S-38.220
Licentiate Course on Signal Processing in Communications,
FALL - 97

The Mobile Radio Channel and how to live with it

Juha Heiskala

Nokia Research Center
Radio Communications Laboratory
P.O. Box 407, FIN-00045 NOKIA GROUP

juha.heiskala@research.nokia.com

Date: 30.10.1997

ABSTRACT

Multipath fading channel is the dominant channel type of current mobile cellular radio networks. To obtain reasonable receiver performance in this type of channel its characteristics must be understood and then suitable countermeasures can be designed. The channel characteristics of a planned mobile radio system must also be taken into consideration in the system design phase. As the type of transmitted signal defines how it behaves in a given channel.

This paper gives an introduction to mobile radio channel modelling and characteristics. It is shown that the mobile radio channel is in general a difficult channel type for the receiver to obtain good performance. Next some methods to improve the receiver performance are introduced. Specifically channel estimation, equalization and diversity. The performance improved achievable with these techniques is discussed.

1. INTRODUCTION

The classical model of a radio channel is the Additive White Gaussian Noise (AWGN) channel model. It consists of statistically independent noise samples additively corrupting the received signal. The noise samples' amplitude has a Gaussian probability density function (pdf) and because the noise samples are independent their autocorrelation function is ideally an impulse. From this it follows that the AWGN channel's power spectral density is flat over all frequencies. So all signal frequencies are identically degraded by the AWGN channel. Also the AWGN channel is usually assumed to be stationary, so that its behaviour does not change with time.

However the AWGN channel does not model mobile radio channel behaviour adequately. The radio channel encountered in current mobile cellular radio networks is a combination of several different propagation paths. This phenomenon is called multipath propagation. The multipath effect creates a channel that is much more difficult to use in communication systems than the AWGN channel. The main reason for this is that the several different propagation paths introduce severe degradation into the received signal. There are two main types of degradation caused by the multipath propagation. The first is intersymbol interface (ISI) caused by different propagation delays of the multipaths, non-constant amplitude and non-linear phase response of the channel. The second type is fading, which is caused by destructive interference of the carrier waves of different propagation paths. The fading causes loss of Signal to Noise Ratio (SNR).

The performance of the receiver in multipath channel depends crucially on how well the degradation caused by the channel can be removed. The ISI effects can be removed by estimating the channel and then equalizing it. If the channel estimate is ideal this effectively removes the ISI totally. However ideal channel estimation is usually not possible, either because the impulse response of the channel is too long for a practical equalizer, or the channel may have non-minimum phase resulting in an unstable ideal equalizer.

The different propagation paths can be used to mitigate the multipath fading. The reason for this is that the propagation paths are usually uncorrelated, which means that it is very unlikely that all of them are faded simultaneously. So if we somehow receive the different propagation paths independently and then combine them, the multipath fading can be cancelled. This technique is called diversity reception. It usually achieves reasonably good signal in most cases. A receiver that uses diversity to combine several independent replicas of the transmitted signal is called a RAKE-receiver.

In this paper the mobile radio channel model characteristics are first discussed in chapter 2. The different sources of multipath propagation and the resulting ISI and fading types. The receiver performance effects of these are also briefly introduced. In chapter 3 the channel estimation problem is discussed. First is given an general view of a system identification problem, which is then viewed from a communications system point of view. Also some algorithms to perform channel estimation are introduced. Chapter 4 deals with diversity, what kind of diversity we can expect and how to take advantage of it. The RAKE-receiver structure is shown and explained. Finally are some performance gains achievable with diversity reception. Chapter 5 gives conclusions about the paper.

2. MOBILE RADIO CHANNEL CHARACTERISATION

Modelling of fading channels dates back to 1950s and 1960s. At that time used channels were over the horizon communications. Nowadays the dominant fading channel is the cellular mobile radio channel. However the same models can be applied in both cases.

2.1 Multipath propagation sources

When a mobile receiver is moving around, it receives signal from different propagation paths, which are generated by three basic sources: reflection, diffraction and scattering [1]. These effects are described below.

