weight d. Substituting (4.4) into (4.1), the upper bound on the word and bit error rate can be expressed as

$$\begin{aligned} P_w \;&\leq\; \sum_{d=d_{free}}^{n} A_d \, P_d(\mathbf{c} \to \hat{\mathbf{c}}) \\ &=\; \sum_{d=d_{free}}^{n} \sum_{i=1}^{k} \binom{k}{i} p(d \mid i) \, P_d(\mathbf{c} \to \hat{\mathbf{c}}) \\ &=\; \sum_{i=1}^{k} \binom{k}{i} E_{d|i} \left\{ P_d(\mathbf{c} \to \hat{\mathbf{c}}) \right\}, \end{aligned} \tag{4.5}$$

and

$$P_e \leq \sum_{i=1}^{k} \frac{i}{k} \binom{k}{i} E_{d|i} \left\{ P_d(\mathbf{c} \to \hat{\mathbf{c}}) \right\}. \tag{4.6}$$

In (4.5) and (4.6), $E_{d|i}\{\cdot\}$ is an expectation with respect to the distribution $p(d \mid i)$. This average upper bound is attractive because relatively simple schemes exist for computing $p(d \mid i)$ from the state transition matrix of the RSC [191, 176]. This information is implicit to the generating function $T(I, D)$ associated with the particular code employed.

4.2.1.1 Evaluation of $P_d(\mathbf{c} \to \hat{\mathbf{c}})$. Under conditions of fast fading, it is generally assumed that the fading is independent in successive signaling intervals. As a result, the sequence of fading amplitudes β_i constitutes an independent and identically distributed (i.i.d.) sequence[2].

Consider the situation where the all-zeros codeword $\mathbf{c} = \mathbf{0} = \mathbf{c_0}$, is transmitted and codeword $\hat{\mathbf{c}} = \mathbf{c_n}$ is received. In addition, a trellis path which re-emerges with the correct all-zero path is considered, having diverged at some point in the past, and differing from the all-zero path in exactly d symbol positions. Define the n-vector

$$\mathbf{S_n} = (S_{n1}, S_{n2}, \ldots, S_{nd}), \tag{4.7}$$

where $S_{ni}, (i = 1, 2, \ldots, d)$ represents the value of the resulting envelope power process in the ith signaling interval where the path differs from the all-zero path. Assuming perfect phase tracking of the phase perturbation process and channel state information (CSI) at either the transmitter or receiver, the conditional pairwise error probability for an incorrect sequence with d error symbols is [191]

$$P_d(\mathbf{c} \to \hat{\mathbf{c}} \mid \mathbf{S_n}) = Q\left(\sqrt{\Gamma_{0c} \sum_{i=1}^{d} s_{ni}}\right), \tag{4.8}$$

where Γ_{0c} is proportional to the effective output signal-to-noise ratio of the coded system (Γ_{0c} should be compared with its uncoded counterpart, Γ_0 defined in Chapter 3). The average error event probability can then be determined by averaging over the random n-vector $\mathbf{S_n}$, resulting in

$$P_d(\mathbf{c} \to \hat{\mathbf{c}}) = E_{\mathbf{S_n}}\left\{Q\left(\sqrt{\Gamma_{0c} \sum_{i=1}^{d} s_{ni}}\right)\right\}, \tag{4.9}$$

where the expectation operator $E_{\mathbf{S_n}}\{\cdot\}$ represents joint expectation with respect to the received signal power components.

Fast fading. For fast fading with perfect CSI, the pairwise error probability is given by

$$P_d(\mathbf{c} \to \hat{\mathbf{c}} \mid \mathbf{S_n}) = Q\left(\sqrt{\Gamma_{0c} \sum_{i=1}^{d} s_{ni}}\right). \tag{4.10}$$

When the pairwise error probability, (4.10), is averaged over (3.28), a multi-dimensional integral given by

$$\begin{aligned} P_d(\mathbf{c} \to \hat{\mathbf{c}}) &= \int_{s_1}\int_{s_2}\cdots\int_{s_d} Q\left(\sqrt{\Gamma_{0c} \sum_{i=1}^{d} s_{ni}}\right) \\ &\times \; p_{S_{n1}}(s_{n1})\, p_{S_{n2}}(s_{n2}) \cdots p_{S_{nd}}(s_{nd})\, ds_{n1}\, ds_{n2} \cdots ds_{nd}, \end{aligned} \tag{4.11}$$

has to be evaluated.

