University of California
Los Angeles

# ON INDOOR WIRELESS CHANNEL CHARACTERIZATION AND THE DESIGN OF INTERFERENCE AWARE MEDIUM ACCESS CONTROL PROTOCOLS FOR PACKET SWITCHED NETWORKS

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Electrical Engineering

by

**George D. Kondylis**

2000

The dissertation of George D. Kondylis is approved.

Kirby Baker

Richard Wesel

Yahya Rahmat-Samii

Gregory Pottie, Committee Chair

University of California, Los Angeles

2000

*To my beloved parents*

# Table of Contents

# List of Figures

# List of Tables

# Acknowledgments

First and foremost, I would like to thank my advisor, Professor Gregory Pottie for his guidance, support and patience during my long years at UCLA. Professor Pottie gave me the independence to do the research that I wanted to, while always providing sound advice and criticism. I thank him also for allowing me, at a period when I did not yet quite know what to do with my research, to work as an intern for Pairgain. This way, I got a respite from the stress of deciding on a research project, while at the same time got new ideas. Professor Pottie gave me freedom on my work while making sure that I was always secure financially, and for this I cannot thank him enough.

I would also like to thank the other members of my committee, Professors Yahya Rahmat-Samii, Richard Wesel and Kirby Baker for their valuable comments during my qualifying exam, and of course for serving in my committee. I especially thank Professor Rahmat-Samii for his interest in my research work and for his excellent courses that taught me so much about electromagnetic theory.

Throughout the years of my studies in UCLA, I was tremendously fortunate to make a few very good friends. They have helped me through difficult times and together we have also had some of the best times. In particular, I would like to thank Dr. Dongsoo Koh and Dr. Franco De Flaviis for their friendship and their invaluable help in collecting the indoor field measurements for validating the new FDTD formulation. The night that we did the measurements will always remain in my mind as a most enjoyable experience!

I would also like to thank Pairgain for partially supporting my research, and for having me as an intern for six months. Especially, working with Pairgain's chief scientist, Dr. George Zimmerman, was a thrilling as well as challenging

experience.

I must also extend my gratefulness to HRL Laboratories for having me as an intern during the last year of my PhD and for providing me with an exciting research environment. My sincere appreciation goes to Son Dao for taking me on board, and making me a member of a fantastic group of researchers and good friends, and to Dr. Dennis Connors for suggesting HRL as an esteemed place to do research. I enjoyed working with Dr. Srikanth Krishnamurthy and I thank him for his help on my research. I also want to thank all the other members of our research group, and in particular my office-mate, Dr. Heungno Lee for his precious advice, friendship and wisdom on physical layer problems, which he shares with me so eagerly. And of course, our new manager, Dr. Bong Ryu, for his trust in me and his patience while I was frantically trying to finish writing my dissertation.

Finally, I want to express my gratefulness to my parents and my wife, Tassoula, for their love and support during my graduate studies. Tassoula has been my pillar of strength here in the United States, while without the unconditional love and support, both ethical and financial, of my parents, I would not have come to America. To me, my parents have always been the models to follow in my life, and if there is something I regret about coming to America, it is that I missed their company for so long. To them, I dedicate this dissertation.

# Vita

| | |
|---|---|
| 1969 | Born, Athens, Greece |
| 1987-1993 | B.S., Electrical Engineering<br>National Technical University, Athens, Greece |
| 1993-1995 | M.S., Electrical Engineering<br>University of California, Los Angeles, California |
| 1993-1996 | Graduate Student Researcher<br>Electrical Engineering Department<br>University of California, Los Angeles |
| 1996-1996 | Research Intern, Pairgain Technologies,<br>Tustin, California |
| 1997-1999 | Graduate Student Researcher<br>Electrical Engineering Department<br>University of California, Los Angeles |
| 1999-1999 | Research Intern, HRL Laboratories,<br>Malibu, California |
| 1999-present | Research Staff Member, HRL Laboratories,<br>Malibu, California |

# Publications

G. D. Kondylis, S. Krishnamurthy, G. J. Pottie, and S. Dao, "Multicasting sustained CBR and VBR traffic in wireless ad-hoc networks," in *Proceedings, IEEE International Conference on Communications, ICC 2000,* New Orleans, LA, June 18-22, 2000.

G. D. Kondylis and G. J. Pottie, "Co-channel interference analysis for the packet reservation multiple access," in *Proceedings, IEEE Globecom 99*, Rio de Janeiro, Brazil, Dec. 5-9, 1999.

G. D. Kondylis and G. J. Pottie, "Dynamic channel allocation strategies for wireless packet access," in *Proceedings, IEEE Vehicular Technology Conference, VTC 99 Fall*, Amsterdam, the Netherlands, Sept. 19-22 1999.

G. D. Kondylis, F. De Flaviis, G. J. Pottie, and Y. Rahmat-Samii, "Indoor channel characterization for wireless communications using reduced finite difference time domain (R-FDTD)," in *Proceedings, International Conference on Electromagnetics in Advanced Applications, ICEAA 1999,* Torino, Italy, Sept. 13-17 1999.

G. D. Kondylis, F. De Flaviis, G. J. Pottie, M. Sironen, and T. Itoh, "Reduced FDTD formulation (R-FDTD) for the analysis of 30 GHz dielectric resonator coupled to a microstrip line," in *Proceedings, IEEE International Microwave Symposium, MTT-S 1999*, Anahaim, CA, June 13-19 1999.

G. D. Kondylis, G. J. Pottie, and F. De Flaviis, "Generalized reduced FDTD formulation (R-FDTD) for the solution of Maxwell equations," in *Proceedings, Asia-Pacific Microwave Conference, APMC 1998,* Yokohama, Japan, Dec. 8-11 1998.

Abstract of the Dissertation

# ON INDOOR WIRELESS CHANNEL CHARACTERIZATION AND THE DESIGN OF INTERFERENCE AWARE MEDIUM ACCESS CONTROL PROTOCOLS FOR PACKET SWITCHED NETWORKS

by

**George D. Kondylis**

Doctor of Philosophy in Electrical Engineering

University of California, Los Angeles, 2000

Professor Gregory Pottie, Chair

This dissertation addresses two distinct topics: The first one has to do with the characterization of the indoor wireless channel. A good understanding of the propagation laws and the statistics obeyed by the indoor wireless channel is very important, especially in view of the rapidly emerging market for home and office wireless network products. The finite difference time domain method (FDTD), very well known in the field of microwave engineering, is used to predict the channel inside a residence. The validity and usability of the method is demonstrated by comparing our predictions against measurements that we did for the same residence. Motivated from the good agreement between the two, we re-formulate FDTD to reduce its memory requirement by 33%. We present the new formu-

lation and discuss experimental results that verify its equivalence to the original FDTD.

The second topic addressed by the dissertation has to do with interference adaptive medium access control (MAC) protocols for packet switched networks. This is also a very important topic for wireless transmission, because new wireless networks will be designed primarily for data and Internet services, with voice traffic also running on top of the popular Internet Protocol. We first study the co-channel interference statistics created by a simple packet switched MAC for micro-cellular environments, and demonstrate that it can be modeled very accurately as Markovian. We present a Markov analysis for the packet dropping probability of voice traffic for this MAC, when co-channel interference becomes the dominant impairment. Subsequently, motivated by this analysis we extend the simple MAC protocol to include dynamic channel allocation (DCA) schemes that adapt resource allocation based on interference measurements. We present two different DCA schemes and study their performance via simulations. We demonstrate through simulations that a single frequency re-use factor is possible when the MAC can do channel assignment on a packet by packet basis, and has the ability to utilize interference measurements and power control. Finally, in the context of an ad-hoc, or in other words infrastructure-less, network we present a MAC protocol that allows for the multicasting of real-time traffic. The problem solved here is that of concurrent bandwidth reservation and route formulation for the multicast mesh. We present a distributed protocol that uses one-hop neighborhood information and study its performance in terms of throughput via simulations.

# CHAPTER 1

# Wireless Channel Characterization and Modeling

## 1.1 Introduction

In wireless communications, information is carried through the air with the use of electromagnetic waves. If this propagation was done in free space, any static receiver would receive the transmitted signal without any kind of distortion, apart from the free space power attenuation. However, the electromagnetic waves used for wireless communications usually propagate in very complex and time varying environments, and receivers are mobile rather than static. For example, outdoor environments include high-rise urban areas, less densely built suburbs, or sparsely built rural areas. Examples of indoor environments include office buildings, warehouses or residences. Time variations in the environment are introduced by movement of the transmitter, receiver, or of humans and objects between them (for example, in a large warehouse, the movement of large vehicles). It easily understood then, that a signal arriving at the receiver has undergone multiple reflections, from many, possibly moving, objects. This introduces two major impairments to the wireless channel, namely *dispersion* and *fading*. Dispersion refers to the spreading of the transmitted signal in time, and is due to multipath (multi-reflection) propagation. Fading refers to the time variation of the received

signal, which is due to the time varying nature of the propagation environment. Such time variations of the received signal result in frequency spreading that is generally called *Doppler spreading*.

To take a concrete example of multipath propagation, consider Figure 1.1. The dark line depicts the magnitude of the received signal inside a residence when the transmitted signal is a very narrow impulse. This means that in essence we are measuring the *channel impulse response*. We see that the received signal has expanded in time, while we can also observe several discrete multipath components that arrive separated in time because of the different paths they have followed to the receiver. In Figure 1.2 the dark line shows the corresponding fre-

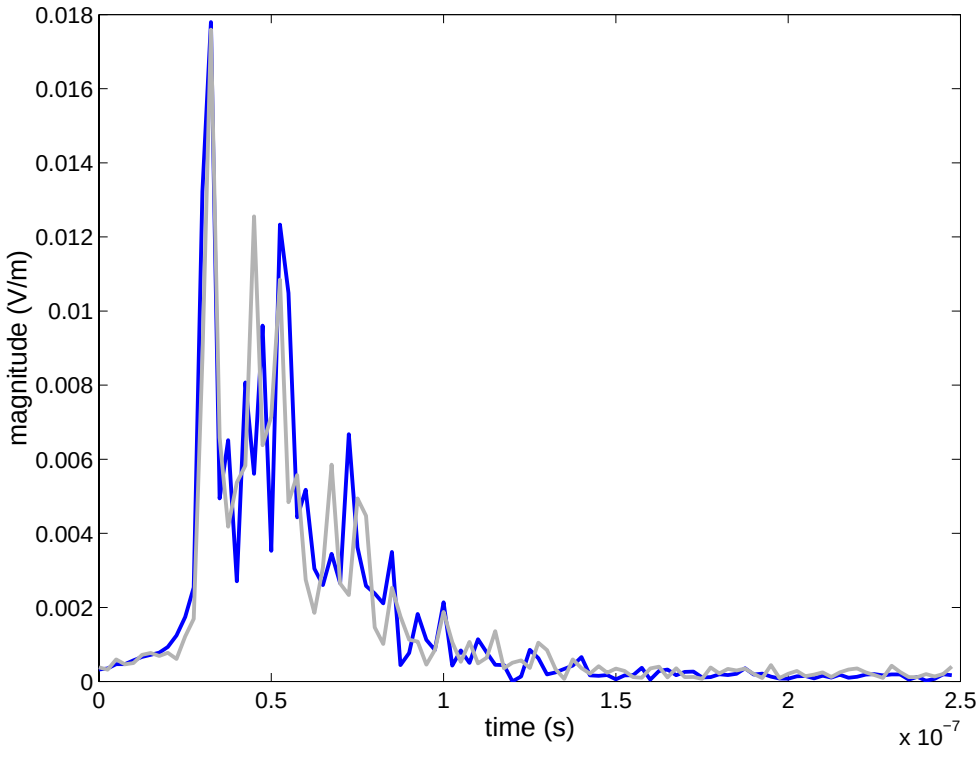



Figure 1.1: Example channel impulse response inside a residence. Depicted is the magnitude of the vertical electric field component.

quency spectrum of the signal in the frequency range of 800-1200 MHz, which is the frequency of interest for some cellular and personal communications systems. Observe how certain frequency components of the signal suffer deep fades. This means that, if a single sinusoid of such a frequency was to be transmitted, all the multipath components would combine destructively (in phase and amplitude) at the receiver position and the received signal would be very close to zero.

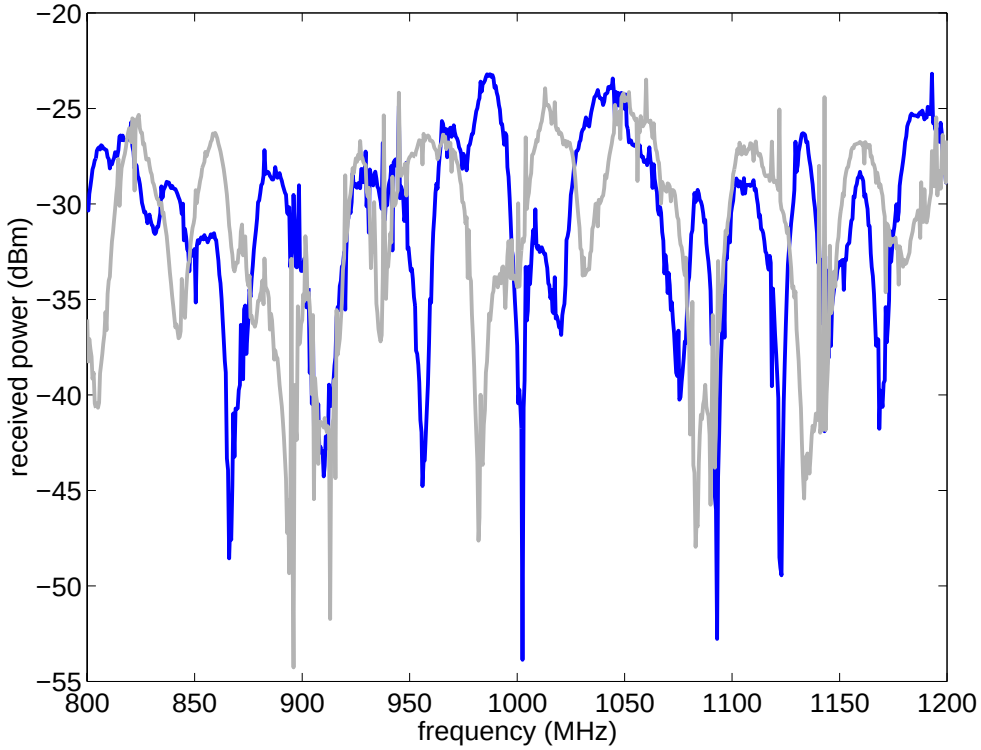



Figure 1.2: Frequency spectrum of the channel impulse response of the figure above.

Due to movement of the receiver, the received signal after a short period of time may look like the one depicted in Figure 1.1 by the light line. We see that some multipath components have increased in magnitude, while others have decreased. The corresponding frequency spectrum is shown in Figure 1.2 by the

light line.

In conclusion, the wireless channel spreads communication waveforms both in time and in frequency, and that is why it is often called *a doubly spread* channel. Whether the communication waveform will be severely distorted because of these impairments depends on both the specific characteristics of the channel and on the signaling rate, or bandwidth, of the communication waveform. In the end, it is the combination of channel characteristics and signal bandwidth that will determine the model best suited for the wireless channel. We discuss these issues in the following section.

## 1.2 Wireless Channel Parameters and Models

The fading multipath channel can be generally modeled as a *linear, time varying* system having an impulse response $\alpha(\tau, t)$. That is, given an input signal $x(t)$, the output $y(t)$ is given by:

$$y(t) = \int_{-\infty}^{+\infty} x(t-\tau)\alpha(\tau, t)\,\mathrm{d}\tau \tag{1.1}$$

which is well known from systems theory as the convolution integral. The interpretation is obvious: $\alpha(\tau, t)$ is the attenuation of the signal components at delay $\tau$ and at time $t$. The input signal $x(t)$ has an equivalent low-pass representation $\tilde{x}(t)$, defined by

$$x(t) = \mathrm{Re}\{\tilde{x}(t)e^{j2\pi f_c t}\} \tag{1.2}$$

where $f_c$ is some reference frequency, usually corresponding to the carrier frequency of the input signal. Substituting $x(t)$ from equation (1.2) into equation (1.1), we obtain:

$$y(t) = \mathrm{Re}\left\{\int_{-\infty}^{+\infty} \left[\alpha(\tau, t)e^{-j2\pi f_c t}\tilde{x}(t-\tau)\,\mathrm{d}\tau\right] e^{j2\pi f_c t}\right\} \tag{1.3}$$

Thus, we arrive at what is really of interest to us, the complex equivalent low-pass channel impulse response:

$$c(\tau, t) = \alpha(\tau, t)e^{-j2\pi f_c \tau} \tag{1.4}$$

The complex phenomena that lead to the wireless channel dispersion and fading preclude any deterministic modeling for $c(\tau, t)$. Rather, statistical methods must be used to describe it. A number of assumptions are made in order to simplify this analysis, beginning from that of a *rich scattering environment*, that is, the existence of a large number of scatterers, which renders $c(\tau, t)$ a complex random process with independent real and imaginary parts. Several statistical descriptions of $c(\tau, t)$ can be made, such as the probability density function of the amplitude and phase of individual multipath components (i.e. for $\tau$ constant), the autocorrelation of individual multipath components as well as the cross-correlation between multipath components (of which we will see a simplification in the following), and so on. Note that, describing $c(\tau, t)$ with a probability density function parameterized with $\tau$ and $t$, which would give the complete description of the channel, is an obviously impossible task. Here, we begin with the autocorrelation function of $c(\tau, t)$, which will be given by:

$$R_c(\tau_1, \tau_2, t_1, t_2) = E\left[c(\tau_1, t_1)c^*(\tau_2, t_2)\right] \tag{1.5}$$

where $*$ denotes complex conjugate. We further assume now that $c(\tau, t)$ is a wide sense stationary (WSS) process in the $t$ variable, that is the autocorrelation function depends only on the difference $t_2 - t_1$ and not on $t_1$ and $t_2$ individually:

$$R_c(\tau_1, \tau_2, t_1, t_2) = E\left[c(\tau_1, t_1)c^*(\tau_2, t_1 + \Delta t)\right] \equiv R_c(\tau_1, \tau_2, \Delta t) \tag{1.6}$$

A brief comment is in order here, with respect to the validity of the wide sense stationarity assumption. Usually, $c(\tau, t)$ will indeed be stationary for short spans

of time, as long the environment around the mobile does not change drastically. However, depending on the mobile's speed, sooner or later the scattering surroundings of the mobile will change, in which case the statistical properties of the channel may also change. Consider, as a simple example, that a user of a cellular phone comes suddenly in direct line of sight with the base station, from which he was previously blocked by a building. The mean of $c(\tau, t)$ will obviously jump from zero, to some non-zero value. Stationarity will only approximately hold also for indoor environments where, as we will see at a later chapter, the wireless channel statistics may change drastically from office to office or room to room.

A final simplifying assumption is that of *uncorrelated scattering* (US). This means that the multipath components that correspond to different time delays (say $\tau_1$ and $\tau_2$) are uncorrelated with each other:

$$R_c(\tau_1, \tau_2, \Delta t) = R_c(\tau_1, \Delta t)\delta(\tau_1 - \tau_2) \tag{1.7}$$

This last equation defines the autocorrelation function of the most widely used wireless channel model, the so called wide sense stationary, uncorrelated scattering channel, or WSSUS channel. In essence, $R_c(\tau, \Delta t)$ gives us the time correlation of the multipath component that arrives at the receiver with delay $\tau$. We must note here that, if we assume that channel variations are due only to the receiver's motion, then we can always associate time correlation with spatial correlation:

$$\Delta r = v\Delta t \tag{1.8}$$

where $\Delta r$ is the distance covered in time $\Delta t$, and $v$ is the receiver's velocity. It is equivalent then to say that $R_c(\tau, \Delta t)$ is the spatial correlation of the multipath component that arrives at the receiver with delay $\tau$. The way to measure this

correlation would be to measure the signal (as magnitude and phase) at delay $\tau$ along a trajectory followed by the receiver, and then correlate this measured signal.

What is more commonly used in the literature rather than $R_c(\tau, \Delta t)$ is its Fourier transform with respect to $\Delta t$, which is called the *scattering function*, $S_c(\tau, \lambda)$. The scattering function gives a measure of the power spectrum of the multipath component at delay $\tau$ and frequency offset $\lambda$, with respect to the carrier frequency $f_c$ (equation (1.4)). Of course, one could also define other Fourier transforms of $R_c(\tau, t)$, for example with respect to the variable $\tau$, and also relate these transforms to the Fourier transforms of the channel $c(\tau, t)$ itself, for a detailed description see [1]. However, they all relate eventually to $S_c(\tau, \lambda)$, which contains all the information and is also intuitively pleasing.

By integrating the scattering function with respect to the Doppler frequency $\lambda$, we obtain the *delay power spectrum*, or *multipath intensity profile* of the channel:

$$S_c(\tau) = \int_{-\infty}^{+\infty} S_c(\tau, \lambda)\, \mathrm{d}\lambda = R_c(\tau, \Delta t = 0) = E\left[|c(\tau, t)|^2\right] \tag{1.9}$$

It is obvious that the multipath intensity profile of the channel gives the average power contained in each multipath component. It could be measured by averaging spatially the power of individual multipath components, for example along a trajectory of the receiver. Since multipath components that arrive earlier have traversed smaller distances from the source, they usually have higher amplitude than components that arrive later, having undergone more reflections along a longer path. For this reason, $S_c(\tau)$ decays with $\tau$. Usually, an exponential model is assumed for this decay:

$$S_c(\tau) = \frac{1}{T_m} \exp(-\tau / T_m) \tag{1.10}$$

where we have assumed unit channel power and $T_m$ is a channel dependent con-

stant, that is often called the *multipath spread* of the channel. There is little experimental backup for this specific model, but it is simple enough since it describes the channel dispersion by one parameter only. In any case, if another model were to be adopted for $S_c(\tau)$, $T_m$ would refer to the time interval where most of the power is concentrated. The inverse of the multipath spread is defined as the *channel coherence bandwidth*, $B_{\text{coh}}$:

$$B_{\text{coh}} = \frac{1}{T_m} \tag{1.11}$$

The coherence bandwidth provides a measure of the width of frequencies that are affected in a similar way by the channel. In other words, if two sinusoids are transmitted with a frequency separation smaller than $B_{\text{coh}}$, their amplitudes will fade in the same way. On the contrary, if they are more than $B_{\text{coh}}$ Hz apart, their amplitude fadings will be mostly uncorrelated. To generalize, if a sufficiently broadband signal is transmitted through the channel, it will be severely distorted by it, since different frequency components will be affected differently. This is the point of view from the frequency domain. In the time domain, the interpretation is that a signal with sufficiently narrow pulses will be affected by the channel because of the well known intersymbol interference (ISI). On the contrary, if the signal bandwidth is much smaller than $B_{\text{coh}}$, the signal may fade in amplitude, but will not be otherwise distorted. Looking at this from the time domain, we see that $T_m$ will be much smaller than the signal's pulse width and therefore, the channel will not induce ISI to the signal.

By integrating the scattering function $S_c(\tau, \lambda)$ with respect to $\tau$, we obtain the *Doppler power spectrum* as:

$$S_c(\lambda) = \int_0^{+\infty} S_c(\tau, \lambda)\, \mathrm{d}\tau \tag{1.12}$$

The Doppler power spectrum gives the power content of the channel at Doppler

frequency shift $\lambda$. It is indeed, the Fourier transform of the autocorrelation function of the channel, with respect to the time variable $t$. The range of frequencies where $S_c(\lambda)$ is essentially non-zero is called the *Doppler spread* $B_D$ of the channel. $B_D$ gives a measure of how fast the channel impulse response varies with time. The inverse of the Doppler spread is called the *coherence time* $T_{\text{coh}}$ of the channel:

$$T_{\text{coh}} = \frac{1}{B_D} \tag{1.13}$$

In complete duality to the coherence bandwidth $B_{\text{coh}}$, the coherence time shows the time interval in which a signal would be affected in a similar fashion by the channel. For example, if two very narrow impulses (delta functions) are transmitted less than $T_{\text{coh}}$ apart, they will experience roughly the same fading. However, if they are more than $T_{\text{coh}}$ apart in time, then the fading will be essentially independent.

Parameters $T_m$ and $B_D$ determine how spread the channel is in time and frequency, respectively. The product $T_m B_D$ is called the *spread factor* of the channel. In general, if $T_m B_D < 1$ the channel is called underspread, otherwise it is called overspread (although notice that it can be overspread or underspread in time or frequency individually). In case $T_m B_D \ll 1$, the channel impulse response can be estimated with great reliability at the receiver and used for the demodulation of the received signal, while it may even be possible to relay the channel estimates to the transmitter, which can optimize the signal transmit power over frequency and time. On the other hand, if $T_m B_D \gg 1$, the channel changes very fast with time, and/or has very severe multipath distortion, for the receiver to be able to estimate it accurately.

We now look at the channel model that we can use in the case that the transmit signal waveform has bandwidth that is much smaller than the channel

coherence bandwidth $B_{\text{coh}}$. The received, equivalent low-pass signal $\tilde{y}(t)$ can be found from equation (1.3), ignoring noise :

$$\tilde{y}(t) = \int_{-\infty}^{+\infty} c(\tau, t)\tilde{x}(t-\tau)\,\mathrm{d}\tau = \int_{-\infty}^{+\infty} C(f,t)\tilde{X}(f)e^{j2\pi ft}\,\mathrm{d}f \tag{1.14}$$

where the second integral is simply the inverse Fourier transform of $\tilde{Y}(f)$. Now, if we assume that the bandwidth of $\tilde{X}(f)$, $W$, is much smaller than $B_{\text{coh}}$, ($W \ll B_{\text{coh}}$, or equivalently $T_m \ll 1/W$), $C(f,t)$ will be approximately constant over the non-zero values of $\tilde{X}(f)$, and equal to $C(0,t)$, since $\tilde{X}(f)$ has low-pass spectrum:

$$\tilde{y}(t) = C(0,t)\int_{-\infty}^{+\infty} \tilde{X}(f)e^{j2\pi ft}\,\mathrm{d}f = C(0,t)\tilde{x}(t) \equiv \alpha(t)e^{j\theta(t)}\tilde{x}(t) \tag{1.15}$$

This approximation is equivalent to saying that the channel impulse response is simply:

$$c(\tau, t) = \alpha(t)e^{j\theta(t)}\delta(\tau) \tag{1.16}$$

or, in other words the channel can be modeled as a single complex tap that multiplies the transmit signal, see Figure 1.3. This model is called the *flat fading* or *frequency non-selective* channel, because all frequency components of the transmit signal get attenuated in exactly the same way. A flat fading channel is said to be *slowly fading*, if its coherence time $T_{\text{coh}}$ is much larger than the symbol duration $T$ of the transmitted signal, i.e. $T_{\text{coh}} \gg T$, or equivalently $B_D \ll 1/T$. This means that a large number of symbols will experience the same attenuation from the channel (and therefore enough time will be given to the receiver to estimate accurately this channel attenuation). If $T_{\text{coh}} \ll T$, or equivalently $B_D \gg 1/T$, the flat fading channel is called *fast fading*. Generally, each symbol from the transmitted signal will be attenuated in an independent way from the channel. This makes it very hard for the receiver to estimate the channel attenuation, although, on the positive side, fast fading offers the opportunity of *time diversity* through channel coding techniques, see [3], [4].

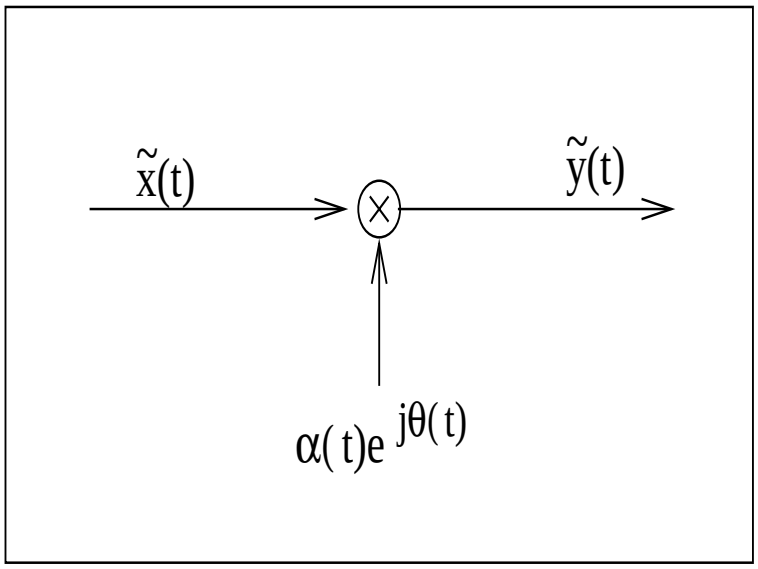



Figure 1.3: The flat fading channel model.