- Reflection path occurs when the signal is reflected from a smooth surface, that has dimensions significantly larger than the signal wavelength.
- Diffraction path is generated when the signal encounters an object larger than the wavelength of the signal. The object causes secondary waves to be formed behind itself. These secondary waves enable communications behind impenetrable objects even without reflected signal paths. These mode of multipath propagation is also called *shadowing*.
- Scattering occurs when the signal encounters a large rough surface. The roughness is compared to the signal wavelength. Scattering occurs also if the signal hits any object about the signal of signal wavelength or less. These kinds of objects cause the signal energy to spread in all directions.

2.2 Fading types

Fading can be roughly divided into two categories: large- and small-scale fading. The large scale fading is caused by large motions of receiver, whereas small-scale fading is caused by an order of half a wavelength changes in the distance between transmitter and receiver.

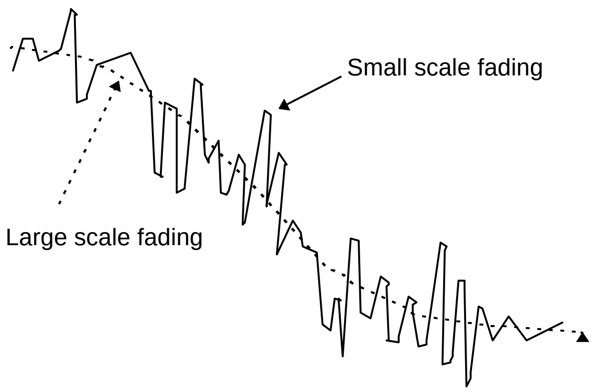

Figure 1 Large- and small scale fading, [1]

The large-scale fading is also called log-normal fading, because its amplitude has a log-normal probability density function. The small-scale fading is usually described as Rayleigh- or Rice-fading, depending on which pdf-describes it.

2.3 Modelling small-scale fading

In this text the Raylaigh-fading is used as the small-scale fading phenomenon. If the different propagation paths are assumed to be uncorrelated then a simple model can be build for the fading phenomenon [1]. With this assumption it is possible to proof that the channel is wide sense stationary, both in time- and frequency-domain. The resulting random process is called Wide Sense Stationary Uncorrelated Scattering (WSSUS) process. The whole model is build from four functions, that describe the fading process in time- and frequency-domains. These functions are:

- Multipath intensity profile function
- Spaced-frequency correlation function
- Spaced-time correlation function
- Doppler power spectrum

First two of these functions describe what kind of the channel is, how it affects the transmitted signal. The latter two functions describe how the channel itself changes as a function of time.

2.3.1 Multipath intensity profile function

The multipath intensity profile $S(\tau)$ function describes how the signal is spread in time. This function shows how the average power of the received signal varies as a function of time delay. The origin point of time-axis is the first received signal component above some threshold value. The support of this function gives the *maximum excess delay* T_m of the signal. If a single impulse is transmitted the T_m is the time delay between the first and last signal component above the threshold level. Figure 2 shows an example of $S(\tau)$.

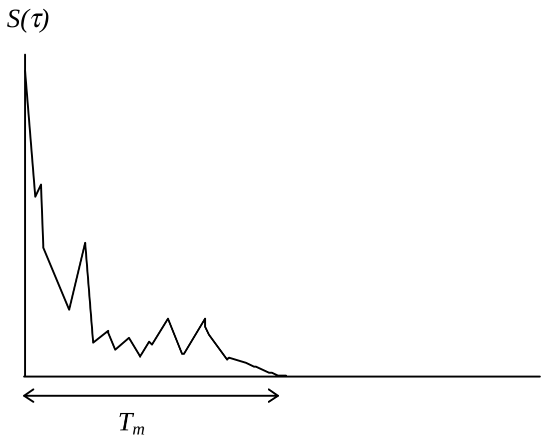

Figure 2 Multipath intensity profile, adapted from [1]

2.3.2 Spaced-frequency correlation function

The spaced-frequency correlation function $|R(\Delta f)|$, which is the Fourier transform of $S(\tau)$ gives information about how equally the channel distorts sinusoids having frequency difference Δf. The $|R(\Delta f)|$ can be viewed as a correlation between these two sinusoids. Width of the support of $|R(\Delta f)|$ is called the *coherence bandwidth* f_0 of the channel. The f_0 and T_m are approximately reciprocally related $f_0 \cong k*1/T_m$, k is a multiplicative constant.
If the is $f_0 < 1/T_s$ the channel is called frequency selective. In this case the signal bandwidth $W \cong 1/T_s$ is wider than f_0, this means that the channel distorts signal frequency components unequally. This causes channel induced ISI, which must be somehow equalized to achieve reasonable receiver performance.
If $f_0 > W$ the channel is called frequency non-selective or flat-fading channel. In this case the channel affects the whole signal spectrum in the same way and thus does not cause ISI.

