CHAPTER FOUR

LINEAR MULTIUSER DETECTORS

The matched filter detector is a linear detector. We will now examine other linear detectors (i.e. detectors that operate on the received samples by means of an arbitrary linear transformation $\mathbf{M}$ as shown in Figure 4.1). The class of linear multiuser detectors discussed in this chapter include the decorrelating detector, the MMSE detector, and the generalized extension of the aforementioned detectors: the optimum linear multiuser detector.

The blind detectors discussed in this dissertation all have mean weight vector solutions that converge to the MMSE solution. It is thus imperative to understand the operation of the MMSE detector, so we can make meaningful comparisons between the blind detectors and the MMSE detector, especially where multipath combining is concerned. In this chapter we will briefly visit the decorrelating detector, after which we will consider the optimum linear and MMSE detectors in more detail.

The performance of the MMSE detector is evaluated by means of performance measures presented in Chapter 3. An extension of the MMSE detector model provided in [31] to the multipath case is also presented. This is done by partially utilizing the derivation in [47].

4.1 THE LINEAR DECORRELATING DETECTOR

Before we discuss the optimum linear and MMSE detectors, let us briefly and qualitatively consider the operation of the linear decorrelating detector. The decorrelating detector is relevant to a certain extent, since the MMSE detector and decorrelating detector perform the same linear transformation when noise is absent from the channel [46], [31]. This means that both detectors exhibit the same asymptotic multiuser efficiency, and both are optimally near-far resistant. The CDMA decorrelating detector first proposed by Schneider [48] is equivalent to the zero-forcing equalizer, as it attempts to

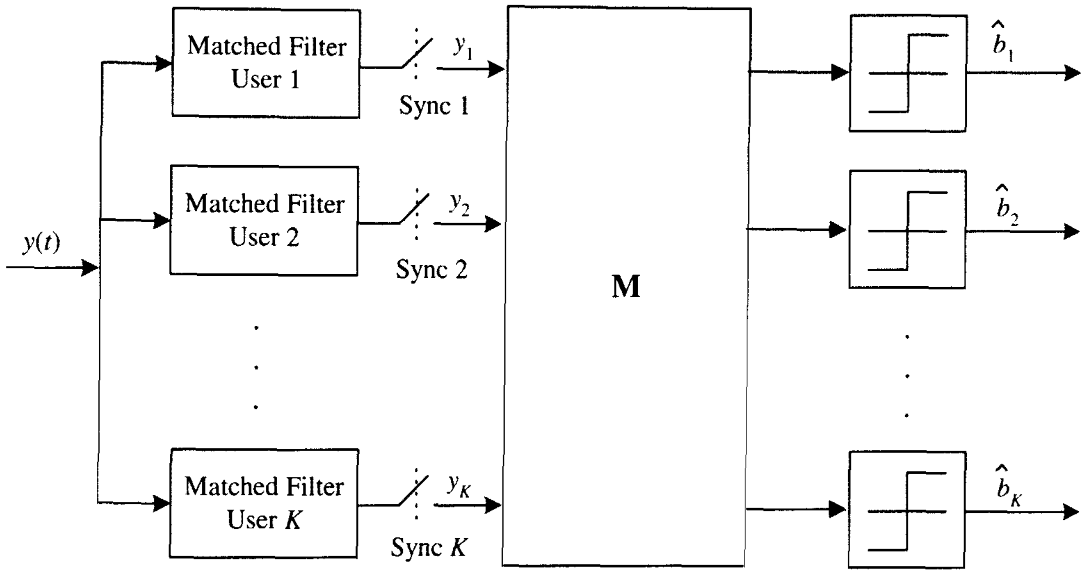

Figure 4.1: Block diagram depicting the structure of the K user linear receiver performing a linear operation $\mathbf{M}$ on the sampled matched filter outputs.

perform a linear inverse correlation matrix operation $\mathbf{M} = \mathbf{R}^{-1}$ on the received signal samples. In some cases, the correlation matrix may be singular, in which case a simple matrix inversion is not possible. A generalized inverse may then be used [46]. Analogous to zero-forcing equalization, noise enhancement may be a problem in the CDMA decorrelating detector. After Schneider, there have been several efforts to realize the decorrelating detector adaptively [49] [50].

Having now briefly visited the decorrelating detector, let us consider a generalized extension of the decorrelating detector.

4.2 THE OPTIMUM LINEAR DETECTOR

Lupas and Verdu [46] extended the MMSE and decorrelating detectors to the optimum linear detector. The class of linear detectors performs a linear transformation on the received signal vector. The optimum linear detector is the detector which maximizes the asymptotic multiuser efficiency for every vector of received amplitudes. In general, it is possible to achieve a certain tradeoff of interference rejection and attenuation of the desired signal component in order to maximize the asymptotic multiuser efficiency within the constraint of linear multiuser detection. Employing the complex vector matrix model of (2.17), let us denote the kth user linear transformation by $\mathbf{t_k}$, with

$$\hat{b}_k = \operatorname{sgn}\left(\mathbf{t}_k^H \mathbf{y}\right), \tag{4.1}$$

where $\mathbf{y}$ is the complex vector of normalized matched filter outputs. Then

$$\mathbf{t}_k^H \mathbf{y} = \sum_{j=1}^{K} A_j b_j \mathbf{t}_k^H \mathbf{r}_j + \mathbf{t}_k^H \mathbf{n}, \tag{4.2}$$

where $\mathbf{r}_j$ is the jth column of the normalized crosscorrelation matrix $\mathbf{R}$. The probability of error achieved by the transformation $\mathbf{t}_k$ can be expressed as

$$P_e^{\mathbf{t}_k} = E\left[Q\left(\frac{A_k \mathbf{t}_k^H \mathbf{r}_k + \sum_{j\neq k} A_j b_j \mathbf{t}_k^H \mathbf{r}_j}{\sigma\sqrt{\mathbf{t}_k^H \mathbf{R} \mathbf{t}_k}}\right)\right], \tag{4.3}$$

where the expectation is with respect to b_j, $j \neq k$. The asymptotic multiuser efficiency of user k is given by the square of the smallest argument of the Q-function normalized by A_k^2/σ^2, i.e.

$$\eta_k(\mathbf{t}_k) = \frac{1}{\mathbf{t}_k^H \mathbf{R} \mathbf{t}_k}\left[\max\left\{0, \mathbf{t}_k^H \mathbf{r}_k - \sum_{j\neq k} \frac{A_j}{A_k}\left|\mathbf{t}_k^H \mathbf{r}_j\right|\right\}\right]^2. \tag{4.4}$$

Due to the presence of the absolute value in (4.4), the maximization of the K-user asymptotic multiuser efficiency entails solving a nonlinear optimization problem that does not permit a closed form solution. Lupas and Verdu [46] presented an algorithm to implement the kth user maximal linear asymptotic multiuser efficiency detector. The authors also presented sufficient conditions for the best linear detector to achieve optimum kth user multiuser efficiency, as well as sufficient conditions for decorrelating detector to be the best kth user linear detector. The computational complexity of the kth user maximal linear asymptotic multiuser efficiency detector is prohibitive for a large number of users when using the algorithm mentioned above.

Although it is not possible to find a closed form solution for the kth user asymptotic multiuser efficiency in a K user channel, it is possible, however, to evaluate a closed form solution in the two user case.

4.2.1 THE TWO USER OPTIMUM LINEAR DETECTOR

We will now examine the optimization of (4.4) with respect to $\mathbf{t}_k^H$ by analyzing the two-user case. As in Chapter 3, we will restrict ourselves to the real domain, as it is instrumental in understanding and visualizing the two user linear case. Without loss of generality, if we let $\mathbf{t}_1 = [1 \;\; x]^H$, the asymptotic multiuser efficiency becomes

$$\eta_1(\mathbf{t}_1) = \left[\max\left\{0, \frac{1 + x\rho - \frac{A_2}{A_1}|x+\rho|}{\sqrt{1 + 2\rho x + x^2}}\right\}\right]^2 = \left[\max\left\{0, g\left(x, \rho, \frac{A_2}{A_1}\right)\right\}\right]^2 \tag{4.5}$$

The value of x that maximizes $g\left(x, \rho, \frac{A_2}{A_1}\right)$ is

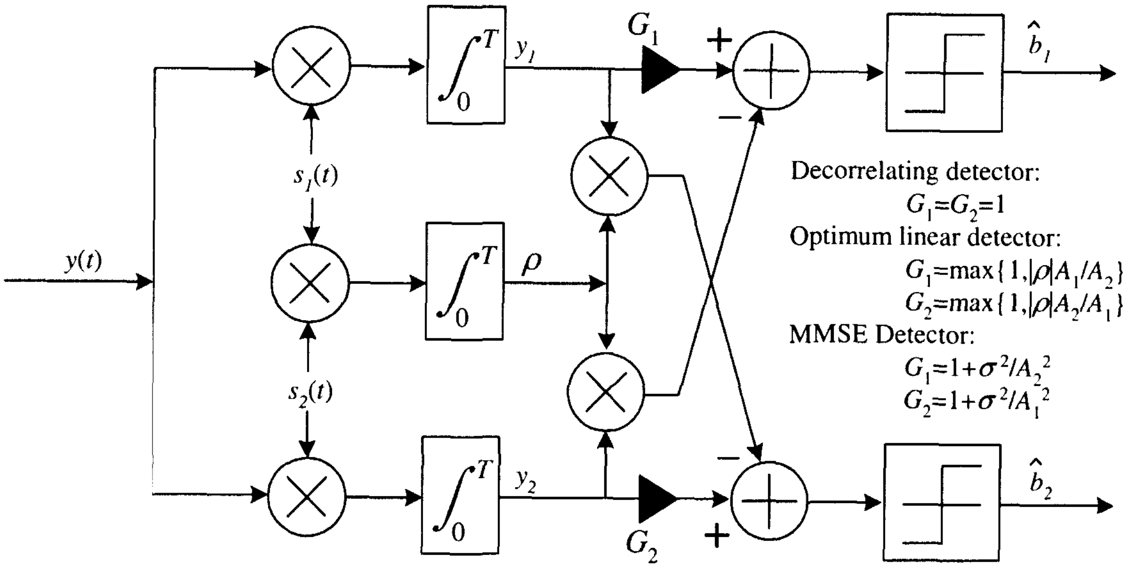

Figure 4.2: Block diagram depicting the structure of the two user linear decorrelating, optimum and MMSE receivers in the real domain.

$$\bar{x} = \begin{cases} -\frac{A_2}{A_1}\text{sgn}(\rho) & \text{if } A_2/A_1 < |\rho| \\ -\rho & \text{otherwise} \end{cases} \tag{4.6}$$

When the relative energy of the interferer is strong enough, i.e. $A_2 \geq A_1|\rho|$, then the decorrelating detector maximizes asymptotic efficiency among all linear transformations. On the other hand, if $A_2 < A_1|\rho|$, then the received signal is correlated with

$$s_1(t) - \frac{A_2}{A_1}\text{sgn}(\rho)s_2(t) \tag{4.7}$$

or equivalently with

$$\frac{A_1}{A_2}|\rho|s_1(t) - \rho s_2(t). \tag{4.8}$$

The optimum linear detector is a compromise solution between the decorrelating detector and the single user matched filter. As the relative power of the interferer decreases the optimum linear detector approaches the matched filter (Figure 4.2).