If the fading powers are independent, the indexes of the differing bit positions are of no importance, since only the incorrect codeword weight matters [191]. The exact evaluation of (4.11) is very difficult. To solve this problem, Hall *et al.*, examined four options [191]. The first option is to simplify (4.11) to a form that can be evaluated through numerical integration [192]. The other three options examined, avoids the problem of numerical integration by seeking closed form upper bounds for $P_d(\mathbf{c} \to \hat{\mathbf{c}})$.

The first option proposed by Hall *et al.* was employed to obtain analytical results. From [192], $Q(x)$ can be expressed in the alternative form written as

$$Q(x) = \frac{1}{\pi} \int_0^{\pi/2} e^{-x^2/(2\sin^2\psi_h)}\, d\psi_h. \tag{4.12}$$

Substituting (4.12) into (4.10), the following expression for the pairwise error probability is found

$$P_d(\mathbf{c} \to \hat{\mathbf{c}} \mid \mathbf{S_n}) = \frac{1}{\pi} \int_0^{\pi/2} \exp\left(-\frac{\Gamma_{0c} \sum_{i=1}^{d} s_{ni}}{2 \sin^2 \phi} \right) d\phi. \tag{4.13}$$

Since all the fading powers are independent, the d-dimensional integral of (4.11) reduces to a product of integrals over each S_{ni}.

Slow fading. Slow fading occurs when the symbol signalling rate is greater than the fading rate. That is, when the effective fading amplitude is assumed to be constant throughout the message sequence, $S_{ni} = \beta_{ni}^2 = \beta^2 = S$. The pdf of the received signal power S is again given by (3.28), where, for slow fading, $\mathbf{S_n} = S$.

It follows from (4.8) that

$$P_d\left(\mathbf{c} \to \hat{\mathbf{c}} \mid S\right) = Q\left(\sqrt{d\, \Gamma_{0c}\, s}\right). \tag{4.14}$$

Using the inequality

$$Q(x) \leq \frac{1}{2} e^{-x^2/2}, \;\; x >> 1, \tag{4.15}$$

the upper bound for the pairwise error probability can be written as

$$P_d(\mathbf{c} \to \hat{\mathbf{c}}) \leq E_{\mathbf{S}} \left\{ \frac{1}{2} \exp\left(-\frac{1}{2}\, d\, \Gamma_{0c} s \right) \right\}. \tag{4.16}$$

Then, from (4.9), the pairwise error probability, averaged over the fading statistics can be written as

$$P_d(\mathbf{c} \to \hat{\mathbf{c}}) \leq \frac{1}{2} \int_s \exp\left(-\frac{1}{2}\, d\, \Gamma_{0c} s \right) p_S(s)\, ds. \tag{4.17}$$

Using (4.6) and (4.13) for fast fading and (4.17) for slow fading, the performance of a space-time convolutional coded system can now be readily calculated.

4.2.2 PCCC Performance

For a turbo code with a fixed interleaver, the construction of A_d (using (4.4)) can only be accomplished through exhaustive search. The latter leads to the proposition of an average upper bound constructed by averaging over all possible interleavers [176]. Therefore, to derive this average performance bound a superfluous interleaver, called the *uniform interleaver*, is used. This interleaver, for a given input block of n bits with input weight i, outputs all $\binom{n}{i}$ distinct permutations with equal probability.

The hyper-trellis transfer function $T(I, D)$ determined by Benedetto *et al.* [193] for turbo codes in three co-decoding configurations has been evaluated in conjunction with continuous, trellis truncated and trellis terminated co-decoding. The latter showed that the performance of the truncated encoder is significantly worse than that of continuous decoding, whereas trellis termination is only slightly worse.

Using (4.6), with $p(d \mid i)$ known, the performance of space-time turbo codes can be evaluated for various channels and transceivers by formulating the conditional pairwise error probability, $P_d(\mathbf{c} \rightarrow \hat{\mathbf{c}})$, for the configuration of interest [194, 195].