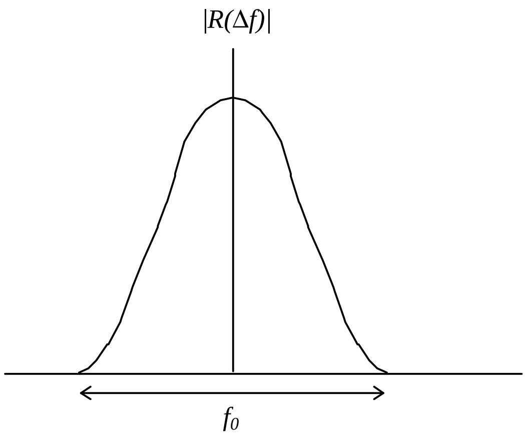

Figure 3 Spaced-frequency correlation function, adapted from [1]

2.3.3 Spaced-time correlation function

The spaced-time correlation function $|R(\Delta t)|$ is a dual function of the $|R(\Delta f)|$ in the time-domain. The $|R(\Delta t)|$ describes how the channel changes as a function of Δt. The width of $|R(\Delta t)|$ is called the *coherence time* T_0 of the channel. If the $T_0 < T_s$ the channel varies during the symbol period causing distortion, this is called fast-fading. When $T_0 > T_s$ the channel is constant during one symbol period, this is called slow-fading.

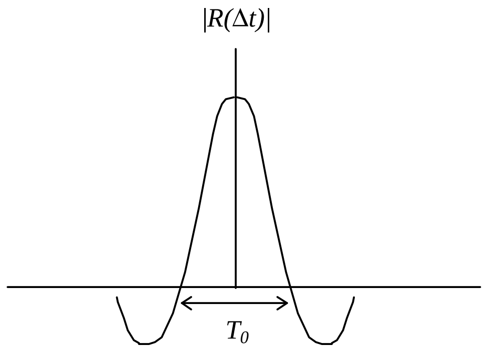

Figure 4 Spaced-time correlation function, adapted from [1]

2.3.4 Doppler power spectrum

The Doppler power spectrum $S(\upsilon)$ is the inverse Fourier transform of the $|R(\Delta t)|$ and it describes how the multipath propagation spreads the spectrum of the transmitted signal. The $S(\upsilon)$ is dual function of the $S(\tau)$ in the frequency domain. Width f_d of the support of $S(\upsilon)$ is called *Doppler spread* or *fading bandwidth* of the channel.
The $S(\upsilon)$ indicates how fast is the fading process of the channel. If $f_d > W$ the channel is called fast fading, which means that the received signal power varies during the symbol period. If $f_d < W$ the channel is called slowly fading and the signal power remains constant during one symbol period.

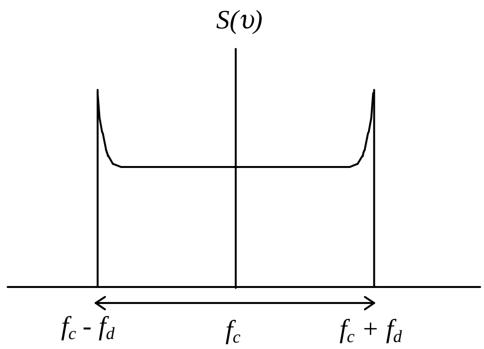

Figure 5 Doppler power spectrum, adapted from [1]

2.4 Channel types

Comparing the symbol length T_s and signal bandwidth $W \cong 1/T_s$ to channel parameters we can identify four different channel types. These are described below.

2.4.1 Slow, frequency non-selective fading

The channel is characterised as slow, frequency nonselective fading channel, if $T_s > T_m$ and $T_s < T_0$, or in the frequency domain if $W < f_0$ and $W > f_d$. In practise this means that all the propagation paths are received within one symbol period and that all the signal frequency components are equally affected by the channel. This is the easiest type of fading channels to deal with.