The maximum asymptotic multiuser efficiency for the two-user case is obtained by substituting (4.6) in (4.5).

$$\eta_1(\bar{\mathbf{t}}_1) = \begin{cases} 1 + \frac{A_2^2}{A_1^2} - 2|\rho|\frac{A_2}{A_1}, & \text{if } A_2/A_1 < |\rho| \\ 1 - \rho^2, & \text{otherwise} \end{cases} \tag{4.9}$$

When $A_2/A_1 < |\rho|$, the near-far resistance of the optimum linear detector is equal to that of the optimum multiuser detector in the high SNR region [46], [31]. On the other hand, when $A_2/A_1 \geq |\rho|$, there is no point, as far as near-far resistance is concerned, in utilizing the values of the received energies.

Note that the optimal linear asymptotic multiuser efficiency detector is optimal only in the high SNR region. It suffers from similar shortcomings as the decorrelating detector with respect to noise enhancement. In a low SNR environment, the single-user matched filter outperforms the optimal linear asymptotic multiuser efficiency detector. It is evident that there is room for improvement, if the noise is taken into account, concerning the performance of the optimal linear asymptotic multiuser efficiency detector. On the one hand, we have the matched filter detector that is optimized for white Gaussian noise. On the other hand, the decorrelating detector mitigates multiuser interference while disregarding the white Gaussian noise. The detector that utilizes information concerning both the SNR and MAI is the MMSE detector.

4.3 THE LINEAR MMSE DETECTOR

The adaptive MMSE detector [51], [52], [53] may solve many of the complexity and assumed knowledge problems associated with many of the other multiuser detector structures. As with matched filtering and de-correlation, the MMSE detection is a linear operation. This has the advantage that the received signal samples can be processed directly, thus simultaneously performing both the function of matched filtering and multiuser detection [7]. The MMSE detector turns the problem of multiuser detection into a problem of linear estimation. This is accomplished by minimizing a mean square error (MSE) cost criterion adaptively. The minimization can be done collectively over all users, or for each user individually.

An important quality of the MMSE detector, is that in addition to multiuser interference cancellation, it can also perform multipath (diversity) combining [54], [47], providing it has adequate filter span and that the channel inverse can be accurately modelled by a finite linear filter. In addition, the MMSE detector is successful at simultaneously mitigating narrow band interference (NBI) and MAI [55], [56]. A drawback of the MMSE detector is that a training sequence is needed to initially determine the CDMA channel conditions. After initial training, the MMSE detector can switch to its own decisions from which the MSE can be determined. This is referred to as *decision directed* mode. The imposition of training sequences implies some system overhead in the form of preamble and midamble bit sequences. The only knowledge required by the receiver is the training sequence of the user of interest. This means that the MMSE detector can be seen as a *single-user detector* capable of *multiuser interference cancellation* [7].

4.3.1 THE MMSE OPTIMIZATION PROBLEM

We start by quantitatively discussing the MMSE detector in terms of the MMSE optimization problem. Note that the same notation is used as in Chapters 2 and 3. The kth user MMSE detector chooses a complex waveform (or linear transformation) c_k of duration T that performs

$$\min_{c_k} E\left[(b_k - \langle y, c_k^* \rangle)(b_k - \langle y, c_k^* \rangle)^*\right] \tag{4.10}$$

and makes the decision

$$\hat{b}_k = \operatorname{sgn}(\langle y, c_k^* \rangle) \tag{4.11}$$

The MMSE linear transformation maximizes the SIR at the output of the linear transformation, i.e.

$$\frac{1}{\min_{c_k} E\left[(b_k - \langle y, c_k^* \rangle)(b_k - \langle y, c_k^* \rangle)^*\right]} = 1 + \max_{c_k} \frac{E\left[(\langle A_k b_k s_k, c_k^* \rangle)(\langle A_k b_k s_k, c_k^* \rangle)^*\right]}{E\left[(\langle y - A_k b_k s_k, c_k^* \rangle)(\langle y - A_k b_k s_k, c_k^* \rangle)^*\right]}. \tag{4.12}$$

In orthogonal representation, we can always express c_k as

$$c_k = c_k^s + c_k^o, \tag{4.13}$$

where c_k^s is spanned by the signature waveforms $s_1, \ldots, s_K$ and c_k^o is orthogonal to the signature waveforms. Them we have

$$E\left[(b_k - \langle y, c_k^* \rangle)(b_k - \langle y, c_k^* \rangle)^*\right] = E\left[(b_k - \langle y, c_k^{s*} \rangle)(b_k - \langle y, c_k^{s*} \rangle)^*\right] + \sigma^2 \|c_k^o\|. \tag{4.14}$$

We will restrict ourselves to c_k spanned by the signature waveforms, i.e. a weighted combination of the matched filter outputs.

To analyze the operation and performance of the MMSE detector, we will start by formulating the vector matrix model of the MMSE detector.

4.3.2 THE MMSE DETECTOR VECTOR MATRIX MODEL

Let us start with the complex vector matrix model defined in (2.17)

$$\mathbf{y} = \mathbf{RAb} + \mathbf{n}, \tag{4.15}$$

where $\mathbf{R}$ is the correlation matrix in Hermitian form, $\mathbf{A}$ is a complex diagonal matrix of the user amplitudes, and $\mathbf{n}$ is a complex valued Gaussian vector with independent real and imaginary components and with a covariance matrix equal to $2\sigma^2\mathbf{R}$. The complex bit vector is denoted by $\mathbf{b}$. The MMSE detector attempts to minimized the MSE or the difference between the actual transmitted bit vector $\mathbf{b}$ and a linear complex transformation $\mathbf{M}$ of the received signal vector $\mathbf{y}$ by adjusting the transformation $\mathbf{M}$. The transformation $\mathbf{M}$ is a $K \times K$ matrix and a K user joint optimization problem. Equivalent to the joint optimization problem, we can also have K uncoupled optimization problems (one for each user), in which case the real error cost function of user k is given by the expected value of the squared error, i.e.

$$J_k = E\left[\left(b_k - \mathbf{m}_k^H \mathbf{y}\right)^2\right], \tag{4.16}$$

where $\mathbf{m_k}$ is the kth column vector of $\mathbf{M}$. In the K user joint optimization problem, the real error cost function J is given by

$$J = E\left[(\mathbf{b} - \mathbf{M}\mathbf{y})^H (\mathbf{b} - \mathbf{M}\mathbf{y})\right] = E\left[\mathbf{e}^H\mathbf{e}\right] \tag{4.17}$$

where $\mathbf{e}$ denotes the complex error vector. Alternatively, the real error cost function is given by the trace of the covariance matrix $\mathbf{J}$ of the error vector, i. e.

$$\begin{aligned} J &= \operatorname{tr}\{\mathbf{J}\} && (4.18)\\ &= \operatorname{tr}\left\{E\left[(\mathbf{b} - \mathbf{M}\mathbf{y})(\mathbf{b} - \mathbf{M}\mathbf{y})^H\right]\right\} && (4.19)\\ &= \operatorname{tr}\{E\left[\mathbf{e}\mathbf{e}^H\right]\}. && (4.20) \end{aligned}$$

To find the complex matrix $\mathbf{M}$ that will minimize the cost function J, we will use the gradient method. This is done by partially differentiating the cost function J with respect to the complex elements of $\mathbf{M}$, equating it to zero, and solving for $\mathbf{M}$. The matrix $\mathbf{M}$ has complex elements, which can be written in the form

$$m_{vw} = x_{vw} + jy_{vw}. \tag{4.21}$$

The definition for the element of the vth row and the wth column of the *complex gradient operator* [57] matrix $\nabla_{\mathbf{M}}$ is given by

$$\nabla_{m_{vw}} = \frac{\partial}{\partial x_{vw}} + j\frac{\partial}{\partial y_{vw}}. \tag{4.22}$$

To solve the MSE cost function optimization problem, we will first apply the gradient operator to the real cost function J. The *complex gradient matrix* $\nabla_{\mathbf{M}}(J)$ is thus given by

$$\nabla_{\mathbf{M}}(J) = \begin{bmatrix} \frac{\partial J}{\partial x_{11}} + j\frac{\partial J}{\partial y_{11}} & \cdots & \frac{\partial J}{\partial x_{1K}} + j\frac{\partial J}{\partial y_{1K}} \\ \vdots & \ddots & \vdots \\ \frac{\partial J}{\partial x_{K1}} + j\frac{\partial J}{\partial y_{K1}} & \cdots & \frac{\partial J}{\partial x_{KK}} + j\frac{\partial J}{\partial y_{KK}} \end{bmatrix}, \tag{4.23}$$

where equation (4.23) represents a natural extension of the customary definition of a gradient for a function of real elements to the more general case of a function of complex elements.[1]

By letting $\nabla_{\mathbf{M}}(J) = \mathbf{0}$ and solving for $\mathbf{M}$, we will have found an expression for $\mathbf{M}$ where the error surface for each user in the K dimensional space has a minimum. To do this, let us first manipulate (4.14)

$$\begin{aligned} J &= \operatorname{tr}\left\{E\left[(\mathbf{b}-\mathbf{My})(\mathbf{b}-\mathbf{My})^H\right]\right\} \\ &= \operatorname{tr}\left\{E\left[(\mathbf{b}-\mathbf{My})(\mathbf{b}^H-\mathbf{y}^H\mathbf{M}^H)\right]\right\} \\ &= \operatorname{tr}\left\{E\left[\mathbf{bb}^H-\mathbf{by}^H\mathbf{M}^H-\mathbf{Myb}^H+\mathbf{Myy}^H\mathbf{M}^H\right]\right\} \\ &= \operatorname{tr}\left\{E\left[\mathbf{bb}^H\right]-E\left[\mathbf{by}^H\right]\mathbf{M}^H-\mathbf{M}E\left[\mathbf{yb}^H\right]+\mathbf{M}E\left[\mathbf{yy}^H\right]\mathbf{M}^H\right\}, \end{aligned} \tag{4.24}$$

where (4.24) follows from the fact that $\mathbf{M}$ is assumed to be constant. Assuming no correlation between the data of different users or between the data and noise vectors, we have

$$E\left[\mathbf{bb}^H\right] = 2\mathbf{I}, \tag{4.25}$$

$$E\left[\mathbf{by}^H\right] = 2\mathbf{AR}, \tag{4.26}$$

$$E\left[\mathbf{yb}^H\right] = 2\mathbf{RA}, \tag{4.27}$$

$$E\left[\mathbf{yy}^H\right] = 2\mathbf{RA}^2\mathbf{R}+2\sigma^2\mathbf{R}. \tag{4.28}$$