The expressions for $P_d(\mathbf{c} \rightarrow \hat{\mathbf{c}})$ derived for convolutional codes, are limited to the case where the output codeword weight, d is fixed. Here, the results is extended to include the performance of turbo codes where the code weight is described in terms of an input-output conditional probability density function (cpdf), $p(d \mid i)$. In Appendix B, the Divsalar cpdf and binomial cpdf have been derived from the constituent encoder state transition matrix, $t(l, i, d)$.

Figure 4.3 shows examples of the Binomial cpdfs given by (B.15) for different code rates, R_c and turbo interleaver size, $N_{tc} = 100$.

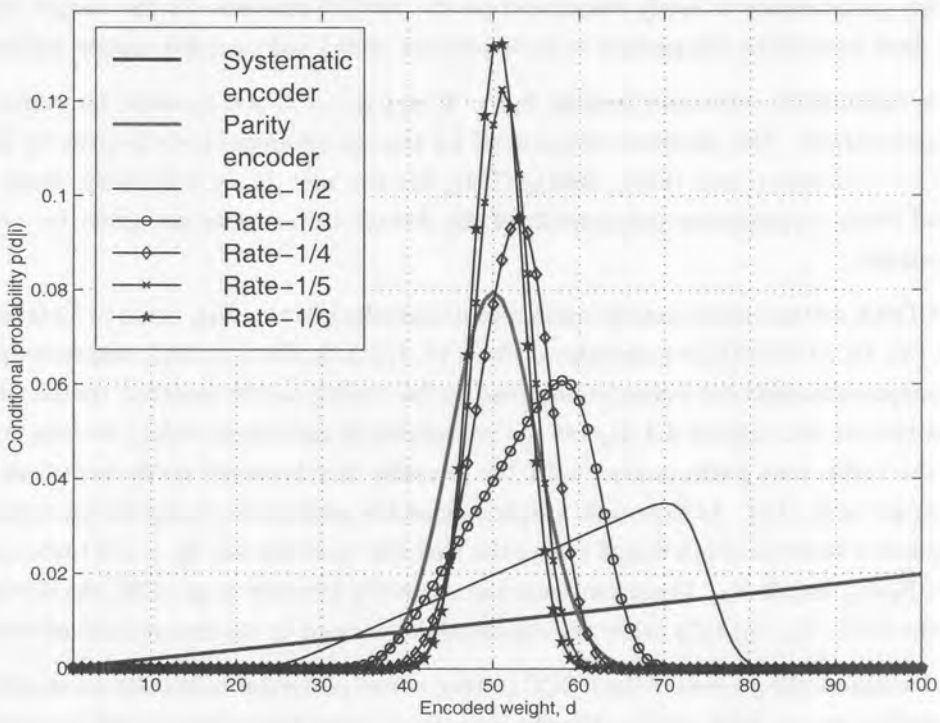

Figure 4.3. Binomial cpdfs, $p(d \mid i)$ for code rate, R_c and turbo interleaver size, $N_{tc} = 100$. (The x-axis is normalized by R_c).

Assuming an AWGN channel, the pairwise BEP of turbo codes may be written as

$$P_d(\mathbf{c} \to \hat{\mathbf{c}}) = E_{d|i}\left\{Q\left(\sqrt{d\,\Gamma_{0c}}\right)\right\}, \tag{4.18}$$

where the conditional expectation $E_{d|i}\{\cdot\}$ is over the evaluated cpdf, $\tilde{p}(d \mid i)$, as explained in Appendix B.

Employing similar arguments to that used in Section 4.2.1.1, the bound given by (4.18) can be extended to include the transmit diversity signalling and multi-path fading channel effects.

From (4.16), for the slow fading channel, the upper bound for the pairwise error probability can be written as

$$P_d(\mathbf{c} \to \hat{\mathbf{c}}) \leq E_{\mathbf{S}}\left\{E_{d|i}\left\{\frac{1}{2}\exp\left(-\frac{1}{2}\,d\,\Gamma_{0c}s\right)\right\}\right\}. \tag{4.19}$$

The conditional expectations, $E_{\mathbf{S}}$ and $E_{d|i}\{\cdot\}$ is calculated over the instantaneous fading power pdf, and the cpdf ($\tilde{p}(d \mid i)$), respectively.

Using (4.13), the bounds of (4.19) can be extended to include fast fading.