2.4.2 Fast, frequency non-selective fading

The channel is characterised as fast, frequency nonselective fading channel, if $T_s > T_m$ and $T_s > T_0$, or in the frequency domain if $W < f_0$ and $W < f_d$. In this case the channel varies during the symbol period. Fast fading is very diffucult to compensate effectively.

2.4.3 Slow, frequency selective fading

The channel is characterised as slow, frequency selective fading channel, if $T_s < T_m$ and $T_s < T_0$, or in the frequency domain if $W > f_0$ and $W > f_d$. This channel type distorts the received symbol in the frequency-domain, not all the frequency components are equally affected. The consequence is that the baseband pulse shape is distorted, which in turn causes intersymbol interference.

2.4.4 Fast, frequency selective fading

The channel is characterised as fast, frequency selective fading channel, if $T_s < T_m$ and $T_s > T_0$, or in the frequency domain if $W > f_0$ and $W < f_d$. This is the worst type of channel for the receiver. The received symbol is spread over many symbol periods and also the channel varies during the symbol peroid. In this kind of channel without any countermeasures reliable communications is not possible, no matter how much power is transmitted.

2.4.5 Bit Error Rate performance

The different channel behaviours cause different receiver bit error rate (BER) performance. Figure 6 shows BER curves for different types of multipath fading channels.

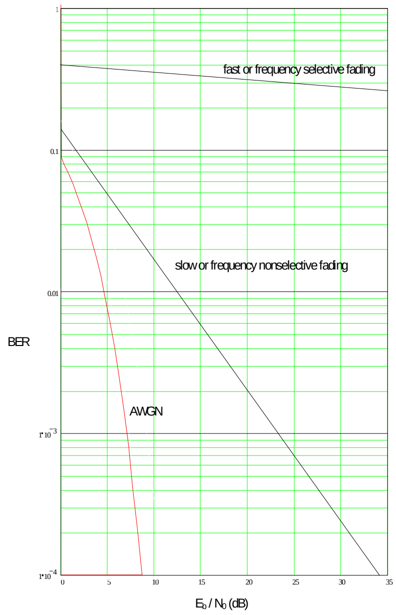

Figure 6 BER curves by channel type, adapted from [2]

From figure 6 we can clearly see, that multipath fading effects must somehow be mitigated to obtain reasonable receiver performance. In fast or frequency selective fading channel the BER performance is miserable, no practical amount of received energy per bit will achieve acceptable performance.

With slow or frequency nonselective fading channel we could possible live with, if absolutely no other option is available. Increasing transmitted power by 10dB we gain one decade in BER performance, however something should be done to obtain performance approaching the AWGN channel. What to do is the objective of next two chapters.

3. CHANNEL ESTIMATION AND EQUALIZATION

The goal of the channel equalization is to cancel the distortion caused by the channel to the transmitted signal as perfectly as possible. In principle this is achieved by estimating the impulse response of the channel and then convolving the received signal with a complex-conjugate, time-reversed version of the estimated impulse response. Ideally the result of the convolution is exactly the transmitted signal. So to successfully equalize the channel induced distortion we need an estimate of the channel, how to obtain this estimate is the subject of this chapter.

3.1 System identification basics

Channel estimation is a special case of a more general system identification problem presented in figure 7.

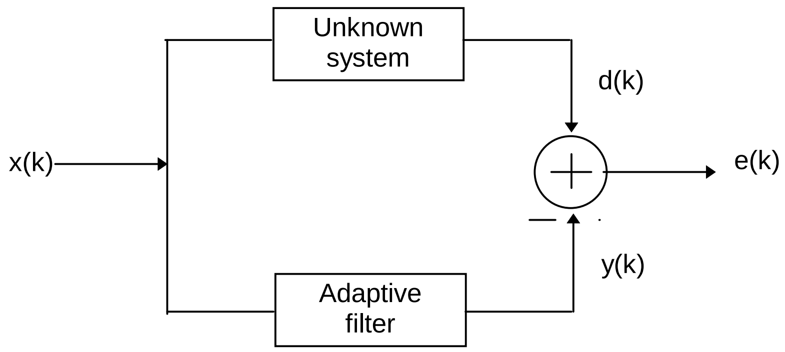

Figure 7 System identification, adapted from [3]

In the system identification problem the x(k) is usually white noise that is fed to both the unknown system and the adaptive filter. The unknown system gives response d(k), called the desired signal and the output of the adaptive filter is y(k). Subtracting these we obtain the error signal e(k). Then the adaptive filter coefficients are updated according to some algorithm that tries to minimize e(k). After the adaptive filter has converged, its impulse response gives us an estimate of the unknown system.
The algorithm that updates the adaptive filter coefficients is the main subject of system identification problem. Many different algorithms are available. They differ in their convergence speed, computational complexity, performance with nonwhite input signal, stability with fixed point arithmetic and so on. Most popular algorithms used in radio receivers are introduced later.