Let us now examine the case where the bandwidth of the transmitted signal is much larger than the coherence bandwidth of the channel, that is, $W \gg B_{\text{coh}}$, or equivalently $T_m \gg 1/W \approx T$. As we mentioned above, the channel will introduce ISI to the transmitted symbol. We are interested in finding a model for the channel that accurately describes this distortion. For this, note that we can approximate the transmitted signal as bandlimited, with a bandwidth of $W$, and use the sampling theorem [1] in order to write $\tilde{x}(t)$ as an infinite sum over its samples:

$$\tilde{x}(t) = \sum_{n=-\infty}^{+\infty} \tilde{x}(n/W)\text{sinc}\left[\pi W\left(t - \frac{n}{W}\right)\right] \tag{1.17}$$

where $\text{sinc}(x) = \sin(x)/x$. The Fourier transform of $\tilde{x}(t)$, $\tilde{X}(f)$ is then:

$$\tilde{X}(f) = \begin{cases} \frac{1}{W}\sum_{-\infty}^{+\infty} \tilde{x}(n/W)e^{-j2\pi fn/W}, & |f| < W/2 \\ 0, \quad |f| > W/2 \end{cases} \tag{1.18}$$

If we substitute equation (1.18) to the second equation of (1.14), we obtain:

$$\begin{aligned}\tilde{y}(t) &= \frac{1}{W}\sum_{n=-\infty}^{+\infty}\tilde{x}(n/W)\int_{-\infty}^{+\infty}C(f,t)e^{j2\pi f(t-n/W)}\,\mathrm{d}f \\ &= \frac{1}{W}\sum_{n=-\infty}^{+\infty}\tilde{x}(n/W)c(t-n/W,t) \\ &= \frac{1}{W}\sum_{n=-\infty}^{+\infty}c(n/W,t)\tilde{x}(t-n/W)\end{aligned} \tag{1.19}$$

where the last equation comes from a simple property of convolution. Now, define the time varying coefficients:

$$c_n(t) = \frac{1}{W}c(n/W,t) \tag{1.20}$$

in which case, we re-write equation (1.19) as:

$$\tilde{y}(t) = \sum_{n=-\infty}^{+\infty}c_n(t)\tilde{x}(t-n/W) \tag{1.21}$$

This last equation shows that the received signal $\tilde{y}(t)$ is the convolution of $\tilde{x}(t)$ with a channel impulse response given by:

$$c(\tau,t) = \sum_{n=-\infty}^{+\infty}c_n(t)\delta(\tau-n/W) \tag{1.22}$$

This is a most intuitively appealing result: the dispersive wireless fading channel can be modeled as a series of very narrow impulses (delta functions), spread apart by (roughly) the transmit signal symbol period, and the complex coefficient of each delta function is just the value of the channel, sampled at that time instant and scaled by (roughly) the symbol period. Of course, $c(\tau,t)$ is usually zero for $\tau<0$ and as we have seen extends up to roughly $\tau=T_m$. Therefore, from equation (1.20) we see that only $\lfloor T_mW\rfloor+1$ coefficients suffice to describe the channel. This channel is called *frequency selective* fading channel and its model is shown in Figure 1.4. It is easy to see also that the flat fading channel is just a special

case of the frequency selective one, where there is only the zeroth coefficient $c_0(t)$ in equation (1.22), compare with equation (1.16). Going back now to the WSSUS wireless channel that we introduced earlier, we see that its interpretation in equation (1.22) would be that coefficients $c_n(t)$ are wide sense stationary random processes, and additionally that they are mutually uncorrelated.

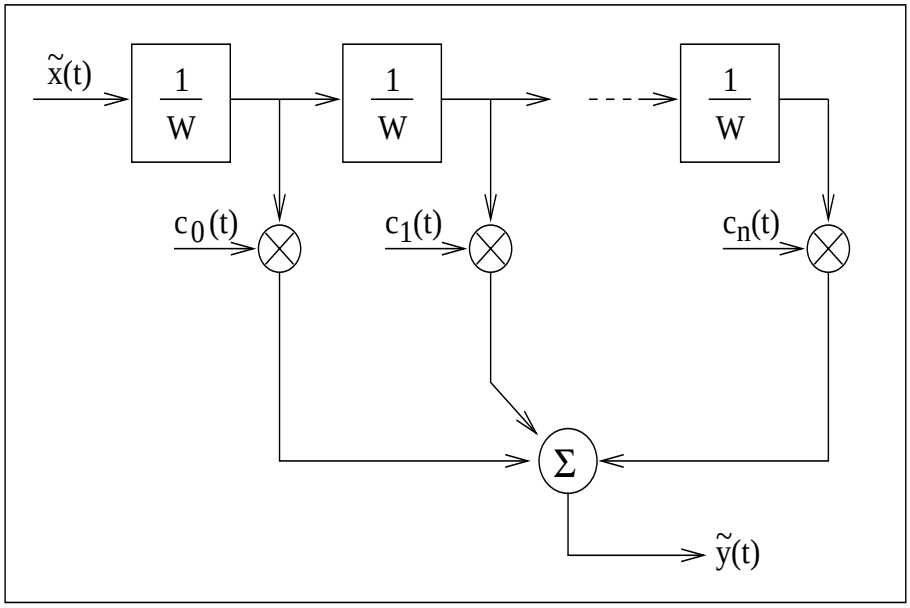



Figure 1.4: The frequency selective channel model.

It is quite interesting now to revisit the autocorrelation function, as defined in equation (1.6) and see what form it takes for the WSSUS channel model of

(1.22):

$$\begin{aligned} R_c(\tau_1,\tau_2,t_1,t_2) &= E\left[c(\tau_1,t_1)c^*(\tau_2,t_2)\right] \\ &= \sum_n \sum_m E\left[c_n(t_1)c_m^*(t_2)\right]\delta\left(\tau_1-\frac{n}{W}\right)\delta\left(\tau_2-\frac{m}{W}\right) \\ &= \sum_n \sum_m E\left[c_n(t_1)c_m^*(t_1+\Delta t)\right]\delta\left(\tau_1-\frac{n}{W}\right)\delta\left(\tau_2-\frac{m}{W}\right) \\ &= \sum_n E\left[c_n(t_1)c_n^*(t_1+\Delta t)\right]\delta\left(\tau_1-\frac{n}{W}\right)\delta\left(\tau_2-\frac{n}{W}\right) \\ &= \delta\left(\tau_1-\tau_2\right)\sum_n E\left[c_n(t_1)c_n^*(t_1+\Delta t)\right]\delta\left(\tau_1-\frac{n}{W}\right) \\ &\equiv \delta\left(\tau_1-\tau_2\right)R_c(\tau_1,\Delta t) \end{aligned} \tag{1.23}$$

where the third line results from stationarity, the fourth from the uncorrelated scattering assumption, and the fifth is just an equivalent representation of the fourth. The scattering function $S_c(\tau,\lambda)$ is the Fourier transform of $R_c(\tau,\Delta t)$ with respect to $\Delta t$:

$$\begin{aligned} S_c(\tau,\lambda) &= \sum_{n=0}^{\lfloor T_m W\rfloor+1} \mathcal{F}\left\{E\left[c_n(t_1)c_n^*(t_1+\Delta t)\right]\right\}\delta\left(\tau-\frac{n}{W}\right) \\ &= \sum_{n=0}^{\lfloor T_m W\rfloor+1} S_{cn}(\lambda)\delta\left(\tau-\frac{n}{W}\right) \end{aligned} \tag{1.24}$$

Hence, the scattering function for the channel model of equation (1.22) is a sum of discrete scattering functions, containing the Doppler spectra of the individual channel coefficients, $S_{cn}(\lambda)$. It is reasonable to assume that the Doppler spectra of all the channel coefficients have the same shape, but possibly different total power. Remembering the exponential multipath profile of equation (1.10) and the definition of $c_n(t)$ from equation (1.20), we can re-write (1.24) as:

$$S_c(\tau,\lambda) = \gamma S_c(\lambda)\sum_{n=0}^{\lfloor T_m W\rfloor+1} e^{-\frac{n}{T_m W}}\delta\left(\tau-\frac{n}{W}\right) \tag{1.25}$$

where $\gamma$ is a normalization constant that keeps the total channel power equal to one.

Let us now focus on the Doppler power spectra $S_c(\lambda)$ and see how we can obtain one of the most commonly used models for it. We have that $S_c(\lambda) = \mathcal{F}\{E[c(t)c^*(t+\Delta t)]\}$, where $E[|c(t)|^2] = 1$. If we go back to the *rich scattering* assumption we have mentioned earlier, that is the assumption of a large number of scatterers, we can write $c(t)$ as:

$$c(t) = \frac{1}{N}\sum_{n=0}^{N} e^{j\phi_n(t)} \tag{1.26}$$

That is, $c(t)$ will just be a normalized sum of complex exponentials, which correspond to the several multipath components that sum up at the receiver with different phases. Now assume that all channel variation comes from receiver mobility, then after time $\Delta t$ the phases of the $N$ multipath components will be:

$$\phi_n(t+\Delta t) = \phi_n(t) + \kappa v \Delta t \cos(\alpha_n) \tag{1.27}$$

where $\kappa = 2\pi/\lambda_0$ is the free-space phase constant, $v$ the receiver's velocity, and $\alpha_n$ the angle between the n-th multipath component's direction and the direction of motion of the receiver. We will further assume that all $\phi_n$ and $\alpha_n$ are mutually independent and that $\alpha_n$ are uniformly distributed in $[0, 2\pi]$. The autocorrelation of $c(t)$ then becomes:

$$E[c(t)c^*(t+\Delta t)] = \frac{1}{2\pi}\int_0^{2\pi} e^{-j\kappa v\Delta t\cos(\alpha_n)}\,\mathrm{d}\alpha_n = J_0(2\pi\frac{v}{\lambda_0}\Delta t) \tag{1.28}$$

where $J_0(x)$ is the Bessel function of first kind and zero order [5]. The Doppler power spectrum $S_c(\lambda)$ is the Fourier transform of the above equation:

$$S_c(\lambda) = \frac{1}{2\pi f_D\sqrt{1-\left(\frac{\lambda}{f_D}\right)^2}} \tag{1.29}$$

where $f_D = \frac{v}{\lambda_0}$ is the maximum Doppler frequency shift at receiver velocity $v$. Equation (1.29) is often called the *Jake's*, or *uniform scattering* model for the

Doppler power spectrum [2]. Combining equation (1.29) with equation (1.25), we have a complete description of the scattering function for the channel model of Figure 1.4, given of course the series of assumptions and simplifications we made along the way.

Finally, we should note that the definitions of fast and slow fading made earlier for flat fading channels, also hold here for the case of frequency selective channels. As seen in equation (1.29), all channel coefficients are usually assumed to have the same Doppler spread ($B_D = f_D$), and therefore fade with the same frequency.

## 1.3 Probability Density Functions for Channel Amplitudes

In the previous section we discussed to some extent the autocorrelation function of the wireless channel, its properties and specific channel models that emerge from these properties in conjuction with the bandwidth of a transmit signal. We will now see some probability density functions (pdf) that model the amplitude of the wireless channel. These pdfs will refer to any multipath component of equation (1.20), or to the component of equation (1.16).

We begin again from the rich scattering environment, and we invoke now the central limit theorem to say that actually $c(\tau, t)$ can be modeled as a complex Gaussian random process, with zero mean. In fact, equation (1.26) gives a zero mean complex Gaussian random process at the limit $N \to \infty$, while a very good approximation is obtained for rather low $N$, say $N = 8$, [2]. This Gaussian approximation means that the phase of $c(\tau, t)$ is uniformly distributed, while the

amplitude $R = |c(\tau, t)|$ is Rayleigh, with pdf:

$$f_R(r) = \frac{2r}{\gamma^2} e^{-r^2/\gamma^2} \tag{1.30}$$

where $\gamma^2 = E(R^2)$ is the mean power of the channel. In the case of the two channel models we presented in the previous section, the frequency non-selective and the frequency selective, this pdf would apply to the individual channel coefficients. In the case of the flat fading channel we would simply have $\gamma = 1$, while in the case of the frequency selective channel, we would have $\gamma_n^2 = \Gamma^2 e^{-n/T_m W}$, where $\Gamma$ is the normalization constant, defined from $\sum_{n=0}^{\lfloor T_m W \rfloor + 1} \gamma_n^2 = 1$. In any case, note that the Rayleigh pdf has only one parameter, $\gamma$.

Another two parameter distribution, that is used often for describing the amplitude of the wireless channel, is the Nakagami distribution [30], whose pdf is defined as:

$$f_R(r) = \frac{2}{\Gamma(m)} \left(\frac{m}{\Omega}\right)^m r^{2m-1} e^{-mr^2/\Omega} \tag{1.31}$$

where $\Omega = E(R^2)$ and $m = \Omega^2/\mathrm{Var}(R^2)$. Since the Nakagami pdf has two parameters, $\Omega$ and $m$, it usually can fit much better experimental (measured) data than the Rayleigh can, although there is not an elegant interpretation for it, as there is for the Rayleigh pdf. We will see that in a following chapter, where we present such measured data. Note also that, the Rayleigh pdf can be obtained from the Nakagami, by setting $m = 1$ in equation (1.31).

A third pdf that has also two parameters and is very closely related to the Rayleigh is the Rician pdf. It is obtained with the same Gaussian approximation on $c(\tau, t)$, only now a non-zero mean is assumed. Obviously, this approximates the rich scattering environment with a line-of-sight (LOS) signal component. The Rician pdf is given by:

$$f_R(r) = \frac{2r}{\gamma^2} e^{-(r^2+s^2)/\gamma^2} I_0\left(\frac{2rs}{\gamma^2}\right) \tag{1.32}$$

where $I_0(x)$ is the Bessel function of second kind and zero order and $s^2$ is the power of the LOS component. Note that, if $s = 0$ then (1.32) reduces to the Rayleigh pdf.

The above density functions describe the amplitude variations of the channel (or channel coefficients according to (1.20)) over small time periods, or equivalently over small spatial displacements of the receiver. However, as we have mentioned earlier, the channel is not usually stationary over longer time periods. For one thing, the receiver may be moving away from, or closer to the transmitter. Therefore, over long time periods the signal strength will vary due to the free space attenuation (such a variation is negligible over small spatial displacements). Additionally, the environment will also change in a random manner that can greatly affect the signal strength, for example obstructions may appear between the transmitter and the receiver, where there were none before, etc. We need to take these long term variations into account, in modeling the channel strength variation. Usually, the approach taken is to decompose channel power variations in three separate contributions: the deterministic free space attenuation, the short term random variations, and a third contribution describing longer term random variations. These three multiplicatively determine the power of the channel. That is, if by $P_c$ we denote the channel power (or the power of the individual coefficients of the model of Figure 1.4), then we can write:

$$P_c = \frac{1}{d^\beta} R^2 L \tag{1.33}$$

where $d$ is the transmitter-receiver distance, $\beta$ is the coefficient of the deterministic power decay, $R$ is the random variable describing short term amplitude variation, and $L$ is the random variable describing longer term amplitude variations. We have already talked about the pdfs that usually describe $R$. $L$ on the other hand is very often called the *shadow fading* component, and varies much

more slowly than $R$, and is modeled as a lognormal random variable. That is, its logarithm is a Gaussian random variable. We can write $L = 10^{\xi/10}$, where $\xi$ is Gaussian, usually with zero mean and standard deviation $\sigma$ which is given in dB. In Figure 1.5 we show what $P_c$ might look like versus $d$, as the receiver moves away from the transmitter. We have assumed $\beta = 3$, $R$ Rayleigh, and $L$ lognormal with standard deviation $\sigma = 6$ dB. In the plot we show individually the deterministic component only (solid line), the combination of the deterministic component with shadow fading (dashed line) and finally the combination of all three components, including Rayleigh fading (dash-dotted line). Observe what a dramatic effect shadow fading can have on the channel power.

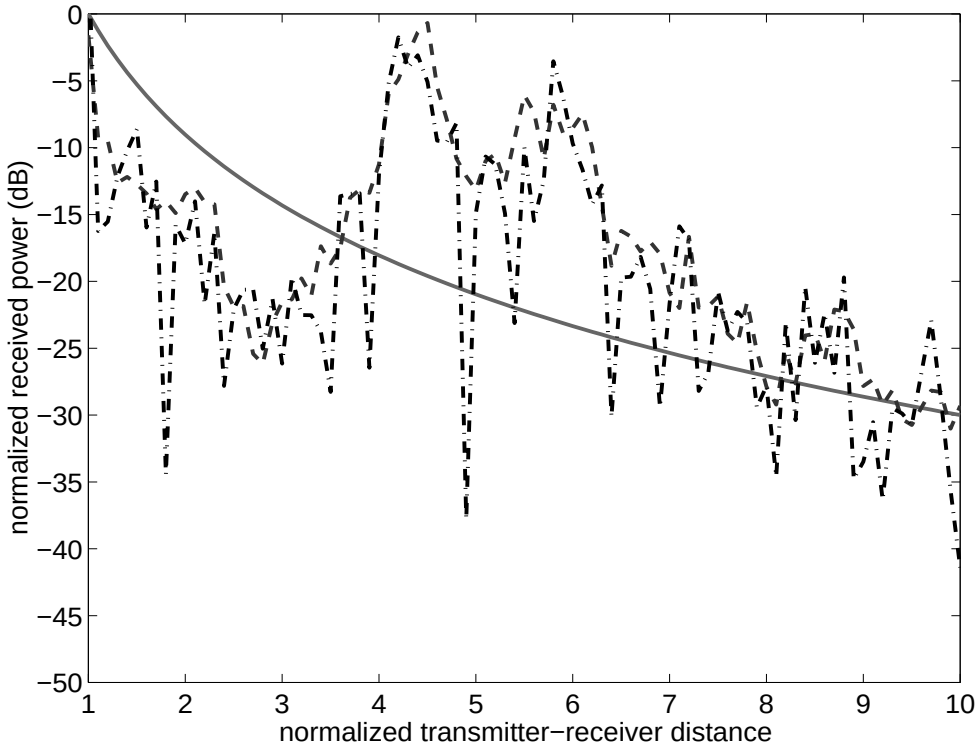



Figure 1.5: Channel power over long time periods (as the receiver moves away from the transmitter).

## 1.4 Summary

In this chapter, we have presented the general characteristics of the wireless channel. We saw that this can be described as a doubly spread channel, since it introduces distortion both in time and in frequency. We begin studying its statistical properties by looking at the autocorrelation function of the low-pass complex equivalent of the channel impulse response, which is a function of both time and time delay. Assuming several approximations and simplifications, we gave equations for this autocorrelation function, as well as its Fourier transform with respect to time, which is called the scattering function. We defined two very important parameters of the scattering function, the multipath spread and the Doppler spread. These two parameters specify how much distortion the channel will introduce in time and frequency to a signal of a given bandwidth. Beginning from a transmit signal's bandwidth, we defined two different channel models, the flat fading and the frequency selective, the former being applicable to a sufficiently narrow-band signal while the latter to a sufficiently broad-band one. We showed the form that the scattering function takes for these channel models. Subsequently, we presented the most commonly used probability density functions that are used to model wireless channel amplitude variations over short time periods. Finally, we discussed the influence of non-stationarity to the wireless channel and we presented a model that captures both short and long term amplitude variations.

For more detailed statistical analysis of fading wireless channels, one can refer to [2], [1] and [6]. For statistical analysis of indoor wireless channels, see [7], [8], and also chapter 3 of this dissertation. Finally, for an information-theoretic tutorial on fading channels see [3].

# CHAPTER 2

# The Finite Difference Time Domain Method in Electromagnetics

## 2.1 Introduction

In the previous chapter we gave an introduction on wireless channel characteristics and modeling. We saw specifically that, deterministic analysis of the wireless channel is not possible because of the randomness inherent both in the scattering environment, and the motion of transmitters and receivers. Statistical methods of modeling should therefore be pursued, and for that we need either measurements of actual channels in various environments (both outdoor and indoor), or accurately predicted channels in such environments. Measurements are certainly preferable, but have obvious disadvantages: they are very cumbersome to perform, need very careful calibration of a large number of parameters, and since they cannot be performed in a controlled environment, they are not repeatable. On the other hand, accurate computer based predictions have relatively little preparation overhead, can be conducted in much larger scales and for many different environments, and are obviously repeatable. Of course, one has to insure first their accuracy, and that can only be done by comparing many "test" predictions with actual measured data.

The dominant method used in computer programs that determine the chan-

nel's impulse response for a specific indoor or outdoor environment is *ray tracing*, [9], [10], [11]. This method is very well known in the electromagnetic community for approximately solving high frequency problems. The basic assumption made here is that the objects scattering the electromagnetic field are electrically large (many wavelengths in their smallest dimension) so that the geometric theory of diffraction (GTD) can be used. The method begins by launching rays from the transmitter to all directions and following the progress of these rays. When a ray hits an object, part of it is reflected and part is transmitted through it, each of the two new rays being determined from the complex reflection and transmission coefficients of the scattering body. In the GTD formulation there are also rays diffracted from edges of bodies, according to the laws of Keller [12]. Ray tracing techniques are very fast, take up relatively small memory and can be easily used with directive antennas to determine the channel's response to sectorized antennas. Their disadvantages are that they have quite complicated structure, especially the 3 dimensional versions, the number of rays that have to be followed explodes exponentially with the number of scattering objects, they have to be run from the beginning for every desired observation point, and they are approximate only, their accuracy decreasing if there are objects that have electrically small dimensions (for example walls that are not modeled as infinitely thin, small windows etc).

In what follows we will describe a different computer prediction method, that is actually also widely known in the electromagnetics community, and is used for solving complex propagation and scattering problems. This method has come to be known as the *Finite Difference Time Domain*, or FDTD method for solving electromagnetics problems. It is nothing else but a method for approximately solving Maxwell's differential equations in the time domain, by discretizing them and transforming them in difference equations, both in the time and the space

domain. The method was first developed by Yee [15] and has since been widely used in the electromagnetics and microwave community for exploring the transient and wideband responses of three dimensional microwave structures. Among its advantages are the simplicity of the formulations, the fact that it is an exact method of solving Maxwell's equations, it can treat fairly complex geometries, and it gives the wideband response of the structure under study. Among the disadvantages are the large memory requirements in the computer simulations and the need for good terminating conditions of the spatial grid in open space problems.

The fact that FDTD is capable of giving the wideband response of a structure, makes it a good candidate for examining the impulse response of a wireless channel. In that respect, FDTD can work in the same way as sounding techniques for experimentally determining the wireless channel [13], [14]; a wideband pulse is launched at the transmitter location and the field is monitored at any desired point. This way, the channel response is obtained with a single run of the program for all the points of the mesh that fills in the space under investigation. This large amount of data that can be obtained with FDTD can be used for doing very accurate statistical analysis of the wireless channel. On the other hand, when the physical space under investigation becomes electrically very large, the memory requirements of FDTD become prohibitively large and the method cannot be used any more. For this reason, FDTD can be useful only for indoor environments. While this is certainly a disadvantage, we should note that the outdoor wireless channel has been studied very thoroughly, because of the advent of cellular telephony [2], [16]. On the other hand, indoor wireless applications have appeared only recently, and therefore the indoor wireless channel has not been researched as extensively. In the next chapter we will present results from the application of FDTD to the determination of the wireless channel inside a

residence.

## 2.2 FDTD Formulation

The FDTD method is formulated by discretizing Maxwell's curl equations over a finite volume and approximating the derivatives with centered difference approximations [15]. Conducting surfaces are treated by setting tangential electric field components to 0. The walls of the mesh, however, require special treatment to prevent reflections from the mesh termination.

Maxwell's curl equations for uniform, isotropic, homogenous and lossless media are written as:

$$\mu \frac{\partial \mathbf{H}}{\partial t} = -\nabla \times \mathbf{E} \tag{2.1}$$

$$\epsilon \frac{\mathbf{E}}{\partial t} = \nabla \times \mathbf{H} \tag{2.2}$$

In order to find an approximate solution to this set of equations, the problem is discretized over a three-dimensional computational domain with appropriate boundary conditions enforced on the conductors and the mesh walls. The six field components are considered to be located on the unit, box-shaped cell of the FDTD mesh, as shown on Figure 2.1 The electric field updating equations can be obtained from discretizing (2.1), i.e. Ampere's law:

$$\begin{aligned} E_x^{n+1}(i,j,k) = {} & E_x^n(i,j,k) \\ & + C_{yi}\left[H_z^{n+1/2}(i,j+1,k) - H_z^{n+1/2}(i,j,k)\right] \\ & - C_{zi}\left[H_y^{n+1/2}(i,j,k+1) - H_y^{n+1/2}(i,j,k)\right], \end{aligned} \tag{2.3}$$

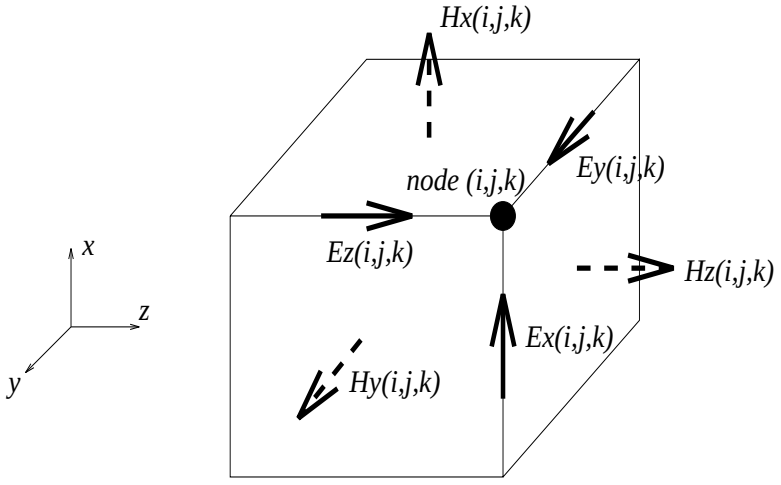



Figure 2.1: The FDTD unit cell

$$\begin{aligned} E_y^{n+1}(i,j,k) &= E_y^n(i,j,k) \\ &+ C_{zi}\left[H_x^{n+1/2}(i,j,k+1) - H_x^{n+1/2}(i,j,k)\right] \\ &- C_{xi}\left[H_z^{n+1/2}(i+1,j,k) - H_z^{n+1/2}(i,j,k)\right], \end{aligned} \tag{2.4}$$

$$\begin{aligned} E_z^{n+1}(i,j,k) &= E_z^n(i,j,k) \\ &+ C_{xi}\left[H_y^{n+1/2}(i+1,j,k) - H_y^{n+1/2}(i,j,k)\right] \\ &- C_{yi}\left[H_x^{n+1/2}(i,j+1,k) - H_x^{n+1/2}(i,j,k)\right], \end{aligned} \tag{2.5}$$

where

$$C_{xi} = \frac{\Delta t}{\varepsilon_i \Delta x}, \qquad C_{yi} = \frac{\Delta t}{\varepsilon_i \Delta y}, \qquad C_{zi} = \frac{\Delta t}{\varepsilon_i \Delta z} \tag{2.6}$$

and the superscript $n$ represents the time-step. Also, finite difference approximations of the Maxwell's equation based on Faraday's law (2.2), give the magnetic

field updating equations:

$$\begin{aligned} H_x^{n+1/2}(i,j,k) &= H_x^{n-1/2}(i,j,k) \\ &\quad - D_y \left[E_z^n(i,j+1,k) - E_z^n(i,j,k)\right] \\ &\quad + D_z \left[E_y^n(i,j,k+1) - E_y^n(i,j,k)\right], \end{aligned} \tag{2.7}$$

$$\begin{aligned} H_y^{n+1/2}(i,j,k) &= H_y^{n-1/2}(i,j,k) \\ &\quad - D_z \left[E_x^n(i,j,k+1) - E_x^n(i,j,k)\right] \\ &\quad + D_x \left[E_z^n(i+1,j,k) - E_z^n(i,j,k)\right], \end{aligned} \tag{2.8}$$

$$\begin{aligned} H_z^{n+1/2}(i,j,k) &= H_z^{n-1/2}(i,j,k) \\ &\quad - D_x \left[E_y^n(i+1,j,k) - E_y^n(i,j,k)\right] \\ &\quad + D_y \left[E_x^n(i,j+1,k) - E_x^n(i,j,k)\right], \end{aligned} \tag{2.9}$$

where

$$D_x = \frac{\Delta t}{\mu \Delta x}, \qquad D_y = \frac{\Delta t}{\mu \Delta y}, \qquad D_z = \frac{\Delta t}{\mu \Delta z}. \tag{2.10}$$

Observe that the magnetic and electric field components are not in fact calculated at the same time instant, but half a time step apart. Also, since the discretization of the differential equations is just of first order, the field quantities at each time step depend only on the ones of the previous time step. Thus, six field elements for each cell have to be stored each time step. As can be seen from (2.6), the coefficients for the electric field updating equations depend on the dielectric constant of the media that occupy each FDTD cell. This can lead to unnecessary memory requirements, if one stores the value of the dielectric constant for each cell. It's obvious that large dielectric bodies (such as walls for example) will be comprised by a very large number of cells, (since the dimensions

of each cell are a fraction of the wavelength as will be discussed later) and thus it's wasteful to store the same number for each of them. Also, most of the space in a wireless channel is occupied by air, whose relative dielectric constant is one. Fortunately, as the field updating equations shown above imply, the spatial updating can begin from any point in space and can be done in separate volumes each time, as long as care is taken not to update the field components twice on adjacent surfaces between two volumes. One can identify different volumes corresponding to different dielectric bodies and assign a single dielectric constant for all the FDTD cells corresponding to each volume. On the transition surfaces of different dielectric bodies, averaging of the dielectric constants is used [18]. With this modification, significant memory savings are possible.

The size of the FDTD cell is determined by the highest frequency of the excitation. It is generally accepted [19] that the maximum size of the cell should be one tenth of the smallest wavelength present in the excitation, or smaller. Also, the decision of the time step is important for the stability of the time update: The maximum time step is limited by the Courante stability condition [19]:

$$\Delta t \leq \frac{1}{v_{\max}} \left( \frac{1}{\Delta x^2} + \frac{1}{\Delta y^2} + \frac{1}{\Delta z^2} \right)^{-1/2}, \tag{2.11}$$

where $v_{\max}$ is the maximum velocity of the electromagnetic wave in the computational domain, usually that is just the light velocity in the vacuum.

Conductors are assumed perfect and infinitely thin and are treated by setting the tangential electric field components that lie on them to zero. It is assumed that the conductors lie exactly on the FDTD computational grid. The tangential electric field components on the conductors are zeroed after the time update of the electric field.