Simplifying the cost function with the above results, we have

$$\begin{aligned} J &= \operatorname{tr}\left\{2\mathbf{I}-2\mathbf{ARM}^H-2\mathbf{MAR}+2\mathbf{M}(\mathbf{RA}^2\mathbf{R}+\sigma^2\mathbf{R})\mathbf{M}^H\right\} \\ &= \operatorname{tr}\left\{2\mathbf{I}-2\mathbf{ARM}^H-2\mathbf{MAR}+2\mathbf{MRA}^2\mathbf{RM}^H+\mathbf{M}\sigma^2\mathbf{RM}^H\right\}. \end{aligned} \tag{4.29}$$

Let us now find the complex gradient matrix $\nabla_{\mathbf{M}}(J)$ of the cost function,

$$\begin{aligned} &\nabla_{\mathbf{M}}(J) \\ &= 2\tfrac{\partial}{\partial \mathbf{M}^*}\left(\operatorname{tr}\{2\mathbf{I}\}-\operatorname{tr}\left\{2\mathbf{ARM}^H\right\}-\operatorname{tr}\{2\mathbf{MAR}\}+\operatorname{tr}\left\{2\mathbf{MRA}^2\mathbf{RM}^H\right\}+\operatorname{tr}\left\{2\mathbf{M}\sigma^2\mathbf{RM}^H\right\}\right) \end{aligned} \tag{4.30}$$

where (4.30) is evaluated in Appendix C. The result of the gradient of the cost function J from Appendix C is given by

$$\nabla_{\mathbf{M}}(J) = -4\mathbf{AR}+4\mathbf{MRA}^2\mathbf{R}+4\mathbf{M}\sigma^2\mathbf{R}. \tag{4.31}$$

[1] Note that the cost function J is *not* analytic, when it is written in terms of complex filter taps. The definition of the derivative of the cost function J with respect to the complex transformation matrix $\mathbf{M}$ requires special attention. This issue is discussed in Appendix C where the relation between derivative and gradient with respect to a complex valued matrix is discussed.

To obtain the minimum on the error surface, the slope or gradient must be set equal to zero, i.e

$$\nabla_{\bar{\mathbf{M}}}(J) = -4\mathbf{AR} + 4\bar{\mathbf{M}}\mathbf{RA}^2\mathbf{R} + 4\bar{\mathbf{M}}\sigma^2\mathbf{R} = \mathbf{0} \tag{4.32}$$

with $\bar{\mathbf{M}}$ the optimum value for the linear transformation $\mathbf{M}$. Solving for $\bar{\mathbf{M}}$ we obtain

$$\bar{\mathbf{M}} = \mathbf{A}^{-1}\left(\mathbf{R} + \sigma^2\mathbf{A}^{-2}\right)^{-1}. \tag{4.33}$$

The MMSE detector outputs the following decision for user k

$$\hat{b}_k = \mathrm{sgn}\left(\frac{1}{A_k}\left[\left(\mathbf{R} + \sigma^2\mathbf{A}^{-2}\right)^{-1}\mathbf{y}\right]_k\right) \tag{4.34}$$

$$= \mathrm{sgn}\left(\left[\left(\mathbf{R} + \sigma^2\mathbf{A}^{-2}\right)^{-1}\mathbf{y}\right]_k\right). \tag{4.35}$$

Note that the dependence of the MMSE detector on received amplitudes is only through the signal-to-noise ratios A_k^2/σ^2 due to the sgn function. Because of this, we can replace the optimum linear transformation in (4.34) with

$$\tilde{\mathbf{M}} = \left(\mathbf{R} + \sigma^2\mathbf{A}^{-2}\right)^{-1}. \tag{4.36}$$

In the formulation of the MMSE detector vector matrix model, we have assumed a great deal less than in the basic CDMA model. We did not assume that the background noise is Gaussian, nor that the bits are binary valued. The only assumptions we made were that the bits were uncorrelated from user to user, that the bit and noise vectors were uncorrelated, and that $E[b_k^2] = 1$.

4.3.3 THE TWO USER MMSE DETECTOR

Once again restricting ourselves to the real domain, in the two user case we have from (4.36)

$$\left(\mathbf{R} + \sigma^2\mathbf{A}^{-2}\right)^{-1} = \left[\left(1 + \frac{\sigma^2}{A_1^2}\right)\left(1 + \frac{\sigma^2}{A_2^2}\right) - \rho^2\right]^{-1}\begin{bmatrix} 1 + \frac{\sigma^2}{A_2^2} & -\rho \\ -\rho & 1 + \frac{\sigma^2}{A_1^2} \end{bmatrix} \tag{4.37}$$

from which the two user MMSE detector follows, as shown in Figure 4.2.

4.3.4 THE LIMITING FORMS OF THE MMSE DETECTOR

The MMSE detector is a compromise between the matched filter detector and the decorrelating detector. To illustrate this, we shall investigate the linear transformation $\mathbf{M}^* = (\mathbf{R} + \sigma^2\mathbf{A}^{-2})^{-1}$ in its limiting forms as $\sigma \to 0$ and $\sigma \to \infty$. On the one hand, if $\sigma \to 0$, then $(\mathbf{R} + \sigma^2\mathbf{A}^{-2})^{-1} \to \mathbf{R}^{-1}$, which means that the MMSE detector approaches the decorrelating detector. On the other hand, if $\sigma \to \infty$, the matrix $(\mathbf{R} + \sigma^2\mathbf{A}^{-2})^{-1}$ becomes strongly diagonal, and the MMSE detector approaches the conventional matched filter detector.

The above results reinforces the statement that the asymptotic multiuser efficiency and near-far resistance of the MMSE detector is equal to that of the decorrelating detector. This is intuitive, as the asymptotic multiuser efficiency and near-far resistance performance measures are evaluated in the limit as $\sigma \to 0$.

4.3.5 THE ASYNCHRONOUS MMSE DETECTOR

The linear time invariant transfer function of the asynchronous MMSE detector for a K user CDMA channel is given by

$$\tilde{\mathbf{M}}_a = \left(\mathbf{R}^H[1]z + \mathbf{R}[0] + \sigma^2\mathbf{A}^{-2} + \mathbf{R}[1]z^{-1}\right)^{-1}. \tag{4.38}$$

This is verified in [31] parallel to the asynchronous decorrelating detector, and is the limiting form of the inverse of the equivalent correlation matrix that we would obtain for a finite frame length (refer to (2.40)). The equivalent correlation matrix for a finite frame length is in the form

$$\mathbf{R}_{a,\mathrm{MMSE}} = \begin{bmatrix} \mathbf{R}[0] + \sigma^2\mathbf{A}^{-2} & \mathbf{R}^H[1] & 0 & \cdots & 0 \\ \mathbf{R}[1] & \mathbf{R}[0] + \sigma^2\mathbf{A}^{-2} & \mathbf{R}^H[1] & \cdots & \vdots \\ 0 & \mathbf{R}[1] & \ddots & \vdots & 0 \\ \vdots & \vdots & \cdots & \mathbf{R}[0] + \sigma^2\mathbf{A}^{-2} & \mathbf{R}^H[1] \\ 0 & \cdots & 0 & \mathbf{R}[1] & \mathbf{R}[0] + \sigma^2\mathbf{A}^{-2} \end{bmatrix}. \tag{4.39}$$

4.3.6 THE WIENER FILTER CHARACTERIZATION OF THE MMSE DETECTOR

For the Wiener filter characterization of the MMSE CDMA detector we return to the synchronous case. To illustrate the operation of the Wiener filter, we will use the model of orthonormal projections as in (2.23). We will limit ourselves to the uncoupled optimization problem, where optimization is done with respect to a single user. Without loss of generality, we consider user 1 as the desired user. We will start by defining a vector $\mathbf{p}$, which is the *cross correlation vector* between the vector $\mathbf{r}$ (2.25) and desired response b_1:

$$\mathbf{p} = E\left[b_1^*\mathbf{r}\right]. \tag{4.40}$$

The optimal vector transformation $\bar{\mathbf{v}}$ that minimizes the mean square error for user 1

$$E\left[e_1 e_1^*\right] = E\left[\left(b_1 - \mathbf{v}^H \mathbf{r}\right)\left(b_1^* - \mathbf{r}^H \mathbf{v}\right)\right] \tag{4.41}$$

can be obtained by setting the gradient equal to the zero-vector, i.e.

$$E\left[b_1^* \mathbf{r} - \mathbf{r}\mathbf{r}^H \mathbf{v}\right] = \mathbf{0}, \tag{4.42}$$

where the gradient of a complex vector is again defined as in the vector case of (C.9). The first term $\mathbf{p} = E[b_1^* \mathbf{r}]$ can be simplified to

$$\mathbf{p} = E\left[b_1^* \mathbf{r}\right] = 2A_1 \mathbf{s}_1, \tag{4.43}$$

from (2.25) and from the fact that the noise and data is uncorrelated and also from the fact that the data of different users is uncorrelated. The second term $E[\mathbf{r}\mathbf{r}^H]$ was derived in (2.27) and is equal to

$$E[\mathbf{r}\mathbf{r}^H] = 2\sigma^2 \mathbf{I} + 2\sum_{k=1}^{K} A_k^2 \mathbf{s}_k \mathbf{s}_k^H. \tag{4.44}$$

Solving for $\mathbf{v}$, we obtain the optimum solution for the linear vector transform

$$\begin{aligned} \bar{\mathbf{v}} &= \left(E\left[\mathbf{r}\mathbf{r}^H\right]\right)^{-1} E\left[b_1^* \mathbf{r}\right] & (4.45)\\ &= \mathbf{C}^{-1}\mathbf{p} & (4.46)\\ &= A_1 \left[\sigma^2 \mathbf{I} + \sum_{k=1}^{K} A_k^2 \mathbf{s}_k \mathbf{s}_k^H\right]^{-1} \mathbf{s}_1, & (4.47) \end{aligned}$$

where $\mathbf{C}$ denotes the covariance matrix of the vector $\mathbf{r}$ divided by 2, and is given by

$$\mathbf{C} = \sigma^2 \mathbf{I} + \sum_{k=1}^{K} A_k^2 \mathbf{s}_k \mathbf{s}_k^H \tag{4.48}$$

Equation (4.46) is an expression of the *Wiener-Hopf* equation [57], [58]. It is beneficial to know the minimum mean-square error achievable with the detector depending on the channel noise. The MMSE is given by

$$
\begin{aligned}
J_{\min} &= E\left[\left(b_1 - \bar{\mathbf{v}}^H\mathbf{r}\right)\left(b_1^* - \mathbf{r}^H\bar{\mathbf{v}}\right)\right] \\
&= E\left[b_1 b_1^*\right] - E\left[b_1^*\bar{\mathbf{v}}^H\mathbf{r}\right] - E\left[b_1\mathbf{r}^H\bar{\mathbf{v}}\right] + E\left[\bar{\mathbf{v}}^H\mathbf{r}\mathbf{r}^H\bar{\mathbf{v}}\right] \\
&= 2 - \left\{\left(E\left[\mathbf{r}\mathbf{r}^H\right]\right)^{-1} E\left[b_1^*\mathbf{r}\right]\right\}^H E\left[\mathbf{r}\mathbf{r}^H\right]\left(E\left[\mathbf{r}\mathbf{r}^H\right]\right)^{-1} E\left[b_1^*\mathbf{r}\right] \qquad (4.49) \\
&= 2 - \left\{\left(E\left[\mathbf{r}\mathbf{r}^H\right]\right)^{-1} E\left[b_1^*\mathbf{r}\right]\right\}^H E\left[b_1^*\mathbf{r}\right] \\
&= 2 - E\left[b_1^*\mathbf{r}\right]^H \left(E\left[\mathbf{r}\mathbf{r}^H\right]\right)^{-1} E\left[b_1^*\mathbf{r}\right] \qquad (4.50) \\
&= 2 - \mathbf{p}^H\mathbf{C}^{-1}\mathbf{p} \qquad (4.51) \\
&= 2 - 2A_1^2\mathbf{s}_1^H\left[\sigma^2\mathbf{I} + \sum_{k=1}^{K} A_k^2\mathbf{s}_k\mathbf{s}_k^H\right]^{-1}\mathbf{s}_1 \qquad (4.52)
\end{aligned}
$$

The expression in (4.51) corresponds to the expression of minimum mean-squared error of the standard Wiener filter as evaluated in [57].

