# Designing Equalizers based on Explicit Channel Models of Direct-Sequence Code-Division Multiple Access Systems

Claes Tidestav
Signal Processing Group, Uppsala University
P.O. Box 27, S–751 03 Uppsala, Sweden
phone: +46–18–183071, fax: +46–18–503611
email: ct@syscon.uu.se

**Abstract**

In this paper, we show how the entire process of transmission and reception in a DS-CDMA system, including the spreading, can be described as a tapped delay line with multiple inputs and a single output. We then demonstrate how this model can be used to design a fractionally spaced decision feedback equalizer with a single input and multiple outputs. In the proposed approach, long codes can be used and channel estimation can be performed in an efficient way. The suggested detector is near-far resistant and the capacity of the system is not limited by intracell interference. Monte Carlo simulations are conducted to demonstrate the superiority of the proposed approach as compared to the conventional detector. The simulation scenario constitute a heavily loaded system with multipath. Situations with and without power control are considered.

# Chapter 1

# Introduction

The second generation mobile telephone systems are now operational and system designers are making plans for the third generation systems. Demands for high capacity and flexible services have resulted in considerable interest in Direct-Sequence Code-Division Multiple Access, or DS-CDMA. DS-CDMA has many advantages: robustness against fading and interference as well as ease of integration of variable rate services, to mention just a few.

Much of the flexibility in the design of a DS-CDMA system relies on the use of *long spreading codes*. Long spreading codes have a period that is much larger than the symbol period. This is in contrast to *short codes*, which have a period that is equal to the symbol period. The use of long codes makes it unnecessary to use code planning. It makes it also extremely easy to supply variable rate services.[1]

DS-CDMA also has problems however. The use of non-orthogonal channels makes the system very sensitive to the fact that the signals from different users arrive at the receiver with very different powers. This phenomenon is known as the *near-far problem*. The common solution, power control in combination with the conventional detector, limits the capacity. It also makes it necessary for the mobile and the base station to exchange power control commands. Finally, it requires the use of amplifiers with a high dynamic range in the mobiles.

To circumvent the near-far problem and thus eliminate the need for power control, alternative detectors have been proposed. These are known as *multiuser detectors*. Multiuser detectors can be divided into two broad classes: those that operate on the outputs from a bank of filters, matched to the spreading codes of each user, and those that do not.

Examples of the first category are the optimal ML detector [1], the multistage detector [2], the decorrelating detector and the MMSE detector [3]. These multiuser detectors are frequently presented as block detectors and multipath prop-

[1]Short codes have the advantage that they can be designed to have low cross correlation. This means that more users can easily share the same bandwidth if the correlation remains low also at the receiver. In practice, this will be the case only under the rather restrictive assumptions of a synchronous system and a channel with little multipath propgation.

agation is rarely considered explicitly. Most papers on multiuser detection treat cases with short codes, although almost all of the detectors could be modified for application in systems with long codes.

A prevalent assumption in the discussion of multiuser detection is perfect synchronization. However, it is shown in [4], that not only detection, but also synchronization is considerably harder in a system without power control. It is also shown in [4] that the decorrelating detector is very sensitive to errors in estimates of the propagation delay.

The second category of multiuser detectors operates directly on the received wideband signal which is sampled at the chip rate. The received signal is passed through an FIR filter followed by a decision device. Some detectors have been suggested to operate with no use of decision feedback [5], while others use decision feedback from decisions concerning one user [6], or from decisions concerning all users [7]. The filters are adaptive and they are updated recursively directly from the received signals. Adaptive detectors that operate on the received wideband signal do not require the propagation delay to be estimated. On the other hand, long codes are impossible to use, since the cyclostationarity in the signals is destroyed when the codes are not periodic with period equal to the symbol period. Also, the adaptive filters are updated on a symbol-by-symbol basis. For low bitrates, the symbols will be long. During such a long bit period, the optimal filter coeffcients can change substantially. It will be hard for the adaptive algorithm to track these changes.

An ideal detector for DS-CDMA would thus be a detector which

- is near-far resistant like the decorrelating detector
- is insensitive to errors in propgation delay estimates like the adaptive detectors that operate directly on the chip sampled signal
- can be used with long codes like the conventional receiver.

In this paper, we show how the standard DS-CDMA system transmission and reception can be reformulated as an equivalent, discrete time, tapped delay line model. This model takes the data symbols as input and has the wideband, chip sampled, signal as output. The spreading operation is represented as linear filtering, which is time-invariant for short codes and time-varying for long codes. This model makes it possible to design equalizers which, directly from the chip sampled signal, filter out the signals of the respective users.

As an example of a detector which has been derived from this model, we subsequently present a generalization of the fractionally spaced decision feedback equalizer (DFE). This DFE has one input and the same number of outputs as the number of users. The DFE is derived to minimize the mean square error (MSE) of the symbol estimates under the assumption of correct past decisions.

We then investigate the performance of the proposed DFE with perfect power control, with average power control and without any power control. For a rather heavily loaded system, the DFE is shown to outperform the conventional detector with a RAKE receiver in the first two cases. In the third case, it is demonstrated that the DFE works without problem in a near-far situation.

The paper is organized as follows. In Chapter 2, the basic system model is presented, as well as our linear filter reformulation of it. The time-varying single-input multiple-output fractionally spaced decision feedback equalizer is presented in Chapter 3. To demonstrate the performance of the DFE, as well as the usefulness of our linear model, the results from Monte Carlo simulations are presented in Chapter 4. Finally, in Chapter 5, conclusions are drawn.

# Chapter 2

# System model

In this chapter, we start by describing the DS-CDMA system as it is usually described. We then, step by step, explain how this system can be reformulated as a completely discrete time model for the single user case. Finally, the entire system is then written as a multiple-input single-output tapped delay line, which will be used in the following chapters.

## 2.1 The DS-CDMA system

We are considering an asynchronous DS-CDMA system with $K$ users. The modulation scheme is BPSK. The symbol period is denoted by $T_s$, whereas $T_c$ represents the duration of a chip. The processing gain, i.e. the ratio $T_s/T_c$, is denoted by $N_c$. Without loss of generality, we set $T_c = 1$.

It is assumed that user $k$ transmits the symbol $d_k(tN_c)$ during the time period $[tN_c, (t+1)N_c[$. Each symbol is spread by a wideband signature sequence, denoted by $c(\cdot)$. For the sake of simplicity, both symbols and chips are assumed to have a rectangular shape. The spreading operation results in a baseband signal, which is shifted up to the carrier frequency and transmitted.

The transmitted passband signal propagates through a wideband channel. The channel is frequency selective. The receiver front end consists of a conventional IQ-stage, where the I and Q signals are down-converted to the baseband. In the following, only equivalent baseband signals and channels will be considered.

The received baseband signal is then passed through a *chip-matched filter*. In the case of rectangular pulse shaping, this corresponds to an integrate-and-dump device, where the integration is over a chip time $T_c$. At the output of the chip-matched filter, the signal is sampled at the chip rate. The resulting complex-valued signal will be denoted by $r(\cdot)$. The receiver front end is depicted in Figure 2.1.