### 2.2.1 Absorbing Boundary Treatment

Due to finite capabilities of computers used to implement the finite difference equations, the FDTD mesh must be limited in the $x, y, z$ directions. The difference equations given above cannot be used to evaluate the field components tangential to the outer boundaries, since they would require the values of field components outside of the mesh. On the other hand, the field on the mesh's boundary has to be updated in a way that gives outgoing waves only, so that there is no spurious reflection coming into the computational domain. The task at hand is basically one of finding a one way wave equation, or an equation that creates waves that propagate on a half space instead of the whole space. The most popular such wave equation was established by Mur [20] and the resulting finite difference equation is:

$$E_0^{n+1} = E_1^n + \frac{v\Delta t - \cos(\theta)\Delta y}{v\Delta t + \cos(\theta)\Delta y}(E_1^{n+1} - E_0^n) \tag{2.12}$$

where $E_0$ represents the tangential electric field components on the mesh wall and $E_1$ represents the tangential electric field components one node inside of the mesh wall. $\theta$ is the angle between the incident wave on the mesh wall and the axis normal to it. Note that during the $n+1$ update of the boundary field, the updated field value one node inside of the boundary is needed. Therefore, we apply the boundary condition (2.12) after having updated the field everywhere inside the boundary nodes of the mesh. For the $z = 0$ absorbing boundary wall of Figure 2.2, equation (2.12) will give:

$$E^{n+1}_{\left\{\begin{smallmatrix} x \\ y \end{smallmatrix}\right\}}(i,j,0) = E^n_{\left\{\begin{smallmatrix} x \\ y \end{smallmatrix}\right\}}(i,j,1) + \frac{v\Delta t - \cos(\theta)\Delta z}{v\Delta t + \cos(\theta)\Delta z}\left(E^{n+1}_{\left\{\begin{smallmatrix} x \\ y \end{smallmatrix}\right\}}(i,j,1) - E^n_{\left\{\begin{smallmatrix} x \\ y \end{smallmatrix}\right\}}(i,j,0)\right) \tag{2.13}$$

where the subscripts $x, y$ denote the tangential field components involved. Equation (2.12) is often referred to as Mur's first order absorbing boundary condition,

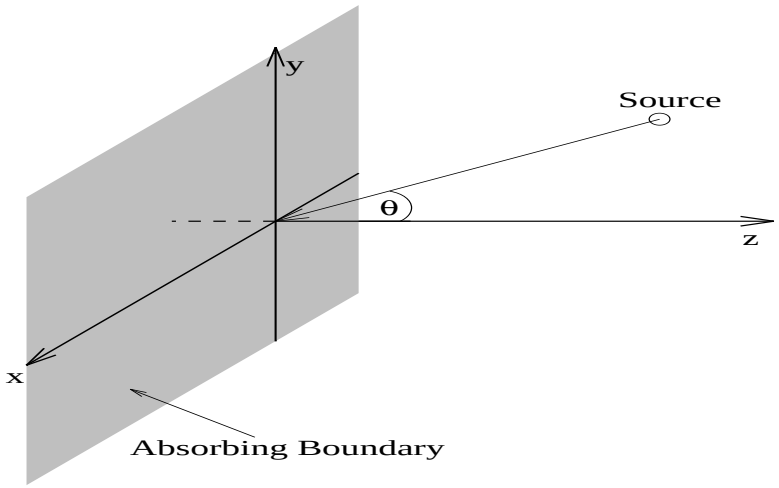



Figure 2.2: The z = 0 (x-y) absorbing boundary.

since higher order conditions can be derived using Mur's method. It actually absorbs completely a wave incident at angle $\theta$ with respect to the normal to the absorbing boundary wall. Waves incident at other angles are absorbed less, the reflection increasing as the actual incident angle differs from $\theta$. For this reason, in the code we used to simulate indoor wireless channels we employed a method for improving the above absorbing conditions, called the super-absorption [21]. In this technique, the same absorbing condition is applied to both the tangential electric and magnetic fields and an iteration is executed on both the $E$ and $H$ fields to reduce the reflection from the boundary. The equations are quite complicated to be presented here, but they have been implemented in the computer code.

Mur's first order absorbing boundary condition together with the super- absorption technique gives sufficiently good spurious reflections suppression, for our purposes. Another good feature of Mur's first order absorbing boundary condition is that it is numerically very stable, in contrast with higher order conditions which suffer from *later time* instabilities. That means that if electrically large

structures have to be investigated, the higher order absorbing boundary conditions tend to be unstable as time progresses [22]. On the other hand, the first order Mur's boundary condition is very robust. Other, more complicated boundary conditions were also considered, like the Berenger perfectly matched layer [23], which achieves extremely low reflection coefficients (of the order of -60dB) over a very wide angle range, but were not implemented due to their excess complexity, very significant memory storage requirements and questionable improvement to our results, within the accuracy that we require (these methods may be very well needed for high precision microwave circuits simulations, but we don't need such level of accuracy in our results).

### 2.2.2 Excitation Pulse

As already mentioned, FDTD uses a wideband pulse as the excitation and in that respect it resembles the experimental sounding techniques used to determine the wireless channel [26]. The pulse used here is the most commonly used one, that is a modulated Gaussian pulse. It's value is given by the equation:

$$g(t) = \exp\left[-(t-t_0)^2/T^2\right] \cdot \sin(\omega_c t) \tag{2.14}$$

where $\omega_c$ is the central frequency of the pulse, used to move the spectrum away from the DC, since first we are not interested in that region and second the absorbing conditions perform poorly there. The delay $t_0$ is introduced so that there is a very smooth transition from zero field to the excitation pulse, thus minimizing numerical errors due to the finite difference approximations. The spectrum of this pulse is:

$$G(\omega) = \exp\left(-\frac{\omega^2 T^2}{2}\right) \tag{2.15}$$

and the 3dB bandwidth,

$$W_{3\text{dB}} = \frac{\sqrt{2\ln 2}}{\pi T} \tag{2.16}$$

Based on (2.16) above and $\omega_c$, one can find the maximum frequency of the Gaussian pulse, and hence decide on the maximum cell size.

The pulse (2.14) is enforced as the value of the electric field at the source point, at the desired direction (for example an $\hat{x}$ directed pulse in Figure 2.2 would imply a vertically placed infinitesimal dipole). All the other initial field components are zero and then they get updated with initial condition the pulse (2.14). This way, the field is created at the source point as a modulated Gaussian pulse and then is propagated through all space, based on the FDTD updating equations. After a sufficient time, depending on the electrical dimensions of the space under consideration, and since the source field is forced to zero, the field in the whole space becomes again zero. The program then has finished and we can have the field history (or profile) in every cell of the computational domain. In that respect, FDTD differs drastically from ray tracing methods that at each run give the field at a single point.

## 2.3 Summary

In this chapter we introduced the finite difference time domain method that is very useful in analyzing microwave circuits. This method is essentially a difference approximation to the differential Maxwell equations, and that is why it gives an exact solution to these equations (within the accuracy of the difference approximation). We presented this method as an alternative to ray tracing techniques that are currently widely used for wireless channel prediction. We discussed the advantages and disadvantages of FDTD with respect to ray tracing. Among the

greatest advantages, we can mention the accuracy of the method and the large number of data that we can obtain, and that are very useful for statistical analysis and modeling of the wireless channel. Its greatest disadvantage is the very large computer memory requirements, that make this method only useful for indoor environments. In the next chapter we will use a memory improved version of FDTD to predict the wireless channel inside a residence. A flowchart of the FDTD algorithm is shown in Figure 2.3 on the next page.

For a very thorough treatise on FDTD and its microwave applications, see [19] and [32]. For discussion on GTD and other high frequency approximations to Maxwell's equations see [12]. Finally, for an extensive presentation of the application of such methods to radio propagation prediction and modeling, see [33].

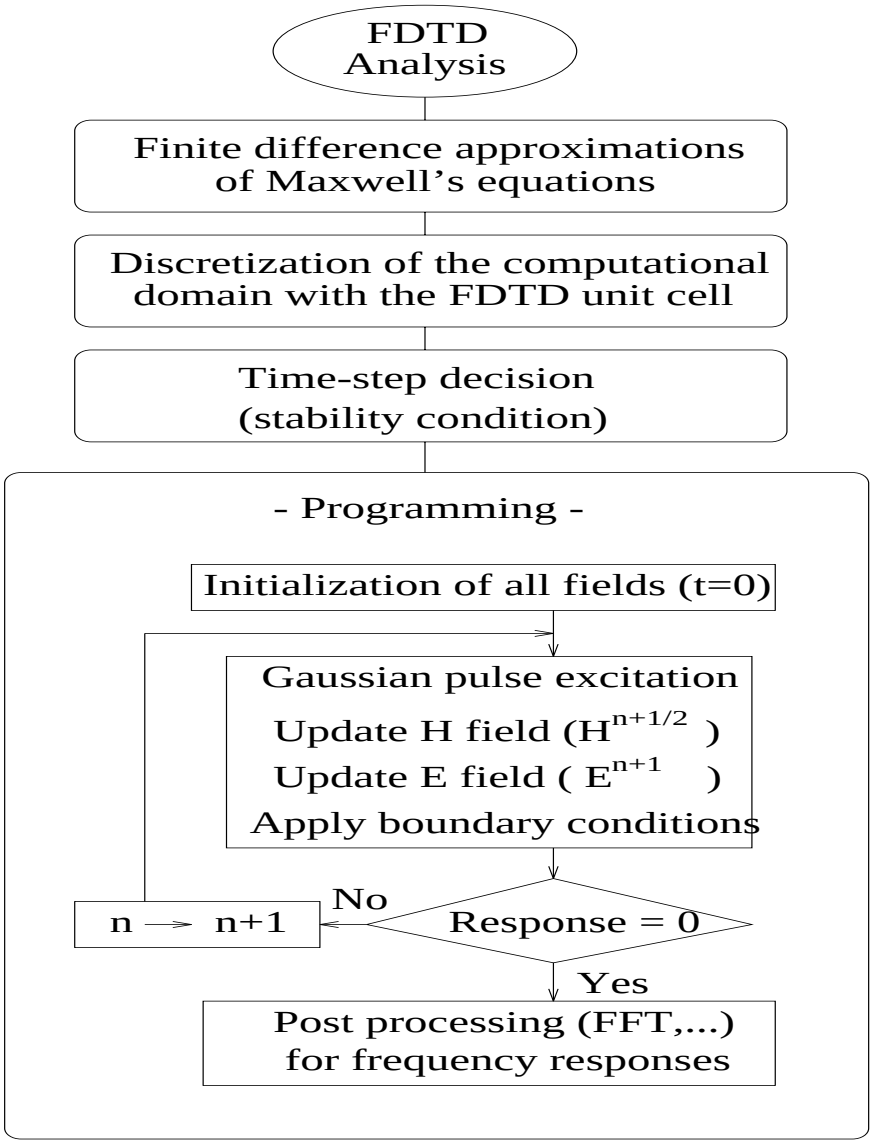



Figure 2.3: The FDTD Flowchart Algorithm

# CHAPTER 3

# Experimental Determination of Indoor Wireless Channel Impulse Response Using FDTD

## 3.1 Introduction

We discussed in earlier chapters how accurate wireless channel prediction and modeling is very important for the design and performance evaluation of high speed wireless communication systems. Factors such as signal strength, fast fading and frequency selectivity of the channel are usually the ones limiting such systems' reliability. Furthermore, with the advent of wireless networking products for the office and home, a great deal of interest has recently gone into characterizing electromagnetic propagation inside buildings. Note that, the indoor radio channel differs from the more traditional outdoor mobile radio channel mainly in two respects: the distances covered are much smaller, and the environment changes more drastically for much smaller separations between transmitter and receiver. On the other hand, mobility is usually limited to pedestrian speeds only. It is also generally agreed that, the statistical models that have been developed for the outdoor channel are not valid for the indoor one. Statistical models that approximate well the real behavior of the electromagnetic field for the indoor wireless channel are therefore needed for the analysis and simulation of tetherless personal communication systems.

We also introduced in the previous chapter the FDTD method as a means for predicting wireless channel impulse responses in indoor environments, and saw its potential advantages and disadvantages with respect to the currently most commonly used technique, ray tracing. In the next sections of this chapter we will present results from using FDTD for the prediction of the indoor channel of a residence, and compare these predictions to measured data that we obtained for that same residence. The basic model we used for the time response of the baseband wireless channel is given by equation (1.22), of chapter 1, with the addition of a large scale path loss multiplicative term. We had noted in that chapter that, if we assume channel variations due only to receiver mobility, then time dependence can always be interpreted equivalently as spatial dependence. The channel model we will use therefore is given as:

$$c(\tau, x, y) = \frac{1}{d^{\beta/2}} \sum_{k=0}^{\lfloor T_m W \rfloor + 1} c_k(x, y) \delta \left( \tau - \frac{k}{W} \right) \tag{3.1}$$

where $d$ is the distance between the transmitter and the observation (receiver) point $(x, y)$, $\beta$ is the large scale path loss exponent, $c_k(x, y)$ are the complex coefficients of individual multipath components (see equation (1.20) of chapter 1), $T_m$ is the multipath spread of the channel and $W$ is the bandwidth over which we are trying to characterize the channel response.

Experimental methods for measuring the impulse response of wireless channels fall under two categories: the time domain ones and the frequency domain ones. In time domain methods (or sounding methods) [13], [14], very short pulses are emitted from the transmitter periodically, and the receiver measures each arriving time profile. Obviously, the period of the pulses should be greater than the multipath delay spread $T_m$, or ISI would be induced at the receiver. Time domain methods are very accurate and are the methods of choice for outdoor channel predictions. However, they need very sophisticated hardware and very

tight synchronization between the receiver and the transmitter. In frequency domain methods [24], [25], the wireless channel is seen as a two-port microwave network, such as the one shown in Figure 3.1. If we assume that the transmitter

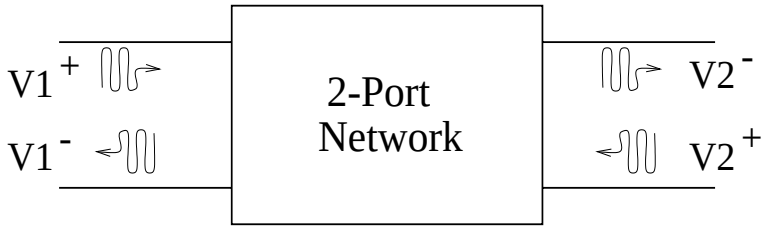



Figure 3.1: A two-port microwave network, representing the wireless channel.

is connected to port 1 and the receiver at port 2, determination of the channel response amounts to determining the frequency response of this two-port device. Specifically, in Figure 3.1 we denote by $V_j^+$ the amplitude of the voltage wave *incident* to port $j$, and by $V_j^-$ the amplitude of the voltage wave *reflected* from port $j$. The *scattering matrix* $\mathcal{S}$ of this microwave device is defined in relation to these incident and reflected voltages as [17]:

$$\begin{bmatrix} V_1^- \\ V_2^- \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} V_1^+ \\ V_2^+ \end{bmatrix} \tag{3.2}$$

The element $S_{ij}$ of matrix $\mathcal{S}$ can be determined as:

$$S_{ij} = \left. \frac{V_i^-}{V_j^+} \right|_{V_k^+=0 \text{ for } k \neq j} \tag{3.3}$$

In other words, $S_{ij}$ is determined by launching an excitation (incident) voltage $V_j^+$ at port $j$ and measuring the reflected (which is the output here) voltage $V_i^-$ at port $i$. The incident voltage on the port without the excitation must be zero, which means that port should be terminated to a matched load to avoid reflections. Returning to wireless channel measurement, if we assume the transmitter

connected to port 1, the receiver to port 2, and the transmitter launches a voltage $V_1^+$, then the channel response will be given by $S_{21}$:

$$S_{21} = \left. \frac{V_2^-}{V_1^+} \right|_{V_2^+=0} \tag{3.4}$$

In this context, the condition $V_2^+ = 0$ has the interpretation that the receive antenna is well matched.

The instrument used to measure the scattering matrix is the vector network analyzer. This is a two-channel instrument, where one channel is used as the transmitter and the other as the receiver. The two-port network is connected between these two channels. In operation, a radio frequency source is used to sweep over a specified range of frequencies, while the instrument takes measurements of the amplitude and phase of all the incident and reflected waves in the two ports. In order to be able to measure correctly, the analyzer needs to do measurements on a *calibration set* of devices first. These devices include a *short-circuit*, an *open-circuit* and a *through-circuit*, whose responses are obvious. In the wireless channel measurements, two broadband antennas are connected to the two channels of the analyzer. It is obvious that the wireless channel measurements will include the responses of the antennas, therefore their impact should be as small as possible (we will discuss this later in some more detail).

## 3.2 Measurement Setup

In order to validate the applicability of FDTD in the prediction of indoor wireless channel impulse responses, a measurement experiment was set up to obtain some real data for comparison with the computer generated prediction. As we mentioned in the previous chapter, FDTD would be applicable for wireless channel prediction in indoor environments, due to its memory limitations, therefore, the

site we used was a residence, whose layout is shown in Figure 3.2.


Figure 3.2: Residence where measurement took place.

First, two discone antennas for the 1 GHz range were built. Discone antennas function as extremely broadband dipoles, and are needed for the broadband measurement of the wireless channel. Obviously, narrow-band dipoles would severely distort the channel response at frequencies other than the short range of frequencies around their resonance. The design of the antennas was based on [27]. Their reflection coefficient was measured in free space and it was below -14 dB at 800 MHz, dropping lower for higher frequencies. This means that the antennas were indeed very well matched for the whole range of frequencies we needed them. However, no attempt was made to measure their pattern over the ranges of frequencies we were interested in, because of the difficulty of the task (too low

frequency for an anechoic chamber). We relied on the data given in [27], where measurements on same type of disconnes show that the almost omnidirectional pattern remains practically unchanged on a bandwidth of more than 2:1, where the lowest frequency corresponds to an electrical height of a quarter wavelength (that frequency for our design was 650 MHz, giving an upper frequency of 1.3 GHz).

The measurements were done in the frequency domain, using a microwave vector network analyzer, following [28]. The measurement setup is shown in Figure 3.3. The transmitted signal from the analyzer was set at a 5 dBm level and was further amplified by a low noise amplifier with a gain of 21 dB, thus giving a total power of around 400 mW, which is a good value for indoor type transmission. The receiver was moved around the house at the points of measurement, with the use of a coaxial cable of length 20 m. The cable had an attenuation of 14 dB at 1 GHz, but was included in the calibration of the network analyzer. Both antennas were placed at a hight of 1.5 m above the floor and they were supported by Styrofoam columns to avoid disturbing their pattern. The position of the transmitter as well as the measurements positions are shown in Figure 3.2. Reception of the antennas was along the vertical axis, that is both antennas were set to receive the vertical electric field component.

The frequency range from 800 MHz to 1.2 GHz was spanned with 801 measurements for each profile. That gives an equivalent time resolution of 2.5 ns and a frequency resolution of 500 KHz. No subsequent filtering was done, so that the maximum time resolution be kept. The whole process was controlled by a personal computer, where the data was also saved. Channel measurements were taken at various points inside the house to determine the large scale path loss, and three sets of more detailed measurements were done for the determination

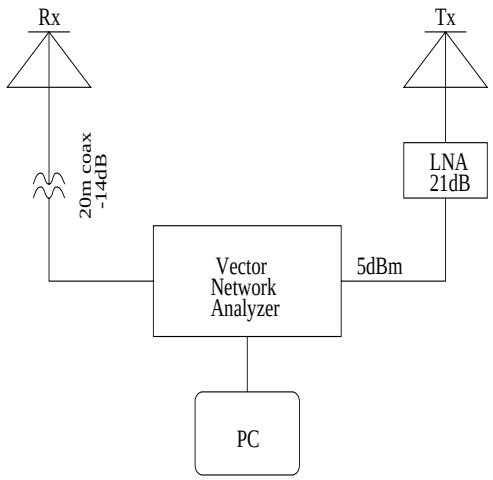



Figure 3.3: Block diagram of measurement setup.

of multipath statistics. In these measurements, three square surfaces of $49 \times 49$ cm were scanned in steps of 7 cm (roughly a quarter wavelength), resulting in 64 measurements per square.

In Figure 3.4 we show one such measurement and specifically the measured power spectral density as well as the phase of a profile. Observe that the phase is linear for most of the frequencies, except for those where the power spectral density has nulls. In Figures 3.5 and 3.6 we plot the magnitude of the frequency response and the corresponding magnitude of the time profile for three points, one from each of the three square areas scanned. The time domain profiles were obtained by simply taking the inverse Fourier transform of the measured frequency domain ones. The frequency selective nature of the channel is obvious from Figure 3.5, as well as from Figure 3.6. One can observe that the point measured in the kitchen, right next to the refrigerator has a lot of reflections, while the one in the living room, very close to the transmitter, has a very clearly defined LOS component and minimal number of reflections.

Finally, we note that no averaging was done to the measurements by taking

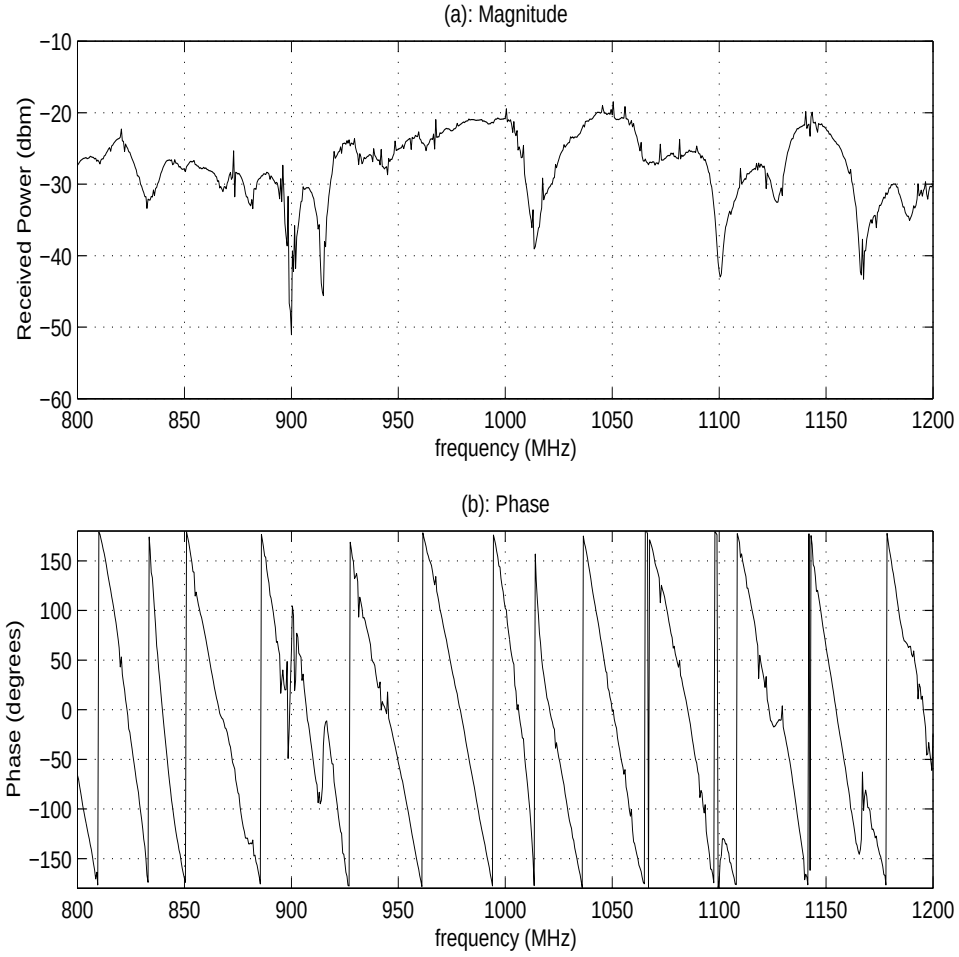



Figure 3.4: An example of a measurement with the vector network analyzer. (a) Magnitude of frequency response in dBm. (b) Phase of frequency response in degrees.

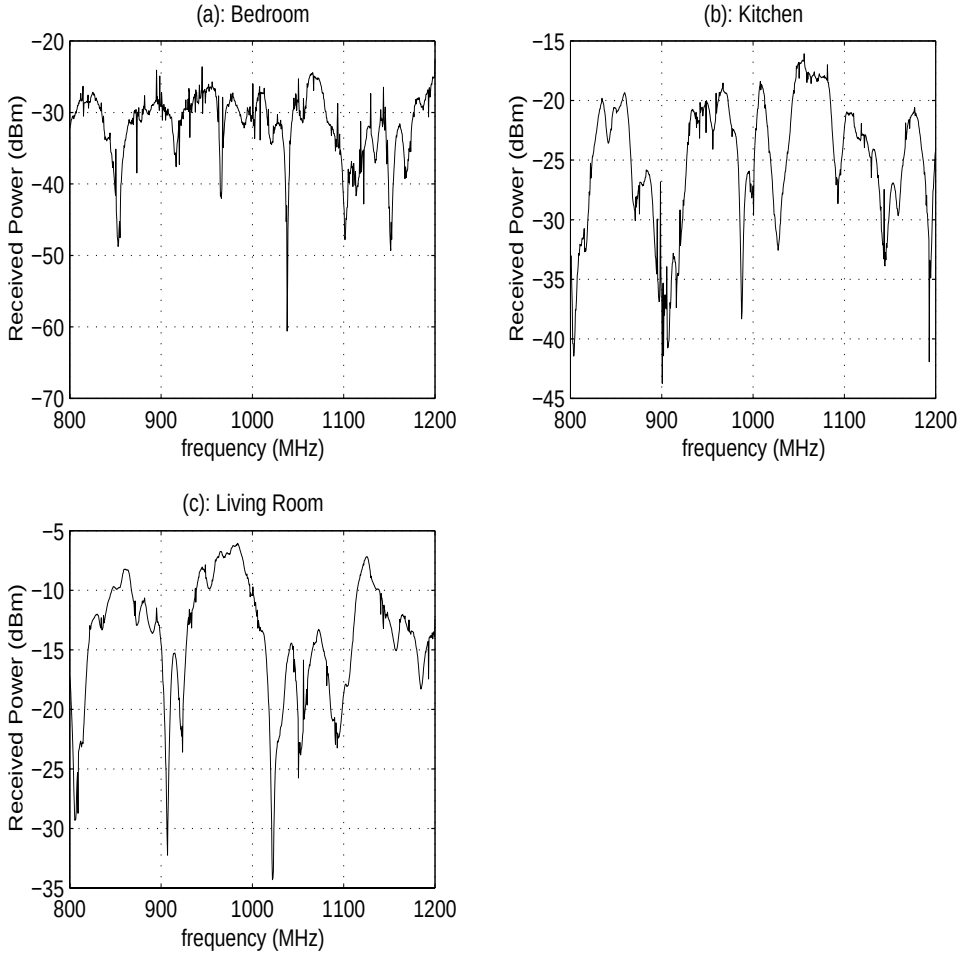



Figure 3.5: Magnitude of frequency response in three points. (a) Bedroom. (b) Kitchen. (c) Living Room.

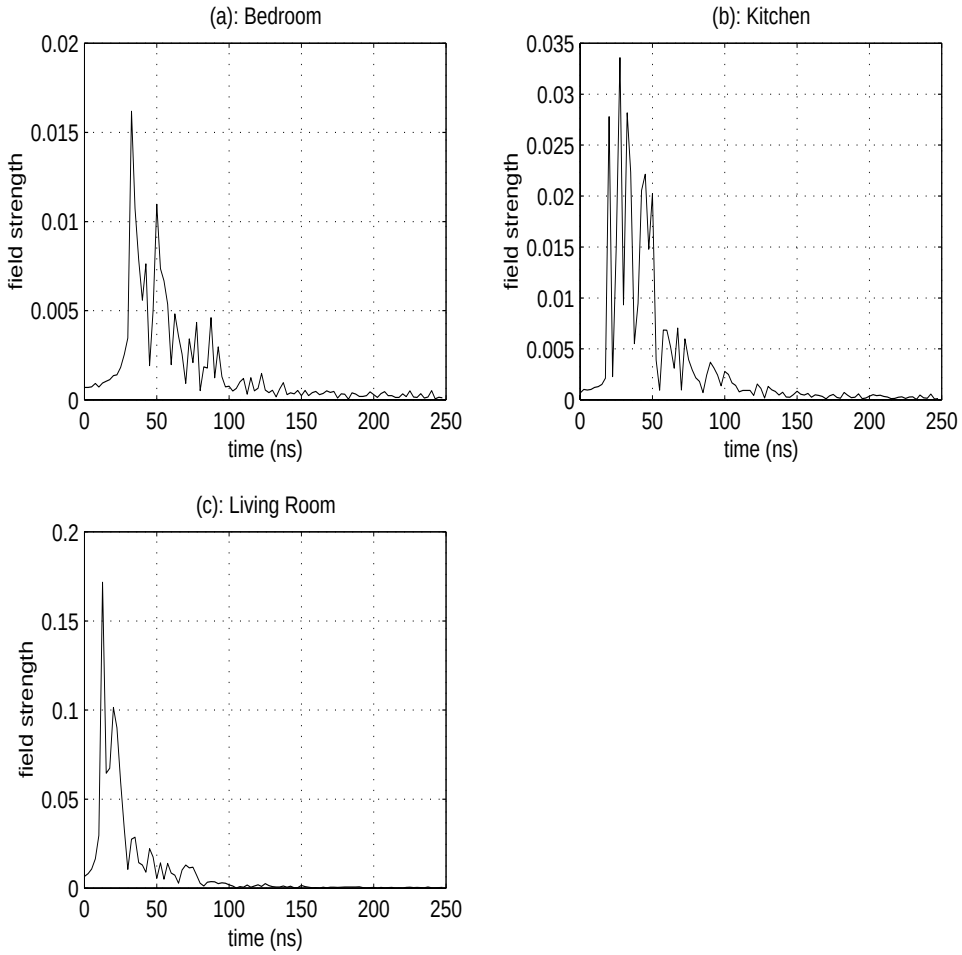



Figure 3.6: Magnitude of time response (profile) in three points. (a) Bedroom. (b) Kitchen. (c) Living Room.

several measurements. Averaging would diminish the spikes evident to the plots of the frequency responses, but we don't expect that they would much affect the time profiles. The measurement equipment was placed in the balcony and the measurements were conducted in the night, to minimize reflections from moving vehicles or people in other apartments. These measurements were considered enough in order to check the validity of our FDTD code. Obviously there are not sufficient for doing very reliable statistics, but having verified the FDTD code with them, we will analyze statistically the channel with the computer generated profiles.

## 3.3 FDTD Prediction of the Wireless Channel

The FDTD method described in the first part was used to predict the wireless channel impulse response in the same residence, shown in Figure 3.2. The house was modeled to the greatest detail possible, with significant metallic objects (like the refrigerator or the stove) and furniture taken into account. The dielectric properties of the walls were chosen such that the results show a good agreement to the experimental measurements. We used a dielectric constant of 6.2 for the ceiling and floor, a value of 5.0 for the outer walls and 4.2 for inner walls. Also, some loss was included in the model, adjusting slightly the updating equations of the electric field, given in chapter 2. The walls were considered having a conductivity of $\sigma = 0.02$. The central frequency was 1 GHz and the cell size for the FDTD was selected at 3 cm (one tenth of the wavelength for the central frequency). The time constant $T$ of the Gaussian pulse of (2.14) of chapter 2 was chosen to be 3 ns, matching closely the temporal resolution of our measurements (2.5 ns). Actually, the 3 dB bandwidth of this pulse is only around 128 MHz, from equation (2.16) of chapter 2, its temporal resolution though corresponds

to a bandwidth of around 200 MHz (i.e. it is about 5 ns). That is broadband enough to make our comparisons. The excitation, along the vertical direction, was placed at the same point as the transmitter in the measurements, and the vertical electrical field component was monitored on a plane 1.5 m above the floor.