4.2.3 Numerical Results

As a benchmark and to illustrate the effect of the interference limited region associated with turbo codes, coded single user performance is firstly considered on the AWGN channel. In the results that follow an approximately fixed bandwidth comparison is made between coded and uncoded system performance.

When coding is considered, constraint lengths $L_{cc} = 9$ and $L_{tc} = 3$ are assumed for convolutional and turbo coding, respectively. The generator polynomial for the convolutional code is given by $(561)_8, (753)_8$ for the rate $R_c = 1/2$ codes, and $(557)_8, (663)_8, (711)_8$ for the rate $R_c = 1/3$ codes, respectively. The feedforward and feedback generator polynomials of the 4-state turbo codes are given by $g_{ff} = 5_8$, and $g_{fb} = 7_8$, respectively.

For the coded CDMA system under consideration a total spreading factor, N_{tot} equal to 32 is assumed. This results in $N = \{32, 16, 10, 8, 6, 5\}$ for code rates of $R_c = \{1, 1/2, 1/3, 1/4, 1/5, 1/6\}$, respectively. Single user performance comparisons between convolutional and turbo coding under identical complexity constraint requirements is carried out. Figure 4.4 depicts the convolutional and turbo coding bounds on the AWGN channel. For the turbo code performance, both the Divsalar and binomial cpdfs have been used in the calculation of P_e given in (4.6). As expected, a tighter bound is achieved by using the binomial cpdf. From Figure 4.4, it is noted that the union bound (using the Divsalar cpdf) for the $R_c = 1/2$ turbo code, diverges at low values of E_b/N_o, for all N_{tc}. Consistent with the results by Divsaler *et al.* [176], the divergence occurs roughly when the SNR (E_b/N_0) falls below the threshold determined by the computational cutoff rate R_o[3].

In an attempt to evaluate the gooness of the PCCC (turbo) codes' performance bounds a computer simulation program was used to obtain BER results. For the simulation performance the system parameters outlined in Table 4.1 were assumed. The BER performance result is also shown on Figure 4.4. In addition to the parameters outlined below, perfect synchronization, channel estimation and CSI are also assumed.

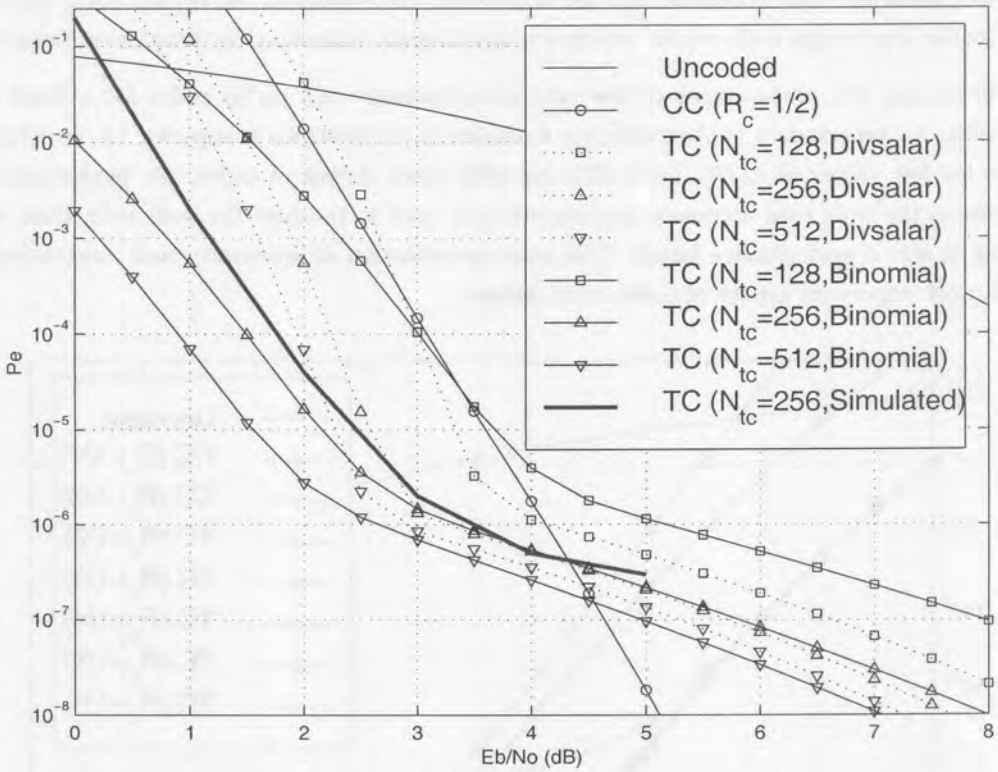

Figure 4.4. CC and TC bounds on AWGN channel with $R_c = 1/2$ and $K = 1$, as a function of turbo interleaver size, N_{tc}.