We will now describe how this standard DS-CDMA system can be recast into a form, in which the entire process of transmission and reception can be represented by a completely discrete time, tapped delay line, model. Based on such a channel

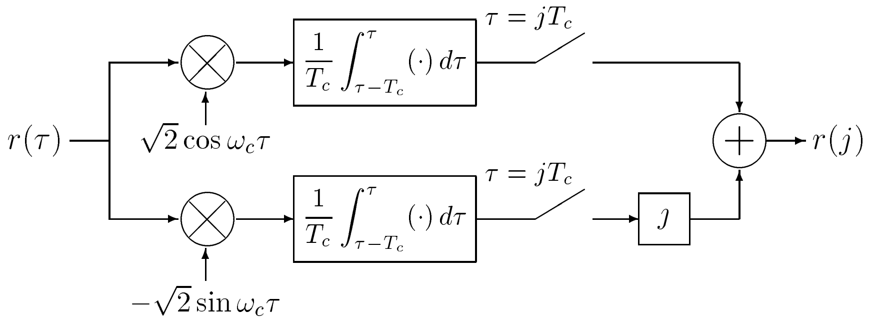



Figure 2.1: The receiver front end.

model it would be possible to construct an equalizer, which restores the transmitted symbol sequence. This tapped delay line model will have the (narrowband) symbol sequence $d_k(tN_c)$ as input at time $tN_c$. The received (wideband) signal $r$ will be the output of the tapped delay line. The impulse response of the model will possibly be rapidly time-varying.

## 2.2 A linear baseband model: the single-user case

Figure 2.2 illustrates the traditional view of the transmitter stage for a single user in a DS-CDMA system. Both the symbol sequences and the signature sequences are presented as *continuous time signals*. To arrive at a discrete time model of

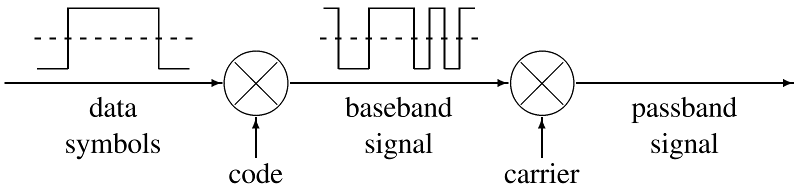



Figure 2.2: The traditional way of viewing the transmission process in a DS-CDMA system.

the system, all pulse shaping is assumed to take place *after* the spreading. This means that the system can be represented as in Figure 2.3.

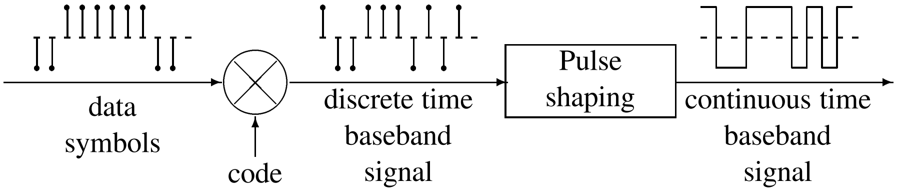



Figure 2.3: The spreading viewed as a discrete time process.

Next, the pulse shaping, the frequency up-conversion, the physical channel, the frequency down-conversion and the chip-matched filter are lumped together and replaced by an equivalent discrete time channel. This equivalent representation will in the following be called the *physical channel*. This means that the transmission and reception can be represented as in Figure 2.4.

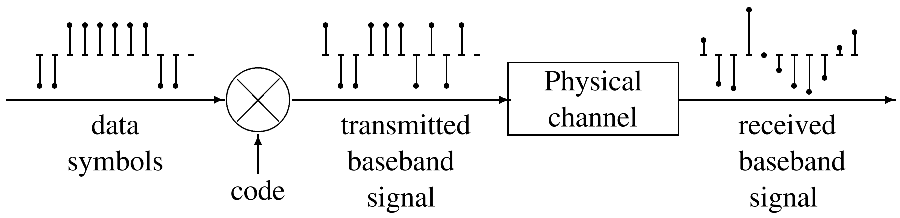



Figure 2.4: A completely discrete time model.

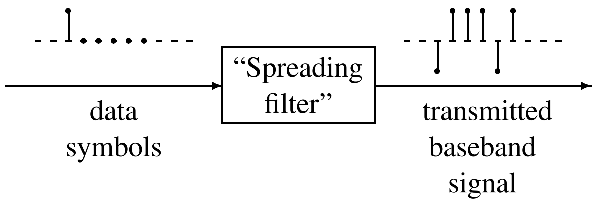
The model is now completely discrete time, but it is not yet a tapped delay line model. In order to arrive at such a model, the spreading must be represented, not by the usual multiplication operation, but by a linear filter. We would then obtain an *equvalent channel*, which can be represented by a convolution of this *spreading filter* and the physical channel.

We shall now show that the spreading process can indeed be interpreted as a filtering operation. We view the symbol sequence as an input sequence and the spread signal as an output sequence of a linear filter. Note that the symbol sequence is considered to be non-zero only every $N_c$ chips. Consider the situation during one symbol period. If the current symbol is $+1$, then the transmitted baseband signal will be the code, whereas if the current symbol is $-1$, the transmitted signal will be the code with opposite sign. The situation is depicted in Figure 2.5.

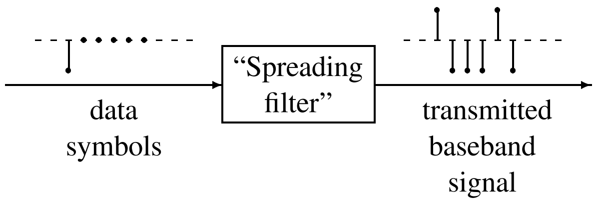
This means that during one symbol period, *the spread signal can be thought of as having been generated by passing the symbol through a discrete time filter with impulse response identical to the code sequence.* The length of the impulse response of the spreading filter is always equal to the processing gain $N_c$.

At this point, it is of some importance to note a difference in the channel model depending on whether short or long codes are employed in the system. If short codes are used, our spreading filter will be *time-invariant*, since the same code is used to spread each symbol.

If, on the other hand, long codes are used, the spreading filter will change every symbol period. This means that even if the physical channel is time-invariant (i.e. if there is no fast fading), the resulting equivalent channel will vary very rapidly! The time-variation is however known, since the code is known.




(a) Input symbol = +1




(b) Input symbol = −1

Figure 2.5: The spreading viewed as a filtering operation.

## 2.3 A multiple-input single-output channel model

According to the preceding section, the entire single-user channel, from the symbol sequence to the chip sampled received baseband sequence, can be represented by a tapped delay line. Since each user is assigned a different spreading code, all users have different channels. The received signal consists of the sum of the outputs from these $K$ channels and some noise, as shown in Figure 2.6.

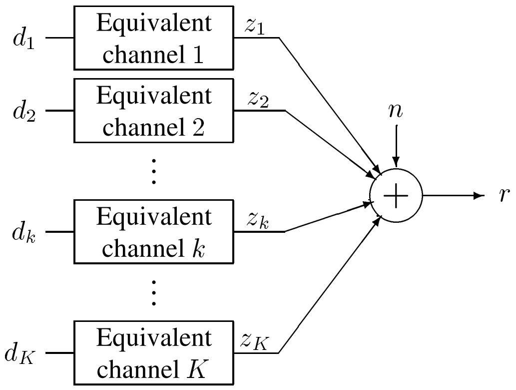



Figure 2.6: A multiple input, single output channel model.

As stated earlier, each tapped delay line is the convolution between the spreading filter and the physical channel. Since the length of the spreading filter is equal to the symbol period, intersymbol interference will be present if any physical channel has more than one tap. Even if all physical channels in the system have only one tap, i.e. if the fading is flat, there will be intersymbol interference if the propagation delays of the different users are not all equal.

The entire channel can thus be represented by a linear model with $K$ inputs and one output. To formulate this mathematically, denote the output of the channel from user $k$ at time $tN_c + j$ as

$$z_k(tN_c + j) = \sum_{i=0}^{m} h^k_{iN_c+j}(tN_c + j) d_k((t-i)N_c) \tag{2.1}$$

where

| | | |
|---|---|---|
| $h^k_{iN_c+j}(tN_c + j)$ | : | tap $iN_c + j$ in the channel from user $k$ at time $tN_c + j$ |
| $m$ | : | maximum extent (over all channels) of intersymbol interference. |

**Remark 1:** The coefficients $h^k_i(tN_c + j)$ represent the entire channel impulse response, including the spreading filter.