The excitation, as given by (2.14), is a modulated Gaussian pulse. Therefore, to obtain the baseband complex channel after the end of the program, each individual profile was demodulated by a sinusoid of 1 GHz frequency and then low pass filtered. In Figures (3.7)(a)-(c), we show the magnitude of three measured profiles from the square surface inside bedroom1, compared with the profiles predicted by the FDTD for the same positions. The profiles are normalized so that they have a peak of one. Therefore, what is of interest is the comparison of secondary multipath components. We can claim a relatively good agreement between measurement and FDTD, given also the potential errors in the exact locations of the measured and predicted profiles.

The model most often used for the power prediction of the wireless channel impulse response is given by equation (1.33) of chapter 1:

$$P_c = \frac{a}{d^{\beta}} R^2 L \tag{3.5}$$

where d is the distance from the transmitter, $L$ is the random power component due to shadowing and $R^2$ is the random power component due to fast fading. The constant $a$ depends on the total transmitted power, while $\beta$ is the large scale path loss exponent. We determine now this exponent, by plotting in log scale (dB) the average power for each measured profile. We also plot the average power as predicted by FDTD for several points inside the house. The results are shown in Figure 3.8. Based on the measured data we also plot the best MSE linear fit which gives us a large scale path exponent of 3. The FDTD data agree very well

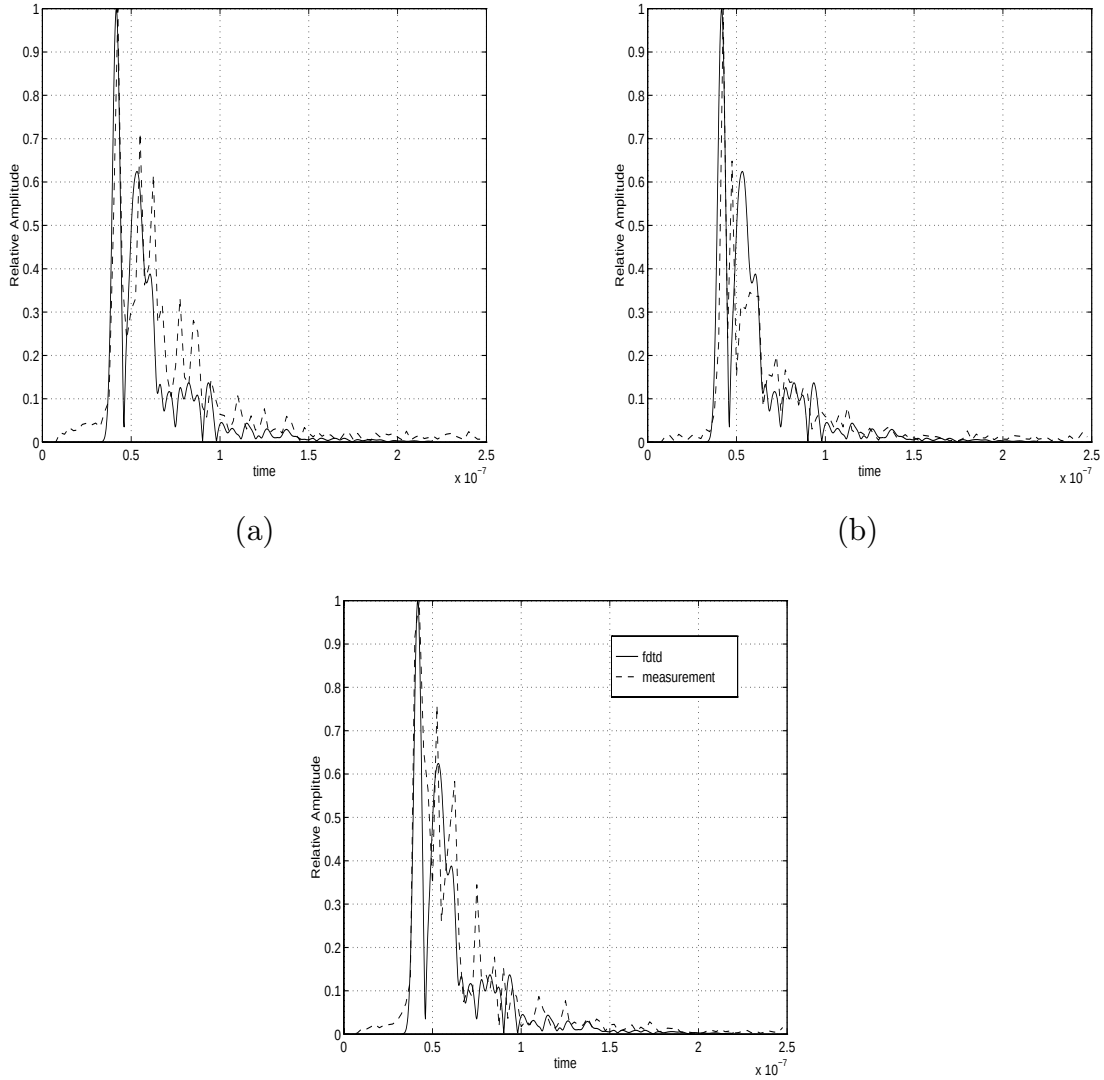



Figure 3.7: Comparison of three measured vs. predicted profiles in Bedroom1

with the measured ones. Notice that for distance of 1 m the average received

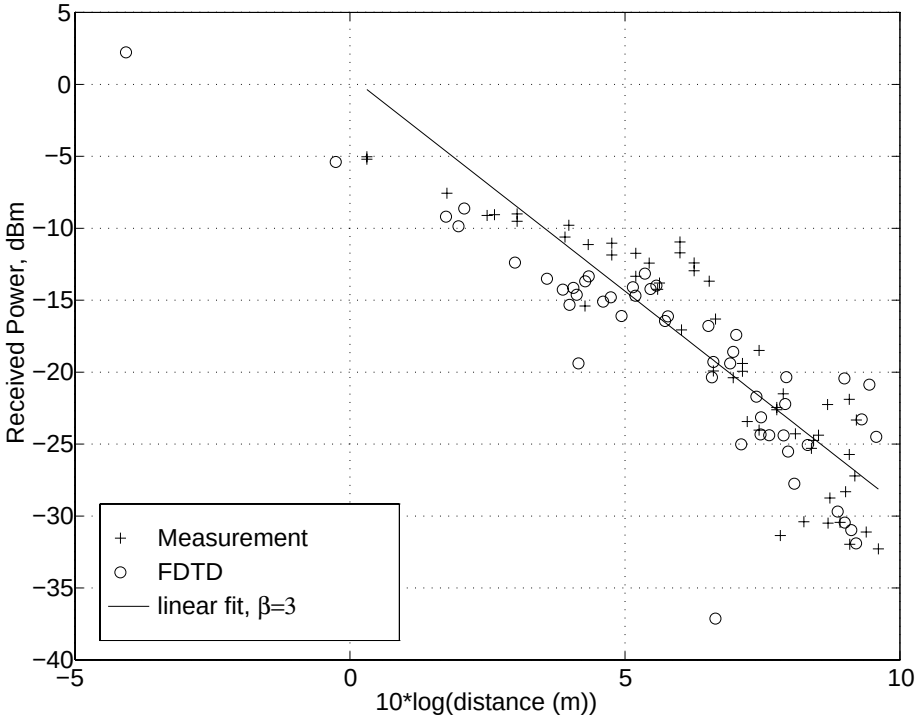



Figure 3.8: Large scale path loss of average power per profile.

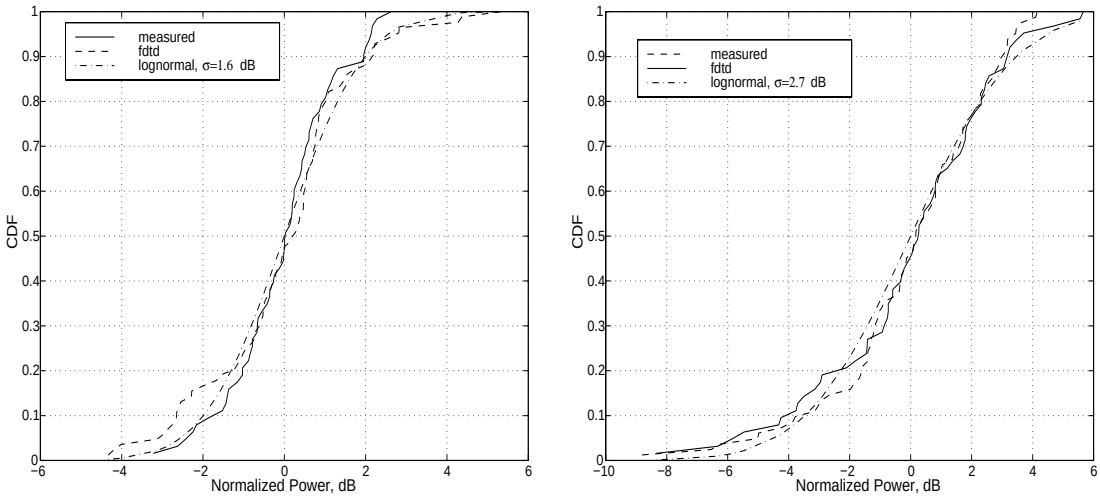
power is about -5 dBm. At 1 m distance from the transmitter the free space path loss, which is given by $\left(\frac{\lambda}{4\pi d}\right)^2$, is -32 dB, which gives an estimate of transmitted power of 27 dBm.

The statistics of the individual multipath components are significant, particularly in the high baud rates of indoor personal communications systems. Therefore, we look now at the statistics of the first and second multipath components of the profiles of the square surface scanned in bedroom1. That is, we obtain statistics for $c_0(x, y)$ and $c_1(x, y)$ of equation (3.1) for bedroom1. The statistics of these profiles are more interesting, since the living room profiles will have dominant LOS and therefore will exhibit Rice distribution (which they indeed do).

Statistics of the profiles of the kitchen give similar results as the bedroom. In Figure 3.9(a) we plot the cumulative distribution function of the power of the first arriving path, both for measured and predicted profiles. The power is plotted in dB because it was found that the best fitting CDF was the lognormal. We have superimposed the lognormal distribution, with a standard deviation of 1.6 dB. The agreement is seen to be relatively good. In Figure 3.9(b) we plot the CDF of the second multipath component, together with the fitted lognormal distribution. The standard deviation is now 2.7 dB, naturally larger than the standard deviation of the first multipath (note that the area scanned for these statistics is very small, only $49 \times 49$ cm and that is why the standard deviations are small). Let us note here that similar values for the standard deviations in individual multipath components have been obtained also elsewhere [7]. The statistics drawn in Figures 3.9(a) and (b) corroborate the fact that for indoor channels the lognormal distribution characterizes better individual multipath components. This result, of course has been observed several times [7].

Another quantity of interest for the wireless channel is the delay power spectrum. The delay power spectrum represents usually the time averaged received power as a function of the delay between the arrivals of different paths. This is nothing else but $S_c(\tau)$ of equation (1.9) of chapter 1. Here, since we have spatial measurements of the wireless channel profiles, we will define the delay power spectrum as the spatial average received power, as a function of the delays of the multipath components. The delay power spectrum is very useful because it gives information about the multipath delay spread ($T_m$) of the channel and the ISI this spread can create.

From a number of profiles, both measured or predicted, we extract the power delay spectrum by dividing the time axis in bins of 2.5 ns, starting from the




(a) (b)

Figure 3.9: Statistics of individual multipath components. (a) First arriving path. (b) Second arriving path.

arrival of the first path. The average received power over all spatial profiles is then calculated for each time bin and plotted against arrival time [29]. We plot in Figure 3.10 the measured delay profile, from the square surface in Bedroom1, as well as the predicted one at the same area. Again a good agreement is observed

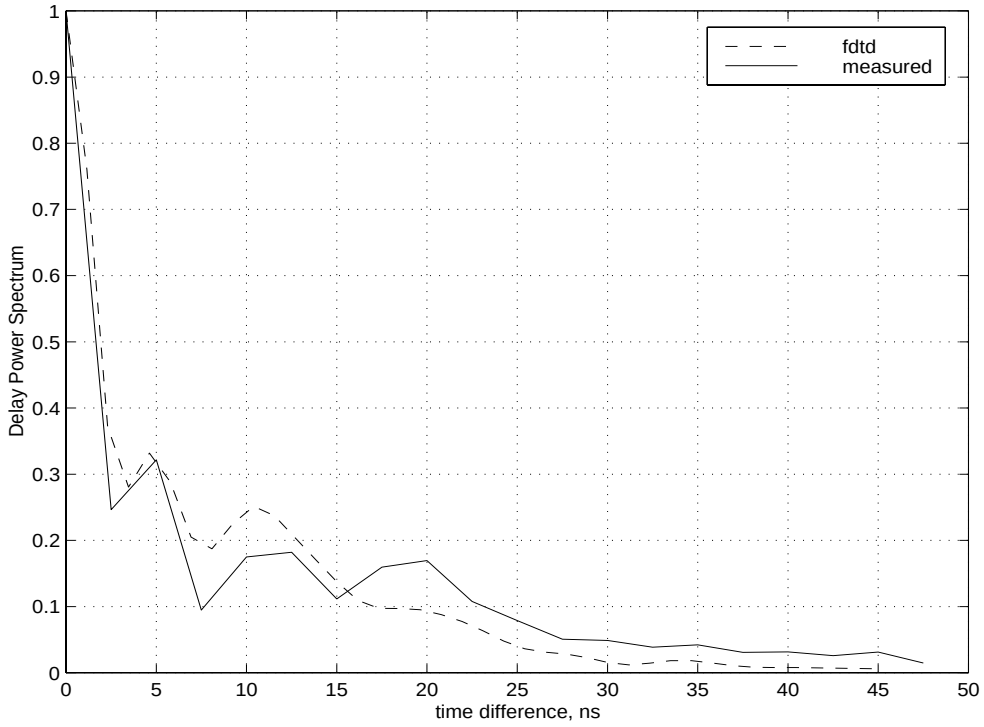



Figure 3.10: Delay power spectrum in Bedroom1

between the measured and predicted curves. The delay profile shows a good agreement with an exponential fitting curve of time constant of 15 ns (i.e. $T_m =$ 15 ns, not shown on the graph). The rms delay spread is defined as [7]:

$$\tau_{\text{rms}} = \left\{ \frac{\sum_k (t_k - \tau_m - t_A)^2 \, \alpha_k^2}{\sum_k \alpha_k^2} \right\}^{1/2} \tag{3.6}$$

where $t_A$ is the arrival time of the first path in a profile and $\tau_m$ is the mean excess

delay defined as:

$$\tau_m = \frac{\sum_k (t_k - t_A)\, \alpha_k^2}{\sum_k \alpha_k^2} \tag{3.7}$$

The rms delays of the channel profiles used to create the delay power spectrum of Figure 3.10 range from 35 ns to 13 ns, giving an average $\tau_{\text{rms}}$ of 20 ns. Similar numerical results for the $\tau_{\text{rms}}$ were obtained from measurements throughout the house and the predicted values from FDTD also agreed well.

We already saw that individual multipath components follow lognormal distributions. Also of interest are the statistics of the channel taps, when a specific baud rate is used. Here, we will use the flat fading channel model, that is we will neglect possible ISI induced by the delay spread of the channel. Given the $\tau_{\text{rms}}$ values found above, with an average of 20 ns, the ISI induced will be negligible for baud rates close to 1 Mbaud/s. Note though, that for higher baud rates (close to 5 Mbaud/s perhaps) ISI can be present and a second or third tap may be needed to model the channel. The statistical analysis given below can then easily be extended to analyze the selective fading channel also.

We model the channel as a single tap channel, at a frequency of 1 GHz. Taking Fourier transform of equation (3.1), we obtain:

$$C(f_0, x, y) = \sum_{k=0}^{\lfloor T_m W \rfloor + 1} c_k(x, y) e^{-j2k\pi f_0 / W} \tag{3.8}$$

where $f_0 = 1$ GHz and we neglected the path loss component. We are therefore interested in the statistics of $|C(f_0, x, y)|$. We assume a baud rate of 1 Mbaud/s and we obtain the channel tap for different locations by Fourier transforming and subsequently filtering the corresponding complex baseband signals. Before that, we subtracted the large scale path loss from the profiles, based on (3.5), so that we have only the fast fading and shadowing components. We also repeated the

process for different central frequencies $f_0$, so that we have the channel performance at different frequencies. With these tap coefficients as a database we can determine both the statistics and the spatial and frequency correlations of the amplitude of the flat fading channel.

From the tap coefficients over different locations, we extracted first the shadowing component, based on equation (3.5). For this, we used a square window of $4\lambda \times 4\lambda$ to average the fast fading. We subtracted then this component to obtain the fast fading component. In Figure 3.11 we plot an example of the fast fading and the shadowing along a straight line that goes through both bedrooms in Figure 3.2.

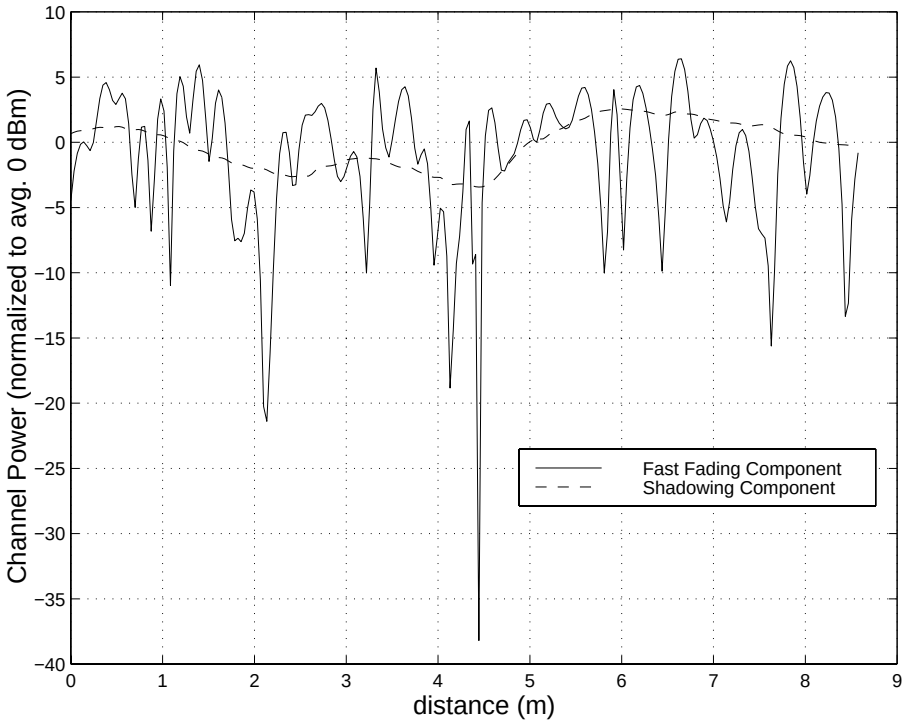



Figure 3.11: An example of the fast fading and the shadowing along a straight line that goes through both bedrooms.

We first plot the shadow fading statistics of the amplitude, in Figure 3.12. We used only our FDTD code for that, because we did not have enough data all over the house to extract the shadowing component. From the plot we see that the predicted shadowing component follows closely a lognormal distribution with a standard deviation of only $\sigma = 1.9$ dB. The low value of the shadowing probably has to do with the limited space of the house. Indeed, we do not expect the average power to change drastically from room to room, except perhaps from the places of direct LOS to areas of no LOS.

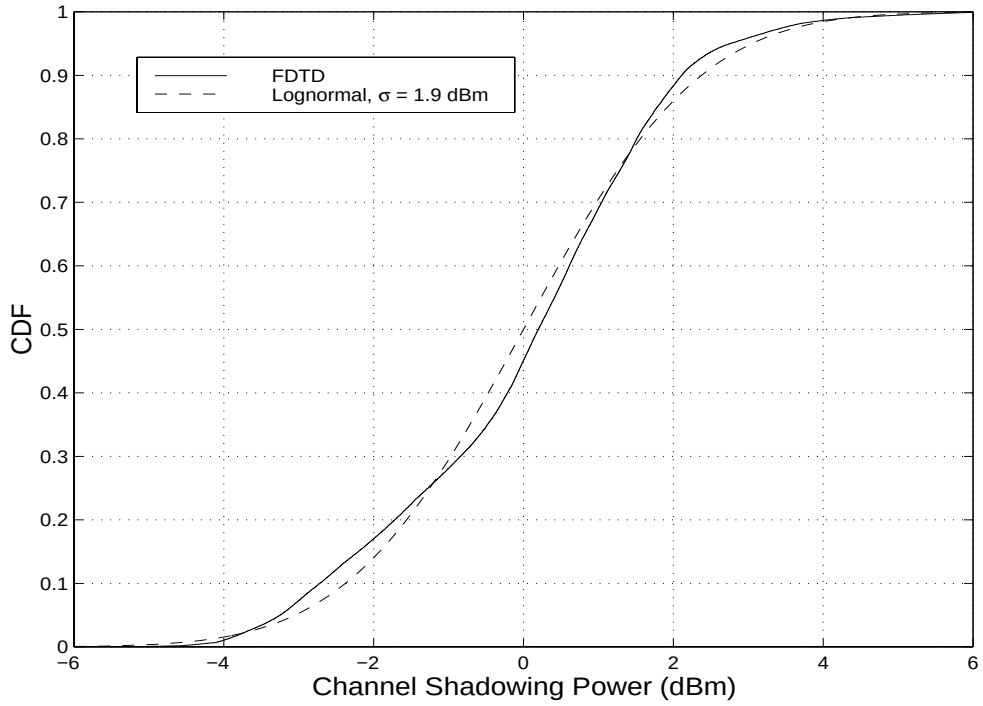



Figure 3.12: Shadowing statistics.

We then compute the fast fading statistics of the amplitude of the single tap channel for the areas where we have detailed measurements. In Figure 3.13 we plot the CDF of the channel envelope for measured data, the FDTD code and

a Nakagami fitting distribution for the square area inside the living room. The agreement of the measured and predicted distributions is evident. The agreement of both to the Nakagami distribution is also obvious. Note that the Nakagami distribution for the channel envelope is given by equation (1.31) of chapter 1. Note also, that there is a relation between the Nakagami and the Rice distributions (the latter given by equation (1.32) in chapter 1) [30], through the parameter $m$ of the Nakagami pdf and the Rice factor $K$, defined as:

$$K = \frac{s^2}{\gamma^2} \tag{3.9}$$

where $s^2$ is the power of the LOS component and $\gamma^2$ the variance of the random component of the Rice pdf (see equation (1.32), chapter 1). $m$ is connected to $K$ through the following equation:

$$m = \frac{(K+1)^2}{2K+1} \tag{3.10}$$

Obviously the higher the $K$ is, the stronger is the LOS component and the larger the $m$. A value of $m = 2.3$ was found to best fit the data in the living room, which gives a value for $K$ of 3.0, or 4.8 dB. This means that the distribution in the living room has a dominant LOS component and can equally well be modeled by a Rice distribution. Of course, this is an expected result, since the transmitter is quite close to the receiving positions.

Next, in Figure 3.14 we plot the CDF for the data collected in the kitchen. The CDF here for the measured and the FDTD predicted magnitude of the channel is Rayleigh. The agreement is again obvious. The measurements inside the kitchen were made right next to the refrigerator and there is obviously no LOS path (see Figure 3.2).

Finally, in Figure 3.15 we plot the CDF for the data collected in the bedroom, as well as the FDTD predicted CDF. The agreement between the two is once more

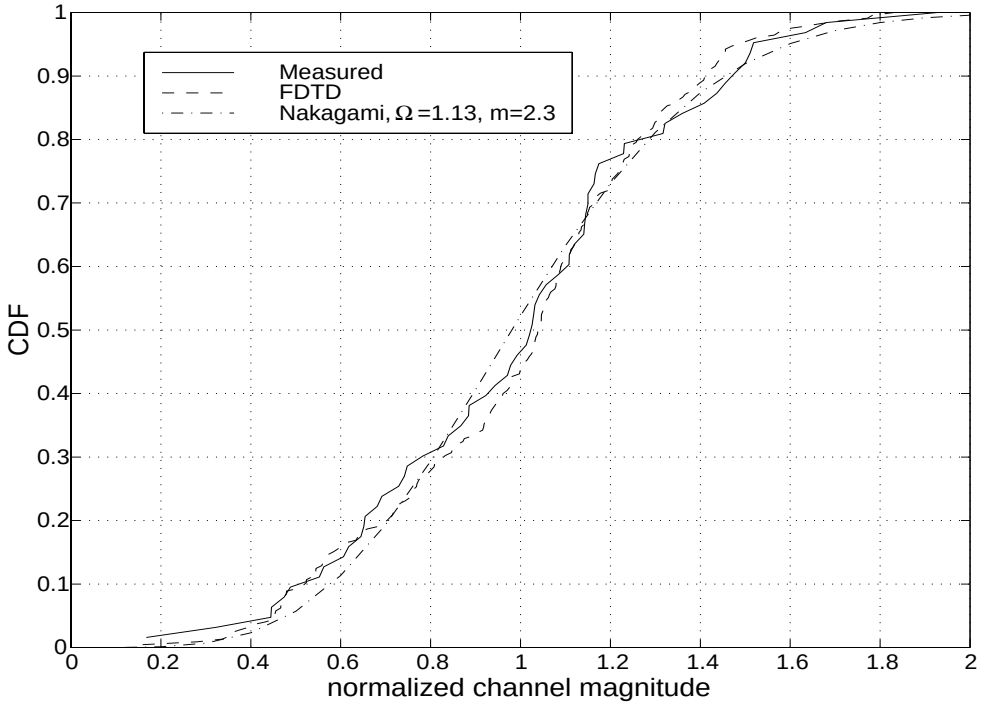



Figure 3.13: Fast fading statistics for single tap channel inside Living Room.

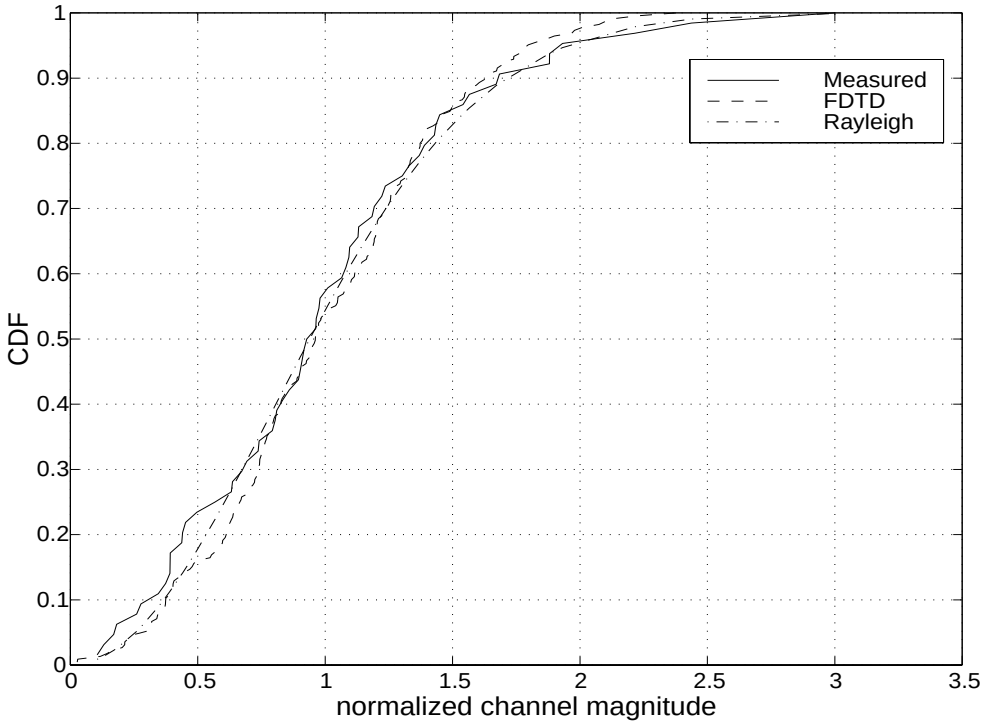



Figure 3.14: Fast fading statistics for single tap channel inside Kitchen.

evident. The best fitting distribution that we found was a Nakagami with a value for the parameter $m$ of 1.24. This value is quite close to 1, where we know that the Nakagami distribution becomes Rayleigh [30], see equation (1.31). Therefore inside the bedroom also the distribution of the flat fading channel is close to Rayleigh.

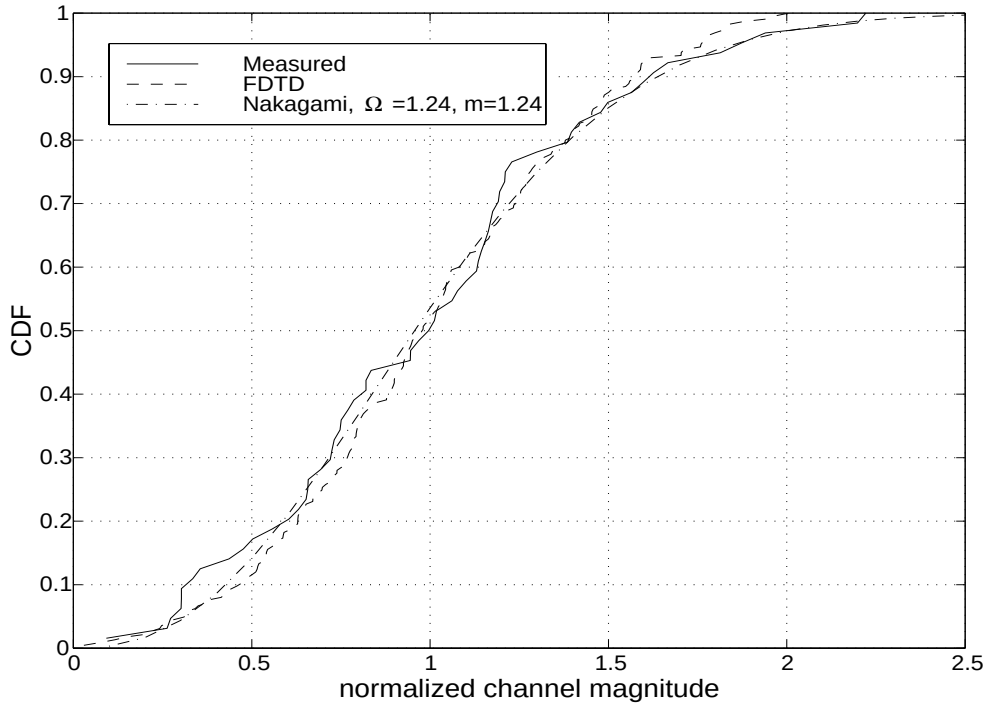



Figure 3.15: Fast fading statistics for single tap channel inside Bedroom1.