4.4 THE MMSE DETECTOR LEAST MEAN SQUARE (LMS) ALGORITHM

From equation (4.47) it can be seen that to determine the optimum solution, a matrix inversion needs to be performed. This is a computationally expensive operation, and other methods need to be considered to avoid this. In addition, mobile channels are time varying, and the detector needs to follow these variations. The LMS algorithm achieves the aforementioned by being simple to implement, being able to learn the channel impulse response adaptively, and being able to follow time channel variations. For correct operation of the LMS algorithm, high certainty data of the desired user must be available at the receiver. This seems like to much to ask, as the data is what we need to determine in the first place. However, this requirement can be fulfilled by sending a *training sequence* to insure initial convergence. After this, the demodulated bits have a high certainty, and can be used by the MMSE detector to follow variations in the channel. The latter mode is referred to as *decision directed* operation. In this way the MMSE detector can be adaptively implemented, but with the disadvantage of some overhead in the form of training sequences.
The operation of the LMS algorithm can be seen as a *feedback control system*. It consists of two basic processes [57], i.e.

- An *adaptive* process which involves the adaptation of the tap weights.
- A *filtering* process which involves the inner product of an input vector with the weight vector, as well as generating an estimation error which actuates the adaptive process.

The LMS algorithm is based on the method of *steepest descent*, which is one of the oldest methods of optimization. To find the minimum value of the mean squared error using the the steepest descent algorithm, we proceed as follows:

1. We begin with an initial value $\mathbf{v}[0]$ for the tap weight vector, which is an arbitrary value.
2. Using the initial or present guess, we compute the gradient vector, the real and the imaginary parts which are defined as the derivative of the mean-squared error $J[n]$, evaluated with respect to the real and imaginary parts of the tap weight vector $\mathbf{v}[n]$ at time n (or the nth iteration).
3. The next guess of the tap weight vector is computed by making a change in the initial or present guess in a direction opposite to that of the gradient vector.
4. Go back to step 2 and repeat.

If the cost function is convex, then the minimum will be found after several iterations of the above algorithm. The distance with which the next guess differs from the current guess is termed the *step size*.

Let us now examine the elements of stochastic gradient descent optimization. Suppose we wish to find the multi-dimensional parameter θ^* that minimizes the function

$$\Psi(\theta) = E\left[g\left(X, \theta\right)\right]. \tag{4.53}$$

For a step size μ, a convex function Ψ and a initial condition θ_0, it would be possible to converge to the global minimum via steepest descent

$$\theta_{j+1} = \theta_j - \mu \nabla \Psi(\theta_j). \tag{4.54}$$

If the step size is arbitrarily small, then eventually θ_j will be close enough to θ^* for all practical purposes. To speed up convergence, the step size can initially be large and progressively decreased as the algorithm converges. Other than the fact that Ψ is convex we did not invoked any structure in (4.53). In order to calculate the expected value, we need to know the distribution of X. This is not so in all cases in practice. Instead, let us assume that the algorithm is allowed to observed an independent sequence $\{X_1, X_2, \ldots\}$ where each of the random variables in the sequence has the same distribution as X. With this information we can estimate the distribution of X and also calculate an approximation to $\nabla\Psi$. This requires too much effort and a simpler approach would be to replace the expected value of the gradient by the immediate (noisy) gradient, i.e.

$$\Psi(\theta) = E\left[g\left(X, \theta\right)\right]. \tag{4.55}$$

This can be justified by the fact that although the immediate negative gradient does not necessarily point in the direction of steepest descent, the average negative gradient of a few iterations does. According to the *law of large numbers*, if the step size is infinitesimally small, the trajectory of the algorithm will very closely track the path of steepest descent. This algorithm is known as the stochastic gradient descent algorithm. In the case where the cost function is a quadratic error cost function, the stochastic gradient algorithm is known as the LMS algorithm. It is important to know that the stochastic gradient algorithm can also be used when the sequence of realizations of X is dependent, subject to the fact that the sequence is also ergodic (the time average of the immediate gradients converges to its expected value).

Applying the stochastic gradient algorithm to the MMSE case (LMS), the linear MMSE detector for user one correlates the received waveform with the signal c_1 that minimizes

$$E\left[(b_1 - \langle y, c_1^* \rangle)^2\right]. \tag{4.56}$$

How does this fit into the stochastic approximation framework that we have derived above? The function $g(X, c_1)$ is our mean square error cost, i.e.

$$g(X, c_1) = (b_1 - \langle y, c_1^* \rangle)^2, \tag{4.57}$$

where X represents the received waveform y and the bit b_1. It is easily verified that (4.57) is strictly convex in c_1. We first will consider the synchronous case, after which we will briefly address the asynchronous case. The independent identically distributed observations used in the stochastic gradient algorithm are $X_j = (b_1[j], y[j])$, where $y[j]$ is the received signal modulated by the jth bit of all the synchronous users. To specify the gradient algorithm of (4.54), all we need to do is evaluate the gradient of $(b_1 - \langle y, c_1^* \rangle)^2$ with respect to c_1, which is equal to

$$2\left(\langle y, c_1^* \rangle - b_1\right) y. \tag{4.58}$$

We thus conclude that, in practice, the update algorithm is simply

$$c_1[j] = c_1[j-1] - \mu\left(\langle y[j], c_1^*[j-1] \rangle - b_1[j]\right) y[j]. \tag{4.59}$$

Since in practice we are working with a finite dimensional vector implementation of the adaptive law, a few things need to be pointed out. If the signature waveforms are known, then the dimensionality of the adaptive vector need not be larger than K. We know that the MMSE receiver does not need to know the transmitted signature vectors. Fortunately, by using a finite dimensional basis known to span all received signature waveforms (such as chip-matched filters), there will be sufficient dimensionality to implement our linear adaptive LMS algorithm. It is furthermore sufficient to sample at the Nyquist rate for approximately band limited chip waveforms in both the synchronous and asynchronous cases. Our LMS adaptation algorithm in finite vector form is then given by

$$\mathbf{v}_1[n] = \mathbf{v}_1[n-1] - \mu\left(\mathbf{v}^H[n-1]\mathbf{r}[n] - b_1[n]\right)\mathbf{r}[n]. \tag{4.60}$$

Global convergence of the LMS algorithm is shown in [31], subject to a sufficient decrease in step size as the algorithm progresses. The maximum step size to ensure convergence at any moment is given by

$$\mu_{\max} = \frac{2}{\sigma^2 + \lambda_{\max}}, \tag{4.61}$$

where $\lambda_{\max}$ is the maximum eigenvalue of $\sum_{k=1}^{K} A_k^2 \mathbf{s}_k \mathbf{s}_k^H$. To retain acceptable performance in the asynchronous case, we need to lengthen the observation window that spans more than one bit period. This implies that the inner product in the penalty function (4.56) is taken over the whole truncated window. This does not affect the convexity of the cost function, allowing the detector to converge to the MMSE solution.

It is expected that the detector will converge to the MMSE solution if the interference is constant. When these parameters are slowly time varying, it is still possible for an adaptive detector to follow these variations. In the case of a new user suddenly being powered on, the decisions might be unreliable in decision directed mode, and the desired user might not converge. In this case, the desired user will then request for the training sequence to be retransmitted. This implies more overhead, and is undesirable. It is for this reason that *blind* multiuser detectors (such as the constant modulus detector) warrant some investigating. Instead of using data (or decision directed) to adapt, the blind detectors utilize the cyclostationarity in the signature waveforms to minimize some given criterion.

4.5 PERFORMANCE OF THE MMSE DETECTOR

In this section we will consider the performance of the MMSE detector, using some of the measures in Chapter 3 to evaluate the detector.

4.5.1 SIGNAL-TO-INTERFERENCE RATIO OF THE MMSE DETECTOR

To derive the SIR of the MMSE detector, we start by defining the covariance matrix of the interference as

$$\mathbf{\Omega} \stackrel{\text{def}}{=} \sigma^2 \mathbf{I} + \sum_{k=2}^{K} A_k^2 \mathbf{s}_k \mathbf{s}_k^H. \tag{4.62}$$

Note that user 1 is excluded from the sum. We can now write the optimum MMSE transformation of (4.47) and the MMSE of (4.52) as

$$\bar{\mathbf{v}} = \frac{A_1}{1 + A_1^2 \mathbf{s}_1^H \boldsymbol{\Omega}^{-1} \mathbf{s}_1} \boldsymbol{\Omega}^{-1} \mathbf{s}_1 \tag{4.63}$$

and

$$J_{\text{min}} = \frac{2}{1 + A_1^2 \mathbf{s}_1^H \boldsymbol{\Omega}^{-1} \mathbf{s}_1}. \tag{4.64}$$

These two results follow from the fact that

$$\left[\boldsymbol{\Omega} + A_1^2 \mathbf{s}_1 \mathbf{s}_1^H\right]^{-1} = \left[1 + A_1^2 \mathbf{s}_1^H \boldsymbol{\Omega}^{-1} \mathbf{s}_1\right]^{-1} \boldsymbol{\Omega}^{-1}, \tag{4.65}$$

which can be proven using the *matrix inversion lemma* [57] or also known as *Woodbury's identity* [7].[2] Using the above results, and remembering that $\bar{\mathbf{v}}$ achieves the maximum output signal to interference ratio of all linear detectors, the SIR of user 1 can be written as

$$\begin{aligned}
\gamma_{c1} &= \frac{E\left[\left(A_1 b_1 \bar{\mathbf{v}}^H \mathbf{s}_1\right)\left(A_1 b_1 \bar{\mathbf{v}}^H \mathbf{s}_1\right)^H\right]}{E\left[\left(\bar{\mathbf{v}}^H \left(\mathbf{r} - A_1 b_1 \mathbf{s}_1\right)\right)\left(\bar{\mathbf{v}}^H \left(\mathbf{r} - A_1 b_1 \mathbf{s}_1\right)\right)^H\right]} && (4.66) \\
&= \frac{E\left[\left(A_1 b_1 \bar{\mathbf{v}}^H \mathbf{s}_1\right)\left(A_1^* b_1^* \mathbf{s}_1^H \bar{\mathbf{v}}\right)\right]}{E\left[\left(\bar{\mathbf{v}}^H \mathbf{r} - A_1 b_1 \bar{\mathbf{v}}^H \mathbf{s}_1\right)\left(\mathbf{r}^H \bar{\mathbf{v}} - A_1^* b_1^* \mathbf{s}_1^H \bar{\mathbf{v}}\right)\right]} \\
&= \frac{2}{E\left[\left(b_1 - \bar{\mathbf{v}}^H \mathbf{r}\right)\left(b_1^* - \mathbf{r}^H \bar{\mathbf{v}}\right)\right]} - 1 && (4.67) \\
&= A_1^2 \mathbf{s}_1^H \boldsymbol{\Omega}^{-1} \mathbf{s}_1, && (4.68)
\end{aligned}$$

From (4.66), it can be seen that the SIR is the expectation of the squared linear transformation of the desired user contribution divided by the expectation of the squared linear transformation of the interferers' contribution.