**Remark 2:** As noted previously, $m = 0$ only if the fading is flat and all users have identical propagation delays.

The received signal will be the noise-corrupted sum of the signals from the $K$ users:

$$r(tN_c + j) = \sum_{k=1}^{K} \sum_{i=0}^{m} h^k_{iN_c+j}(tN_c + j) d_k((t-i)N_c) + n(tN_c + j) \ , \tag{2.2}$$

where $n(tN_c + j)$ is assumed to be wide sense stationary, zero mean noise with autocorrelation function

$$E\left[n(i)n^H(j)\right] = \psi_{i-j} \ . \tag{2.3}$$

Introduce the vector of symbols transmitted by all users at time $tN_c$

$$d(tN_c) \triangleq \begin{pmatrix} d_1(tN_c) & \dots & d_K(tN_c) \end{pmatrix}^T$$

and define

$$h^{tN_c+j}_{iN_c+j} \triangleq \begin{pmatrix} h^1_{iN_c+j}(tN_c + j) & \dots & h^K_{iN_c+j}(tN_c + j) \end{pmatrix} \ . \tag{2.4}$$

Then the received signal (2.2) can be rewritten as

$$r(tN_c + j) = \sum_{i=0}^{m} h^{tN_c+j}_{iN_c+j} d((t-i)N_c) + n(tN_c + j) \ . \tag{2.5}$$

Equation (2.5) is the desired $K$-input single-output channel model, which relates the sequence of symbol vectors to the chip sampled output sequence.

To make use of this model to design a DS-CDMA detector, two problems have to be solved:

- the estimation of the channel coefficient vector (2.4) for $i = 0, 1, \ldots, m$ and $j = 0, 1, \ldots, N_c - 1$ and

- the design of an equalizer under the assumption that the impulse response coefficients (2.4) are known. An equalizer is necessary, since $m \neq 0$ in general.

When performing the channel estimation, we use the a priori information that a part of the channel, i.e. the spreading filter, is known and unnecessary to identify. To estimate the remaining part of the channel impulse response, we suggest the use of the transmitted chip sequence as regressors to identify only the physical channel. The complete channel is then obtained by convolving the physical channel with the spreading filter. A propagation delay will result in leading zeros in the estimated impulse response of the channel.

One good candidate for an equalizer is the decision feedback equalizer (DFE). To be applicable in this context, a DFE has to be fractionally spaced, because the received (information bearing) signal has a bandwidth, which is considerably larger than the symbol rate. Another possible detector is a linear equalizer. The performance of the linear equalizer is worse than the performance of the DFE and the complexity of the two equalizers is comparable. Another conceivable alternative would be an ML optimal approach, based on the Viterbi algorithm. Such a detection scheme would have a complexity that is exponential in the number of users. Due to the worse performance of the linear equalizer and the higher complexity of the optimal ML detector, we will in the following only investigate the fractionally spaced decision feedback equalizer.

# Chapter 3

# The single-input $K$-output fractionally spaced DFE

Our proposed decision feedback equalizer operates directly on the chip sampled signal $r(\cdot)$. The outputs from the equalizer are the data symbols of the $K$ users. This means that the proposed DFE will be a single-input multiple-output equalizer.

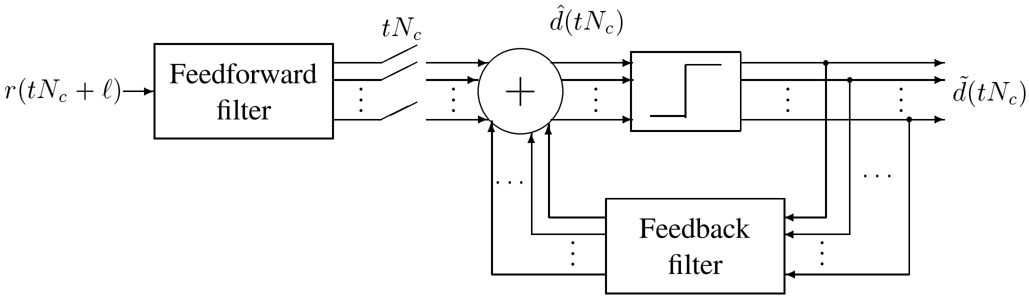



Figure 3.1: The single-input multiple-output fractionally spaced decision feedback equalizer.

The structure of the DFE is depicted in Figure 3.1. The scalar, chip sampled signal $r(\cdot)$ is used as input to the feedforward filter. The $K$ outputs of the feedforward filter are sampled at the symbol rate. The influence of previously detected symbols are removed by subtracting the $K$ outputs of the feedback filter from the corresponding $K$ outputs of the feedforward filter. The resulting signal (vector) $\hat{d}(tN_c)$ is an estimate of the data symbols of the $K$ users. This estimate is then passed through a decision device, where hard decisions are made.

The feedforward filter is specified by its coefficient matrices $\mathbf{S}_i^{tN_c}$ of dimension $K|1$, whereas the feedback filter is specified by its coefficient matrices $\mathbf{R}_i^{tN_c}$ of dimension $K|K$. Here the subscript represents the position of the coefficient in the impulse response of the filter. The superscript indicates the time instant when the filter is used. The estimator can then be written:

$$\hat{d}(tN_c|tN_c+\ell) = \sum_{i=0}^{\ell} \mathbf{S}_i^{tN_c} r(tN_c+\ell-i) - \sum_{i=1}^{m} \mathbf{R}_i^{tN_c} \tilde{d}((t-i)N_c) \tag{3.1}$$

The DFE estimates the symbol vector that was transmitted at time $tN_c$, given data up to time $tN_c+\ell$. Above, the $K$-vector $\tilde{d}((t-i)N_c)$ represents decisions on symbols which have previously been estimated. The design variable $\ell$ is known as the *smoothing lag* or *decision delay*. The larger the smoothing lag is, the better the DFE performs, but increasing the smoothing lag above the length of the channel impulse response results in little improvement.

The coefficient matrices of the two filters are adjusted so that the mean square error of the symbol vector estimate $\hat{d}(tN_c|tN_c+\ell)$ is minimized. The cost function that should be minimized is thus

$$J = E||d(tN_c) - \hat{d}(tN_c|tN_c+\ell)||^2 \ . \tag{3.2}$$

For the derivation of the optimal filters, we are assuming the presence of correct decisions on previous symbol vectors and of error-free estimates of the channel coefficients (2.4).