Parameter	Simulation value
Spreading sequence length	No spreading
Operating environment	AWGN channel
Code type and rate	4-state TC, $R_c = 1/2$
Interleaver	S-type, $N_{tc} = 256$
Decoder	Iterated MAP, serial configuration

Table 4.1. System parameters for BER simulation of PCCC.

Comparing the simulation curve with the performance bound, it is clear that the binomial cpdf results in an improved bound, with a slight divergence around the cutoff rate threshold. Care should be taken for values of E_b/N_o less than the cutoff rate, the performance bounds based on the binomial cpdf behaves as a lower bound to the code's simulated performance. At high SNR rations, the performance bounds based on the binomial cpdf and Divsalal cpdf become converge and both compares well with the simulated BER curve.

Focusing the attention back to the theoretical bounds, the turbo code performance reflects the expected interleaver gain in the waterfall region[4] of the performance curve. This provides an effective way to decrease the BEP without invoking any changes in the system configuration.

At high SNR ratios, the design of the interleaver in a practical system is therefore of great importance. At high values of SNR, turbo codes will perform well with almost any interleaver provided that the two

(or more) RSC encoders receive inputs that are sufficiently uncorrelated. At higher SNR, performance is dominated by the low weight code words, which are significantly influenced by interleaver design

Figure 4.5 illustrates the performance of low rate convolutional- and turbo codes for a fixed interleaver size, $N_{tc} = 256$. As expected, it is observed that turbo code performance is superior to convolutional code performance for low values of E_b/N_o (< 5 dB). As with other classes of codes, the performance of turbo codes improve as the code rate decreases. If puncturing is used to increase the code rate, then the manner of puncturing is also a performance factor. The joint optimization of interleaver and puncturing matrix is perhaps the most important aspect of turbo code design.