3.2 Channel estimation

Channel estimation causes some special problems into system identification. First we do not have the desired signal directly available. This is a communications system level problem. If it is known that a channel equalizer will be needed in a receiver, the system designers must include somekind of mechanism for obtaining the desired signal at the receiver. There is two main approaches for achieving this, training sequences and pilot channel or symbols. Training sequences are known data that is inserted into the transmitted data. The receiver knows the location of this training information and the channel estimation is the performed as the training sequence is received.
Pilot channel or symbols are known data sent continuously during the transmission at a known location. The receiver uses this pilot information to estimate the channel.
If the channel is time-variant, as a mobile radio channel is, the channel estimate must be continuously updated. If the system uses pilot information to obtain channel estimate, continuos updating does not differ from the initial training. If training data is used the desired signal must be obtained from somewhere after the training period has elapsed. The only alternative are the decisions we have made, we trust that the decisions are correct and use them as the desired signal. This is illustrated in figure 8, the equalizer structure has a switch to use either training signal or decisions as the desired signal.

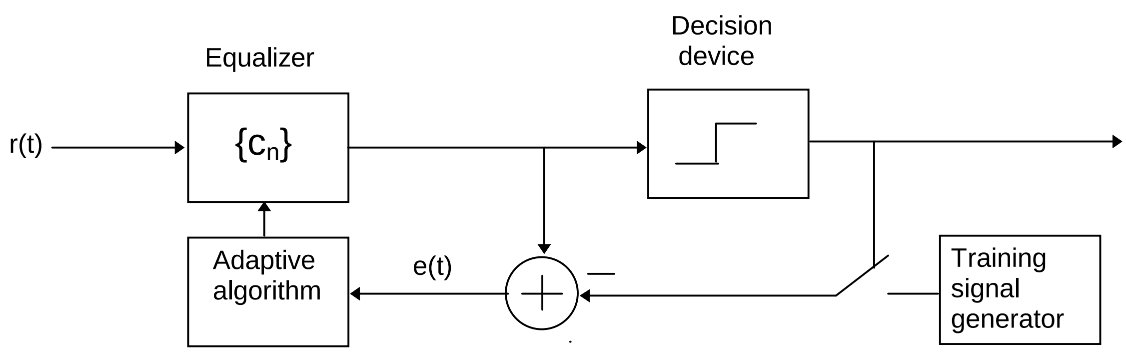

Figure 8 Adaptive equalizer, adapted from [4]

Crucial property of the adaptation process is that how fast channel variations it can follow. If the process is too slow our channel estimate is out-of-date and the equalizer output is perhaps even more distorted signal than the input. The spaced-time correlation function gives us an indication about how fast the channel will change and thus we can design the adaptation process so that it is capable of following the channel variations.

3.3 Adaptive algorithms

There exists a large number of algorithms for adapting the channel estimate [3]. However one is above the others in terms of popularity, the Least Mean Square (LMS) algorithm. The LMS algorithm was developed in 1960s and ever since it has been the dominant adaptive algorithm [4]. Its popularity has many reasons:

- The LMS algorithm has well developed theory and thus its performance can be reliably predicted.
- The conditions for guaranteed convergence of the algorithm are available.
- It is numerically stable in fixed point implementations.
- Low computational complexity, of order O(N), with equalizer length N.

Major drawback of the LMS is its somewhat slow convergence. And the property that the convergence speed is dependent on the input signal statistics, specifically on the eigenvalue spread of input signal correlation matrix [3]. The LMS performs best if the eigenvalue spread is one and the performance decreases with increasing eigenvalue spread.
During the last few years the Recursive Least Squares (RLS) adaptive algorithm has become a viable solution. Especially if very fast convergence is needed. And the speed of the convergence is independent of the eigenvalue spread of the input correlation matrix [3]. However, this speed does not come without cost. The RLS is computationally quite complex and can be unstable when implemented in fixed-point arithmetic. Anyway if fast convergence is needed RLS should be considered as an alternative for LMS.