The main conclusion to be made from the distributions of the flat fading channel above is that there does not seem to exist a unique distribution that characterizes every area. Areas where strong LOS components exist have of course Ricean distributions, while areas with no LOS exhibit either Rayleigh of Nakagami with variable $m$ parameter. This result is corroborated by FDTD predicted statistics in other areas also. Having verified the validity of our FDTD

code in generating both accurate individual profiles and statistics, we also plot now the statistics of an area $70 \times 70$cm inside bedroom number 2 (see Figure 3.2). First we plot the CDF of the first two individual multipath components. In Figure 3.16 we plot the CDF of the first arriving path. The best fitting curve we found was a Gaussian on the magnitude of the multipath component. Although a Gaussian distribution on the magnitude implies Rician behavior, we found that the Rician distribution did not fit the data as well. The second multipath, shown in Figure 3.17 has a completely different behavior, showing a lognormal fit with a standard deviation of 3.5 dB. Finally, the equivalent flat fading channel (single tap) in Figure 3.18 shows yet another behavior, being Nakagami distributed in this case with parameter $m = 1.29$. The small value of $m$ indicates that the single tap channel has a more Rayleigh behavior, which is quite interesting, given that the main multipath component does not at all have such a distribution.

### 3.3.1 Correlation Statistics for Flat Fading

The very fine resolution of FDTD, which gives us channel profiles every $\lambda/10$, permits us to conduct some further investigations in the nature of the several correlations that exist between these profiles, as well as between the frequency components of the same profile.

We begin with the frequency correlation of the shadowing and fast fading components. These correlations of course, are connected to the power spectrum profile of the channel and they show us how effective frequency diversity (e.g. frequency hopping) can be. Since we have very small delay spreads, we expect a large decorrelation bandwidth, which is roughly inversely proportional to the delay spread. Using the notation of equation (3.5), let $L(f_i)$ be the shadowing power component at a given frequency of the channel profile. Then, the frequency

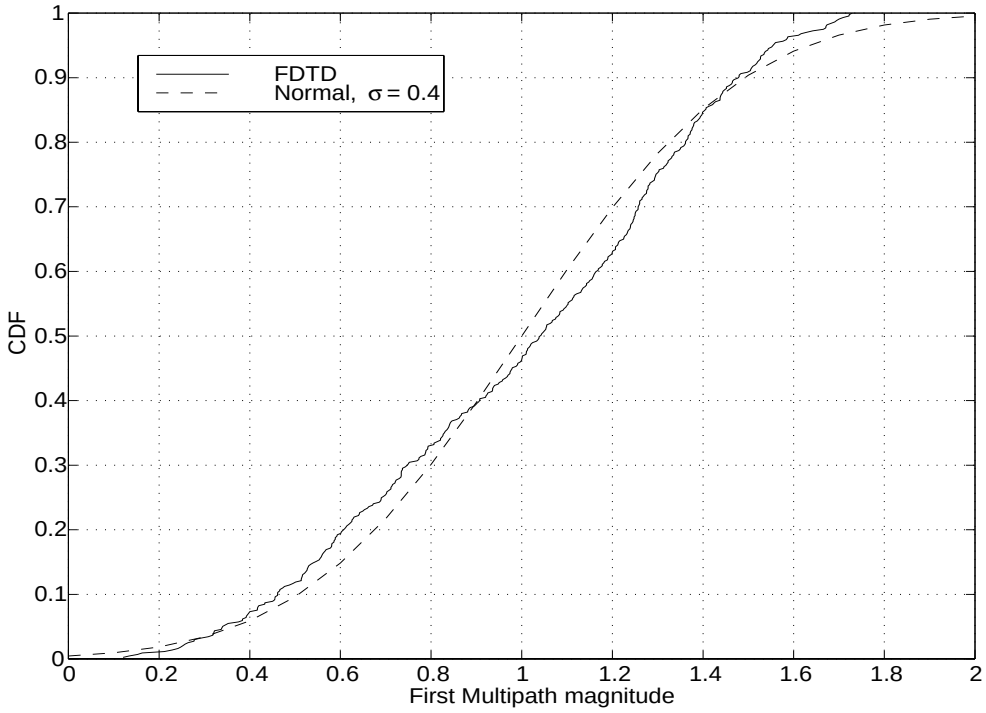



Figure 3.16: First Multipath Component CDF, Bedroom2.

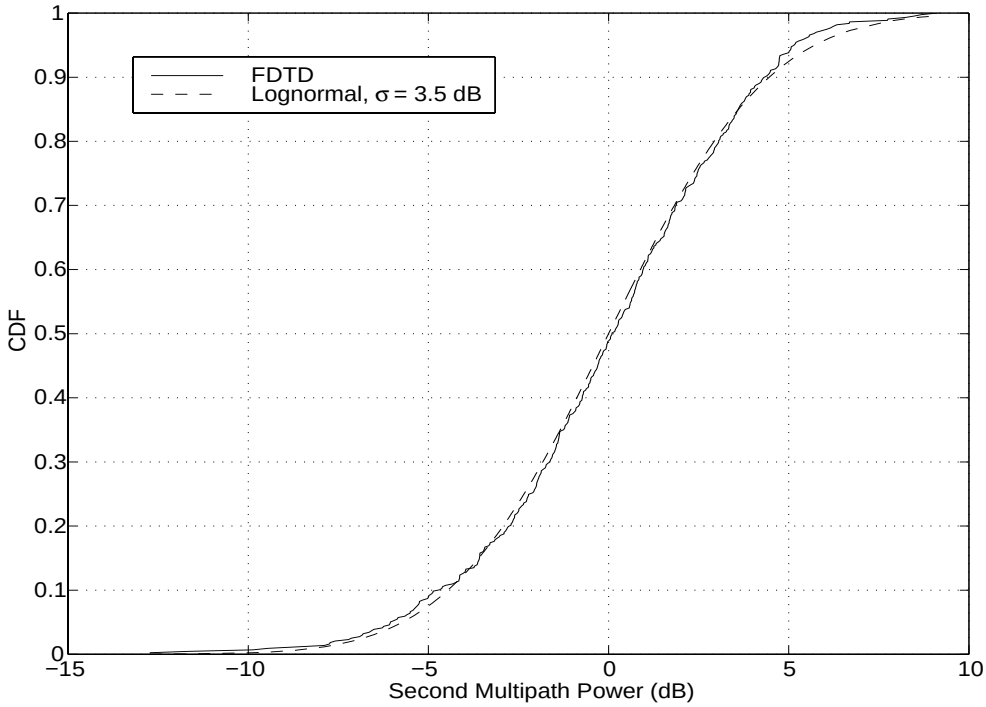



Figure 3.17: Second Multipath Component CDF, Bedroom2.

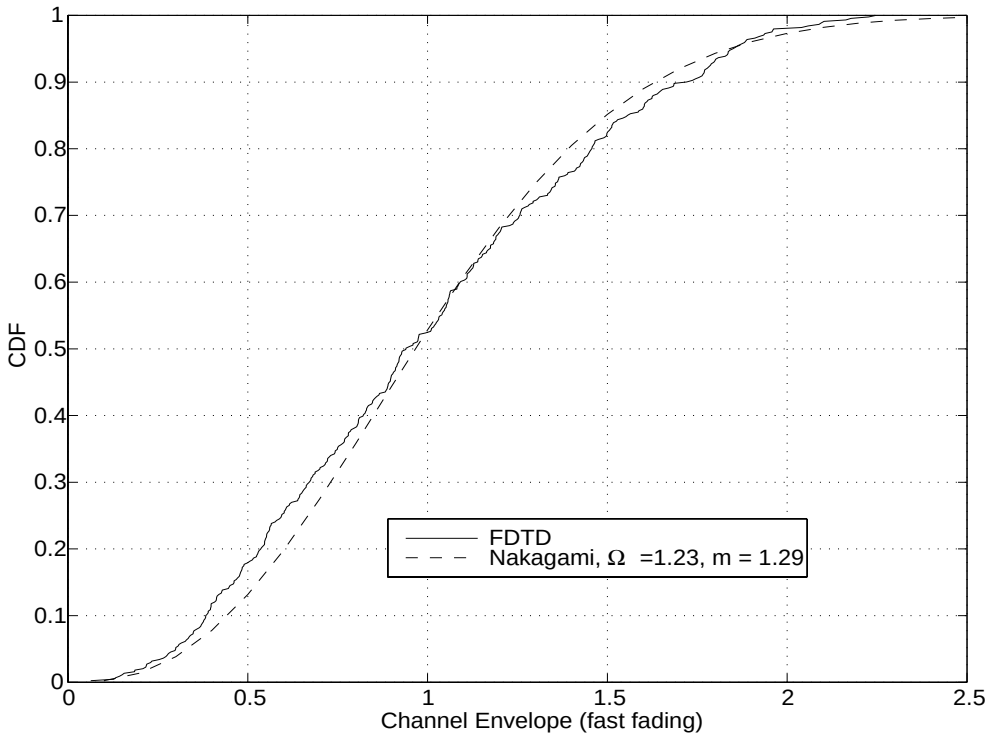



Figure 3.18: Fast fading statistics for single tap channel inside Bedroom2.

correlation for the shadowing will be given by:

$$\Phi_L(\Delta f) = E\left\{L(f_i) \cdot L(f_i + \Delta f)\right\} \tag{3.11}$$

where the averaging is done over space (different spatial profiles) and we assume the shadowing is a stationary process over the frequencies. Actually, we compute the autocovariance of shadowing, which is the above minus the mean. The same things can be said for the frequency correlation of the fast frequency component, giving rise to the frequency correlation for the fast fading, denoted by $\Phi_R(\Delta f)$.

In Figure 3.19 we plot the above autocovariances for both the shadowing and fast fading. The frequency correlation of the fast fading component is much smaller than for shadowing, although still it's quite large. We need about 5 MHz frequency separation to get to a correlation coefficient of 0.6. The shadowing frequency correlation is much more persistent, even at 20 MHz we are still at a value of 0.8 (the last part of this plot, where the correlation seemingly increases again should probably be ignored). This is expected, and shadowing is usually modeled as flat.

We proceed now to find the spatial correlation of shadowing and of the fast fading component of the channel. The spatial correlation for both these components is very important, especially for high data rates that render the channel very slowly fading. The correlation for both components is evaluated as (we give the equation of shadowing only, since the fast fading is identical):

$$\Phi_R(\Delta x, \Delta y) = E\left\{R^2(x, y) \cdot R^2(x + \Delta x, y + \Delta y)\right\} \tag{3.12}$$

where the averaging is done spatially and again we assume that we actually calculate autocovariances, removing the means from the above equation. The two-dimensional autocovariance was calculated for the shadowing component of the Bedroom1 and is plotted in Figure 3.20. For the y and z directions refer

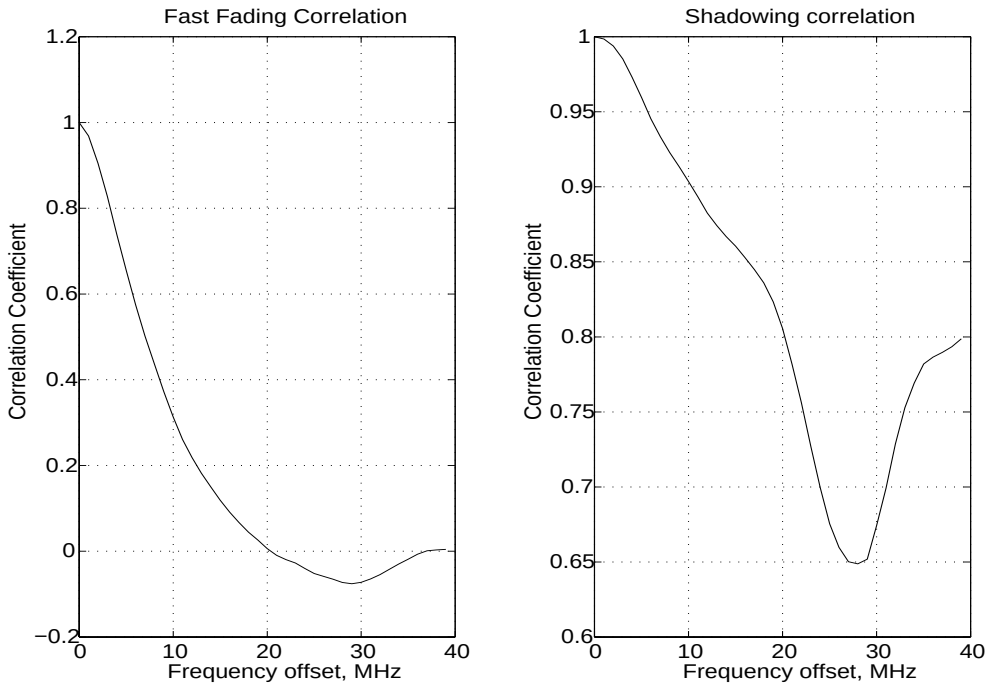



Figure 3.19: The frequency correlation for the shadowing and fast fading components of the flat fading channel.

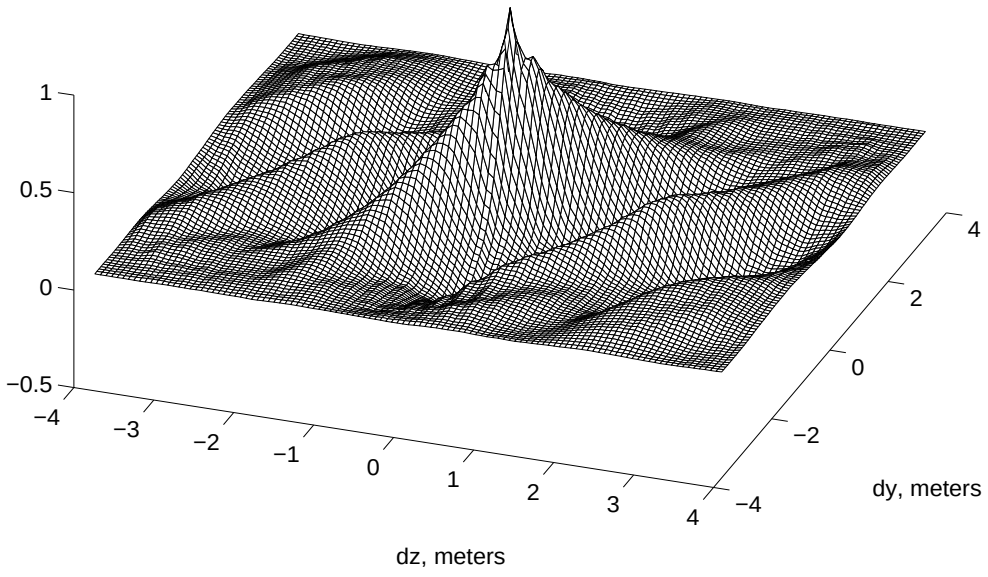



Figure 3.20: The spatial correlation of shadowing.

to Figure 3.2. It's obvious from these two figures that there is a very strong correlation towards the transmitter, and much less so in directions normal to the direction of the transmitter. This result shows that for small areas, assumptions of isotropic scattering are not really valid. Of course, the alternative would be a model that would change depending on the relative position of the receiver and the transmitter, and that's probably prohibitively complex. In any case, we decided to neglect this spatial selectivity in the correlation and model it as isotropic. Given the fact that shadowing is lognormal (i.e. normal in dB), we can easily make a 2-dimensional filter to model the shadowing correlation. We choose to make an AR filter because it's much shorter that a FIR for the same accuracy of approximation for the correlation. Refer to Figure 3.21 for an explanation. We create two series of Gaussian random numbers with an appropriate variance

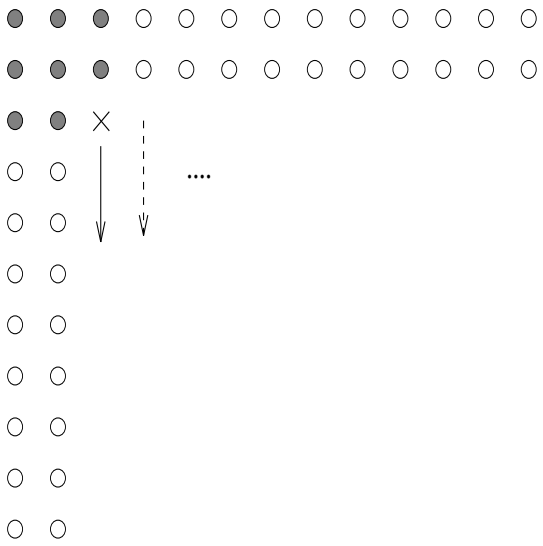



Figure 3.21: Construction of correlated shadowing from a 2-D AR model.

(of which more below) for each direction of the 2D space. These are the small circles in Figure 3.21, and they are spaced in an appropriate distance so that the sampling theorem for the spatial correlation is satisfied. The new correlated

Gaussian variable, shown as an $\times$, is then created by the 8 numbers shown as filled circles, and the procedure continues in the direction indicated. Of course, one has to be relatively away from the boundary to obtain numbers with the appropriate correlation. The variance for the Gaussian random variables should be the error from the solutions of the standard Yule-Walker equations for the AR coefficients. An alternative approach is that of a 2D FIR filter, but that would need a much larger number of taps. In this approach, independent Gaussian random variables are created on a grid (with appropriate spacing) on the 2D space. Then, a shifting 2D window with the FIR coefficients is used to generate a correlated variable in it's center. The positive side of this method is that there is no need to be away from the boundary to get correct correlation statistics.

In Figure 3.22 we plot the one-dimensional autocovariance of shadowing together with the models described above. Note that the actual two-dimensional autocovariance is the same at any direction. The spacing we used was $\lambda/2$. The AR model does not approximate quite as well the actual autocovariance, as the FIR filter, but on the other hand has only 8 coefficients versus 121.

It is interesting to note that the above model for the shadowing can be combined with frequency correlation, seen in Figure 3.19. For each point in space (each profile) we can generate a sequence of correlated Gaussian random variables according to the frequency correlation and then run on top of that the spatial correlation procedure described above, for each set of initially spatially uncorrelated random numbers. This way we get a set of frequency and spatially correlated channels. It is obvious though that the whole procedure becomes rather tedious, but nevertheless is doable.

Finally, we present results obtained from the FDTD code for the complex autocorrelation coefficient of the single tap channel given by equation (3.8). We

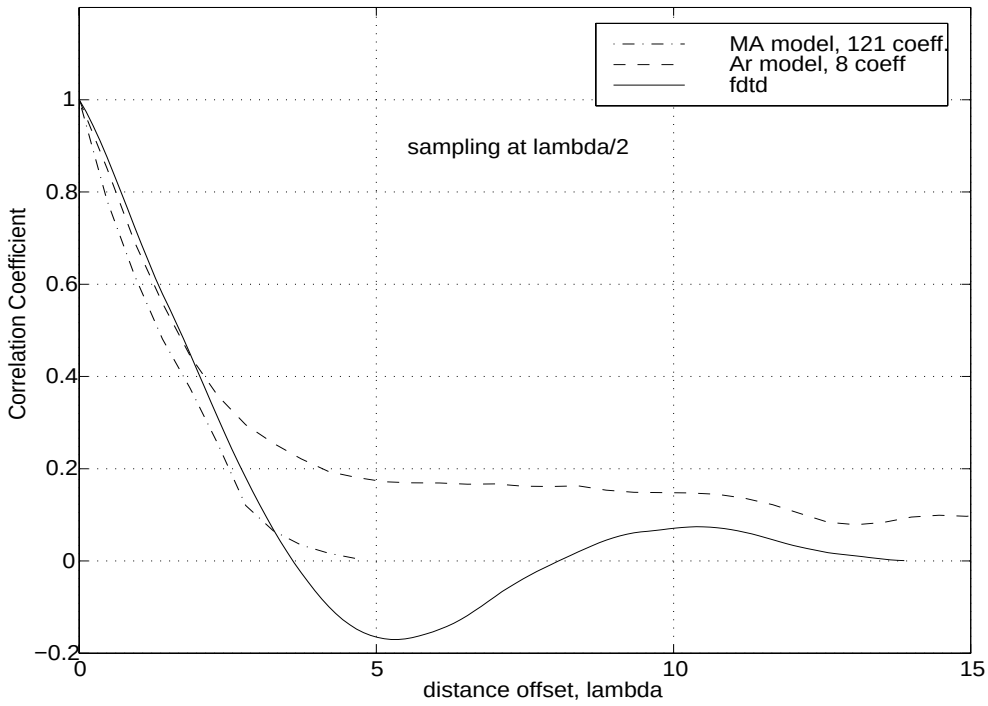



Figure 3.22: One dimensional shadow correlation and approximations with AR an MA filters.

define the elementary spatial step $\Delta r = \sqrt{\Delta x^2 + \Delta y^2}$, and calculate the auto-correlation coefficient as:

$$\rho_{\mathrm{C}}(k\Delta r) = \frac{E\left[C(f_0, r)C^*(f_0, r + k\Delta r)\right]}{\sqrt{E\left[|C(f_0, r)|^2\right]}} \tag{3.13}$$

where $*$ denotes complex conjugate.

Before we give our results for the fast fading correlation, we return briefly to the well known Jakes model for the correlated flat fading channel [2]. As is well known, a component wave which has a relative angle $\phi$ with the direction of motion of the receiver, will exhibit a frequency shift that can be defined through the following equation:

$$f\Delta t = \frac{v\Delta t}{\lambda_0}\cos(\phi) = \frac{\Delta r}{\lambda_0}\cos(\phi) \tag{3.14}$$

where $\lambda_0$ is the wavelength at the carrier frequency. It can be proved than that, equation (3.13) for the channel autocorrelation can be written as:

$$\rho_{\mathrm{C}}(k\Delta r) = \alpha(k\Delta r) - j\beta(k\Delta r) \tag{3.15}$$

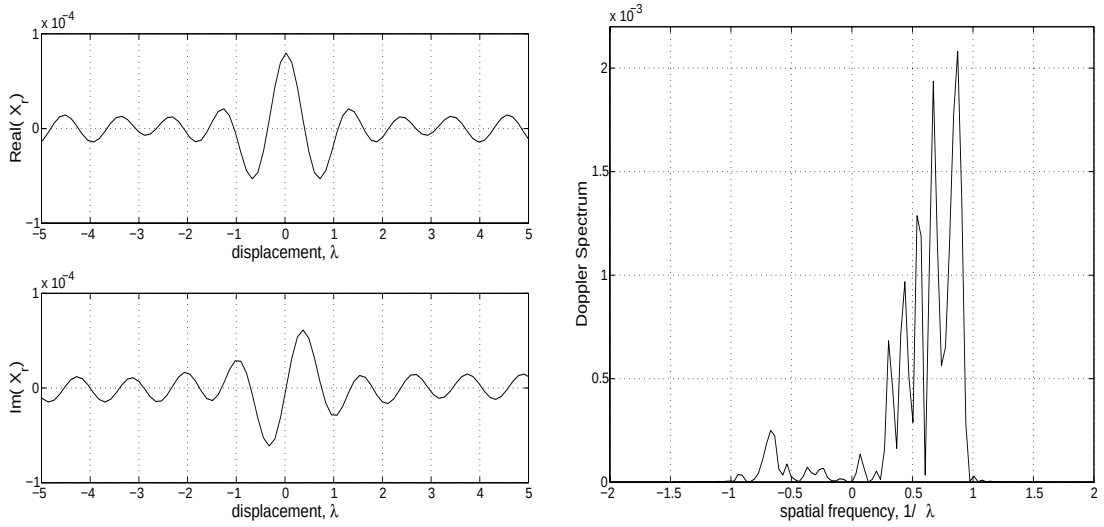
where:

$$\alpha(k\Delta r) = E\left\{\cos\left[\frac{k\Delta r}{\lambda_0}cos(\phi)\right]\right\} \tag{3.16}$$

$$\beta(k\Delta r) = E\left\{\sin\left[\frac{k\Delta r}{\lambda_0}cos(\phi)\right]\right\} \tag{3.17}$$

and the expectations are taken over the probability distribution of the arrival angles of the individual wave components. In the case of isotropic scattering (Jakes model), (3.17) is identically zero, while (3.16) gives the well known Bessel function type of correlation, with a spectrum of:

$$S_C(\xi) = \frac{v}{2\pi\lambda_0\sqrt{1 - (\xi\lambda_0)^2}} \tag{3.18}$$




(a) (b)

Figure 3.23: Complex correlation of fast fading and Doppler spectrum. (a) Real and imaginary parts of correlation. (b) Resulting Doppler spectrum.

where $\xi$ is the spatial frequency and it is always less than $\lambda_0^{-1}$. All the above is a direct transformation from the very well known temporal correlation to spatial correlation.

In Figure 3.23 we plot the real and imaginary parts of the autocorrelation (i.e. $\alpha$ and $\beta$ of equations (3.16) and (3.17) above), and also the spectrum of the correlation of the channel. The data are taken from one straight lines along the z-direction inside Bedroom1 (see Figure 3.2).

Two very important observations are made from these plots: first, $\beta(k\Delta r)$ is not identically zero here, although $\beta(0)$ is zero, which indicates that the in-phase and quadrature channel components are independent as in the Jakes model. The other, more prominent observation, is that the spectrum of the correlation lacks

contributions from scattering waves at angles larger than $90^o$. Observe from (3.14) that negative frequencies correspond to angles between incoming waves and the direction of motion of the receiver, larger than $\pi/2$. The interpretation of this is shown in Figure 3.24. There are no wave components coming from wide angles, or in other words, most of the energy comes from the vicinity of the transmitter. It makes sense that the situation for indoor wireless channels will be quite different from the cellular channel, where the assumption of uniform (isotropic) scattering is valid due to the large distances from the transmitter. Unfortunately, modeling such a situation can be very tedious, since the Doppler spectrum will depend on the relative position of the transmitter and receiver and the motion of the receiver.

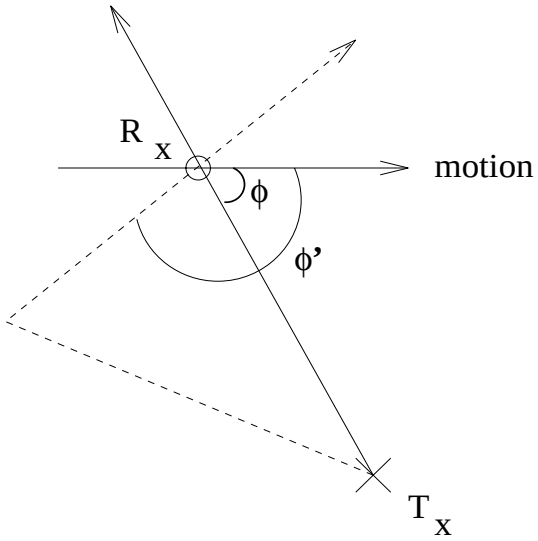



Figure 3.24: Illustration of scattering waves. Components having angles $\phi' > 90^o$ would create negative frequency shifts.

## 3.4 Summary

A new method was introduced for predicting the channel profiles for indoor wireless communication purposes. The basis is the finite-difference time-domain method in electromagnetics, a numerical and very powerful method to solve complex scattering problems. Advantages of the method are the remarkable simplicity with respect to much more complex ray tracing methods, the fact that it's not a high frequency approximation but an exact solution of Maxwell's equations and the fact that it produces the channel profiles in every point inside the three-dimensional area we are interested in. Its major disadvantage is the extremely high amount of computer memory needed for the application of the method (for the house we showed in Figure 3.2 we needed 256 Mbytes of computer memory).

In order to check the validity of our method we did a series of measurements and then ran our code for the same residence. The results are very encouraging, the agreement is very good in all cases for the probability distributions of individual components, as well as equivalent channel taps for a given baud rate. The main result from the statistical analysis of a single tap (flat fading) channel is that the distributions vary inside the residence, depending on how much of a LOS component exists. There is a range of distributions from Rayleigh, Nakagami to Rician. The individual components are usually lognormally distributed, unless a very strong LOS components is present. It's worth noting that though the multipath components may be lognormal, the equivalent channel taps are not.

Statistical analysis was also performed for the correlations of the channel taps, both in frequency and in space. We found the two-dimensional correlation of shadowing and devised an AR model to approximate this correlation. We also examined the correlation properties of mutlipath fading components, where we saw that the Doppler spectrum can differ dramatically from the well known Jakes

model, due to the fact that most of the scattering comes from the vicinity of the source.

Very slow fading channels (i.e. for pedestrian speeds) have been modeled in the literature often as Markov channels. This modeling relies heavily on the first order Markovian assumption for a Rayleigh channel, which has been verified in the literature [31]. The fading has been assumed invariably to follow the Jakes model. It would be interesting to test the Markovian assumption for the fading correlations we obtained above for our channel (see Figure 3.23) and if this holds, we could devise a Markov model for our channel. The simplest thing would be to use an autoregressive model for the amplitude, or we could discretize it and use a Markov chain model. The coefficients of either model will depend on the fading rate, which in turn depends on the pedestrian speed and the baud rate.

# CHAPTER 4

# A Memory Efficient Formulation of the Finite Difference Time Domain Method for the Solution of Maxwell Equations

## 4.1 Introduction

In the previous two chapters we introduced the FDTD method for solving electromagnetic scattering problems, and we applied it to the statistical characterization of an indoor wireless channel. The use of FDTD is very attractive for the electromagnetic analysis of complicated geometries, due mainly to its algorithmic simplicity. However, the computational requirements are high, and computer memory can become a limitation for electrically large bodies. We saw in chapter 2 that this is the biggest drawback in using FDTD extensively for wireless channel characterization. The memory limitation comes from the fact that the whole computational domain has to be discretized with rectangular cells having size of tenths of a wavelength, in order to avoid numerical dispersion which will lead to numerical inaccuracy [32].