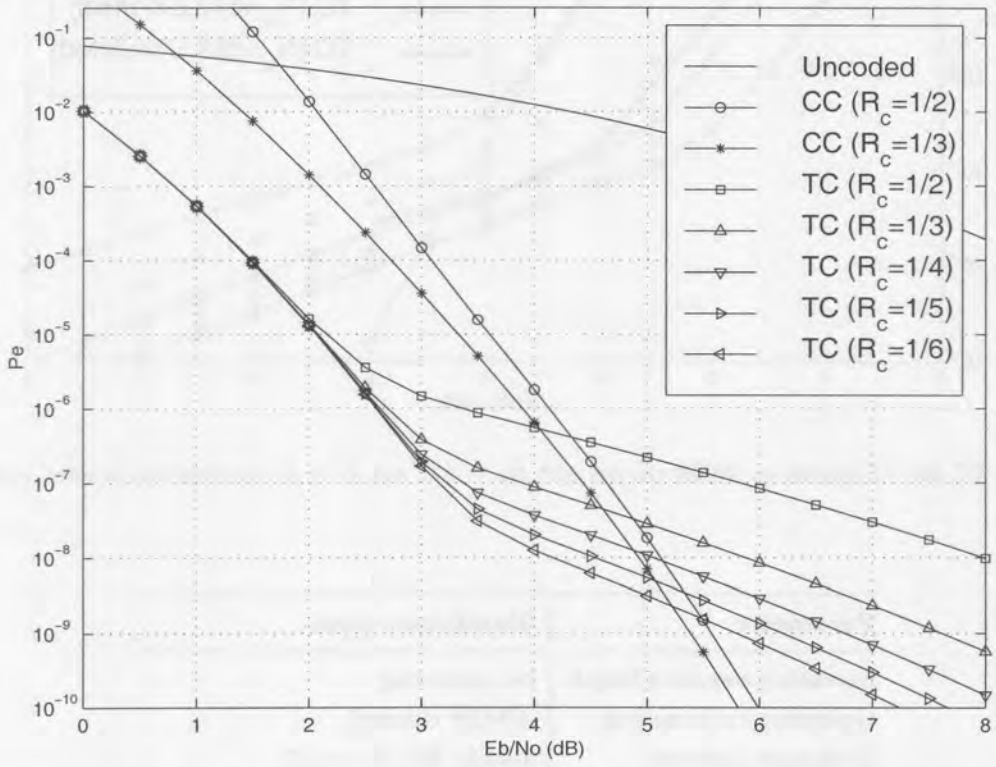

Figure 4.5. CC and TC bounds on AWGN channel with $N_{tc} = 256$ and $K = 1$, as a function of code rate.

It is noted that for a fixed interleaver size, decreasing the code rate does not significantly affect the waterfall performance region. Decreasing the code rate, however, causes the error floor region to be lowered, which may be attributed to the stronger code structure of the low rate codes. Since this region occurs at higher SNR values, the actual weight spectrum becomes more important in influencing the performance.

Since the performance of a MF-based CDMA receiver is interference limited, the uncoded BEP region of importance for a turbo code is roughly $10^{-4} < P_e < 10^{-2}$. This is the focus BEP range and coding should provide acceptable performance in this region.

For a fair comparison to an uncoded system under equal throughput and bandwidth conditions, the spreading sequence length, N, of the coded system must be shortened by a factor of $1/R_c$. This results in a degradation due to the MAI since it is well-known that the normalized cross-correlation between any two spreading sequences is proportional to the Welsh-bound, given as $1/\sqrt{N}$. A trade-off between the greater distance properties of low rate codes and increased cross-correlation effects (due to shorter sequence lengths) is fundamental to the success of coded CDMA.

As another means to investigate multiuser performance, let us define the system load as the quantity $V = K/N_{tot}$, where $N_{tot} = N/R_c$. The system load is therefore the number of active users normalized to the overall spreading factor. Figure 4.6 depicts the system load for low rate convolutional- and turbo coding. It is clear that the system load using turbo codes are substantially higher than for convolutional coding. Another interesting effect is that the coding gain for low rate turbo codes are reduced as the system load increases. This is due to the error floor effect as seen in Figures 4.4 and 4.5. It should be noted that low rate turbo coding will only exhibit this behavior at relatively high E_b/N_0. If the system is operated at low E_b/N_0, low rate turbo coding provides an increase in system capacity.