3.4 Equalizer structures

The equalizer structure of figure 8 has a couple of drawbacks. First it operates at symbol rate, which usually means that aliasing is introduced [4]. The aliasing occurs, because usually the signal bandwidth is larger than the half of the symbol rate, thus the sampling theorem is violated. Secondly full advantage is not taken of the decisions.
The first drawback is fixed with a so called fractionally-spaced equalizer (FSE). Its basic idea is that the equalizer operates at a rate higher than the symbol rate, thus aliasing is not introduced. The reduction to the symbol rate is done after the equalizer. Fractionally-spaced equalizer usually outperforms the symbol rate equalizer [4]. The better performance comes at the cost of higher computational complexity, usually two or four times more arithmetic operations per unit time.
The second drawback is fixed with a so called decision-feedback equalizer (DFE), which is presented in figure 9.

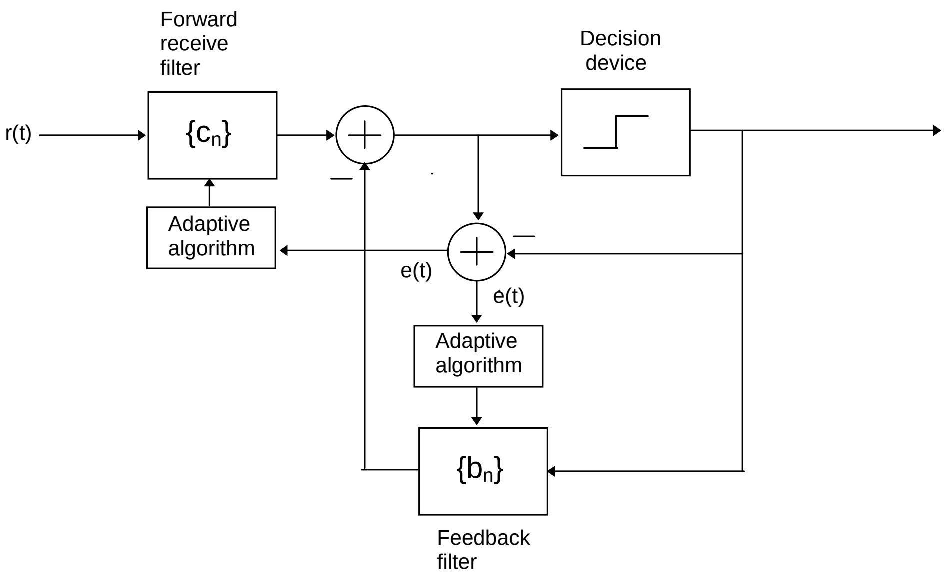

Figure 9 Decision-feedback equalizer, adapted from [4]

The idea of the DFE is to cancel the pre- and post-symbol ISI separately. This is can be seen from the figure 9. The forward receive filter cancels the presymbol distortion. The feedback filter cancels the postsymbol distortion, based on the made decision. The advantage of this structure is the fact that decisions are not contaminated by noise, thus the postsymbol interference cancellation does not enhance noise.

3.5 BER performance

The equalization process is able to cancel the effect of slow, frequency selective fading [2]. So with the aid of channel equalization the BER curve should approach the slow- or frequency nonselective-fading channel. But there is still some ground to gain to achieve the AWGN channel curve, how to get closer to that curve is the topic of the next chapter.

3.6 CDMA aspects of channel equalization

CDMA systems are inherently resistive against interference and ISI and multipath propagation are a form of interference. Thus channel equalization does not necessarily give directly much performance gain over the decorrelation process of normal CDMA receiver [2]. However many of the current advanced CDMA receiver algorithms require knowledge about the channel to perform multiple-access interference cancellation. This channel state information (CSI) can be obtained with the presented channel estimation algorithms.
Also current cellular CDMA networks include pilot channels to obtain synchronization these pilot channels can be used for the channel estimation.

4. DIVERSITY

Diversity receiver coherently combines several statistically independent replicas of the transmitted signal. These replicas can be intentionally generated to mitigate the multipath fading channel or they may be generated by the channel itself. Obvious example being the different propagation paths that are independent.
These independent replicas of the transmitted signal can be used to compensate the slow- or frequency nonselective channel [2], thus moving the multipath BER curve closer to the AWGN channel curve.