4.5.2 ASYMPTOTIC MULTIUSER EFFICIENCY AND NEAR-FAR RESISTANCE OF THE MMSE DETECTOR

Since the operation of the decorrelating detector and that of the MMSE detector are identical in a noiseless environment, they have the same asymptotic multiuser efficiency and near-far resistance. The asymptotic multiuser efficiency of the MMSE (or decorrelating) detector is given in [46] by

[2]The matrix inversion lemma states that for positive definite square matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{D}$ related by $\mathbf{A} = \mathbf{B}^{-1} + \mathbf{C}\mathbf{D}^{-1}\mathbf{C}^H$, the inverse of $\mathbf{A}$ is given by $\mathbf{A}^{-1} = \mathbf{B} - \mathbf{B}\mathbf{C}(\mathbf{D} + \mathbf{C}^H\mathbf{B}\mathbf{C})^{-1}\mathbf{C}^H\mathbf{B}$.

$$\bar{\eta}_k = 1 - \mathbf{a}_k^H \mathbf{R}_k^+ \mathbf{a}_k \tag{4.69}$$

$$= \frac{1}{R_{kk}^+}, \tag{4.70}$$

where $\mathbf{R}^+$ is the *Moore-Penrose* generalized inverse and denotes the inverse of a singular (or non-singular) square matrix $\mathbf{R}$.[3] The subscript k of $\mathbf{R}_k$ denotes the removal of the kth row and the kth column from the matrix $\mathbf{R}$. The vector $\mathbf{a}_k$ is the kth column of $\mathbf{R}$ with the kth entry removed and contains the correlations between the kth user and all other users. The value R_{kk}^+ is the element of the kth row and kth column of the generalized inverse of $\mathbf{R}$.

From Figure (4.3), it is evident that the asymptotic multiuser efficiency defined in (4.70) does not depend on the amplitude of the interfering user. This implies that the MMSE detector's asymptotic multiuser efficiency and near-far resistance are all exactly equal.
In the asynchronous case, Verdu [31] shows the near-far resistance to be

$$\bar{\eta}_k = \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\mathbf{R}^H[1]e^{j\omega} + \mathbf{R}[0] + \mathbf{R}[1]e^{-j\omega} \right]_{kk}^{+} d\omega \right)^{-1}. \tag{4.71}$$

Lupas and Verdu showed in [46] that the near-far resistance of the MMSE, optimum linear and decorrelating detector is equal to that of the optimum (non-linear) multiuser detector if the desired user is linearly independent from the other users.

4.5.3 BEP OF THE MMSE DETECTOR

The decorrelating detector is only an optimization with respect to interference, whereas the MMSE detector is an optimization with respect to the combined contribution of noise and interference. This effectively means that the MMSE transformation will inevitably allow some residual multiuser interference to remain. The consequence of this is that the derivation of the MMSE detector BEP is similar to that of the single user matched filter. As in the case of the single user matched filter, the decision statistic depends on the sum of a Gaussian random variable (due to AWGN) and a binomial random variable (due to residual multiple access interference). In the synchronous case, the first user MMSE decision statistic can be written as

$$\left(\tilde{\mathbf{M}}\mathbf{y}\right)_1 = \left(\left(\mathbf{R} + \sigma^2 \mathbf{A}^{-2}\right)^{-1} \mathbf{y} \right) \tag{4.72}$$

$$= B_1 \left(b_1 + \sum_{k=2}^{K} \beta_k b_k \right) + \sigma \hat{n}_1, \tag{4.73}$$

[3] A generalized inverse $\mathbf{C}$ of a matrix $\mathbf{B}$ is any matrix that satisfies: $\mathbf{CBC} = \mathbf{C}$ and $\mathbf{BCB} = \mathbf{B}$. The Moore-Penrose generalized inverse is the unique inverse for which $\mathbf{BC}$ and $\mathbf{CB}$ are symmetric. It follows that if $\mathbf{B}$ is a square non-singular matrix, then its Moore-Penrose generalized inverse is $\mathbf{B}^{-1}$.

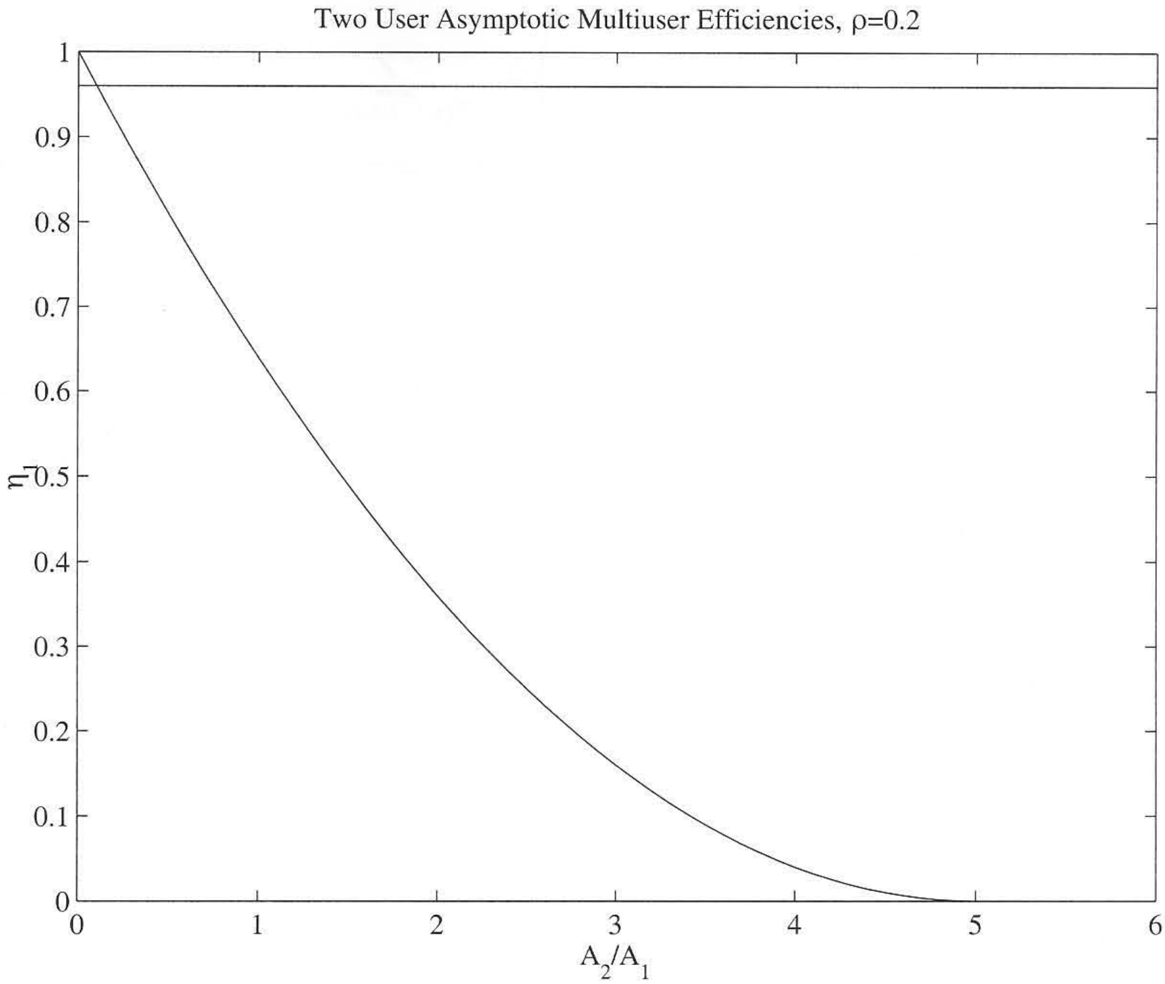

Figure 4.3: Asymptotic Multiuser Efficiencies of the Matched Filter, Decorrelating and MMSE Detectors.

where

$$B_k = A_k \left(\tilde{\mathbf{M}}\mathbf{R}\right)_{1k}, \tag{4.74}$$

$$\beta_k = \frac{B_k}{B_1}, \tag{4.75}$$

$$\hat{n}_1 \sim \mathcal{N}\left(0, \left(\tilde{\mathbf{M}}\mathbf{R}\tilde{\mathbf{M}}\right)_{11}\right). \tag{4.76}$$

The symbol β_k denotes a measure of the residual interference of the kth interferer, and is termed the *leakage coefficient*. The Gaussian noise random variable is denoted by $\hat{n}_1$ and the binomial random variable is denoted by the sum in (4.73). The probability of error is given by

$$P_e(\sigma,1) = 2^{1-K} \sum_{b_1,\ldots,b_K \in \{-1,1\}^{K-1}} Q\left(\frac{A_1}{\sigma}\frac{(\tilde{\mathbf{M}}\mathbf{R})_{11}}{\sqrt{(\tilde{\mathbf{M}}\mathbf{R}\tilde{\mathbf{M}})_{11}}}\left(1+\sum_{k=2}^{K}\beta_k b_k\right)\right). \tag{4.77}$$

We face a similar problem as in the case of exact computation of the single user BEP, in that the number of computations grow exponentially with the number of active users. This is further complicated by the computation of the leakage coefficients. We will now apply the Gaussian approximation method to the MMSE case.