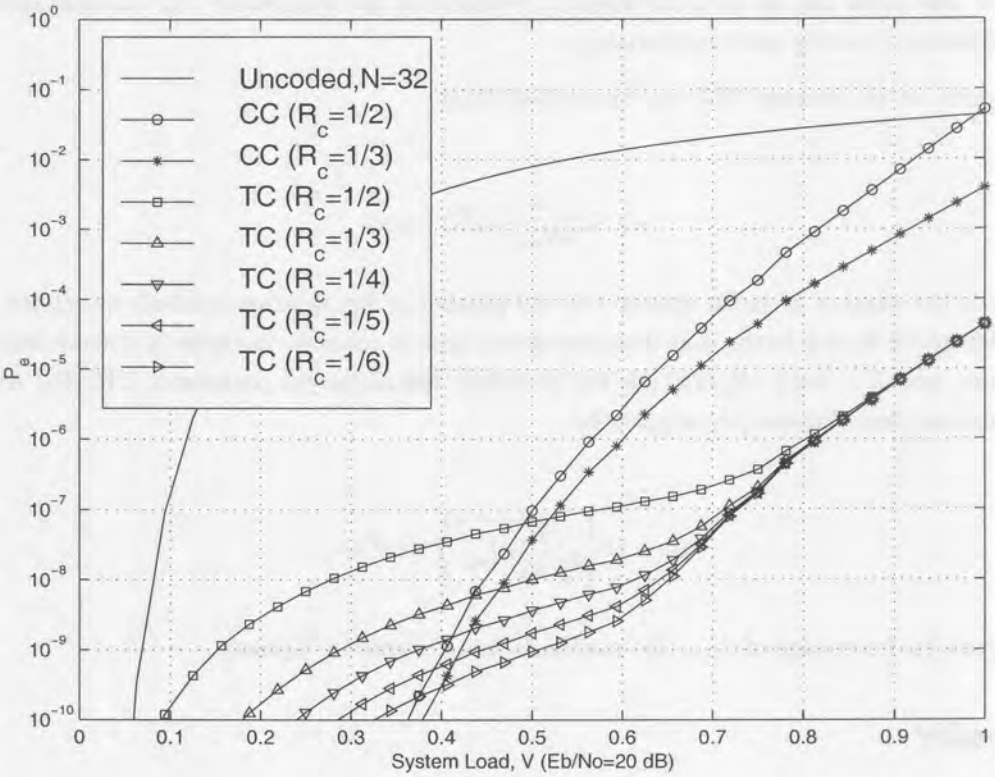

Figure 4.6. Analytical AWGN system load – low rate convolutional and turbo coding ($N_{tc} = 256$).

4.2.4 *Trellis Coded Modulation (TCM) Performance*

The sequence of trellis encoder output symbols has a very carefully controlled structure which enables the detection and correction of transmission errors in a multiple transmit/receive antenna signalling scenario. For the space-time encoder, these symbols should be designed in such a way that the combined spatial and temporal properties will guarantee maximum diversity. The channel coding may be either convolutional or turbo coded. In Chapter 6, details concerning the application and performance of space-time trellis coded modulation to cellular CDMA will be given.

In the analysis of uncoded and coded systems in multiple antenna transmission scenarios two measures of performance are commonly employed, namely SNR improvement and mutual information [196, 197]. While the two metrics are closely related, they have important differences. SNR characterizes the performance of typical uncoded systems, while mutual information measures the maximum rate of reliable communication achievable with coded systems.

In the case of ideal CSI, the transmitter or receiver is assumed to have exact knowledge of the fading channel conditions, and the decoding metric is then ML. On the other hand, when no CSI is available at the receiver, the decoding metric is no longer ML and this introduces an additional weakness into the generalized bounding procedure for TCM [21, 198, 199]. The latter situation occurs when the transmitter arrays are used in point-to-point scenarios in which the transmitter has no knowledge of the channel parameters.

It has been shown by Narula *et al.* [196], that when no information about the channel parameters is provided, beamforming cannot be used to achieve channel capacity. In addition, it was shown that when the CSI is perfect, the SNR-based design and mutual-information-based design become equivalent. All the analysis carried out in this thesis rely on accurate channel knowledge at the transmitter (for transmit diversity) and receiver (for receive diversity and beamforming).

An upper bound on the average BEP can be obtained from

$$P_e \leq \frac{k_0}{bN_{trel}} T(D) \mid_{D=Z}, \tag{4.20}$$

where N_{trel} is the number of trellis states; b is the number of information symbols associated with each branch in the trellis; k_0 is a factor that depends on the type of channel, the type of demodulation and the code structure; and $Z = \exp\{-E_s/4N_o\}$ is the so-called Bhattacharyya parameter [21]. For AWGN with optimum coherent demodulation, k_0 is given by

$$k_0 = Q\left\{\sqrt{\frac{E_s}{N_0} d_{free}^2}\right\} D^{-d_{free}^2}, \tag{4.21}$$

i.e., to compute k_0, knowledge of d_{free} the minimum free distance is required.

4.3 SUMMARY

This chapter classified, defined and discussed forward error correction techniques, including classical convolutional and turbo, and trellis codes for cellular CDMA.

The BEP performance of these codes has been addressed by the derivation of analytical average upper bounds based on the union bound and code weight distributions. Some numerical results were presented. In the following sections, the derived upper bounds will be used to evaluate the performance of space-time coded cellular CDMA systems over the channels with independent and correlated fading.

Notes

1. In this thesis the 'cc' refers to a convolutional code constraint length, the subscript 'tc' will be used to denote the turbo codes' constraint lengths.
2. In an fully interleaved scheme, fast fading is created as the de-interleaving mechanism creates a virtually memoryless channel [200].
3. The cutoff rate is defined as $E_s/N_o = R_c E_b/N_o < -\ln(2^{1-R_c} - 1)$ for a code with rate R_c [129].
4. The "waterfall region" is defined, as the part of the performance curve where the BEP decreases rapidly with increased SNR. The region where the BEP performance changes very slowly, with increased SNR, is defined as the "error floor region". (It is actually incorrect to call this a error floor, as the BEP still improves as the SNR increases).