Introduce the matrices

$$\Theta_S^{tN_c} \triangleq \begin{pmatrix} \mathbf{S}_0^{tN_c} & \dots & \mathbf{S}_\ell^{tN_c} \end{pmatrix}^H \tag{3.3a}$$

$$\Theta_R^{tN_c} \triangleq \begin{pmatrix} \mathbf{R}_1^{tN_c} & \dots & \mathbf{R}_m^{tN_c} \end{pmatrix}^H \tag{3.3b}$$

of dimensions $(\ell+1)|K$ and $mK|K$ respectively. We also define the matrices

$$\mathcal{F}_t \triangleq \begin{pmatrix} \bar{\beta}_0^{t+q} & \dots & & \bar{\beta}_q^{t+q} \\ 0 & \beta_0^{t+q-1} & \dots & \beta_{q-1}^{t+q-1} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \beta_0^t \end{pmatrix}$$

$$\mathcal{G}_t \triangleq \begin{pmatrix} \bar{\beta}_{q+1}^{t+q} & \dots & \bar{\beta}_m^{t+q} & 0 & \dots & 0 \\ \beta_q^{t+q-1} & \dots\dots\dots & \beta_m^{t+q-1} & & \ddots & \vdots \\ \vdots & & & & \ddots & 0 \\ \beta_1^t & \dots\dots\dots\dots\dots & & & & \beta_m^t \end{pmatrix}$$

where

$$\beta_i^t \triangleq \begin{pmatrix} h_{iN_c+N_c-1}^{tN_c+N_c-1} \\ h_{iN_c+N_c-2}^{tN_c+N_c-2} \\ \vdots \\ h_{iN_c}^{tN_c} \end{pmatrix} \quad \text{and} \quad \bar{\beta}_i^t \triangleq \begin{pmatrix} h_{iN_c+p}^{tN_c+\ell} \\ h_{iN_c+p-1}^{tN_c+\ell-1} \\ \vdots \\ h_{iN_c}^{tN_c+qN_c} \end{pmatrix} \quad i = 0, 1, \dots, q \ , \tag{3.4}$$

of dimensions $N_c|K$ and $(p+1)|K$ respectively. The integers $p$ and $q$ are implicitely defined by

$$\ell \triangleq qN_c + p$$

For the calculation of the DFE, the following matrices are also required:

$$h_t \triangleq \begin{pmatrix} h_\ell^{tN_c+\ell} \\ h_{\ell-1}^{tN_c+\ell-1} \\ \vdots \\ h_0^{tN_c} \end{pmatrix} \quad \text{and} \quad \Psi \triangleq \begin{pmatrix} \psi_0 & \psi_1 & \dots & \psi_\ell \\ \psi_{-1} & \psi_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \psi_1 \\ \psi_{-\ell} & \dots & \psi_{-1} & \psi_0 \end{pmatrix} .$$

The following result describes how the matrices (3.3a) and (3.3b) are calculated:

**Theorem 1** Consider the time-varying fractionally spaced DFE described by (3.1) and the time-varying channel model (2.5). Assuming correct past decisions, the matrix coefficients $\mathbf{S}_i^{tN_c}$ and $\mathbf{R}_i^{tN_c}$ that minimize (3.2) are obtained as follows:

1. Solve the system of linear equations

$$\left(\mathcal{F}_t\mathcal{F}_t^H + \Psi\right)\Theta_S^{tN_c} = h_t \tag{3.5}$$

for the matrix coefficients $\mathbf{S}_i^{tN_c}$.

2. Perform the matrix multiplication

$$\Theta_R^{tN_c} = \mathcal{G}_t^H \Theta_S^{tN_c} \tag{3.6}$$

to obtain the matrix coefficients $\mathbf{R}_i^{tN_c}$.

■

**Proof:** See Appendix A.

**Remark 1** The time-varying nature of the channel is explicitly taken into account by including *future* channel models in the calculation of optimal filter coefficients. If long codes are used, the channel will in fact change completely every symbol period. Since future values of the code are known, this is not a problem.

**Remark 2** Note that the linear system of equations (3.5) is solved for multiple right hand sides simultaneously, thereby only requiring one LU factorization.

**Remark 3** Unlike most multiuser detectors, the proposed DFE is a symbol-by-symbol detector, making it possible to implement it without modification.

The computational complexity of calculations necessary to optimize the equalizer coefficients is rather high. If long codes are employed, the DFE has to be recalculated for every symbol estimation. If short codes are used, the rate at which the equalizer has to be updated is determined by the fading.

The number of multiplications necessary to calculate a new DFE are presented below. For ease of reference, we recapitulate the necessary notation:

- The number of users $K$
- The processing gain $N_c$
- The smoothing lag $\ell = qN_c + p$
- Maximum extent of intersymbol interference $m$

1. Calculation of $\mathcal{F}_t\mathcal{F}_t^H + \Psi$:
$$\frac{q(q+1)N_cK}{2}\left(\ell + 1 - \frac{N_c}{6}(q-1)\right) \text{ multiplications}$$
2. Factorization of the matrix $\mathcal{F}_t\mathcal{F}_t^H + \Psi$:
$$\frac{(\ell+1)^3}{6} + (\ell+1)^2 - \frac{7(\ell+1)}{6} \text{ multiplications}$$
3. Calculation of the feedforward filter coefficients based on (3.5):
$$(\ell+1)^2K \text{ multiplications}$$
4. Calculation of the feedback filter coefficients based on (3.6):
$$\frac{mK^2(q+2)(2m-q-1)}{2} \text{ multiplications}$$

Using the DFE to estimate a transmitted symbol vector then requires:
$$(\ell+1)K + mK^2 \text{ multiplications}$$

Notice that the complexity is not exponential in any system parameters and that the dominant complexity (item 2) is not directly dependent on the number of users.[1]

[1] The special structure of the various matrices have been taken into account to reduce the computational complexity.

# Chapter 4

# Simulations

## 4.1 Simulation conditions

We will now investigate the utility of our channel model, as well as the usefulness of our proposed DFE with the aid of Monte Carlo simulations. In the simulations, the modulation that is employed is BPSK, both for symbols and codes. We are using *long* codes. We are considering the processing gain $N_c = 8$, and $K = 5$ users in the system, which represents a rather heavily loaded system.

All physical channels have four taps, not including the propagation delays. The amplitudes of the taps are Rayleigh distributed, whereas the phases are uniformly distributed in the interval $[0, 2\pi[$. All taps fade independently, and during transmission of a burst, they are time invariant. The propagation delays of the users are uniformly distributed in the interval $[0, N_c - 1]$ (in units of the chip period). Only delays that are integer multiples of the chip period are considered. This is not a serious restriction however, because delays that are fractions of the chip period will result in interchip interference, which is considered anyway.

A more serious restriction in the simulation study is the assumption that we know the channel impulse response. In the cases where the DFE is compared to other detectors, these detectors also of course utilize the exact knowledge of the channel impulse response, so the comparison is fair.

Although the DFE can improve the detection by taking the color of the noise into account through the presence of the autocorrelation function $\psi_{i-j}$ in (3.5), we have in this simulation chosen the noise to be white and gaussian with power spectral density $N_0$.

We are considering three scenarios. They differ in the kind of power control that is used. The three kinds of power control we consider are

1. Perfect power control

2. Average power control (explained in Section 4.3)

3. No power control

The results of the simulations are presented in the following.

## 4.2 Perfect power control

With perfect power control, the single-input $K$-output fractionally spaced DFE is compared to the conventional detector with a RAKE receiver. The RAKE has four fingers. The $E_b/N_0$, which is identical for the five users, varies between 5 and 20 dB and the average BER of the users is estimated.

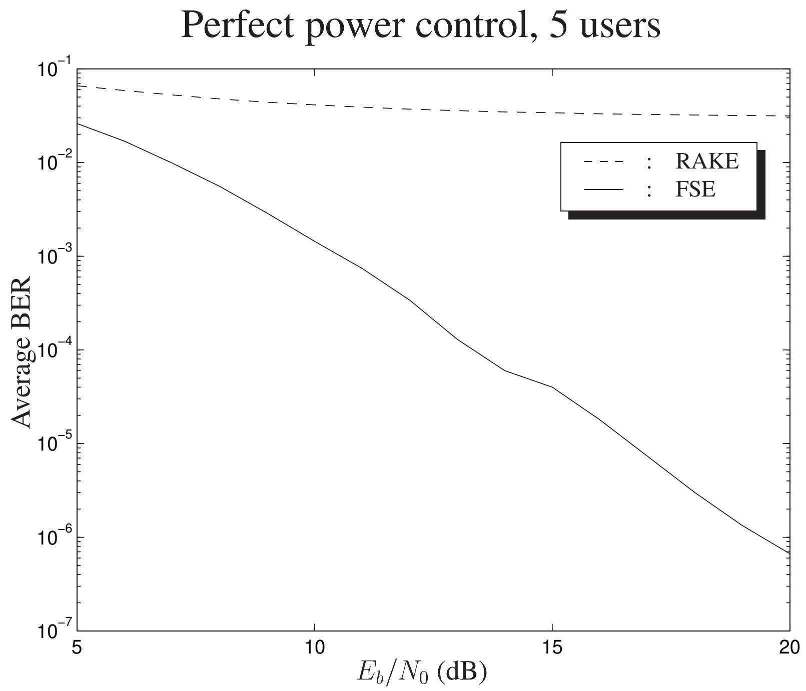



Figure 4.1: Average BER as a function of $E_b/N_0$ for DFE (solid) and conventional detector with RAKE (dashed). Perfect power control is used.