4.5.3.1 GAUSSIAN APPROXIMATION OF THE MMSE DETECTOR BEP

The Gaussian approximation method is surprisingly accurate when applied to the BEP of the MMSE detector. This is done by replacing the multiple access interference by a Gaussian random variable with identical variance, i.e. $Q(\text{SIR}_1)$. We can use (4.68) together with (A.6) in Appendix A:

$$E\left[Q\left(\mu+\lambda X\right)\right] = Q\left(\frac{\mu}{\sqrt{1+\lambda^2}}\right), \tag{4.78}$$

where X is unit normal,

$$\mu = \frac{A_1}{\sigma}\frac{(\tilde{\mathbf{M}}\mathbf{R})_{11}}{\sqrt{(\tilde{\mathbf{M}}\mathbf{R}\tilde{\mathbf{M}})_{11}}} \tag{4.79}$$

and

$$\lambda^2 = \mu^2 \sum_{k=2}^{K} \beta_k. \tag{4.80}$$

Let us verify the accuracy of the approximation on an intuitive basis. We will qualitatively evaluate the deviation from Gaussianity of the decision statistic for the two limiting cases of $\sigma \to 0$ and $\sigma \to \infty$. As $\sigma \to 0$, the leakage coefficients disappear, removing the contribution of the binomial random variable. On the other hand, as $\sigma \to \infty$, the Gaussian noise contribution at the output of the transformation dominates the multiple access interference. In both cases, the decision statistic appears asymptotically Gaussian. The accuracy of the MMSE Gaussian approximation method is verified by several analytical results in [59]. Figure 4.4 depicts the accuracy of the Gaussian approximation BEP for the MMSE detector when compared with the exact calculated BEP. In [59] it also showed that the MMSE BEP is upper bounded by the decorrelating detector BEP.

Another expression of the Gaussian approximated BEP in terms of $J_{\min}$ is given in [7] as

$$P_e(\sigma,1) \approx Q\left(\sqrt{\frac{1-J_{\min}}{J_{\min}}}\right). \tag{4.81}$$

A further approximation of (4.81) is given by

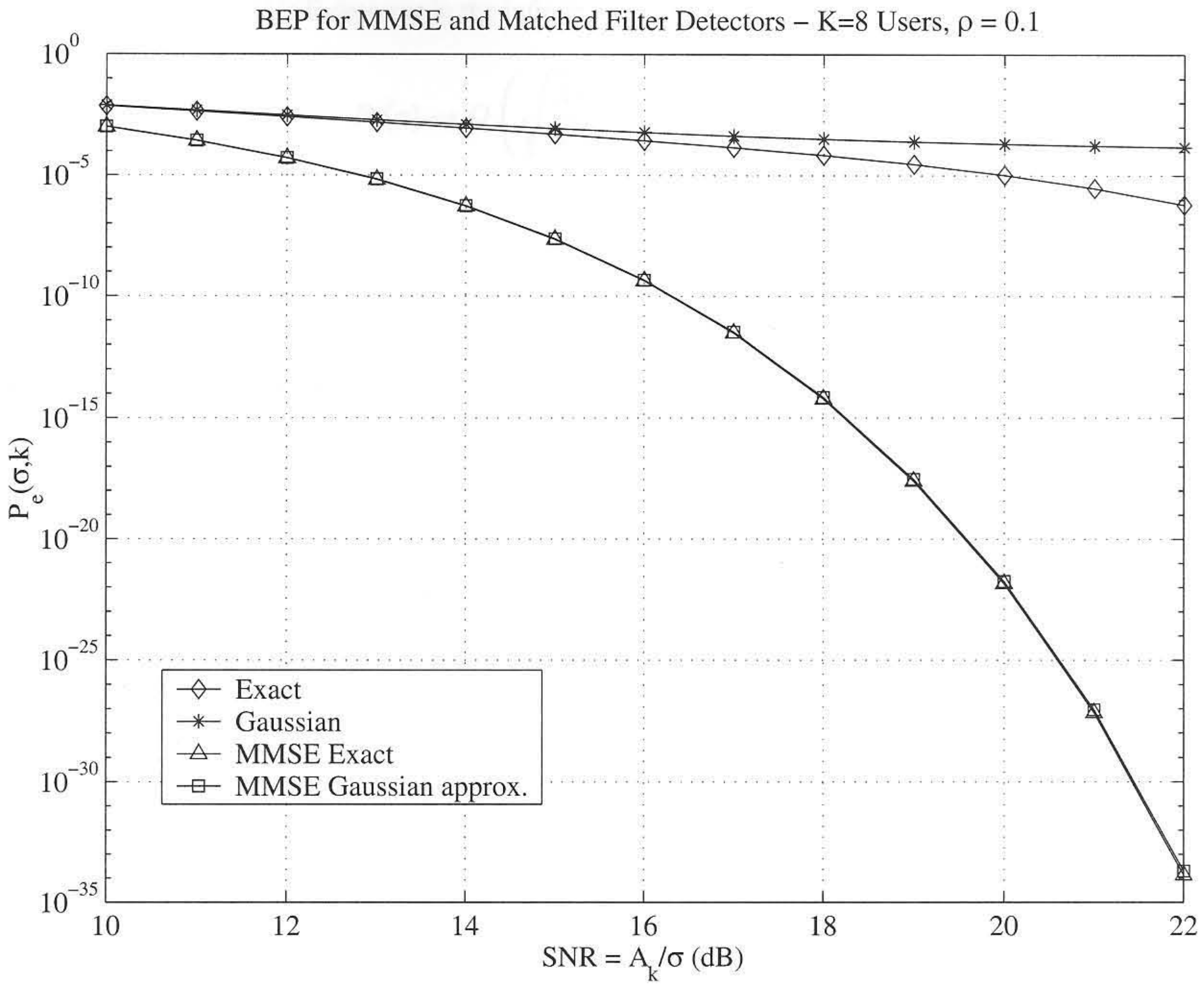

Figure 4.4: BEP graph comparing the exact and Gaussian approximated curves of the MF and MMSE detectors.

$$P_e(\sigma, 1) \approx Q\left(\frac{1}{\sqrt{J_{\min}}}\right). \tag{4.82}$$

4.5.3.2 INFINITE USER LIMIT OF THE MMSE DETECTOR BEP

The infinite user limit BEP is of interest when we consider averaging over random binary sequences. The derivation is rather involved, and we will supply only the result as stated in [31]. It is assumed that all the users have equal power. If the ratio of the number of users to the spreading gain is, or converges to, a constant

$$\beta = \lim_{K\to\infty} \frac{K}{N}, \quad \beta \in (0, +\infty), \tag{4.83}$$

then the BEP of the MMSE detector in the infinite user limit ($K \to \infty$) is given by

$$P_e(\sigma) \to Q\left(\sqrt{\frac{A^2}{\sigma^2} - \frac{1}{4}\mathcal{F}\left(\frac{A^2}{\sigma^2}, \beta\right)}\right), \tag{4.84}$$

where

$$\mathcal{F}(x, z) \stackrel{\text{def}}{=} \left(\sqrt{x\left(1+\sqrt{z}\right)^2+1} - \sqrt{x\left(1-\sqrt{z}\right)^2+1}\right)^2. \tag{4.85}$$

4.5.4 POWER TRADEOFF REGIONS OF THE MMSE DETECTOR

Using the results in the previous sections, we can now determine the power tradeoff regions of the MMSE (and related blind) detectors in the real two user scenario. In Figure 4.5 it can be seen that for all but very high cross correlation values ρ, the SNR needed to attain a BEP of less than 3×10^{-5} for both users is slightly above 12dB for a two user system. If we compare this to the matched filter case in Figure 3.6, we find that the SNR needed does not increase along with the interfering user's amplitude. This means that the MMSE detector is effective in mitigating the near far problem, and the interferer's power has no effect on the desired user's bit rate.

4.5.5 MMSE DETECTOR PERFORMANCE IN MULTIPATH CHANNELS

Having looked at the performance of the MMSE detector in synchronous (non-multipath) channels, we will now consider how the detector operates in multipath channels. An extensive evaluation of the performance of the MMSE detector in a multipath environment was done in [47]. We will follow a similar approach using an asynchronous version of our orthonormal projection model in (2.25). We are interested in the performance of the MMSE detector both in terms of minimum mean-square error and BEP. Concerning our derivation, the following important assumptions are made:

1. The received signal window length is equal to one symbol period;
2. No multipath component is later than one symbol period;
3. The receiver is synchronized to the first multipath component;

With this in mind, we can visualize the multipaths of user k as depicted in Figure 4.6.
Within the received signal window, any two multipath components have a correlated part due to the present bit, and an uncorrelated part due to the preceding bit of the later path. The correlated part can be seen as part of the desired signal and a useful diversity component. The uncorrelated part belongs to the preceding bit, which can be viewed as interference. Using our existing model, we will now derive the MMSE in the case of multipath.

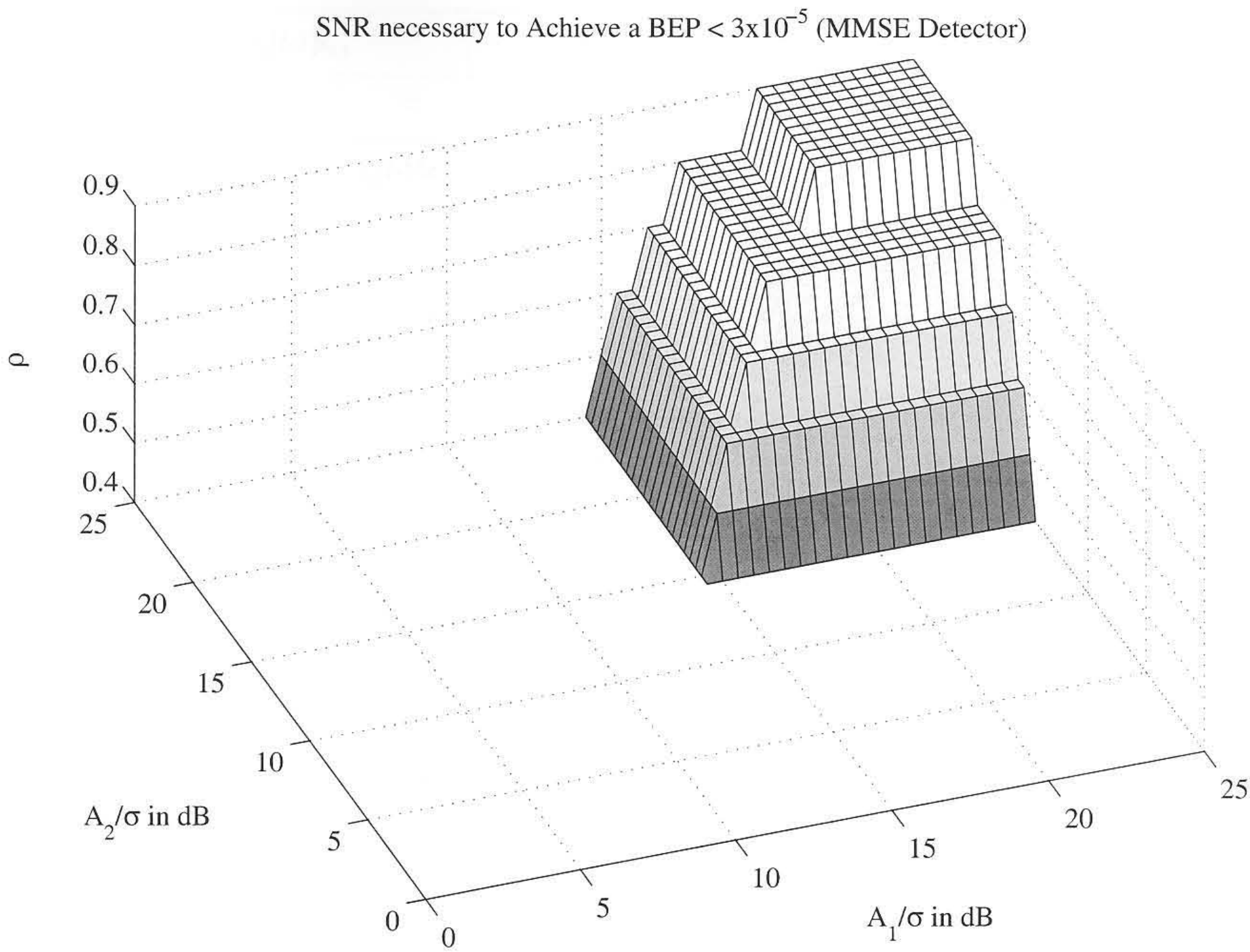

Figure 4.5: Regions of signal-to-noise ratios to attain a BEP of 3×10^{-5} for both users using a MMSE detector.