In this chapter, we modify the original FDTD formulation to reduce computer memory requirements, allowing a 50% increase in the computational volume for a given computer memory size, with moderate increase in computation and code complexity. To our knowledge, only one research group has worked on

this problem in the past. In [39], the use of divergence free electric field regions was introduced, combined with the scalar wave equation in order to achieve this goal. In their formulation, the authors subdivided the computational domain into divergence free and non-free regions. Normal FDTD was applied in regions of discontinuities, conductors, sources, and dielectric interfaces, while the scalar wave equation, requiring four memory elements per cell, instead of the usual six of FDTD, was used in electric field divergence free regions. Even though this approach reduces memory requirements, its implementation can be complicated to program, due to the necessity of having sub-regions. More importantly, the memory reduction for this technique is only achieved in some specific geometries, such as planar structures, or other cases where large homogeneous regions exist.

In our new method, which we call Reduced-Finite Difference Time Domain (R-FDTD), we eliminate the necessity of subdividing the computational domain into sub-regions, maintaining the advantage of reducing the number of required field components to four, while being also able to treat easily conductors and source regions. This is achieved with the use of the divergence free nature of the electric displacement instead of the electric field as in [39]. Conductors and source regions are properly treated by calculating the induced charges, which are then used in the divergence of $\mathbf{D}$ ($\nabla \cdot \mathbf{D} = \rho$). In our formulation, although we store only four field components over the whole domain and the induced charge wherever conductors are present, we can always reconstruct the two missing field components. This way, standard absorbing boundary conditions, such as Mur [20] or PML [23] can be implemented as will be shown in the numerical examples section. Here lies one significant merit of R-FDTD, that is, its close resemblance to regular FDTD. For large electromagnetic problems, where memory can really become the decisive factor in using time domain methods, the transition from FDTD to R-FDTD requires only moderate code modifications.

In the following sections, we begin with the formulation for divergence free regions, where no conductors or sources are present, and explain the new algorithm for the two-dimensional case, together with a numerical example demonstrating the equivalence with FDTD. Next we extend R-FDTD to the three-dimensional case and present the algorithm in detail. In the third and fourth sections, we make further extensions for the treatment of conductors and source regions. Subsequently, we present numerical results where the perfect agreement with FDTD confirms the validity of this new technique. In the last section, we conclude with some discussion of the merits and drawbacks of the new method, as well as potential areas of application.

## 4.2 The New Formulation in Charge Free Regions

The standard Yee algorithm [15] for the solution of Maxwell equations is based on their discretization in space and time. Starting from system (4.1), the time marching solution is obtained using a leap-frog scheme [40] to propagate from each component of $\mathbf{E}$ to $\mathbf{H}$ and vice versa.

$$\begin{aligned} \frac{\partial}{\partial t}\mathbf{D} &= \nabla \times \mathbf{H} \\ \frac{\partial}{\partial t}\mathbf{B} &= -\nabla \times \mathbf{E} \end{aligned} \tag{4.1}$$

In this scheme, all six scalar field components are used explicitly and therefore must be stored over the whole computational domain, due to the presence of first derivatives in time in (4.1). In charge free regions though, these components are not independent, but are linked through the two flux equations. This link can be obtained by taking the divergence on both sides of Ampère's and Faraday's

equations of system (4.1), in which case we obtain [12]:

$$\nabla \cdot \frac{\partial}{\partial t}\mathbf{D} = 0 \tag{4.2}$$

$$\nabla \cdot \frac{\partial}{\partial t}\mathbf{B} = 0 \tag{4.3}$$

Upon approximating the time derivatives with central differences in time in (4.2) and with forward differences in time in (4.3), we rewrite the above equations as follows:

$$\nabla \cdot \left(\mathbf{D}^{n+1/2} - \mathbf{D}^{n-1/2}\right) = 0 \tag{4.4}$$

$$\nabla \cdot \left(\mathbf{B}^{n+1} - \mathbf{B}^{n}\right) = 0 \tag{4.5}$$

Assuming that initially (at time $t = 0$, $n = 0$) all the field components are zero over the whole computational domain, we obtain:

$$\nabla \cdot \mathbf{D}^{n+1/2} = 0 \tag{4.6}$$

$$\nabla \cdot \mathbf{B}^{n+1} = 0 \tag{4.7}$$

Note that the assumption of zero initial fields is generally true everywhere except the source region, which will be treated in detail later on in this paper (see section 4.4).

Equations (4.6) and (4.7) demonstrate a spatial dependence between the components of vectors $\mathbf{D}$ and $\mathbf{B}$ respectively. Notice that, because $\mathbf{D} = \epsilon\mathbf{E}$ and $\mathbf{B} = \mu\mathbf{H}$, this spatial dependence exists regardless of the properties of the medium, which are included in equations (4.6), (4.7). We use this dependence in (4.6) and (4.7) to link spatially two of the field components to the other four, and this way we reduce the number of field components needed in FDTD, from six to four. Similarly, in the two-dimensional case we reduce the number of components from three to two. In the following, we begin with the simpler to

explain two-dimensional case and we continue with the three-dimensional extension. Throughout the paper we assume that the magnetic permeability coincides with that of free space ($\mu = \mu_0$). Our results extend directly to mediums with magnetic properties ($\mu \neq \mu_0$), by a parallel formulation.

### 4.2.1 The Two-Dimensional Formulation

Consider at first the TE two-dimensional case, that is the only field components are $E_x, E_y$ and $H_z$, as shown in Figure 4.1. Equation (4.6) can be used to link $E_x$ and $E_y$, and upon discretization in space and reordering can be written as:

$$E_y^{n+1/2}(i,j) = \frac{\epsilon_{i,j-1}}{\epsilon_{i,j}} E_y^{n+1/2}(i,j-1) - \frac{\Delta y}{\Delta x} \frac{\epsilon_{i,j} E_x^{n+1/2}(i,j) - \epsilon_{i-1,j} E_x^{n+1/2}(i-1,j)}{\epsilon_{i,j}} \tag{4.8}$$

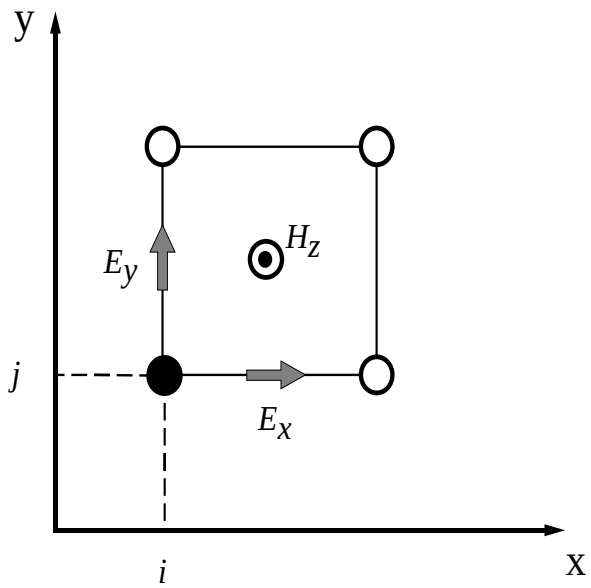



Figure 4.1: Definition of the elementary cell for the two-dimensional TE case.

Equation (4.8) can be incorporated in a standard FDTD algorithm, and upon

a proper update scheme, only $E_x$ needs to be stored over the whole computational domain. To explain how this is done, assume that $E_x^{n+1/2}$ has already been calculated with the regular FDTD update equation over the whole computational domain and we want to continue with the update of $H_z^{n+1}$, without having to store $E_y^{n+1/2}$ everywhere. The FDTD update equation for $H_z^{n+1}$ is [15]:

$$H_z^{n+1}(i,j) = H_z^n(i,j) + \frac{\Delta t}{\Delta y \mu_0}\left[E_x^{n+1/2}(i,j+1) - E_x^{n+1/2}(i,j)\right] - \frac{\Delta t}{\Delta x \mu_0}\left[E_y^{n+1/2}(i+1,j) - E_y^{n+1/2}(i,j)\right] \quad (4.9)$$

Therefore, in order to proceed with the $H_z^{n+1}(i,j)$ field update, $E_y^{n+1/2}(i,j)$ and $E_y^{n+1/2}(i+1,j)$ are needed. These can be spatially updated using (4.8), assuming we have already calculated $E_y^{n+1/2}(i,j-1)$ for all $i$. In other words, the update of $H_z^{n+1}(i,j)$ can be done *one $j$ at a time* and for all $i$, with the prior update of $E_y^{n+1/2}(i,j)$ through (4.8) for all $i$ and having stored only $E_y^{n+1/2}(i,j-1)$ for all $i$. Since $E_y^{n+1/2}(i,j-1)$ was needed only for the calculation of $H_z^{n+1}(i,j-1)$ and is not used from $j$ on, we can use the same memory locations where we stored $E_y^{n+1/2}(i,j-1)$ for all $i$ to store the new values, $E_y^{n+1/2}(i,j)$. The spatial update of $E_y^{n+1/2}$ is also illustrated in Figure 4.2, where a vector of length $N_x$ (the size of the computational domain along $x$) for temporary storage of $E_y^{n+1/2}(i,j-1)$ is employed. Obviously, in order to begin the algorithm, we need to know $E_y^{n+1/2}(i,1)$, which can be obtained with the normal FDTD update equation for $E_y$:

$$E_y^{n+1/2}(i,1) = E_y^{n-1/2}(i,1) - \frac{\Delta t}{\Delta x \epsilon_{i,1}}\left[H_z^n(i,1) - H_z^n(i-1,1)\right] \quad (4.10)$$

Below, we present the algorithm of the new proposed formulation, skipping boundary condition implementation, which will be discussed later in detail. Note that $v_E(i)$ is the vector where $E_y(i,j)$ is stored temporarily and the computational

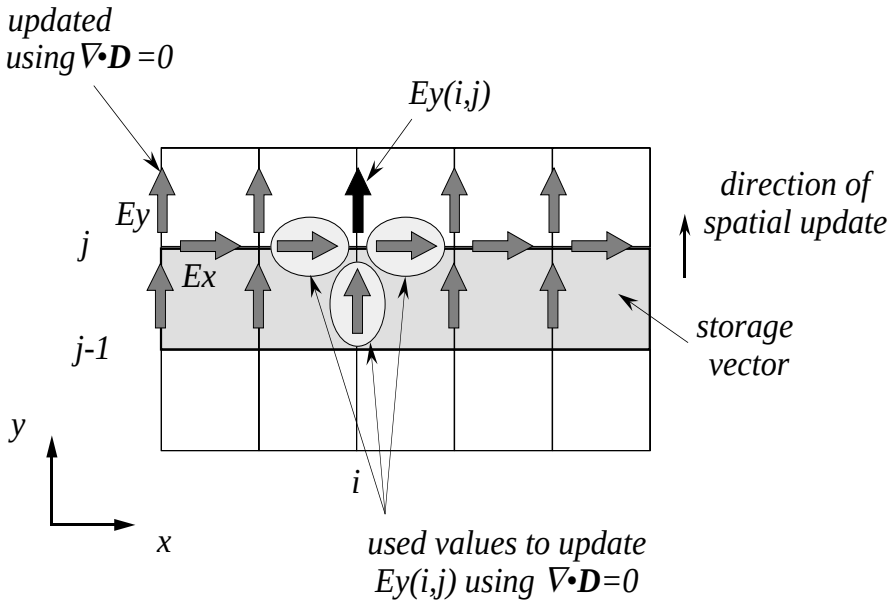



Figure 4.2: Spatial field update mechanism.

domain extends for $1 \leq i \leq N_x$, $1 \leq j \leq N_y$.

The pseudo-code shows the essence of the algorithm: by updating $H_z(i,j)$ one $j = \textit{constant}$ each time and for all $i$, we are able to calculate $E_y(i,j)$ only locally for that $j$ and store it in the same vector we stored $E_y(i,j-1)$, which is no longer needed. Note that this approach as opposed to [39], can treat any dielectric discontinuity or inhomogeneity naturally (since $\mathbf{D} = \epsilon\mathbf{E}$ and $\mathbf{B} = \mu\mathbf{H}$ for charge free regions are always divergence free vectors) while maintaining the

generality of FDTD.

Pseudo-code for 2-D case

1. $E_x$ field update. Use regular FDTD:

for i=1:$N_x - 1$

for j=2:$N_y - 1$

$$E_x^{n+1/2}(i,j) = E_x^{n-1/2}(i,j) + \frac{\Delta t}{\Delta y \epsilon_{i,j}} \left[ H_z^n(i,j) - H_z^n(i,j-1) \right] \quad \text{(FDTD)}$$

end j

end i

2. $H_z$ field update.

a. Get $E_y^{n+1/2}(i,1)$ from regular FDTD and store it in $v_E(i)$.

b. Update $H_z$ for all i and j:

for j=1:$N_y - 1$

for i=1:$N_x - 1$

$$H_z^{n+1}(i,j) = H_z^n(i,j) + \frac{\Delta t}{\Delta y \mu_0} \left[ E_x^{n+1/2}(i,j+1) - E_x^{n+1/2}(i,j) \right] - \frac{\Delta t}{\Delta x \mu_0} \left[ v_E(i+1) - v_E(i) \right] \quad \text{(FDTD)}$$

end i

Update spatially $E_y^{n+1/2}(i,j+1)$ for next $j$ iteration:

for i=2 : $N_x - 1$

$$v_E(i) = \frac{\epsilon_{i,j}}{\epsilon_{i,j+1}} v_E(i) - \frac{\Delta y}{\Delta x \epsilon_{i,j+1}} \Big[ \epsilon_{i,j+1} E_x^{n+1/2}(i,j+1) - \epsilon_{i-1,j+1} E_x^{n+1/2}(i-1,j+1) \Big] \quad (\nabla \cdot \mathbf{D} = 0)$$

end i

end j

The TM two-dimensional case (that is the field components now are $H_x$, $H_y$, $E_z$) can be formulated along the same lines, by using equation (4.7) instead of (4.6). Upon discretization in space and solving, for example, for $H_y^{n+1}(i,j)$, (4.7)

is written as (assuming $\mu = \mu_0$):

$$H_y^{n+1}(i, j+1) = H_y^{n+1}(i, j) - \frac{\Delta y}{\Delta x}\left[H_x^{n+1}(i+1, j) - H_x^{n+1}(i, j)\right] \tag{4.11}$$

Equation (4.11) can be used in a similar fashion as equation (4.8), so that only $H_x(i, j)$ needs to be stored everywhere. The formulation is straightforward and will be skipped.

As a validation of the proposed new algorithm, we present a numerical experiment for the internal electric field of a uniform, circular dielectric cylindrical scatterer. The cylinder is assumed to be infinite in the z-direction. The incident radiation is a TE plane wave with respect to the cylinder symmetry axis. Because there is no variation of either scatterer geometry or incident field in the z-direction, this problem may be treated as a two-dimensional one. The cylinder has a diameter of 7.5 mm and a relative dielectric constant of 4. The plane wave was excited as a series of modulated Gaussian pulses along the $x$-direction. The pulse has a width of 100 ps, corresponding to a bandwidth of 10 GHz and the modulation frequency was 20 GHz. Since the divergence of the electric field is zero everywhere on the boundary of the computational domain, we use the boundary condition proposed in [39], where first order Mur boundary condition is applied to the normal derivative of the normal field component (in our case, $E_x$ on the boundaries $x = 1$ and $x = N_x$). This way, $E_y(1, j)$ and $E_y(Nx, j)$ can still be calculated through $\nabla \cdot \mathbf{E} = 0$.

Results are shown in Figure 4.3, where the amplitude of the 20 GHz component of the field inside the cylinder along its diameter perpendicular to the plane wave is graphed. The plot is compared with the exact solution (continuous line) as in [41], and show perfect agreement with FDTD and with the analytical solution, within the error due to the stair-casing of the cylinder below 2%. The particular geometry was chosen for easy comparison with the canonical solution.

However, since no restrictions are made in the computational domain, the algorithm can be used for any arbitrary geometry. This is a substantial improvement on what was presented in [39], where the domain has to be subdivided into homogeneous sub-regions. Consequently it did not allow the treatment of scattering problems, where the whole region of interest is inhomogeneous.

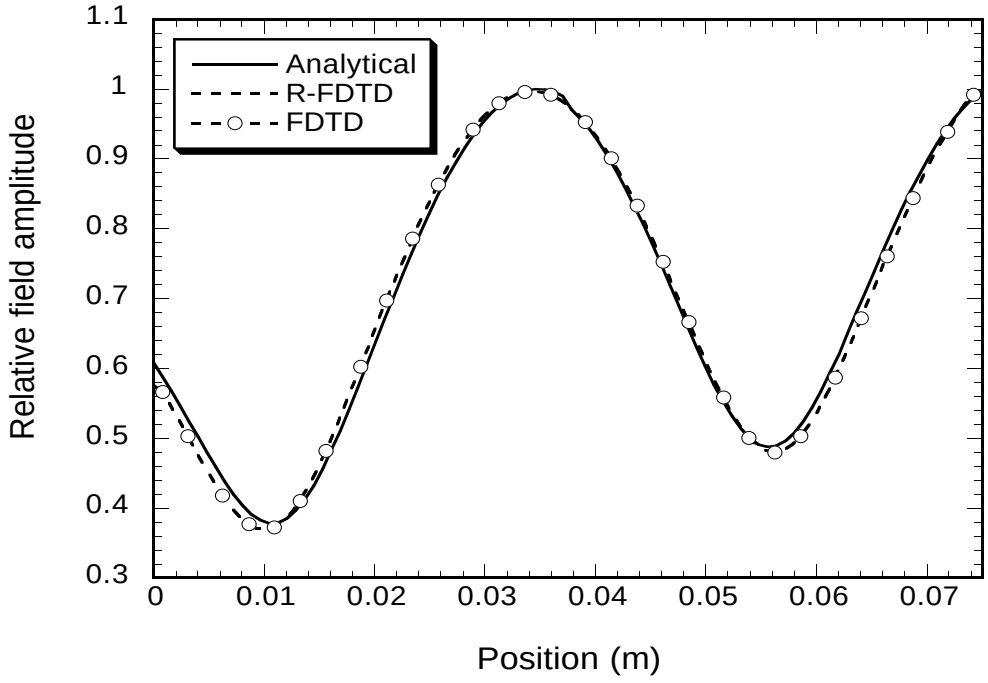



Figure 4.3: Comparison between FDTD, R-FDTD and analytical for the field inside the cylinder.

### 4.2.2 Extension to the Three-Dimensional Formulation

The extension of our technique to the three-dimensional problem can be carried out in a similar fashion as in the two-dimensional case, with the addition that both (4.6) and (4.7) are used simultaneously. Because of the spatial dependence of the electric and magnetic fields through these two equations, the total number

of variables for the three-dimensional formulation is reduced from six required in regular FDTD to four. In principle, one can choose independently the components of **E** and **H** which will be only locally updated and not stored, but it will simplify calculations if we choose components of the same direction for **E** and **H**. This choice has also advantages for the conductor treatment that we will present later.

Without loss of generality, consider the case where $E_y$ and $H_y$ are the components that are only locally calculated, and they are not stored in the whole domain. The update equations for them are derived by discretization of (4.6)-(4.7), based on the elementary cell definition shown in Figure 4.4, and are given below:

$$E_y^{n+1/2}(i,j,k) = \frac{\epsilon_{i,j-1,k}}{\epsilon_{i,j,k}} E_y^{n+1/2}(i,j-1,k) - \\ - \frac{\Delta y}{\Delta x} \frac{\epsilon_{i,j,k} E_x^{n+1/2}(i,j,k) - \epsilon_{i-1,j,k} E_x^{n+1/2}(i-1,j,k)}{\epsilon_{i,j,k}} - \\ - \frac{\Delta y}{\Delta z} \frac{\epsilon_{i,j,k} E_z^{n+1/2}(i,j,k) - \epsilon_{i,j,k-1} E_z^{n+1/2}(i,j,k-1)}{\epsilon_{i,j,k}} \quad (4.12)$$

$$H_y^{n+1}(i,j+1,k) = H_y^{n+1}(i,j,k) - \frac{\Delta y}{\Delta x} \left[ H_x^{n+1}(i+1,j,k) - H_x^{n+1}(i,j,k) \right] - \\ - \frac{\Delta y}{\Delta z} \left[ H_z^{n+1}(i,j,k+1) - H_z^{n+1}(i,j,k) \right] \quad (4.13)$$

Beginning with the updates of $E_x^{n+1/2}$, $E_z^{n+1/2}$, one realizes that these can be done one $j =$*constant plane* at a time, with the prior spatial update of $H_y^n$ through (4.13). This is so, since $E_x^{n+1/2}(i,j,k)$ depends on $H_y^n(i,j,k)$, $H_y^n(i,j,k-1)$, and $E_z^{n+1/2}(i,j,k)$ depends on $H_y^n(i,j,k)$, $H_y^n(i-1,j,k)$, [15]. In the same way, the updates of $H_x^{n+1}$, $H_z^{n+1}$ can also be done one $j =$*constant plane* at a time, with the prior spatial update of $E_y^{n+1/2}$ through (4.12).

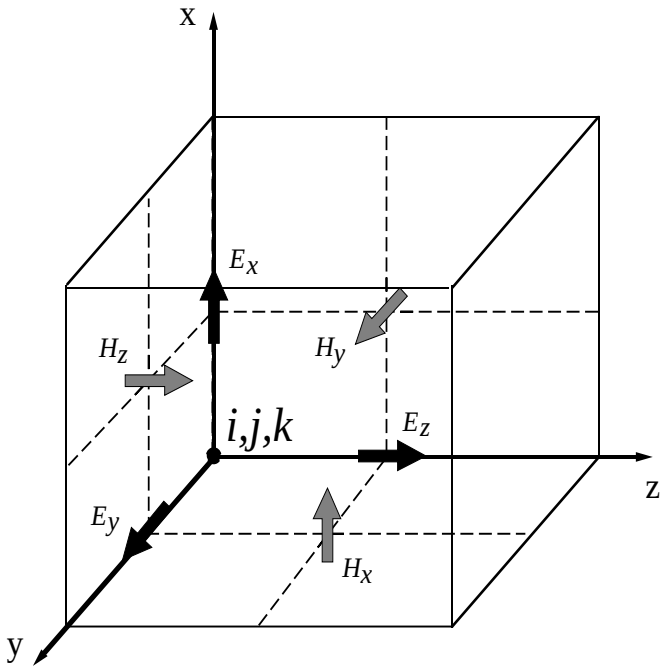



Figure 4.4: Definition of the elementary cell for the three-dimensional case.

As in the two-dimensional formulation, a storage of the $y$ components of the fields at the previous location along $y$ is required. For the three-dimensional case, this means that a two-dimensional array having size of $N_x \times N_z$ must be used for each one of the removed components. The extra memory requirement for these arrays is not significant, and since the choice of which component to remove is arbitrary, we can always choose to remove the field components in the direction for which the corresponding arrays are minimized. Below, we present the pseudo-code for the three dimensional formulation. $v_H(i,k)$ and $v_E(i,k)$ are the arrays where $H_y(i,j,k)$ and $E_y(i,j,k)$ are stored temporarily. The computational domain extends for $1 \leq i \leq N_x$, $1 \leq j \leq N_y$, $1 \leq k \leq N_z$.

Pseudo-code for the 3-D case

1. $E_x, E_z$ field updates.

a. Get $H_y^n(i,2,k)$ from regular FDTD and store it in $v_H(i,k)$.

b. Update $E_x$ and $E_z$ for all i,k and j:

for j=$2 : N_y - 1$

for i=$1 : N_x - 1$, k=$1 : N_z - 1$

$$E_x^{n+1/2}(i,j,k) = E_x^{n-1/2}(i,j,k) + \frac{\Delta t}{\Delta y \epsilon_{i,j,k}} [H_z^n(i,j,k) - H_z^n(i,j-1,k)] - \frac{\Delta t}{\Delta z \epsilon_{i,j,k}} [v_H(i,k) - v_H(i,k-1)] \quad \text{(FDTD)}$$

$$E_z^{n+1/2}(i,j,k) = E_z^{n-1/2}(i,j,k) + \frac{\Delta t}{\Delta x \epsilon_{i,j,k}} [v_H(i,k) - v_H(i-1,k)] - \frac{\Delta t}{\Delta y \epsilon_{i,j,k}} [H_x^n(i,j,k) - H_x^n(i,j-1,k)] \quad \text{(FDTD)}$$

end k, i

Update spatially $H_y^n(i,j+1,k)$ for next $j$ iteration:

for i=$2 : N_x - 1$, k=$2 : N_z - 1$

$$v_H(i,k) = v_H(i,k) - \frac{\Delta y}{\Delta x} [H_x^n(i+1,j,k) - H_x^n(i,j,k)] - \frac{\Delta y}{\Delta z} [H_z^n(i,j,k+1) - H_z^n(i,j,k)] \quad (\nabla \cdot \mathbf{H} = 0)$$

end k, i

end j

Pseudo-code for the 3-D case, continued

2. $H_x, H_z$ field updates.

a. Get $E_y^{n+1/2}(i,1,k)$ from regular FDTD and store it in $v_E(i,k)$.

b. Update $H_x$ and $H_z$ for all i,k and j:

for j=$1 : N_y - 1$

for i=$1 : N_x - 1$, k=$1 : N_z - 1$

$$H_x^{n+1}(i,j,k) = H_x^n(i,j,k) + \frac{\Delta t}{\Delta z \mu_0}\left[v_E(i,k+1) - v_E(i,k)\right] - \frac{\Delta t}{\Delta y \mu_0}\left[E_z^{n+1/2}(i,j+1,k) - E_z^{n+1/2}(i,j,k)\right] \quad \text{(FDTD)}$$

$$H_z^{n+1}(i,j,k) = H_z^n(i,j,k) + \frac{\Delta t}{\Delta y \mu_0}\left[E_x^{n+1/2}(i,j+1,k) - E_x^{n+1/2}(i,j,k)\right] - \frac{\Delta t}{\Delta x \mu_0}\left[v_E(i+1,k) - v_E(i,k)\right] \quad \text{(FDTD)}$$

end k, i

Update spatially $E_y^{n+1/2}(i,j+1,k)$ for next $j$ iteration:

for i=$2 : N_x - 1$, k=$2 : N_z - 1$

$$v_E(i,k) = \frac{\epsilon_{i,j,k}}{\epsilon_{i,j+1,k}} v_E(i,k) - \frac{\Delta y}{\Delta x \epsilon_{i,j+1,k}}\Big[\epsilon_{i,j+1,k} E_x^{n+1/2}(i,j+1,k) - \epsilon_{i-1,j+1,k} E_x^{n+1/2}(i-1,j+1,k)\Big] - \frac{\Delta y}{\Delta z \epsilon_{i,j+1,k}} \times \quad (\nabla \cdot \mathbf{D} = 0)$$
$$\times\Big[\epsilon_{i,j+1,k} E_z^{n+1/2}(i,j+1,k) - \epsilon_{i,j+1,k-1} E_z^{n+1/2}(i,j+1,k-1)\Big]$$

end k, i

end j

Again, note the innovative feature of the formulation: for example, $H_y^n$ is calculated in the **E** field update part of the algorithm, one $j$ =*constant plane* at a time and is stored in the same memory locations as for the previous $j$. Correspondingly, $E_y^{n+1/2}$ is calculated in the **H** field update part, also one $j$ =*constant plane* at a time and is also stored "in place".

As in the FDTD formulation, the stability condition for this new numerical technique is imposed by the Courant condition [40], which requires that in one

time step the propagating wave must not travel more than one computational cell.

## 4.3 Conductor Treatment

In the presence of conductors, the divergence of the electric flux is not zero anymore, but rather equals the induced charge on the conductor[1] ($\rho$), and equation (4.6) is not satisfied. In order to use our method in the presence of a conductor, the induced charge must be found. Indeed, in the following we proceed to show how this can be done through the use of the regular FDTD update equations, independently of our new algorithm, and then we discuss the extension of R-FDTD to include regions with conductors. Although we treat in detail electric conductors, all the discussions extend easily to the case of magnetic conductors with a parallel analysis.

Revisiting equation (4.4), let us denote by $\Delta\mathbf{D}^{n+1/2}(i,j,k)$ the difference approximation to $\nabla \cdot \mathbf{D}^{n+1/2}$. Based on the elementary cell definition shown in Figure 4.4, $\Delta\mathbf{D}^{n+1/2}(i,j,k)$ can be written as:

$$\begin{aligned}\Delta\mathbf{D}^{n+1/2}(i,j,k) = &\frac{\epsilon_{i,j,k}E_x^{n+1/2}(i,j,k) - \epsilon_{i-1,j,k}E_x^{n+1/2}(i-1,j,k)}{\Delta x} + \\ &+ \frac{\epsilon_{i,j,k}E_y^{n+1/2}(i,j,k) - \epsilon_{i,j-1,k}E_y^{n+1/2}(i,j-1,k)}{\Delta y} + \\ &+ \frac{\epsilon_{i,j,k}E_z^{n+1/2}(i,j,k) - \epsilon_{i,j,k-1}E_z^{n+1/2}(i,j,k-1)}{\Delta z}\end{aligned} \tag{4.14}$$

Then, equation (4.4) in discrete-space, discrete-time form will be written as:

$$\Delta\mathbf{D}^{n+1/2}(i,j,k) = \Delta\mathbf{D}^{n-1/2}(i,j,k) \tag{4.15}$$

[1]The induced charge on a PEC is in reality a surface charge. However, due to the discretization in FDTD, it will appear as a volume charge, dispersed inside the cells comprising the conductor. An equivalent surface charge can always be obtained by multiplying $\rho$ by the size of the cell in the direction normal to the surface of the conductor.

It is crucial for the rest of the development to realize that (4.15) is always satisfied after $E_x$, $E_y$ and $E_z$ have been updated from the usual FDTD equations, since these are naturally divergence free (this can also be shown by direct substitution of the regular FDTD **E** field update equations in (4.14)). In the absence of conductors or sources, $\Delta\mathbf{D}^{n+1/2}(i,j,k) = 0$, which is nothing else but the discretized version of (4.6).