The results are shown in Figure 4.1. The RAKE is clearly inferior to the DFE at all investigated signal-to-noise-ratios. This is natural, since the codes are not orthogonal at the receiver. The interference limited nature of the conventional receiver is clearly visible: the BER approaches an error floor as $E_b/N_0 \to \infty$.

The DFE on the other hand is not limited by the interference from the other users. The BER just decreases as the SNR is increased. The influence from other users is removed by the DFE, which decorrelates the signals. The performance of the DFE is then limited, not by interference from other users within the same cell, but by the interference from users in other cells, which will have to be considered as (colored) noise.

## 4.3 Average power control

The demand of strict power control is abandoned in this simulation setup. We are only assuming that the received signals have the same *average* power. This means that the power control will not have to compensate for the fast fading, but only for the shadowing loss. Apart from this, the simulation conditions are the same as in Section 4.2.

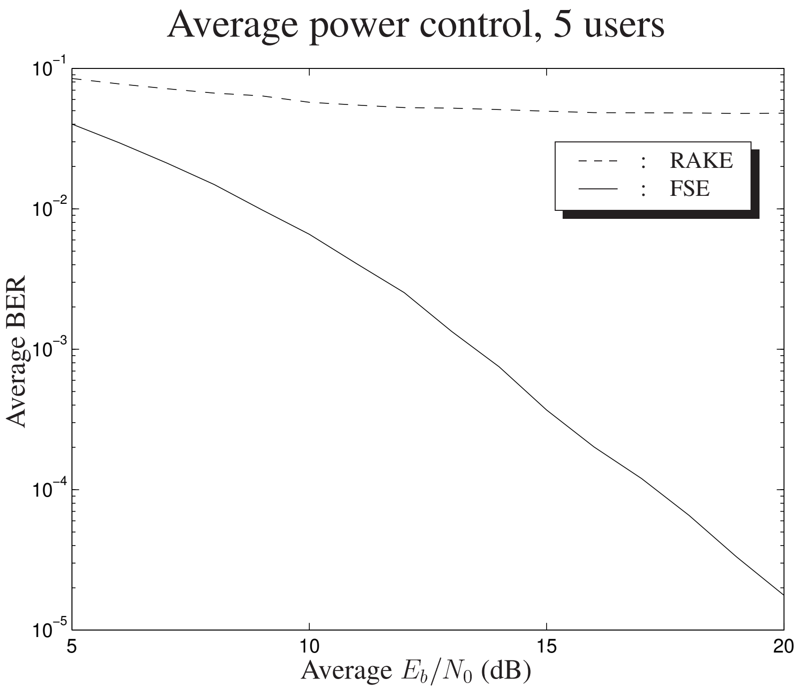



Figure 4.2: Average BER as a function of average $E_b/N_0$ for DFE (solid) and conventional detector with RAKE (dashed). Average power control is used.

The results are shown in Figure 4.2. Both the RAKE and the FSE perform worse in this case, as compared to the results of Figure 4.1. The BER for the RAKE is about 50% higher for the entire range of SNR:s compared to the case with perfect power control. The BER for the FSE is about twice what it was for the case of perfect power control for low SNR:s. For high SNR:s, the BER is an order of magnitude higher. The important property of the DFE being a noise limited, rather than an interference limited, detector is however preserved. Still, there is no visible error floor.

## 4.4 No power control

To investigate the performance of the DFE in a situation where the near-far problem is present, we have performed a simulation where not even the average power of the users are equal. To emphasize the near-far problem, four of the users have the same average power, whereas the fifth user has an average power that is 10, 20 or 30 dB lower. The weaker user (user 1) has an average $E_b/N_0$ between 5 and 20 dB.

The performance measure in this case is the BER of the weaker user, because the number of errors for the stronger users will not be noticeable. Another difference in this case is that we have not included the RAKE in the comparison. This is because such a comparison would obviously be unfair: the conventional receiver cannot deal with large near-far effects.

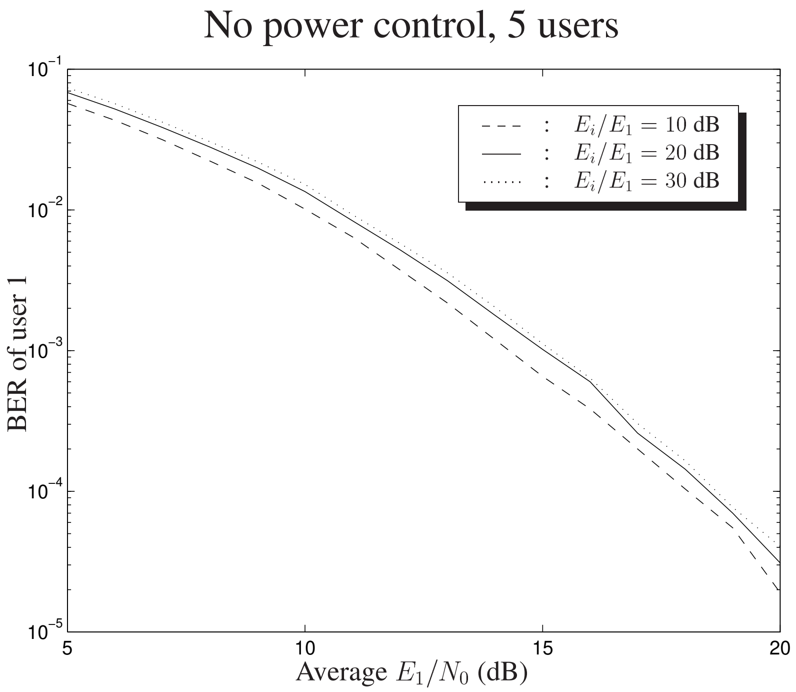



Figure 4.3: Average BER as a function of average $E_b/N_0$ for the weaker user (user 1). No power control is used.

The result is shown in Figure 4.3. We see that the near-far effect has an impact on the DFE, but the effect is small: for a power difference of 10 dB, the deteriation is around 1 dB for the entire investiagted range of SNR:s as compared to average power control.

# Chapter 5

# Conclusions

In this paper, we have described the complete transmission and reception process in a DS-CDMA system, including the spreading, by a completely discrete time tapped delay line model. The model takes as the input the symbols from all users in the system. The output of the model is the received signal, sampled at the chip rate. By using this discrete time model, the problem of detecting the signals of the users has been recast as a deconvolution problem, similar to problems which have been investigated for decades. Also, the reformulation makes it possible to perform channel estimation in a way that reduces problems with fast fading.

To demonstrate the usefulness of the reformulation, we have subsequently derived a fractionally spaced decision feedback equalizer based on the model. This DFE detects all signals simultaneously, thereby making it possible to operate in an environment without power control. In contrast to many other multiuser detectors, it is derived as a symbol-by-symbol detector, making it well suited for implementation. Another advantage is that due to the fast sampling, the DFE only requires rough synchronization.