Remember from (2.59) that when the received signal window is one symbol long, we have

$$y(t)[i] = \sum_{k=1}^{K} \sum_{p=0}^{P-1} A_{k,p} b_k[i] s(\tau - \tau_p) \exp(-j\theta_{k,p}) + \sigma n(t). \tag{4.86}$$

Formulating an asynchronous version of equation (2.25), we have

$$\mathbf{r} = \sum_{k=1}^{K} \sum_{p=0}^{P-1} \left(\tilde{A}_{k,p} b_k[i] \mathbf{s}_{k,p}^{R} + \tilde{A}_{k,p} b_k[i-1] \mathbf{s}_{k,p}^{L} \right) + \sigma \mathbf{m}, \tag{4.87}$$

where the term $\tilde{A}_{k,p}$ refers to the complex amplitude due to the phase term $\theta_{k,p}$ in (4.86), P is the number of resolvable multipaths, and

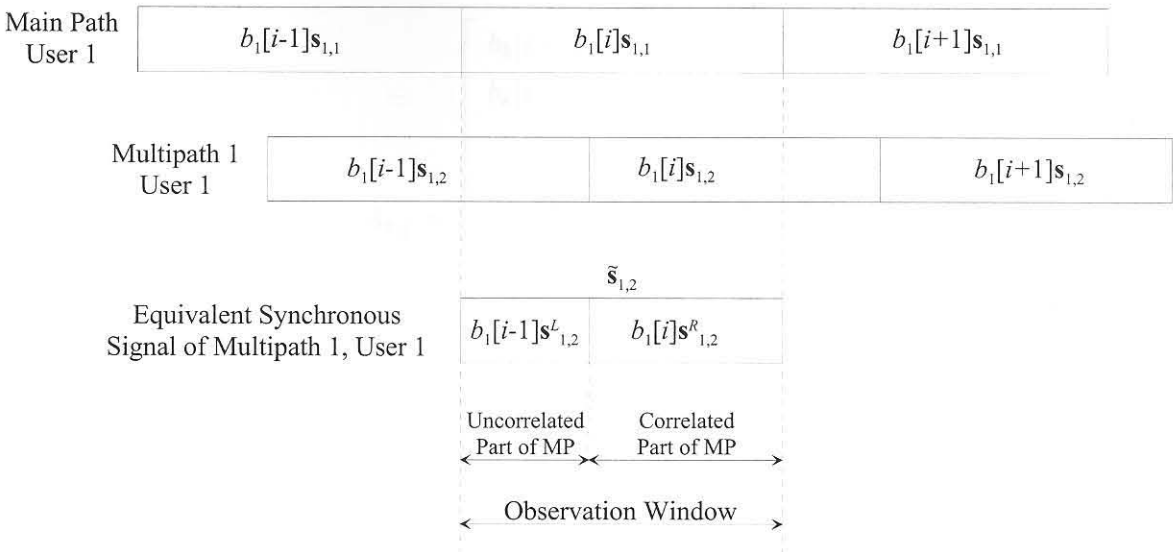

Figure 4.6: Depiction of the equivalent synchronous multipath model of a CDMA channel.

$$\mathbf{s}_{k,p}^{L} = \begin{bmatrix} s_{k(L-D_p+1)} \\ s_{k(L-D_p+2)} \\ \vdots \\ s_{kL} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{4.88}$$

$$\mathbf{s}_{k,p}^{R} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ s_{k1} \\ s_{k2} \\ \vdots \\ s_{k(L-D_p)} \end{bmatrix}. \tag{4.89}$$

The symbol s_{kl} defined in (2.23) denotes the the projection of the lth orthonormal signal on the signature waveform of user k. The symbol D_p denotes the delay of the pth multipath. We can let

$$\tilde{\mathbf{s}}_{k,p} = \begin{bmatrix} b_k[i-1]s_{k(L-D_p+1)} \\ b_k[i-1]s_{k(L-D_p+2)} \\ \vdots \\ b_k[i-1]s_{kL} \\ b_k[i]s_{k1} \\ b_k[i]s_{k2} \\ \vdots \\ b_k[i]s_{k(L-D_p)} \end{bmatrix}, \tag{4.90}$$

then

$$\mathbf{r} = \sum_{k=1}^{K}\sum_{p=0}^{P-1} \tilde{A}_{k,p}\tilde{\mathbf{s}}_{k,p} + \sigma\mathbf{m}. \tag{4.91}$$

Analogous to (2.27) we have the covariance matrix

$$E\left[\mathbf{r}\mathbf{r}^H\right] = 2\sigma^2\mathbf{I} + E\left[2\sum_{k=1}^{K}\sum_{p=0}^{P-1}\left|\tilde{A}_{k,p}\right|^2 \tilde{\mathbf{s}}_{k,p}\tilde{\mathbf{s}}_{k,p}^H\right] \tag{4.92}$$

$$= E\left[2\left|\tilde{A}_{1,1}\right|^2 \mathbf{s}_1\mathbf{s}_1^H + 2\sum_{p=1}^{P-1}\left|\tilde{A}_{1,p}\right|^2 \tilde{\mathbf{s}}_{1,p}\tilde{\mathbf{s}}_{1,p}^H + 2\sum_{k=2}^{K}\sum_{p=0}^{P-1}\left|\tilde{A}_{k,p}\right|^2 \tilde{\mathbf{s}}_{k,p}\tilde{\mathbf{s}}_{k,p}^H\right] + 2\sigma^2\mathbf{I}$$

$$= E\left[2\left|\tilde{A}_{1,1}\right|^2 \mathbf{s}_1\mathbf{s}_1^H + 2\sum_{p=1}^{P-1}\left|\tilde{A}_{1,p}\right|^2 \left(b_1[i-1]\mathbf{s}_{1,p}^L + b_1[i]\mathbf{s}_{1,p}^R\right)\left(b_1[i-1]\mathbf{s}_{1,p}^{L\ H} + b_1[i]\mathbf{s}_{1,p}^{R\ H}\right)\right.$$

$$\left. +2\sum_{k=2}^{K}\sum_{p=0}^{P-1}\left|\tilde{A}_{k,p}\right|^2 \left(b_k[i-1]\mathbf{s}_{k,p}^L + b_k[i]\mathbf{s}_{k,p}^R\right)\left(b_k[i-1]\mathbf{s}_{k,p}^{L\ H} + b_k[i]\mathbf{s}_{k,p}^{R\ H}\right)\right] + 2\sigma^2\mathbf{I}$$

$$= 2\sigma^2\mathbf{I} + 2\left|\tilde{A}_{1,1}\right|^2 \mathbf{s}_1\mathbf{s}_1^H + 2\sum_{p=1}^{P-1}\left|\tilde{A}_{1,p}\right|^2 \mathbf{s}_{1,p}^R\mathbf{s}_{1,p}^{R\ H} + 2\sum_{p=1}^{P-1}\left|\tilde{A}_{1,p}\right|^2 \mathbf{s}_{1,p}^L\mathbf{s}_{1,p}^{L\ H}$$

$$+E\left[2\sum_{p=1}^{P-1}\left|\tilde{A}_{1,p}\right|^2 b_1[i-1]b_1[i]\mathbf{s}_{1,p}^L\mathbf{s}_{1,p}^{R\ H}\right] \tag{4.93}$$

$$+E\left[2\sum_{p=1}^{P-1}\left|\tilde{A}_{1,p}\right|^2 b_1[i]b_1[i-1]\mathbf{s}_{1,p}^R\mathbf{s}_{1,p}^{L\ H}\right] \tag{4.94}$$

$$+2\sum_{k=2}^{K}\sum_{p=0}^{P-1}\left|\tilde{A}_{k,p}\right|^2 \mathbf{s}_{k,p}^L\mathbf{s}_{k,p}^{L\ H} + 2\sum_{k=2}^{K}\sum_{p=0}^{P-1}\left|\tilde{A}_{k,p}\right|^2 \mathbf{s}_{k,p}^R\mathbf{s}_{k,p}^{R\ H}$$

$$+E\left[2\sum_{k=2}^{K}\sum_{p=0}^{P-1}\left|\tilde{A}_{k,p}\right|^2 b_k[i-1]b_k[i]\mathbf{s}_{k,p}^L\mathbf{s}_{k,p}^{R\ H}\right] \tag{4.95}$$

$$+E\left[2\sum_{k=2}^{K}\sum_{p=0}^{P-1}\left|\tilde{A}_{k,p}\right|^2 b_k[i]b_k[i-1]\mathbf{s}_{k,p}^R\mathbf{s}_{k,p}^{L\ H}\right] \tag{4.96}$$

$$= \underbrace{2\sigma^2\mathbf{I}}_{A} + \underbrace{2\left|\tilde{A}_{1,1}\right|^2 \mathbf{s}_1\mathbf{s}_1^H}_{B} + \underbrace{2\sum_{p=1}^{P-1}\left|\tilde{A}_{1,p}\right|^2 \mathbf{s}_{1,p}^R\mathbf{s}_{1,p}^{R\ H}}_{C} + \underbrace{2\sum_{p=1}^{P-1}\left|\tilde{A}_{1,p}\right|^2 \mathbf{s}_{1,p}^L\mathbf{s}_{1,p}^{L\ H}}_{D}$$

$$\underbrace{+2\sum_{k=2}^{K}\sum_{p=0}^{P-1}\left|\tilde{A}_{k,p}\right|^2 \mathbf{s}_{k,p}^L\mathbf{s}_{k,p}^{L\ H}}_{E} + \underbrace{2\sum_{k=2}^{K}\sum_{p=0}^{P-1}\left|\tilde{A}_{k,p}\right|^2 \mathbf{s}_{k,p}^R\mathbf{s}_{k,p}^{R\ H}}_{F}. \tag{4.97}$$

Note that we retain the expected value in (4.92), since the dependence on the current and previous bit is contained in $\tilde{\mathbf{s}}_{k,p}$. When we expand (4.92), the terms (4.93), (4.94), (4.95) and (4.96) become zero, as consecutive bits are uncorrelated. The terms in (4.97) require some further explanation. The term A denotes the sum of the AWGN due to all the multipaths. The terms B and C denote the contributions of the synchronous first multipath and the correlated parts of the other paths respectively of user 1.