Now, to fix ideas, assume that there is a conductor segment at cell $(i_0, j_0, k_0)$, extending onto the $xz$-plane. That means that $E_x(i_0, j_0, k_0)$ and $E_z(i_0, j_0, k_0)$ are always zero. We will examine in detail the effect of setting $E_x(i_0, j_0, k_0)$ to zero, the results for $E_z(i_0, j_0, k_0) = 0$ follow by a simple rotation of coordinates. In regular FDTD, $E_x(i_0, j_0, k_0)$ is first updated with the usual update equation, as for all the other $(i, j, k)$, and later set to zero:

**step 1:** Usual FDTD update:

$$\begin{aligned} E_x^{n+1/2}(i_0,j_0,k_0) = E_x^{n-1/2}(i_0,j_0,k_0) \\ + \frac{\Delta t}{\Delta y \epsilon_{i_0,j_0,k_0}} \left[H_z^n(i_0,j_0,k_0) - H_z^n(i_0,j_0-1,k_0)\right] \\ - \frac{\Delta t}{\Delta z \epsilon_{i_0,j_0,k_0}} \left[H_y^n(i_0,j_0,k_0) - H_y^n(i_0,j_0,k_0-1)\right] \end{aligned} \quad (4.16)$$

Note here that, from the previous time update $E_x^{n-1/2}(i_0, j_0, k_0) = 0$, and therefore $E_x^{n+1/2}(i_0, j_0, k_0)$ depends only on the magnetic field. Call $E_{0x}^{n+1/2}(i_0, j_0, k_0)$, this temporary value of $E_x^{n+1/2}(i_0, j_0, k_0)$.

**step 2:** Later in the code: $E_x^{n+1/2}(i_0, j_0, k_0) = 0$.

Now, notice that after step 1 above, but before step 2, $\Delta\mathbf{D}^{n+1/2}(i, j, k)$ satisfies equation (4.15) for all $(i, j, k)$, since only the regular FDTD equations have been involved before step 2. After step 2, $\Delta\mathbf{D}^{n+1/2}(i, j, k)$ will be affected at cells

$(i_0, j_0, k_0)$ and $(i_0 + 1, j_0, k_0)$ (i.e. at the two ends of vector $E_x^{n+1/2}(i_0, j_0, k_0)$), as can be seen by equation (4.14). Specifically, bringing to the right hand side of (4.15) the term to be zeroed in step 2, we can write:

$$\Delta\mathbf{D}^{n+1/2}(i_0, j_0, k_0) = \Delta\mathbf{D}^{n-1/2}(i_0, j_0, k_0) - \frac{\epsilon_{i_0,j_0,k_0}}{\Delta x} E_{0x}^{n+1/2}(i_0, j_0, k_0) \quad (4.17)$$

$$\Delta\mathbf{D}^{n+1/2}(i_0 + 1, j_0, k_0) = \Delta\mathbf{D}^{n-1/2}(i_0 + 1, j_0, k_0) + \frac{\epsilon_{i_0,j_0,k_0}}{\Delta x} E_{0x}^{n+1/2}(i_0, j_0, k_0) \quad (4.18)$$

Equations (4.17), (4.18) provide recursive relations to calculate $\Delta\mathbf{D}^{n+1/2}(i_0, j_0, k_0)$ and $\Delta\mathbf{D}^{n+1/2}(i_0 + 1, j_0, k_0)$, beginning at $n = 0$ with $\Delta\mathbf{D}^{-1/2}(i_0, j_0, k_0) = \Delta\mathbf{D}^{-1/2}(i_0 + 1, j_0, k_0) = 0$ and $E_{0x}^{1/2} = 0$. The obvious solutions are:

$$\Delta\mathbf{D}^{n+1/2}(i_0, j_0, k_0) = -\frac{\epsilon_{i_0,j_0,k_0}}{\Delta x} \sum_{m=1}^{n} E_{0x}^{m+1/2}(i_0, j_0, k_0) \quad (4.19)$$

$$\Delta\mathbf{D}^{n+1/2}(i_0 + 1, j_0, k_0) = \frac{\epsilon_{i_0,j_0,k_0}}{\Delta x} \sum_{m=1}^{n} E_{0x}^{m+1/2}(i_0, j_0, k_0) \quad (4.20)$$

Since $\Delta\mathbf{D}^{n+1/2}(i, j, k)$ is just the discrete-time, discrete-space approximation to $\nabla\cdot\mathbf{D}$, it is clear that the imposition of $E_x(i_0, j_0, k_0) = 0$ created induced charges at cells $(i_0, j_0, k_0)$ and $(i_0 + 1, j_0, k_0)$, which are given by the right hand side of equations (4.19), (4.20). The induced charges resulting from $E_z(i_0, j_0, k_0) = 0$ can be obtained by (4.19), (4.20) by interchanging the $x$ for the $z$ coordinate, and $\Delta\mathbf{D}^{n+1/2}(i, j, k)$ due only to $E_z(i_0, j_0, k_0) = 0$ will not be zero anymore for cells $(i_0, j_0, k_0)$, $(i_0, j_0, k_0 + 1)$. At these cells the following equations hold:

$$\Delta\mathbf{D}^{n+1/2}(i_0, j_0, k_0) = -\frac{\epsilon_{i_0,j_0,k_0}}{\Delta z} \sum_{m=1}^{n} E_{0z}^{m+1/2}(i_0, j_0, k_0) \quad (4.21)$$

$$\Delta\mathbf{D}^{n+1/2}(i_0, j_0, k_0 + 1) = \frac{\epsilon_{i_0,j_0,k_0}}{\Delta z} \sum_{m=1}^{n} E_{0z}^{m+1/2}(i_0, j_0, k_0) \quad (4.22)$$

It is obvious now that the charge at cell $(i_0, j_0, k_0)$ due to both $E_x$ and $E_z$ being zero, will actually be the sum of the right hand sides of equations (4.19) and (4.21).

Now, we can introduce a new variable, $\rho(i,j,k)$, for the PEC cells only, that represents the induced charge of that cell. In the example analyzed above, equation (4.19) implies that the charge induced to cell $(i_0, j_0, k_0)$, due to the imposition of $E_x^{n+1/2}(i_0, j_0, k_0) = 0$, will be updated based on the equation:

$$\rho^{n+1/2}(i_0, j_0, k_0) = \rho^{n-1/2}(i_0, j_0, k_0) - \frac{\epsilon_{i_0,j_0,k_0}}{\Delta x} E_{0x}^{n+1/2}(i_0, j_0, k_0) \tag{4.23}$$

while, from (4.20), the charge induced to cell $(i_0 + 1, j_0, k_0)$ and due to $E_x^{n+1/2}(i_0, j_0, k_0) = 0$ will be updated by:

$$\rho^{n+1/2}(i_0 + 1, j_0, k_0) = \rho^{n-1/2}(i_0 + 1, j_0, k_0) + \frac{\epsilon_{i_0,j_0,k_0}}{\Delta x} E_{0x}^{n+1/2}(i_0, j_0, k_0) \tag{4.24}$$

Similar equations can be written for the charges induced by $E_z^{n+1/2}(i_0, j_0, k_0) = 0$. Two remarks are in order here: first, observe that $E_{0x}^{n+1/2}(i_0, j_0, k_0)$ is nothing else but $E_x^{n+1/2}(i_0, j_0, k_0)$ as calculated in step 1 above (we just called it a different name, so that it does not get confused with the zero value of $E_x^{n+1/2}(i_0, j_0, k_0)$ at step 2). This quantity is calculated within the regular FDTD algorithm and therefore it is already available to us; second, equations (4.23)-(4.24) do not give the *total* induced charge in cells $(i_0, j_0, k_0)$ and $(i_0 + 1, j_0, k_0)$, but only the part due to $E_x^{n+1/2}(i_0, j_0, k_0) = 0$. If the PEC extends to cell $(i_0 - 1, j_0, k_0)$ for example, there will be another contribution to cell $(i_0, j_0, k_0)$ from $E_x^{n+1/2}(i_0 - 1, j_0, k_0) = 0$, as implied by equation (4.20). Equation (4.23) then will become:

$$\begin{aligned}\rho^{n+1/2}(i_0, j_0, k_0) &= \rho^{n-1/2}(i_0, j_0, k_0) + \\ &+ \frac{1}{\Delta x}\left[\epsilon_{i_0-1,j_0,k_0} E_{0x}^{n+1/2}(i_0 - 1, j_0, k_0) - \epsilon_{i_0,j_0,k_0} E_{0x}^{n+1/2}(i_0, j_0, k_0)\right]\end{aligned} \tag{4.25}$$

As a concrete example of the charge calculation, consider a conductor plate on the $xz$-plane, at $j = j_c$ and extending for $i_0 \le i \le i_1$, $k_0 \le k \le k_1$. After having updated $E_x^{n+1/2}(i,j,k)$ and $E_y^{n+1/2}(i,j,k)$ everywhere, we update the charge and then set the tangential fields on the conductor to zero, as follows:

1. Charges due to $E_x$ and setting $E_x^{n+1/2}$ to zero:

for $k = k_0 : k_1$,

$$\rho^{n+1/2}(i_0, j_c, k) = \rho^{n-1/2}(i_0, j_c, k) - \frac{\epsilon_{i_0, j_c, k}}{\Delta x} E_x^{n+1/2}(i_0, j_c, k)$$

$$\rho^{n+1/2}(i_1, j_c, k) = \rho^{n-1/2}(i_1, j_c, k) + \frac{\epsilon_{i_1-1, j_c, k}}{\Delta x} E_x^{n+1/2}(i_1 - 1, j_c, k)$$

for $i = i_0 + 1 : i_1 - 1$,

$$\rho^{n+1/2}(i, j_c, k) = \rho^{n-1/2}(i, j_c, k) + \frac{1}{\Delta x} \Big[ \epsilon_{i-1, j_c, k} E_x^{n+1/2}(i - 1, j_c, k) - \epsilon_{i, j_c, k} E_x^{n+1/2}(i, j_c, k) \Big]$$

$$E_x^{n+1/2}(i - 1, j_c, k) = 0$$

end $i$

$$E_x^{n+1/2}(i_1 - 1, j_c, k) = 0$$

end $k$

2. Charges due to $E_z$ and setting $E_z^{n+1/2}$ to zero:

for $i = i_0 : i_1$,

$$\rho^{n+1/2}(i, j_c, k_0) = \rho^{n-1/2}(i, j_c, k_0) - \frac{\epsilon_{i, j_c, k_0}}{\Delta z} E_z^{n+1/2}(i, j_c, k_0)$$

$$\rho^{n+1/2}(i, j_c, k_1) = \rho^{n-1/2}(i, j_c, k_1) + \frac{\epsilon_{i, j_c, k_1-1}}{\Delta z} E_x^{n+1/2}(i, j_c, k_1 - 1)$$

for $k = k_0 + 1 : k_1 - 1$,

$$\rho^{n+1/2}(i, j_c, k) = \rho^{n-1/2}(i, j_c, k) + \frac{1}{\Delta z} \Big[ \epsilon_{i, j_c, k-1} E_z^{n+1/2}(i, j_c, k - 1) - \epsilon_{i, j_c, k} E_z^{n+1/2}(i, j_c, k) \Big]$$

$$E_z^{n+1/2}(i, j_c, k - 1) = 0$$

end $k$

$$E_z^{n+1/2}(i, j_c, k_1 - 1) = 0$$

end $i$

Note from this example, that the charges on the edges of the conductor are updated differently from the charges inside the conductor, due to the fact that the field components normal to the edges, but external to the conductor, are not

zero. Based on this example, the generalization for PECs of any orientation is straightforward.

Having found the charge in an independent way as explained above, we can now use it to extend the reduced FDTD algorithm to include electric conductors. Of course, we need to introduce a new variable, the charge, for the conductor cells. The number of conductor cells is usually very small with respect to the total number of cells in the whole domain and therefore there is practically a negligible increase in memory requirements. The charge is updated using the appropriate equation between (4.23)-(4.25) and along the lines of the example described above. One problem arises when a conductor extends in the direction of that electric field component that is spatially updated. For example, in the case we update spatially $E_y$ and $H_y$, and a conductor lies along the $y$-direction, we cannot use equations like (4.23)-(4.25), since we do not have $E_{0y}^{n+1/2}(i_0, j_0, k_0)$ (which is $E_y^{n+1/2}(i_0, j_0, k_0)$ before being set to zero). But, as we have already mentioned, since $E_y^{n-1/2}(i_0, j_0, k_0) = 0$, $E_y^{n+1/2}(i_0, j_0, k_0)$ will only depend on the magnetic field, and the regular FDTD update would be:

$$E_y^{n+1/2}(i_0, j_0, k_0) = \frac{\Delta t}{\Delta z \epsilon_{i_0, j_0, k_0}} \left[H_x^n(i_0, j_0, k_0) - H_x^n(i_0, j_0, k_0 - 1)\right] - \frac{\Delta t}{\Delta x \epsilon_{i_0, j_0, k_0}} \left[H_z^n(i_0, j_0, k_0) - H_z^n(i_0 - 1, j_0, k_0)\right] \quad (4.26)$$

Since we have $H_x^n$ and $H_z^n$ everywhere, we can use the right hand side of equation (4.26) in the place of $E_{0y}^{n+1/2}(i_0, j_0, k_0)$, to update the charge for conductors that extend in the $y$-direction.

The charge $\rho$ found in the way exhibited above is simply added to the left hand side of equation (4.12) for the conductor cells. Observe also that we do not have to set $E_y$ explicitly to zero whenever it is tangential to a PEC; it will be forced to zero (i.e. to a very small numerical value) by equation (4.12) with the

addition of $\rho$.

As a final remark, we would like to note that equations (4.23)-(4.25) are nothing else but discrete forms of the differential equation of the conservation of charge, namely:

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot \mathbf{J} \tag{4.27}$$

where $\mathbf{J}$ is the induced current[2]. Indeed, consider Ampère's law:

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J} \tag{4.28}$$

Across a conductor, the tangential electric field components are zero and therefore:

$$\mathbf{J} = (\nabla \times \mathbf{H})_t \tag{4.29}$$

where subscript $t$ refers to the tangential components. Therefore, equation (4.27) becomes:

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\nabla \times \mathbf{H})_t = \nabla \cdot (\nabla \times \mathbf{H})_n \tag{4.30}$$

where $n$ denote the normal to the conductor surface component of $\nabla \times \mathbf{H}$. Equations (4.23)-(4.25) can be derived anew from discretization of (4.30). We preferred to show how they can be obtained directly from FDTD in order to demonstrate their actual application and to avoid technical difficulties with the edges of the conductors.

### 4.3.1 An Alternative Conductor Treatment

In this section we outline an alternative conductor treatment formulation, which can be adopted whenever FDTD is formulated for uniaxial dielectric media. This

[2]Again, here the induced current on the conductor appears as a volume current. For the case of a conductor on the $xz$-plane, an equivalent surface current can be obtained by multiplying $\mathbf{J}$ by $\Delta y$.

is the case, for example, when high dielectric constant materials are used and an average dielectric constant must be taken at the interfaces, in order to maintain accuracy of results [43]. For uniaxial dielectric materials, the electric displacement can be written as:

$$\mathbf{D} = \begin{pmatrix} \epsilon_x & 0 & 0 \\ 0 & \epsilon_y & 0 \\ 0 & 0 & \epsilon_z \end{pmatrix} \mathbf{E} \tag{4.31}$$

so that each electric field component sees a different dielectric constant. This allows us to treat conductors as infinite dielectric constant materials, by setting to infinity those dielectric constant components that are tangential to conductors. This results in the annulment of the corresponding tangential electric field components. This second approach has an advantage over the first one, since metals are really treated as dielectrics, so no particular formulation is needed as before. Of course, by doing so, extra memory is used to store the three dielectric values instead of the usual one. On the other hand, for uniaxial dielectric media these three components for the dielectric constant are needed, so no extra burden is added.

For practical implementation in the computer code, we identified that relative dielectric constant of the order of $10^8$ gives identical results as treating the metal with the first approach, over all the frequency ranges we examined. Larger values of relative dielectric constant do not improve the precision of results, while they can lead to numerical instability.

## 4.4 Source Treatment

The standard way to excite the electromagnetic field in FDTD is to impose a time varying field distribution in a certain space region. As can be expected, **D** does

not maintain its divergence free property, though as we shall see, in the case of homogeneous media, this happens only for these cells where the imposed electric flux is terminated. In any case, we will be able to compute the appropriate correction term in advance (i.e. before the start of the time iteration) and use it in $\nabla \cdot \mathbf{D}$. The analysis follows the same lines as for the conductor treatment, i.e. finding a recursive relation for $\Delta \mathbf{D}^{n+1/2}$ (which is again the discrete-time, discrete-space approximation to $\nabla \cdot \mathbf{D}$) and solving it to obtain the equivalent charge induced by the source. To fix ideas, consider the case where a soft electric field source along $x$ is imposed on the plane $k = k_0$, for $i_0 \leq i \leq i_1$, $j_0 \leq j \leq j_1$:

$$E_x^{n+1/2}(i,j,k_0) = E_x^{n+1/2}(i,j,k_0) + \exp\left(-\frac{(n-n_0)\Delta t}{T}\right)^2 \qquad \begin{matrix} i_0 \leq i \leq i_1 \\ j_0 \leq j \leq j_1 \end{matrix} \tag{4.32}$$

Note here that the imposed field flux ends at $i = i_1 + 1$, see also Figure 4.5.

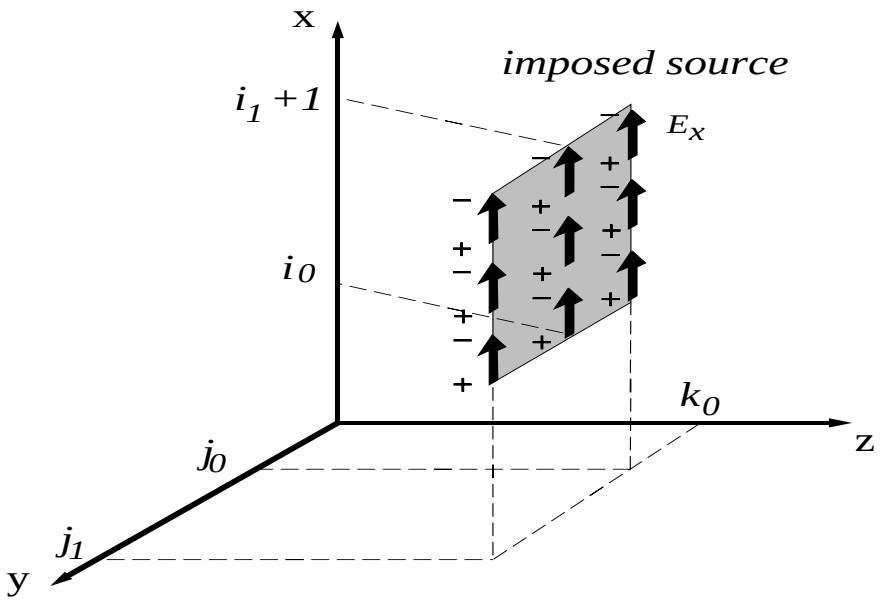



Figure 4.5: An imposed source field on a portion of the $xy-$plane.

We need now to determine the equation satisfied by $\Delta \mathbf{D}^{n+1/2}$ in the source

region, so that we can use it in our algorithm for the spatial update of one electric field component. Note that, just before the application of the source in the code, $\Delta\mathbf{D}^{n+1/2}$ satisfies as usual equation (4.15). It is straightforward to show that after imposing the soft source according to (4.32), the following equations hold:

1. For $i_0 < i \leq i_1$, $j_0 \leq j \leq j_1$:

$$\Delta\mathbf{D}^{n+1/2}(i,j,k_0) = \Delta\mathbf{D}^{n-1/2}(i,j,k_0) + \\ + \frac{\epsilon_{i,j,k_0} - \epsilon_{i-1,j,k_0}}{\Delta x} \exp\left(-\frac{(n-n_0)\Delta t}{T}\right)^2 \quad (4.33)$$

2. For $i = i_0$, $j_0 \leq j \leq j_1$:

$$\Delta\mathbf{D}^{n+1/2}(i,j,k_0) = \Delta\mathbf{D}^{n-1/2}(i,j,k_0) + \frac{\epsilon_{i_0,j,k_0}}{\Delta x} \exp\left(-\frac{(n-n_0)\Delta t}{T}\right)^2 \quad (4.34)$$

since there is no imposed field in $E_x^{n+1/2}(i_0 - 1, j, k_0)$.

3. For $i = i_1 + 1$, $j_0 \leq j \leq j_1$:

$$\Delta\mathbf{D}^{n+1/2}(i,j,k_0) = \Delta\mathbf{D}^{n-1/2}(i,j,k_0) - \frac{\epsilon_{i_1,j,k_0}}{\Delta x} \exp\left(-\frac{(n-n_0)\Delta t}{T}\right)^2 \quad (4.35)$$

since there is no imposed field in $E_x^{n+1/2}(i_1 + 1, j, k_0)$.

One can note here that in cases of homogeneous media, where $\epsilon$ is constant, equation (4.33) implies that $\Delta\mathbf{D}^{n+1/2}(i,j,k_0) = 0$ for $i_0 < i \leq i_1$, $j_0 \leq j \leq j_1$. That is, $\Delta\mathbf{D}^{n+1/2}(i,j,k_0)$ is non zero only where the imposed electric flux is terminated, that is for $i = i_0$ and $i = i_1 + 1$, as shown by equations (4.34), (4.35) (see also Figure 4.5). In the general case, we can use the recursive equations (4.33)-(4.35) with the initial condition $\Delta\mathbf{D}^{-1/2} = 0$ to write the following equation

for $\Delta\mathbf{D}^{n+1/2}$, which holds for all $j_0 \leq j \leq j_1$, $k = k_0$ and the $i$ specified:

$$\Delta\mathbf{D}^{n+1/2}(i,j,k_0) = \begin{cases} \frac{\epsilon_{i,j,k_0} - \epsilon_{i-1,j,k_0}}{\Delta x} \sum_{m=1}^{n} \exp\left(-\frac{(m-n_0)\Delta t}{T}\right)^2 & \text{if } i_0 < i \leq i_1, \\ \frac{\epsilon_{i_0,j,k_0}}{\Delta x} \sum_{m=1}^{n} \exp\left(-\frac{(m-n_0)\Delta t}{T}\right)^2 & \text{if } i = i_0, \\ -\frac{\epsilon_{i_1,j,k_0}}{\Delta x} \sum_{m=1}^{n} \exp\left(-\frac{(m-n_0)\Delta t}{T}\right)^2 & \text{if } i = i_1 + 1 \end{cases} \tag{4.36}$$

The terms on the right hand side of (4.36) can be calculated for each $n$ before the beginning of the time update of the fields, and then used at the appropriate time iteration. In the most common case where the source is imposed in a homogeneous medium, the right hand side of equation (4.36) simplifies greatly and one needs to store only one of the two lower terms for each $n$, since the first term is zero and the other two are just opposite. A vector of length equal to the number of time iterations is needed. Then, at time update $n$, the $n$-th term of this vector gets added or subtracted accordingly, to the cells defined by equation (4.36).

Finally, we would like to note that, in the case of homogeneous media, the fact that the divergence of $\mathbf{D}$ is non-zero only wherever the imposed field flux is terminated, can be explained by equivalent charges that are induced at these points, as can be seen in Figure 4.5: in the intermediate cells of the source region, the charges cancel each other, but not so at the termination of the imposed $E_x$ field. The charges induced by the source keep accumulating according to equation (4.36).

## 4.5 Numerical Results

As a validation of this new formulation, two numerical results have been generated for different types of structures and they are compared with FDTD and/or previously published results. In both cases, perfect agreement is obtained with

FDTD, confirming the exact equivalence of R-FDTD with FDTD.

In the first example, a microstrip patch antenna printed on duroid ($\epsilon_r = 2.2$) dielectric substrate is analyzed with this new technique. In this formulation the $E_y$ and $H_y$ field components have been eliminated. The dimensions of the microstrip patch antenna are shown in Figure 4.6, the total mesh dimensions are $60 \times 100 \times 16$ cells, giving a grid size of $\Delta x = 0.389$ mm, $\Delta y = 0.4$ mm, $\Delta z = 0.265$ mm. The excitation is obtained through a Gausssian pulse under the strip conductor having width T=15 ps, corresponding to a maximum frequency of 33 GHz. A standard procedure is used to extract the S parameters [43]. The reflection coefficient obtained is shown in Figure 4.7 for both regular FDTD and our approach (the two curves overlap completely) and is in perfect agreement with the one presented in [18]. First order MUR absorbing conditions were used for this first example. The implementation of absorbing conditions in the reduced formulation is done by calculating all the three field components at the absorbing interfaces.

In the second example, a microstrip coupled dielectric resonator is analyzed with R-FDTD and FDTD and results are compared with measured data. The measurement was performed on a cylindrical resonator of diameter of 2.26 mm and height of 0.91 mm, made of a perovskite based on Ba,Zn, Ta-oxide (see Figure 4.8). The dielectric resonator (manufactured by Trans-Tech, model D8733-0089-036) has a relative dielectric permittivity of 30.15, and a Q factor at 10 GHz of 12200. The resonator is placed on top of a dielectric substrate, the substrate is made of 124 $\mu$m-thick alumina (relative dielectric constant $\epsilon_r = 9.8$) with a ground metal (gold plated) on the other side as shown in Figure 4.8. A 9 mm long offset conductor of 124 $\mu$m width is used to couple the field to the resonator. The whole substrate was enclosed on a rectangular waveguide whose dimensions

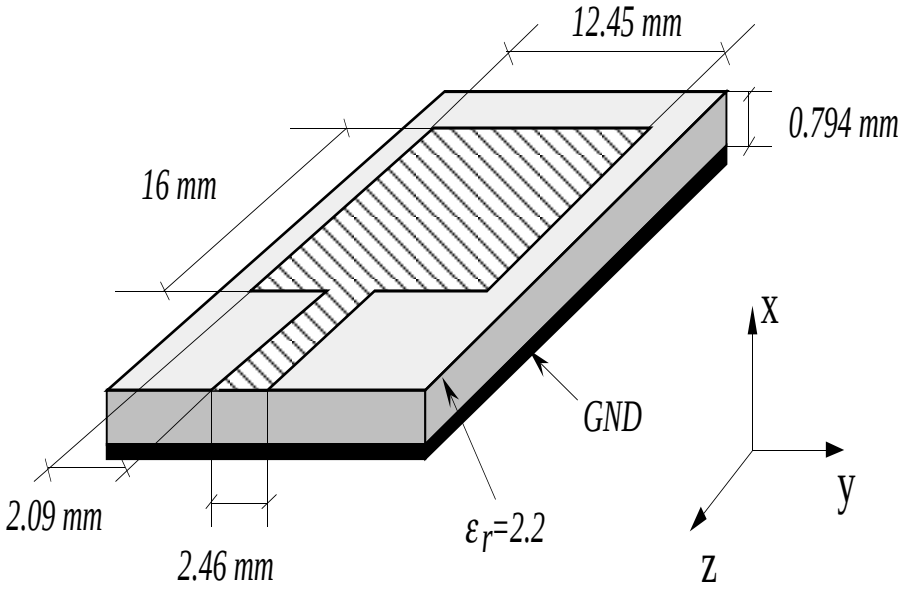



Figure 4.6: Schematic of the rectangular patch antenna analyzed.

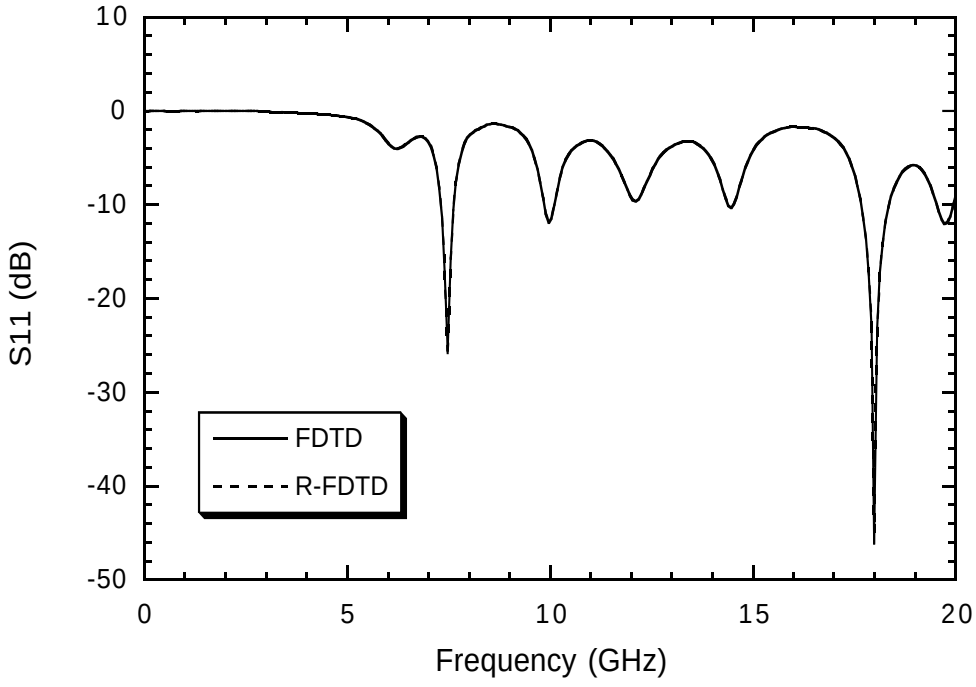



Figure 4.7: The magnitude of the simulated reflection coefficient ($|s_{11}|$) for the microstrip patch antenna shown in Figure 4.6.

are $2.5 \times 4.6 \times 7.5$ mm, to reduce radiation losses (see Figure 4.8). The resonator was placed directly on the substrate (no spacer was used) with one edge exactly at the edge of the microstrip line (no overlapping). The measurement of the resonance frequency was done using a network analyzer (HP 8510C) with coaxial probes (Picoprobe Model 40 A). Wrap around type ground transitions for the probes were done with silver epoxy. By using the TRL [42] calibration technique, 25 dB return loss sensitivity was achieved over the band going from 20-40 GHz. The total length of the microstrip from the two reference planes was 7.5 mm. The measured transmission coefficient for the proposed structure is reported in Figure 4.9. The desired resonance frequency is obtained at 29.862 GHz, also a box resonance is observed at 26.075 GHz and two higher resonances respectively at 36.875 GHz and 37.963 GHz.