The DFE is in general time-varying, which is a requirement for application in systems where long codes are employed. The use of such non-orthogonal codes are commonly acknowledged to lower the capacity of the system. Since the DFE effectively decorrelates the users, a large part of this capacity loss could be eliminated.

The simulations in Chapter 4 show that

- The DFE outperforms the conventional detector with a four-finger RAKE in a system with a spreading gain of 8, five users and perfect power control.
- The DFE also outperforms the conventional detector with a RAKE with four fingers for the same system with only average power control.
- The DFE works well in the same system when four of the users are 10, 20 or 30 dB stronger than the fifth.

The amount of computations that has to be performed to update the equalizer is substantial, but it is not exponential in any system parameter. In fact, the complexity of the most time-consuming part of the calculation is not directly dependent on the number of users.

# Bibliography


[1] S. Verdú. *Optimum Multi-user Signal Detection*, PhD thesis, University of Illinois, Dept. of Electrical Engineering, Urbana, IL, September 1984.

[2] M. K. Varanasi and B. Aazhang, "Multistage detection in asynchronous code-division multiple-access communications," *IEEE Transactions on Communications*, vol. 38, no. 4, pp. 509–519, April 1990.

[3] R. Lupas. *Near-far Resistant Linear Multiuser Detection*, PhD thesis, Princeton University, Princeton, NJ, January 1989.

[4] S. Parkvall, *Asynchronous Direct-Sequence Code-Division Multiple Access Systems: Propagation Delay Estimation and Sensitivity Analysis.* Licentiate thesis, Signal Processing, Dept. of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, October 1994.

[5] U. Madhow and M. L. Honig, "MMSE interference suppression for direct-sequence spread spectrum CDMA," *IEEE Transactions on Communications*, vol. 42, no. 12, pp. 3178–3188, December 1994.

[6] M. Abdulrahman. *DFE for Interference and Multipath Suppression in a CDMA System*, PhD thesis, Department of Systems and Computer Engineering, Carlton University, Ottawa, Canada, January 1994.

[7] P. B. Rapajic and B. S. Vucetic, "Adaptive receiver structures for asynchronous CDMA systems," *IEEE Journal on Selected Areas in Communications*, vol. 12, no. 4, pp. 685–697, May 1994.

# Appendix A

# Derivation of the time-varying single input-multiple output fractionally spaced decision feedback equalizer

Suppose an $N$ times oversampled time-varying FIR-channel is given

$$y(tN+k) = \mathbf{B}_k(tN+k)d(t) + \cdots + \mathbf{B}_{k+mN}(tN+k)d(t-m) + v(tN+k) \quad 0 \le k < N$$
$$Ed(t)d^H(\tau) = I\delta_{t\tau} \qquad Ev(t)v^H(\tau) = \psi_{t-\tau} \qquad Ed(t)v^H(\tau) = 0 \ . \tag{A.1}$$

Note that there are basically *two* sampling rates in the system: the output is sampled $N$ times for every input sample that enters the channel. The channel has $M$ inputs and one output. This means that the dimensions of the various entities are

$$y(t) = \underbrace{\left(\ \bullet\ \right)}_{1} \}1 \qquad d(t) = \underbrace{\left(\ \bullet\ \right)}_{1} \}M \qquad v(t) = \underbrace{\left(\ \bullet\ \right)}_{1} \}1$$
$$\mathbf{B}_n(t) = \underbrace{\left(\ \bullet\ \right)}_{M} \}1 \qquad \psi_n = \underbrace{\left(\ \bullet\ \right)}_{1} \}1 \ .$$

The objective is to estimate the symbols of all users at time $t$, given channel output measurements up to time $tN+\ell$ and previous decisions of the symbols of all users up to time $t-1$. Here $\ell$ is a design variable, the so-called *smoothing lag*.

The estimator to be used is a single input-multiple output fractionally spaced decision feedback equalizer:

$$\hat{d}(t|tN+\ell) = \begin{pmatrix}\mathbf{S}_0(t) & \dots & \mathbf{S}_\ell(t)\end{pmatrix} \begin{pmatrix} y(tN+\ell) \\ \vdots \\ y(tN) \end{pmatrix} - \begin{pmatrix}\mathbf{R}_1(t) & \dots & \mathbf{R}_m(t)\end{pmatrix} \begin{pmatrix} \tilde{d}(t-1) \\ \vdots \\ \tilde{d}(t-m) \end{pmatrix}$$

$$= \Theta_S^H(t) \mathrm{y}_{tN+\ell} - \Theta_R^H(t) \tilde{\mathrm{d}}_{t-1} \quad , \tag{A.2}$$

where

$$\Theta_S^H(t) = \begin{pmatrix}\mathbf{S}_0(t) & \dots & \mathbf{S}_\ell(t)\end{pmatrix} \tag{A.3a}$$

$$\Theta_R^H(t) = \begin{pmatrix}\mathbf{R}_1(t) & \dots & \mathbf{R}_m(t)\end{pmatrix} \tag{A.3b}$$

and

$$\mathrm{y}_{tN+\ell} = \begin{pmatrix} y(tN+\ell) & \dots & y(tN) \end{pmatrix}^T$$

$$\tilde{\mathrm{d}}_{t-1} = \begin{pmatrix} \tilde{d}(t-1) & \dots & \tilde{d}(t-m) \end{pmatrix}^T \quad .$$

The dimensions of the additional matrices in (A.3a) and (A.3b) are

$$\mathbf{S}_i(t) = \underbrace{\begin{pmatrix} \bullet \end{pmatrix}}_{1} \Big\} M \qquad \mathbf{R}_i(t) = \underbrace{\begin{pmatrix} \quad \bullet \quad \end{pmatrix}}_{M} \Big\} M \quad .$$

The problem is now to calculate the coefficients $\mathbf{S}_i(t)$ and $\mathbf{R}_i(t)$ so that the mean square error of the symbol (vector) estimate $\hat{d}(t|tN+\ell)$ is minimized. Denote the estimation error by $\varepsilon(t)$:

$$\varepsilon(t) = d(t) - \hat{d}(t|tN+\ell) \quad . \tag{A.4}$$

In order to minimize the expected value of $\varepsilon^H(t), \varepsilon(t)$, we adjust the estimator coefficients so that $\varepsilon(t)$ is orthogonal to all signals which the estimate $\hat{d}(t|tN+\ell)$ may be based upon, i.e. $\mathrm{y}_{tN+\ell}$ and $-\tilde{\mathrm{d}}_{t-1}$:

$$E\left[ \begin{pmatrix} \mathrm{y}_{tN+\ell} \\ -\tilde{\mathrm{d}}_{t-1} \end{pmatrix} \varepsilon^H(t) \right] = 0 \quad . \tag{A.5}$$