4.1 What kind of diversity

Different diversity replicas can be obtained in many ways:

- time diversity
- frequency diversity, (inherit in CDMA)
- space diversity
- polarization

Time diversity can be intentionally generated by using interleaving. The multipath propagation gives us another source of time diversity. In this chapter we assume that the multipath propagation is the source of diversity.
Space diversity can be obtained by using multiple antennas located so that they receive independent components of the transmitted signal. Polarization diversity is obtained by using antennas having different polarazation.

4.2 RAKE-receiver

In a multipath propagation channel the output from the CDMA decorrelator might look like in figure 10. Now if the receiver just selects the path p_2 as the received symbol significant portion of the signal energy is lost in the another propagation path components. The goal of the RAKE-receiver is to coherently receive all these multipath components and to optimally combine them. It is very unlikely that all the independent paths would have faded simultaneously, thus this combining process will usually significantly increase the signal-to-noise ratio (SNR).

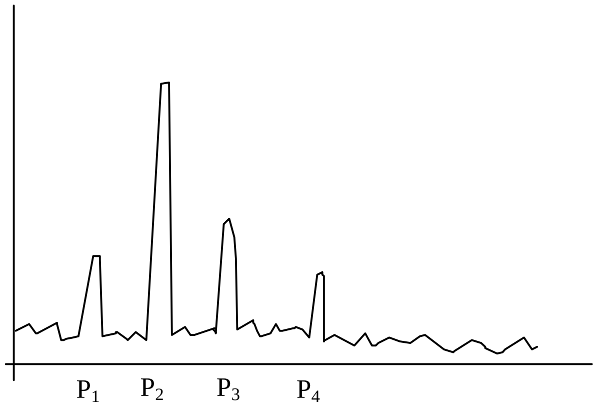

Figure 10 Multipath profile

Figure 11 shows the structure of the RAKE receiver. The goal of the receiver is to estimate the different propagation delays and then assign each propagation path to one branch of the RAKE receiver. The adaptive delays of the receiver are adjusted so that all the received paths arrive simultaneously at the combiner. The path resolution accuracy is $1/W$ [5]. For example the Interim Standard-95 (IS-95) has a bandwidth of about 1.25 MHz, so propagation paths separated by atleast about 1 μs can be resolved and used as diversity. Propagation delays shorter than 1 μs are not resolved and cause frequency selective fading. The good diversity resolution is one advantage of CDMA systems, that usually have very large signal bandwidths compared to another multiple access schemes.