These are the signals of interest. Term D denotes the uncorrelated parts of the other multipaths of user 1. The term E denotes the correlated parts of the multipaths of the remaining users, while term F denotes the uncorrelated parts of the multipaths of the remaining users. The terms E and F can both be seen as the contribution of multiple access interference.

The cross correlation vector $\mathbf{p}$, following the same reasoning as in (4.43), is given by the correlation between the vector $\mathbf{r}$ and desired response b_1, i.e

$$\begin{aligned}
\mathbf{p} &= E\left[b_1^*[i]\mathbf{r}\right] \\
&= E\left[b_1^*[i]\left(\sum_{k=1}^{K}\sum_{p=0}^{P-1}\left(\tilde{A}_{k,p}b_k[i]\mathbf{s}_{k,p}^R + \tilde{A}_{k,p}b_k[i-1]\mathbf{s}_{k,p}^L\right) + \sigma\mathbf{m}\right)\right] \\
&= E\left[\sum_{k=1}^{K}\sum_{p=0}^{P-1}\tilde{A}_{k,p}b_k[i]b_1^*[i]\mathbf{s}_{k,p}^R\right] + E\left[\sum_{k=1}^{K}\sum_{p=0}^{P-1}\tilde{A}_{k,p}b_k[i-1]b_1^*[i]\mathbf{s}_{k,p}^L\right] + E\left[b_1^*[i]\sigma\mathbf{m}\right] \\
&= 2\tilde{A}_{1,1}\mathbf{s}_1 + 2\sum_{p=1}^{P-1}\tilde{A}_{1,p}\mathbf{s}_{1,p}^R, \qquad (4.98)
\end{aligned}$$

where (4.98) follows from the fact that the ith bit of user 1 is uncorrelated with the bits of the other users, the previous bits of user 1 and the other users, and the AWGN. If we let

$$\mathbf{g}_1 = \mathbf{s}_1 + \sum_{p=1}^{P-1}\frac{\tilde{A}_{1,p}}{\tilde{A}_{1,1}}\mathbf{s}_{1,p}^R \qquad (4.99)$$

then

$$\mathbf{p} = 2\tilde{A}_{1,1}\mathbf{g}_1. \qquad (4.100)$$

Calculating the optimum solution for the vector transform $\bar{\mathbf{v}}$ for the MMSE multipath case, similar to (2.45), we have

$$\begin{aligned}
\bar{\mathbf{v}} &= \left(E\left[\mathbf{r}\mathbf{r}^H\right]\right)^{-1}E\left[b_1^*\mathbf{r}\right] \\
&= 2\tilde{A}_{1,1}\Bigg[2\sigma^2\mathbf{I} + 2\left|\tilde{A}_{1,1}\right|^2\mathbf{s}_1\mathbf{s}_1^H + 2\sum_{p=1}^{P-1}\left|\tilde{A}_{1,p}\right|^2\mathbf{s}_{1,p}^R\mathbf{s}_{1,p}^{R\;H} + 2\sum_{p=1}^{P-1}\left|\tilde{A}_{1,p}\right|^2\mathbf{s}_{1,p}^L\mathbf{s}_{1,p}^{L\;H} \\
&\quad + 2\sum_{k=2}^{K}\sum_{p=0}^{P-1}\left|\tilde{A}_{k,p}\right|^2\mathbf{s}_{k,p}^L\mathbf{s}_{k,p}^{L\;H} + 2\sum_{k=2}^{K}\sum_{p=0}^{P-1}\left|\tilde{A}_{k,p}\right|^2\mathbf{s}_{k,p}^R\mathbf{s}_{k,p}^{R\;H}\Bigg]^{-1}\mathbf{g}_1 \qquad (4.101)
\end{aligned}$$

Similar to the non-multipath case, we express the minimum mean-square error as

$$\begin{aligned} J_{\min} &= 2 - \mathbf{p}^H \mathbf{C}^{-1} \mathbf{p} \\ &= 2 - 2\left|\tilde{A}_{1,1}\right|^2 \mathbf{g}_1^H \left[\sigma^2 \mathbf{I} + \left|\tilde{A}_{1,1}\right|^2 \mathbf{s}_1 \mathbf{s}_1^H + \sum_{p=1}^{P-1} \left|\tilde{A}_{1,p}\right|^2 \mathbf{s}_{1,p}^R \mathbf{s}_{1,p}^{R\ H} + \sum_{p=1}^{P-1} \left|\tilde{A}_{1,p}\right|^2 \mathbf{s}_{1,p}^L \mathbf{s}_{1,p}^{L\ H} \right. \\ &\left. + \sum_{k=2}^{K}\sum_{p=0}^{P-1} \left|\tilde{A}_{k,p}\right|^2 \mathbf{s}_{k,p}^L \mathbf{s}_{k,p}^{L\ H} + \sum_{k=2}^{K}\sum_{p=0}^{P-1} \left|\tilde{A}_{k,p}\right|^2 \mathbf{s}_{k,p}^R \mathbf{s}_{k,p}^{R\ H} \right]^{-1} \mathbf{g}_1 \end{aligned} \quad (4.102)$$

4.5.5.1 SIGNAL-TO-INTERFERENCE RATIO OF THE MMSE DETECTOR IN A MULTIPATH CHANNEL

We define the interference covariance matrix in the multipath case as

$$\tilde{\boldsymbol{\Omega}} \overset{\text{def}}{=} \sigma^2 \mathbf{I} + \sum_{p=1}^{P-1} \left|\tilde{A}_{1,p}\right|^2 \mathbf{s}_{1,p}^L \mathbf{s}_{1,p}^{L\ H} + \sum_{k=2}^{K}\sum_{p=0}^{P-1} \left|\tilde{A}_{k,p}\right|^2 \mathbf{s}_{k,p}^L \mathbf{s}_{k,p}^{L\ H} + \sum_{k=2}^{K}\sum_{p=0}^{P-1} \left|\tilde{A}_{k,p}\right|^2 \mathbf{s}_{k,p}^R \mathbf{s}_{k,p}^{R\ H}, \quad (4.103)$$

and the covariance matrix of the desired component as

$$\tilde{\mathbf{S}}_1 \tilde{\mathbf{S}}_1^H \overset{\text{def}}{=} \mathbf{s}_1 \mathbf{s}_1^H + \sum_{p=1}^{P-1} \frac{\left|\tilde{A}_{1,p}\right|^2}{\left|\tilde{A}_{1,1}\right|^2} \mathbf{s}_{1,p}^R \mathbf{s}_{1,p}^{R\ H}, \quad (4.104)$$

where

$$\tilde{\mathbf{S}}_1 = \begin{bmatrix} \mathbf{s}_1 & \mathbf{s}_{1,1}^R & \cdots & \mathbf{s}_{1,P-1}^R \end{bmatrix} \quad (4.105)$$

is a $L \times P$ matrix. Since the product $\tilde{\mathbf{S}}_1 \tilde{\mathbf{S}}_1^H$ is not a scalar, we cannot use the same simplification as (4.63) and (4.64) by using (4.65). We can simply write the optimum MMSE transformation in a multipath channel as

$$\tilde{\mathbf{v}} = 2\tilde{A}_{1,1} \left[\tilde{\boldsymbol{\Omega}} + \left|\tilde{A}_{1,1}\right|^2 \tilde{\mathbf{S}}_1 \tilde{\mathbf{S}}_1^H \right]^{-1} \mathbf{g}_1, \quad (4.106)$$

and the minimum mean-square error as

$$J_{\min} = 2 - 2\left|\tilde{A}_{1,1}\right|^2 \mathbf{g}_1^H \left[\tilde{\boldsymbol{\Omega}} + \left|\tilde{A}_{1,1}\right|^2 \tilde{\mathbf{S}}_1 \tilde{\mathbf{S}}_1^H \right]^{-1} \mathbf{g}_1. \quad (4.107)$$

The SIR of user 1 in the multipath case is given by

$$\gamma_{c1} = \left|A_{1,1}\right|^2 \mathbf{g}_1^H \tilde{\boldsymbol{\Omega}}^{-1} \mathbf{g}_1, \quad (4.108)$$

where $\mathbf{g}_1^H \mathbf{g}_1$ is the gain due multipath. The loss due to the uncorrelated part of the multipaths of user 1, as well as the multiple access interference is contained in $\tilde{\boldsymbol{\Omega}}$.

4.5.5.2 BEP OF THE MMSE DETECTOR IN A MULTIPATH CHANNEL

Evaluating the exact BEP of the MMSE detector in a multipath environment is even more computationally expensive than in the non-multipath case. To evaluate the BEP of the MMSE detector in a multipath environment we will simply use the approximation in (4.81) and (4.82), i.e.

$$P_e(\sigma, 1) \approx Q\left(\sqrt{\frac{1 - J_{\min}}{J_{\min}}}\right), \tag{4.109}$$

and

$$P_e(\sigma, 1) \approx Q\left(\frac{1}{\sqrt{J_{\min}}}\right). \tag{4.110}$$

4.6 SUMMARY

In this chapter a rigorous analysis of the MMSE detector, within the context of linear detectors, is undertaken. The blind detectors explored in this dissertation have the same vector weight solutions as the MMSE detector. This necessitates a thorough understanding of the operation and performance of the MMSE detector.

The linear decorrelating detector is introduced in the first section of this chapter. The linear decorrelating detector bears a close resemblance to the MMSE detector, as it performs the same operation as the MMSE detector in the noise free case. The linear multiuser detection optimization problem is then generalized to the finding of the best linear detector. The K user case does not permit a closed form solution to this optimization problem. The two user case is subsequently examined, which does permit a closed form solution.

Following this general view of linear multiuser detectors, focus is then shifted to the operation of the joint linear MMSE detector. The MMSE optimization problem is presented, and is solved through use of the complex valued MMSE detector vector matrix model. The two user MMSE detector is briefly considered. The noise limiting forms of the MMSE detector is then discussed, with focus on the relation between the linear decorrelating detector and the linear MMSE detector. The asynchronous linear MMSE detector model is briefly presented. The Wiener characterization of the linear MMSE detector is subsequently considered, where optimization is reduced from joint optimization, to optimization with respect to only one of the users. In the ensuing section the LMS algorithm for the linear MMSE detector is derived.

The rest of the chapter focuses on the performance of the linear MMSE detector based on the criteria stated in Chapter 3. The performance criteria considered include SIRs, asymptotic multiuser efficiency, BEP and power tradeoff regions. In the case of BEP, the Gaussian approximation method and the infinite user limit for the MMSE detector is also considered. The following section contains the extension of the model in Verdu [31] to the complex valued multipath case. Certain assumptions are made, and the expressions for SIR and BEP are derived for the multipath channel.