Insert $\varepsilon(t)$ from (A.4) and $\hat{d}(t|tN+\ell)$ from (A.2) into the orthogonality relation (A.5)

$$E\left[ \begin{pmatrix} \mathrm{y}_{tN+\ell} \\ -\tilde{\mathrm{d}}_{t-1} \end{pmatrix} d^H(t) \right] = E\left[ \begin{pmatrix} \mathrm{y}_{tN+\ell} \\ -\tilde{\mathrm{d}}_{t-1} \end{pmatrix} \left( \Theta_S^H(t) \mathrm{y}_{tN+\ell} - \Theta_R^H(t) \tilde{\mathrm{d}}_{t-1} \right)^H \right]$$

$$E\left[ \begin{pmatrix} \mathrm{y}_{tN+\ell}\mathrm{y}_{tN+\ell}^H & -\mathrm{y}_{tN+\ell}\tilde{\mathrm{d}}_{t-1}^H \\ -\tilde{\mathrm{d}}_{t-1}\mathrm{y}_{tN+\ell}^H & \tilde{\mathrm{d}}_{t-1}\tilde{\mathrm{d}}_{t-1}^H \end{pmatrix} \begin{pmatrix} \Theta_S(t) \\ \Theta_R(t) \end{pmatrix} \right] = E\left[ \begin{pmatrix} \mathrm{y}_{tN+\ell} \\ -\tilde{\mathrm{d}}_{t-1} \end{pmatrix} d^H(t) \right]$$

$$\begin{pmatrix} E\left[\mathrm{y}_{tN+\ell}\mathrm{y}_{tN+\ell}^H\right] & -E\left[\mathrm{y}_{tN+\ell}\tilde{\mathrm{d}}_{t-1}^H\right] \\ -E\left[\tilde{\mathrm{d}}_{t-1}\mathrm{y}_{tN+\ell}^H\right] & E\left[\tilde{\mathrm{d}}_{t-1}\tilde{\mathrm{d}}_{t-1}^H\right] \end{pmatrix} \begin{pmatrix} \Theta_S(t) \\ \Theta_R(t) \end{pmatrix} = \begin{pmatrix} E[\mathrm{y}_{tN+\ell} d^H(t)] \\ -E[\tilde{\mathrm{d}}_{t-1} d^H(t)] \end{pmatrix} \quad . \tag{A.6}$$

Assume all previous decisions to be correct, i.e. assume that $\tilde{d}(t-j) = d(t-j)$, $1 \leq j \leq m$. Also, utilize the assumption of uncorrelated symbols to simplify (A.6) to

$$\begin{pmatrix} E[\mathrm{y}_{tN+\ell}\mathrm{y}_{tN+\ell}^H] & -E[\mathrm{y}_{tN+\ell}\mathrm{d}_{t-1}^H] \\ -E[\mathrm{d}_{t-1}\mathrm{y}_{tN+\ell}^H] & I \end{pmatrix} \begin{pmatrix} \Theta_S(t) \\ \Theta_R(t) \end{pmatrix} = \begin{pmatrix} E[\mathrm{y}_{tN+\ell}d^H(t)] \\ 0 \end{pmatrix} \quad , \tag{A.7}$$

where

$$\mathrm{d}_{t-1} = \begin{pmatrix} d(t-1) & \dots & d(t-m) \end{pmatrix}^T \quad .$$

In order to proceed, we need an expression for $\mathrm{y}_{tN+\ell}$. We set the smoothing lag $\ell = pN + r$, $0 \leq r < N$, and split $\mathrm{y}_{tN+\ell}$ into $p+1$ subvectors:

$$\mathrm{y}_{tN+\ell} = \begin{pmatrix} \bar{\mathrm{y}}_{p+1} \\ \bar{\mathrm{y}}_p \\ \bar{\mathrm{y}}_{p-1} \\ \vdots \\ \bar{\mathrm{y}}_1 \end{pmatrix} \quad ,$$

where

$$\bar{\mathrm{y}}_j = \begin{cases} \begin{pmatrix} y(tN+jN-1) & y(tN+jN-2) & \dots & y(tN-(j-1)N) \end{pmatrix}^T & \text{for } j = 1, 2, \dots, p \\ \begin{pmatrix} y(tN+pN+r) & y(tN+pN+r-1) & \dots & y(tN+pN) \end{pmatrix}^T & \text{for } j = p+1. \end{cases}$$

By using (A.1), the subvector $\bar{\mathrm{y}}_j$ can be written

$$\bar{\mathrm{y}}_j = \begin{pmatrix} \mathbf{B}_{N-1}(tN+jN-1) & \mathbf{B}_{2N-1}(tN+jN-1) & \dots & \mathbf{B}_{mN-1}(tN+jN-1) \\ \mathbf{B}_{N-2}(tN+jN-2) & \mathbf{B}_{2N-2}(tN+jN-2) & \dots & \mathbf{B}_{mN-2}(tN+jN-2) \\ \vdots & \vdots & & \vdots \\ \mathbf{B}_0(tN+(j-1)N) & \mathbf{B}_N(tN+(j-1)N) & \dots & \mathbf{B}_{mN}(tN+(j-1)N) \end{pmatrix} \times$$
$$\times \begin{pmatrix} d(t+j-1) \\ d(t+j-2) \\ \vdots \\ d(t+j-m-1) \end{pmatrix} + \begin{pmatrix} v(tN+jN-1) \\ v(tN+jN-2) \\ \vdots \\ v(tN+jN-N) \end{pmatrix}$$

for $j = 1, 2, \ldots, p$, and

$$\bar{\mathrm{y}}_{p+1} = \begin{pmatrix} \mathbf{B}_r(tN+\ell) & \mathbf{B}_{r+N}(tN+\ell) & \ldots & \mathbf{B}_{r+mN}(tN+\ell) \\ \mathbf{B}_{r-1}(tN+\ell-1) & \mathbf{B}_{r-1+N}(tN+\ell-1) & \ldots & \mathbf{B}_{r-1+mN}(tN+\ell-1) \\ \vdots & \vdots & & \vdots \\ \mathbf{B}_0(tN+pN) & \mathbf{B}_N(tN+pN) & \ldots & \mathbf{B}_{mN}(tN+pN) \end{pmatrix} \times$$

$$\times \begin{pmatrix} d(t+p) \\ d(t+p-1) \\ \vdots \\ d(t+p-m) \end{pmatrix} + \begin{pmatrix} v(tN+pN+r) \\ v(tN+pN+r-1) \\ \vdots \\ v(tN+pN) \end{pmatrix}$$

for $j = p+1$. Define

$$\beta_i(t) = \underbrace{\begin{pmatrix} \mathbf{B}_{iN+N-1}(tN+N-1) \\ \mathbf{B}_{iN+N-2}(tN+N-2) \\ \vdots \\ \mathbf{B}_{iN}(tN) \end{pmatrix}}_{M} \left.\vphantom{\begin{pmatrix} a \\ a \\ \vdots \\ a \end{pmatrix}}\right\} N$$

and

$$\bar{\beta}_i(t) = \underbrace{\begin{pmatrix} \mathbf{B}_{iN+r}(tN+\ell) \\ \mathbf{B}_{iN+r-1}(tN+\ell-1) \\ \vdots \\ \mathbf{B}_{iN}(tN+pN) \end{pmatrix}}_{M} \left.\vphantom{\begin{pmatrix} a \\ a \\ \vdots \\ a \end{pmatrix}}\right\} r+1 \quad .$$