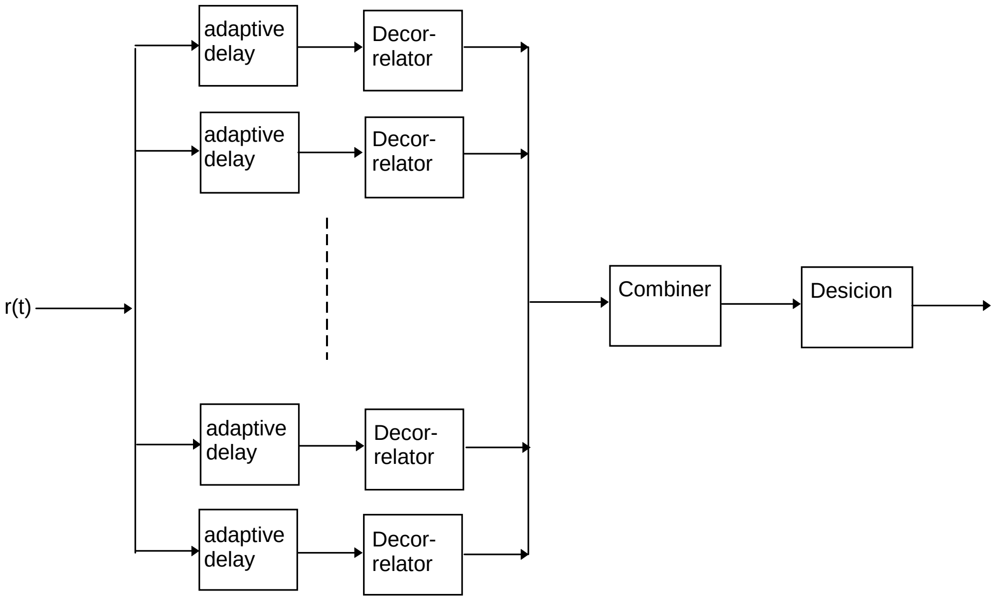

Figure 11 RAKE receiver

The combiner is an essential component of the RAKE-receiver. Its job is to generate as good signal as possible from the received diversity components. The optimal method to combine the diversity paths is called *Maximal ratio combiner* (MRC) [6]. The MRC generates the decision variables by multiplying each diversity path k by a factor $\alpha_k e^{j\phi k}$. After the multiplication the different paths are added together to generate the final decision variable. The α_k and ϕ_k are the channel attenuation and phase shift. The α_k factor gives stronger signal components more weight in the decision variable than the weaker ones. The $e^{j\phi k}$ factor corrects the phase shift of the channel.

4.3 Diversity performance gain

Figure 12 shows how the required E_b/N_0 for achieving a constant BER behaves as a function of received diversity paths [5]. From the figure we can clearly see that diversity is an essential method to achieve good receiver performance in fading multipath channel.

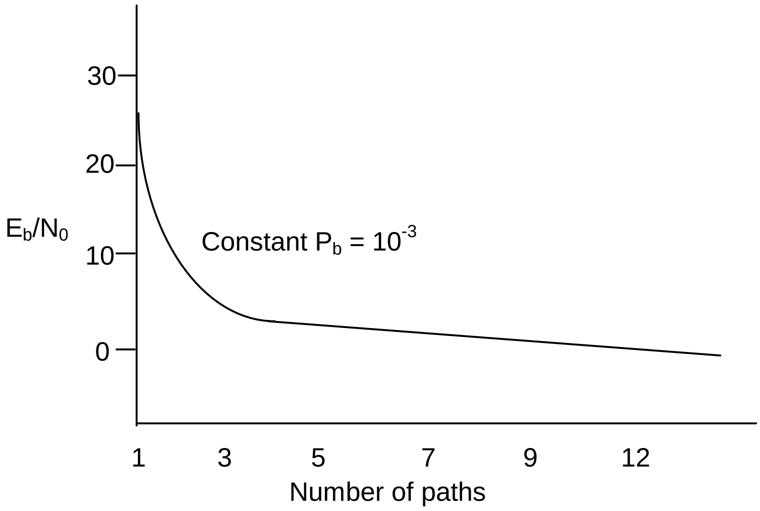

Figure 12 Required E_b/N_0 for constant BER, adapted from [5]

5. CONCLUSIONS

This paper first gives an introduction to mobile radio channel characteristics and the fading channel types a mobile receiver must be able to live with. It is shown that in a fading channel the BER performance of the receiver is bad or unacceptable without any attempt to mitigate the fading.

Then the channel estimation and equalization is introduced as a method to improve the receiver performance operating in a fading channel. The basic channel estimation algorithms and equalizer structures are discussed. The performance improvement achieved is shown to significantly improve BER of the receiver. Then are some comments about CDMA and channel equalization.

The next chapter shows that diversity is a powerful method the improve receiver performance in fading multipath channel. Some different diversity sources are introduced and an optimal method to combine different diversity paths in briefly discussed. As a final point is shown how the required E_b/N_0 for a constant BER behaves as function of the number of diversity paths.

REFERENCES

[1] Sklar, Bernard, “Rayleigh Fading Channels in Mobile Digital Communication Systems Part I: Characterization”, IEEE Communications Magazine, July 1997, pp. 90-100

[2] Sklar, Bernard, “Rayleigh Fading Channels in Mobile Digital Communication Systems Part II: Mitigation”, IEEE Communications Magazine, July 1997, pp. 102-109

[3] Diniz, Paolo S. R., “Adaptive Filtering, Algorithm and Practical Implementation”, Kluwer Academic Publishers, Norwell, Massachusetts, 1997

[4] Qureshi, S. U. H, “Adaptive Equalization”, Proceedings of the IEEE, Vol. 73, No. 9, September 1985

[5] Turin George L, Introduction to Spread-Spectrum Antimultipath Techniques and Their Application to Urban Digital Radio, Proceedings of the IEEE, Vol. 68 No. 3 March 1980

[6] Proakis, John, “Digital Communications”, New York, McGraw-Hill, 2ed. 1989