By utilizing these definitions, the subvector $\bar{\mathrm{y}}_j$ may be written

$$\bar{\mathrm{y}}_j = \begin{pmatrix} \beta_0(t+j-1) & \beta_1(t+j-1) & \ldots & \beta_m(t+j-1) \end{pmatrix} \begin{pmatrix} d(t+j-1) \\ d(t+j-2) \\ \vdots \\ d(t+j-1-m) \end{pmatrix} +$$

$$+ \begin{pmatrix} v(tN+jN-1) \\ v(tN+jN-2) \\ \vdots \\ v(tN+jN-N) \end{pmatrix} \tag{A.8}$$

for $j = 1, 2, \ldots, p$, and

$$\bar{\mathrm{y}}_{p+1} = \begin{pmatrix}\bar{\beta}_0(t+p) & \bar{\beta}_1(t+p) & \ldots & \bar{\beta}_m(t+p)\end{pmatrix} \begin{pmatrix} d(t+p) \\ d(t+p-1) \\ \vdots \\ d(t+p-m) \end{pmatrix} + \begin{pmatrix} v(tN+pN+r) \\ v(tN+pN+r-1) \\ \vdots \\ v(tN+pN) \end{pmatrix} \tag{A.9}$$

for $j = p+1$. By employing (A.8) and (A.9), the complete measurement vector $\mathrm{y}_{tN+\ell}$ can be written

$$\mathrm{y}_{tN+\ell} = \begin{pmatrix} \bar{\beta}_0(t+p) & \bar{\beta}_1(t+p) & \ldots & \bar{\beta}_m(t+p) & 0 & \ldots & 0 \\ 0 & \beta_0(t+p-1) & \beta_1(t+p-1) & \ldots & \beta_m(t+p-1) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \beta_0(t) & \beta_1(t) & \ldots & \beta_m(t) \end{pmatrix} \times$$

$$\times \begin{pmatrix} d(t+p) \\ \vdots \\ d(t-m) \end{pmatrix} + \begin{pmatrix} v(tN+\ell) \\ v(tN+\ell-1) \\ \vdots \\ v(tN) \end{pmatrix} =$$

$$= \begin{pmatrix} \bar{\beta}_0(t+p) & \bar{\beta}_1(t+p) & \ldots & \bar{\beta}_p(t+p) \\ 0 & \beta_0(t+p-1) & \ldots & \beta_{p-1}(t+p-1) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & \beta_0(t) \end{pmatrix} \begin{pmatrix} d(t+p) \\ \vdots \\ d(t) \end{pmatrix} +$$

$$+ \begin{pmatrix} \bar{\beta}_{p+1}(t+p) & \ldots & 0 \\ \beta_p(t+p-1) & \ldots & \vdots \\ \vdots & & 0 \\ \beta_1(t) & \ldots & \beta_m(t) \end{pmatrix} \begin{pmatrix} d(t-1) \\ \vdots \\ d(t-m) \end{pmatrix} + \begin{pmatrix} v(tN+\ell) \\ v(tN+\ell-1) \\ \vdots \\ v(tN) \end{pmatrix} . \tag{A.10}$$

Define

$$\bar{\mathrm{d}}_{t+p} = \begin{pmatrix} d(t+p) & \dots & d(t) \end{pmatrix}^T \tag{A.11}$$

$$\mathrm{v}_{tN+\ell} = \begin{pmatrix} v(tN+\ell) & v(tN+\ell-1) & \dots & v(tN) \end{pmatrix}^T \tag{A.12}$$

$$\mathcal{F}(t) = \begin{pmatrix} \bar{\beta}_0(t+p) & \dots & & \bar{\beta}_p(t+p) \\ 0 & \beta_0(t+p-1) & \dots & \beta_{p-1}(t+p-1) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \beta_0(t) \end{pmatrix} \tag{A.13}$$

$$\mathcal{G}(t) = \begin{pmatrix} \bar{\beta}_{p+1}(t+p) & \dots & \bar{\beta}_m(t+p) & 0 & \dots & 0 & 0 \\ \beta_p(t+p-1) & \dots & & \beta_m(t+p-1) & \ddots & & \vdots \\ \vdots & & & & & \ddots & 0 \\ \beta_1(t) & \dots & & & & & \beta_m(t) \end{pmatrix} \tag{A.14}$$

By using these and earlier definitions, (A.10) can be more compactly expressed as:

$$\mathrm{y}_{tN+\ell} = \mathcal{F}(t)\bar{\mathrm{d}}_{t+p} + \mathcal{G}(t)\mathrm{d}_{t-1} + \mathrm{v}_{tN+\ell} \tag{A.15}$$

The expectations in (A.7) can now readily be evaluated:

$$\begin{aligned} E[\mathrm{y}_{tN+\ell}\mathrm{y}^H_{tN+\ell}] &= \mathcal{F}(t)E[\bar{\mathrm{d}}_{t+p}\bar{\mathrm{d}}^H_{t+p}]\mathcal{F}^H(t) + \mathcal{G}(t)E[\mathrm{d}_{t-1}\mathrm{d}^H_{t-1}]\mathcal{G}^H(t) + E\mathrm{v}_{tN+\ell}\mathrm{v}^H_{tN+\ell} \\ &= \mathcal{F}(t)\mathcal{F}^H(t) + \mathcal{G}(t)\mathcal{G}^H(t) + \Psi \end{aligned} \tag{A.16a}$$

$$E[\mathrm{y}_{tN+\ell}\mathrm{d}^H_{t-1}] = \mathcal{F}(t)E[\bar{\mathrm{d}}_{t+p}\mathrm{d}^H_{t-1}] + \mathcal{G}(t)E[\mathrm{d}_{t-1}\mathrm{d}^H_{t-1}] + E[\mathrm{v}_{tN+\ell}\mathrm{d}^H_{t-1}] = \mathcal{G}(t) \tag{A.16b}$$

$$\begin{aligned} E[\mathrm{y}_{tN+\ell}d^H(t)] &= \mathcal{F}(t)E[\bar{\mathrm{d}}_{t+p}d^H(t)] + \mathcal{G}(t)E[\mathrm{d}_{t-1}d^H(t)] + E[\mathrm{v}_{tN+\ell}d^H(t)] \\ &= \begin{pmatrix} \bar{\beta}_p(t+p) \\ \beta_{p-1}(t+p-1) \\ \vdots \\ \beta_0(t) \end{pmatrix} = \begin{pmatrix} \mathbf{B}_\ell(t+\ell) \\ \mathbf{B}_{\ell-1}(t+\ell-1) \\ \vdots \\ \mathbf{B}_0(t) \end{pmatrix} \triangleq h(t) \end{aligned} \tag{A.16c}$$

Above, frequent use has been made of the assumption that $E[d(t)d^H(\tau)] = \delta_{t\tau}I$. We have also defined

$$\Psi = \begin{pmatrix} \psi_0 & \psi_1 & \dots & \psi_\ell \\ \psi_{-1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & & \psi_1 \\ \psi_{-\ell} & \dots & \psi_{-1} & \psi_0 \end{pmatrix}$$

By inserting (A.16a)–(A.16c) into the normal equations (A.7), we obtain

$$\begin{pmatrix} \mathcal{F}(t)\mathcal{F}^H(t) + \mathcal{G}(t)\mathcal{G}^H(t) + \Psi & -\mathcal{G}(t) \\ -\mathcal{G}^H(t) & I \end{pmatrix} \begin{pmatrix} \Theta_S(t) \\ \Theta_R(t) \end{pmatrix} = \begin{pmatrix} h(t) \\ 0 \end{pmatrix}$$

or

$$\left(\mathcal{F}(t)\mathcal{F}^H(t) + \mathcal{G}(t)\mathcal{G}^H(t) + \Psi\right)\Theta_S(t) - \mathcal{G}(t)\Theta_R(t) = h(t) \tag{A.17a}$$

$$-\mathcal{G}^H(t)\Theta_S(t) + \Theta_R(t) = 0 \tag{A.17b}$$

By eliminating $\Theta_R(t)$ from (A.17a) by use of (A.17b), we finally get

$$\left(\mathcal{F}(t)\mathcal{F}(t)^H + \Psi\right)\Theta_S(t) = h(t) \tag{A.18a}$$

$$\Theta_R(t) = \mathcal{G}^H(t)\Theta_S(t) \tag{A.18b}$$

The solution to the linear system of equations (A.18a) and (A.18b) are the (matrix) coefficients of the MSE optimal single input-multiple output fractionally spaced decision feedback equalizer.

**Remark 1** Note that the linear system of equations (A.18a) and (A.18b) are solved for multiple right hand sides simultaneously, thereby only requiring one LU factorization.

**Remark 2** In principle, the derivation is valid also for a DFE with multiple inputs. Such a DFE could occur if multiple antennas were employed for space diversity.