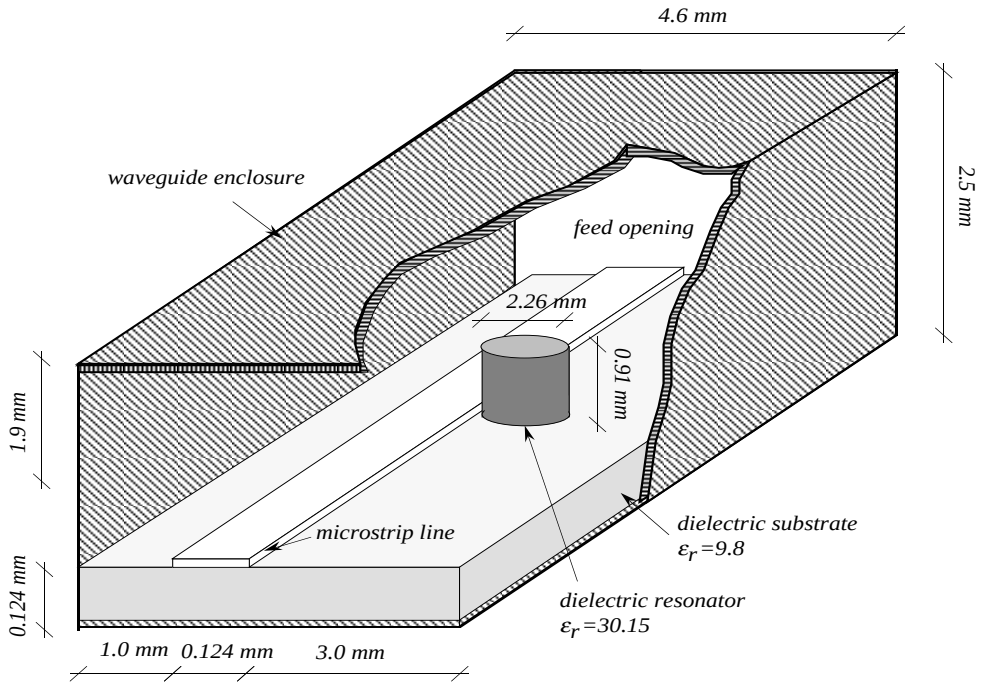



Figure 4.8: Geometry of the microstrip coupled dielectric resonator analyzed.

To simulate the described structure the whole domain was discretized with

uniform grid having dimensions of $100 \times 60 \times 166$ respectively along $x, y$, and $z$ directions. The corresponding grid size was $\Delta x = 45.2$ $\mu$m, $\Delta y = 41.33$ $\mu$m and $\Delta z = 45.2$ $\mu$m, this choice allows to best fit the resonator height and the substrate dimensions with the grid. The excitation is obtained through a Gaussian pulse under the strip conductor having width T=10 ps, corresponding to a maximum frequency of 50 GHz. Reflection conditions are used on the metal side walls of the waveguide, and perfectly matched absorbing boundary conditions (PML) [23] are used at the two waveguide ends. As previously, all field components are calculated and stored inside the PML. Standard procedure is used to extract the transmission parameters [43]. The transmission coefficient obtained is shown in Figure 4.9 for normal FDTD, our formulation (R-FDTD), and is compared with measured data. The two calculated curves overlap completely, while some difference is observed with the measured data for the highest resonance frequencies.

## 4.6 Summary

We have presented a new formulation of FDTD that reduces computer memory requirements by 33%, by eliminating global storage of one electric and one magnetic field component. Using the fact that in charge free, magnetically homogeneous media both **D** and **H** are solenoidal, we linked spatially one **E** and one **H** component to the other two. This way, the spatially updated components need to be stored only locally. The formulation follows closely that of FDTD, as can be seen from the pseudo-codes presented for the two and three-dimensional cases, and therefore retains most of its simplicity.

Conductors are treated by calculating the charge induced to them by the electromagnetic field, and subsequently using it in the divergence of **D**. The in-

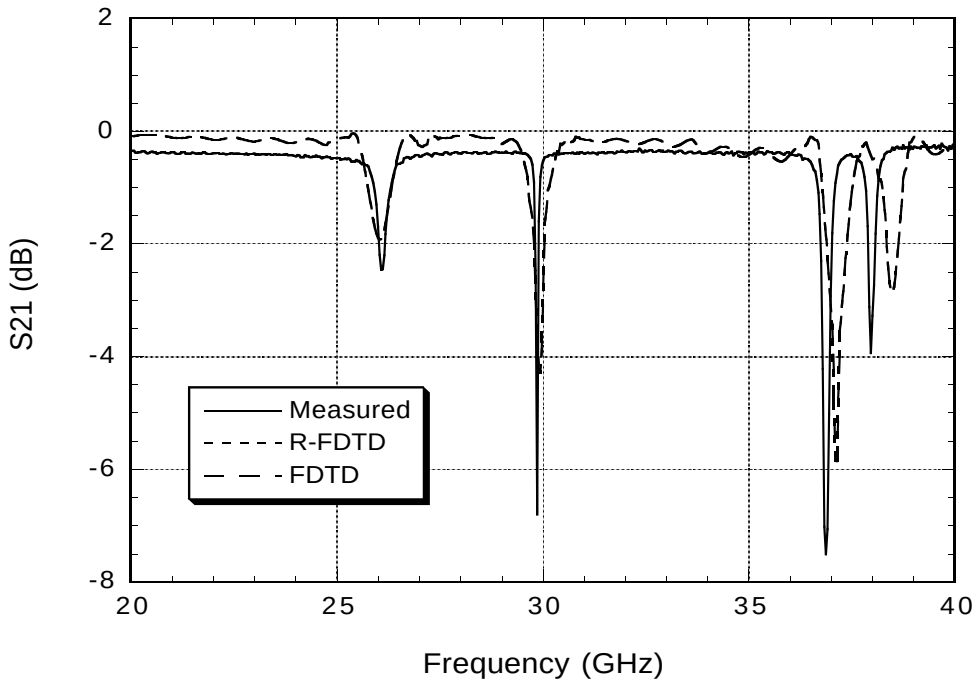



Figure 4.9: Comparison of measured and simulated data for the structure illustrated in Figure 4.8.

duced charge is essentially calculated through the continuity equation, although we derived the update equations directly from FDTD, since this way their implementation is more clear. A treatment very similar to the one used for conductors yields also the equivalent charge induced by the source, and therefore we were able to extend our method in source regions as well. A simpler formulation for the treatment of conductors can be used when three dielectric constant values are stored at each cell location, as in the case of uniaxial dielectric material. With these extensions, for conductors and sources, the new formulation becomes general and can be used in the place of regular FDTD, whenever memory savings are critical. In terms of numerical results, it has proven practically equivalent to FDTD for all the cases we examined.

It can be argued that, due to the existence of the three-dimensional matrix $\epsilon_{i,j,k}$, the actual memory reduction in our formulation is not 33% but rather 28.6%. In applications where memory is the limiting factor though, it does not make much sense to use a matrix like that, where very large sub-blocks (corresponding to different dielectric bodies) will contain the same dielectric constant. With an increase in coding complexity, FDTD can be formulated to do without $\epsilon_{i,j,k}$. One simple method is to update the fields inside individual dielectric bodies first and then on the interfaces. Our method extends rather easily to such an update scheme also.

A 33% reduction in memory translates to a 50% increase in the computational domain volume, under the same memory. If all three dimensions are increased by the same amount, this leads to a 14.5% increase in each dimension. For two and one dimensions the increase would be 22.5% and 50% respectively. On the other hand, the new formulation is computationally more demanding than FDTD: as it can be checked by the algorithms, FDTD requires 24 additions

and 18 multiplications per cell, while R-FDTD requires 24 additions and 24 multiplications. This increase in multiplications comes from the spatial update of the **E** field component through the divergence of **D**.

We would also like to stress here the fact that for all the electromagnetic problems which do not involve metals, like dielectric scatterers and dielectric waveguides, the application of this novel proposed technique has almost no increase of complexity, compared to the standard FDTD formulation, since charges do not need to be calculated on conductors.

In view of the memory savings described above, but also the disadvantages (more computations and a slight increase in coding complexity), the new method we propose should find applications in large electromagnetic problems where memory is the liming factor.

# CHAPTER 5

# Packet Switched Systems for Wireless Personal Communications

## 5.1 Introduction

Telecommunications de-regulation and the very fast pace of development of advanced digital communication technology will bring many new services to residential and business customers. For example, cable television companies have moved towards distribution of Internet services via high speed modems, which have the capability of providing two-way high bandwidth data transmission. Telephone companies are currently deploying digital subscriber lines, offering high data rates, either in symmetrical or asymmetrical fashion (up to 6 Mbits/s for asymmetrical digital subscriber lines), over ordinary telephone lines [49]. Another competing solution is wireless provisioning of high bandwidth services, through local multipoint distribution services (LMDS/MMDS), where a network of base stations (BS) will provide coverage to a number of business and home buildings [50].

Distribution of this large bandwidth inside buildings will most probably be tetherless. The advantages of wireless over wired distribution are obvious, reduced installation cost, easier reconfiguration of the network, anywhere access, etc. Such wireless products for local area networks have already appeared, or

are about to appear in the market and are based on very recent standardization efforts, [48], [47]. However, microwave spectrum is expensive, so efficient strategies for reusing frequencies and managing co-channel interference are critically important. To support a user data rate of 10 Mb/s in an interference limited wireless environment, a bandwidth of several megahertz is needed for time division multiple access (TDMA). In contrast to narrowband cellular networks where radio spectrum is divided into multiple channel sets, which are reused only in relatively distant cells, broadband wireless networks must reuse frequency very aggressively, ideally reusing the same frequency in every cell. This is especially true in indoor environments, where frequency planning is inherently difficult [6]. Co-channel interference must be dealt with and its statistics must be very well understood if this goal is to be achieved. On the other hand, medium access control (MAC) protocols that are designed taking co-channel interference into account have been little researched. It seems though imperative that multiple access protocols that adapt to the interference statistics are needed for the next generation wireless services.

Access to the radio interface is a key issue in a the efficient design of wireless systems. Currently implemented and emerging systems provide many users with simultaneous access to the same wireless medium using frequency division, time division and code division. Even if fixed channel assignment is dominant in the present implementations, packet oriented techniques are getting increasing attention for the next generation wireless systems.

The intrinsic advantages of packet multiplexing are those well known also for wired networks. Traffic is usually bursty in nature and bandwidth efficiency can be increased by multiplexing sources, that is using the same channel during silence periods. Packet multiplexing is also attractive for integrating different

services that have different delay and bandwidth requirements. Additionally, packet access techniques seem appropriate for aggressive reuse of bandwidth with little or no frequency planning.

For the reasons outlined above, current research in packet switched protocols is intense. A significant such family of access protocols evolved from the initial proposal of the packet reservation multiple access (PRMA) protocol [51], [52]. They follow the same general trends, with some notable exceptions which we will describe later:

- Available slots are assigned in a random fashion among active terminals and this assignment usually stays constant for the duration of a burst. As a result, independent of the contention mechanism, during reservation periods the system operates like a fixed TDMA scheme, among the mobile terminals (MT) that have a reservation. This approach comes to a large extent from the voice-only orientation of these protocols, where speech packets are generated at a fixed rate.

- Since the design is made for a single cell, power control is almost never included as a protocol feature.

- Re-transmission of packets is rarely considered [69]. The reason here is two fold: packets can only be dropped during the contention period (at least in a single cell environment) and therefore packet dropping probability is negligible for the loading region, where the protocols are stable; the applications considered are usually real-time (e.g. voice or video) and it is deemed too time consuming to re-transmit packets. However, we will later see that for indoor access the delays involved in packet transmission are small enough to allow re-transmissions.

As we have already mentioned, co-channel interference has not been taken into account for the design of most current packet switched protocols. Therefore, if we take the point of view of interference generation we can easily come to the following conclusions, with respect to the above described design philosophies:

- The behavior of a packet switched protocol can be dramatically different in a cellular environment from the one analyzed in a single cell case. Indeed, numerous simulative analyses [67], [57], [63] have shown that performance takes a big hit in cellular environments, to the point of making the single cell design irrelevant. The only way around this problem is the introduction of large reuse distances, which reduces bandwidth efficiency and also necessitates frequency planning, something not attractive for personal communication services.

- The constant assignment of reservation slots to active terminals offers no flexibility in case that a long hold interferer appears. Obviously, retransmission does not help either in this case and the only viable alternative is the reassignment of the terminal to a less interfered slot.

- Power adaptation can be a useful tool towards the reduction of co-channel interference but, as already mentioned, has hardly been studied for packet switched protocols (for two recent exceptions see [76] and [77]). Its application must necessarily be based on different criteria than the ones already known for circuit switched traffic [82], [81]. Sources now alternate from inactive to active states and interference will, in general change fast. Power control ideally should be applied on a packet by packet basis, but the rules are not clear. (For example, it should depend on the number of times a packet is hit by interference, or retransmission number, on the remaining lifetime of the packet, etc.).

- The integration of different services in one protocol presents new difficulties from an interference point of view. Different sources have widely different statistics and therefore will generate different interference dynamics. Consider the case of voice and video. Voice is generated in bursts and during a talk-spurt packets are transmitted at a constant rate. Mean silence and activity periods are 1.35 sec and 1.0 sec respectively. Video on the other hand has a constant minimum bandwidth requirement with random excursions to higher bandwidth. Mean hold times are much larger than that of voice. Interference generated by a video session can therefore be more stationary. Which service would suffer more from a random assignment of bandwidth? Note also here that probably video would benefit more from power control, since its interference can be more predictable, if video sources are segregated to their own bandwidth.

## 5.2 Some Packet Access Protocols

In the following section we will describe some packet switched medium access control protocols for short range wireless communications. We will use some of these protocols as our basis for designing interference-adaptive MACs in the next chapters.

### 5.2.1 Packet Reservation Multiple Access (PRMA)

The first and more simple contention based channel access protocol for local wireless networks is packet reservation multiple access (PRMA) [51], [52] which is an adaptation to the cellular environment of the reservation ALOHA protocol, conceived for satellite systems. PRMA uses a slotted channel structure with time

slots organized in frames, each containing a fixed number of slots. Usually, only voice traffic is considered. Voice packets are generated periodically with a period matching the frame period, that is there is one packet per frame from an active mobile. Mobile terminals classify each slot as either *reserved* or *available*, based on information broadcasted by the BS. A terminal that holds a reservation for a given slot has exclusive use of that slot in subsequent frames, until the end of the talk-spurt. At the end of a talk-spurt the slot becomes available again for terminals to contend for access.

To transmit a packet, a terminal must verify two conditions: the current time slot must be available, and the terminal has permission to transmit. Permission is granted according to a pseudo-random number generator, locally at each mobile and independently of other terminals. Each terminal attempts to transmit the initial packet of the talk-spurt until the BS acknowledges its successful reception, or until the packet is discarded by the terminal because it has been delayed too long, usually after two frame periods for voice packets. If a terminal drops the first packet of a burst, it continues to contend for a reservation to send subsequent packets. The packet dropping mechanism then is clear: packets get dropped in the beginning of a talk-spurt only, when they have exceeded their lifetime. It is important to keep in mind though that this is the case only for a single cell situation. In a cellular environment there will be another dropping mechanism, that of interfered packets which can appear anywhere in the talk-spurt. In this case, the percentage of interfered packets completely dominates the first dropping mechanism, which becomes irrelevant for system performance.

#### 5.2.1.1 Performance Analysis of PRMA (Single Cell Case)

We present below an analysis of PRMA based on the *equilibrium point theory* [79]. Equilibrium point theory facilitates the solution of multi-dimensional Markov chains, by assuming that the system modeled by the Markov chain is in equilibrium, that is the expected traffic offered to the system (e.g. in terms of packets per second) equals the expected traffic served by the system. We can equivalently say that at the equilibrium point, the expected number of terminals entering a state of the Markov chain equals the expected number of terminals leaving that state. Finally we note that the analysis presented below was first reported in [52]. It is repeated here because we will use some of its tools and results in the next chapter, where we will expand the method to apply it for a multi-cell, co-channel interference environment.

PRMA acts as a statistical multiplexer, multiplexing different MTs on the same bandwidth (TDMA slots). This is made possible by the intermittent nature of voice activity. Specifically, voice generation is modeled as a two state Markov process, where one state corresponds to a silence period and the other to a talking period [85]. This is shown in Figure 5.1, with the two states denoted by SIL and TLK respectively. If a silence period has a mean time $t_{\mathrm{SIL}}$ and a talking period has a mean time $t_{\mathrm{TLK}}$[1], then the probability that a talk-spurt ends in a time slot of duration $\tau$ is:

$$\gamma = 1 - \exp\left\{-\frac{\tau}{t_{\mathrm{TLK}}}\right\} \tag{5.1}$$

which is also the transition probability from state TLK to state SIL. Similarly, the probability that a silence period ends in a time slot of duration $\tau$ is:

$$\sigma = 1 - \exp\left\{-\frac{\tau}{t_{\mathrm{SIL}}}\right\} \tag{5.2}$$

[1]Typical values are $t_{\mathrm{SIL}} = 1.35$ s and $t_{\mathrm{TLK}} = 1.0$ s.

which is also the transition probability from state SIL to state TLK.

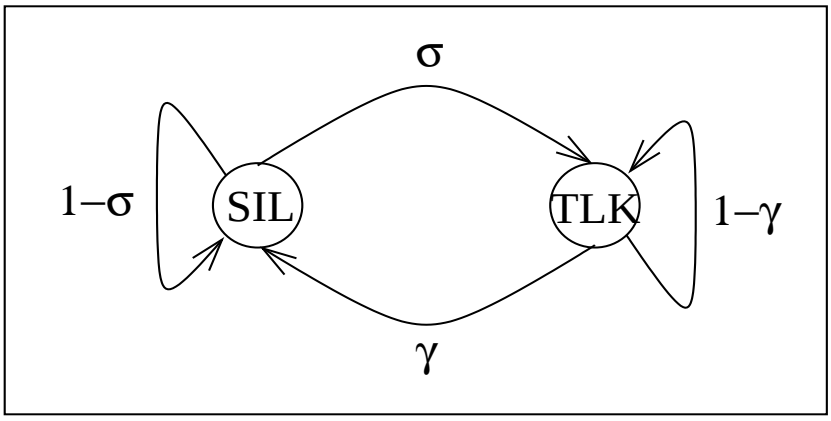



Figure 5.1: Voice activity Markov chain model.

Now, let us assume that a TDMA frame has $N$ slots, there are $M$ MTs in the cell, and each MT accesses the BS with a probability $p$, independently from other MTs[2]. We will approximate the single cell PRMA system as a Markov process, with states:

$$\{\text{SIL}, \text{CON}, \text{RES}_0, \text{RES}_1, \ldots, \text{RES}_{N-1}\} \tag{5.3}$$

as shown in Figure 5.2. These correspond to silence, contention, and holding a reservation for the $i$-th slot, respectively. The state variables now are:

$$\{S, C, R_0, R_1, \ldots, R_{N-1}\} \tag{5.4}$$

which correspond to the number of terminals in silence, contention and having a reservation for the $i$-th slot respectively. Note here that, since only one terminal can hold a reservation for a specific slot, $R_i$ is a binary random variable. Solving for the stationary probabilities of this Markov chain with $N + 2$ states is prohibitively complex, but instead one can resort to an *equilibrium point analysis*,

[2]Typical values for a PRMA system are: $N = 20$, $p = 0.3$, $M =$ 30-40, while a frame period is 16 ms.

[79]. The equilibrium point is defined as the values of the state variables, for which their expected change is zero. In terms of the PRMA system, the equilibrium point corresponds to such a distribution of terminals among the states, that the rate at which terminals arrive in a state exactly equals the rate at which terminals leave that state. We denote the equilibrium point as:

$$\{s, c, r_0, r_1, \ldots, r_{N-1}\} \tag{5.5}$$

Note here that, since $R_i$ is a binary random variable, $r_i$ corresponds to the probability that the $i$-th slot is reserved.

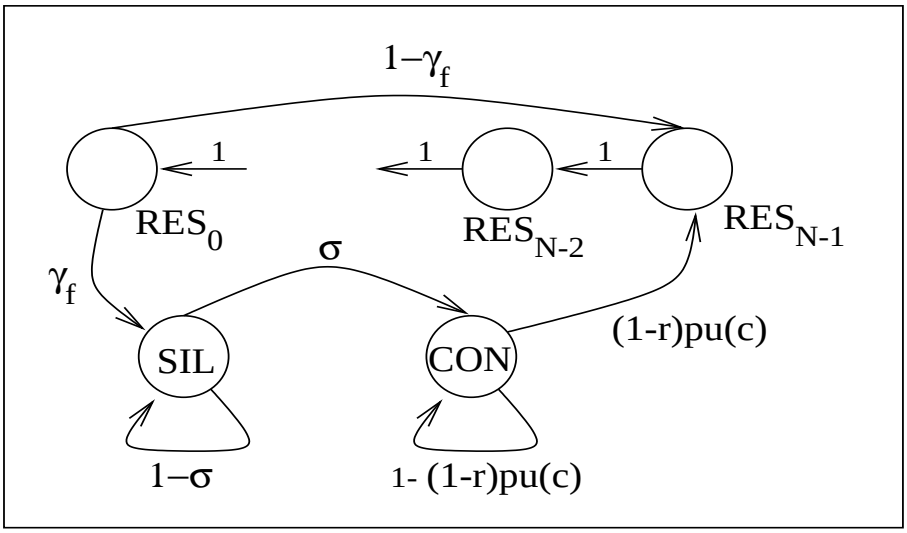



Figure 5.2: A Markov chain model for the PRMA (single cell) system.

The transition probabilities of the Markov model of Figure 5.2 can now be determined as follows[3]:

- A silent terminal enters contention, (i.e. SIL → CON) at the end of a silence period, which happens with probability $\sigma$ in any slot.

[3]Transitions on the Markov model of Figure 5.2 occur at the end of every slot.

- From contention, a terminal obtains a reservation for a slot (i.e. CON$\rightarrow$ RES$_{N-1}$) when all of the following conditions are met: the slot is empty, probability $(1-r)$; the terminal has the right to contend for the slot, probability $p$; all the other terminals in contention (i.e. $c-1$ terminals in equilibrium) do not have the right to contend for the same slot, probability $u(c)$, where:

$$u(c) = \begin{cases} (1-p)^{c-1}, & c \geq 1 \\ 1, & c < 1 \end{cases} \tag{5.6}$$

- From a reservation state, the terminal will deterministically go through all other $N-1$ reservation states (i.e. RES$_i$ $\rightarrow$ RES$_{i-1}$), since it hold that reservation for at least one frame.

- From a reservation state, the terminal returns to the silent state (i.e. RES$_0$ $\rightarrow$ SIL) if the talk-spurt ended during the most recent frame, probability:

$$\gamma_f = 1 - (1-\gamma)^N \approx N\gamma \tag{5.7}$$

These transition probabilities are all shown in Figure 5.2. We have to note here that we have neglected the probability that a talk-spurt ends before a terminal can obtain a reservation. That would have added a transition from state CON to SIL, but assuming a stable PRMA system most terminals will acquire a reservation before dropping all the packets in a talk-spurt.

In order to find the equilibrium point values $s$, $c$ and $r_i$, we will equate the in and out flows at each state of Figure 5.2. First, we note that, since all slots are equivalent, $r_i = r_j \equiv r$ for all $i$ and $j$, that is, the probability that a slot is reserved is the same for all $N$ slots. The equilibrium equation at state RES$_{N-1}$

gives us:

$$r(1-\gamma_f) + c(1-r)pu(c) = r \tag{5.8}$$

while the one at state SIL gives us:

$$r\gamma_f = s\sigma \tag{5.9}$$

Taking a third equilibrium equation for state CON would only give us a linear combination of (5.8) and (5.9), but instead we know that the total number of terminals in all states must equal $M$:

$$s + c + Nr = M \tag{5.10}$$

We now have a set of three equations for the three unknowns, and by eliminating $s$ and $r$, we obtain:

$$c + \left(N + \frac{\gamma_f}{\sigma}\right) \frac{cpu(c)}{cpu(c) + \gamma_f} = M \tag{5.11}$$

Equation (5.11) is an equation with a single unknown, the number of contending terminals at equilibrium, $c$. We can solve this equation for any number of terminals $M$, slots per frame $N$ and permission probability $p$. In Figure 5.3 we present a graphical solution, where we plot the left hand side of (5.11) versus $c$. If we interpret the vertical axis as the total number of terminals $M$, and for some $M$ we find the intersection with the curve, the corresponding value on the horizontal axis is the equilibrium point $c$, of the number of contending terminals $C$. We have plotted four curves for four different values of $p$. From these curves we observe that for some values of $M$, there are actually three equilibrium points $c$. As explained in [52], two of these are stable, while the third is unstable. What we want to stress out is that we want the PRMA system to operate with such a number of terminals $M$, such that the values of $c$ are less than one. The reason for

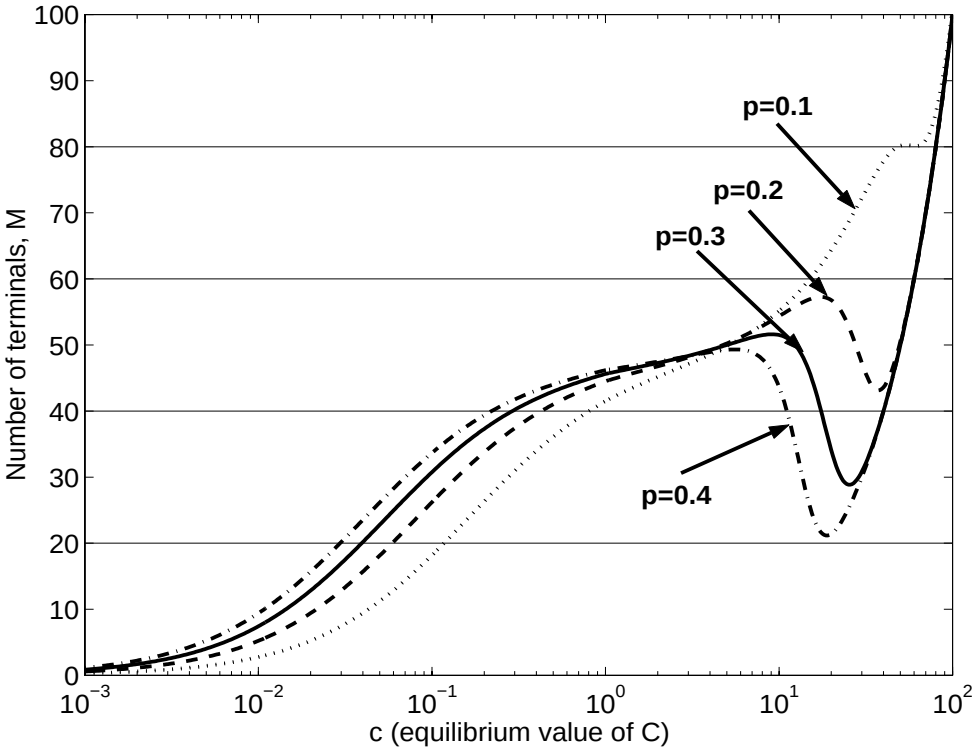



Figure 5.3: Graphic solution of equation (5.11), for several values of $p$.

this is the following: $c$, the equilibrium value of $C$ actually denotes the "average" number of contending terminals that *attempt to access* the channel at any empty slot. Since an attempt by more than one terminals to access the channel results in a collision, $c$ should be less than one in order to have an acceptable packet dropping probability. We will see that also below, when we compute the packet dropping probability. Figure 5.3 also shows that decreasing the access probability $p$, has the effect of moving the equilibrium points of the system that are greater than one to much higher values of $M$. On the other hand, for $p = 0.1$, the smallest equilibrium point is moved closer to one for lower values of $M$. Therefore, there is a trade-off in the choice of $p$. We want to choose it in such a way as to have a single equilibrium point for high enough $M$, but this point has to be less than one.

Let us now for a moment leave the equilibrium point calculations and turn to the computation of the packet dropping probability, which is the main performance metric for delay sensitive traffic (i.e. traffic where the delay is strictly bounded). For general values of $S$, $R_i$ and $C$, a terminal in state CON obtains a reservation if all of the following conditions are met: the time slot is not reserved, probability $1 - R/N$, where $R = \sum_{i=0}^{N-1} R_i$ is the number of reserved slots in a frame; the terminal has permission to transmit, probability p; the other $C - 1$ contending terminals do not have probability to contend, probability $(1-p)^{C-1}$. Therefore, the probability that a contending terminal does not obtain a reservation is given by:

$$v(C, R) = 1 - \left(1 - \frac{R}{N}\right) p(1-p)^{C-1} \tag{5.12}$$

Now, as we have already mentioned, terminals drop packets that have exceeded their lifetime without having being transmitted successfully to the BS, which happens at the beginning of a talk-spurt and before getting a reservation. Let

us denote the maximum packet lifetime by $D$ slots[4]. The terminal, upon ending a silence period, will have to wait $j$ slots before obtaining a reservation. The probability that the wait period is $j$ slots is given by the geometric distribution:

$$P_W(j) = (1-v)v^{j-1} \tag{5.13}$$

In a talk-spurt of $L$ packets, no packets are dropped if $j \leq D$. After waiting $D$ slots, the terminal will start dropping packets, starting from the oldest one. The terminal will drop packets every $N$ slots, as long as it is still waiting for a reservation. If, after $D + (L-1)N$ slots the terminal has not yet obtained a reservation, it drops the last packet of the talk-spurt. To summarize, if we call $n_{\text{drop}}(j)$ the number of packets dropped as a function of waiting time, then:

$$n_{\text{drop}}(j) = \begin{cases} 0, & 1 \leq j \leq D; \\ k & (0 < k < L), \quad D + (k-1)N + 1 \leq j \leq D + kN; \\ L, & D + (L-1)N + 1 \leq j. \end{cases} \tag{5.14}$$

Using (5.13) and (5.14) we can now calculate the probability that there are $n_{\text{drop}} = k$ packets dropped in a talk-spurt of $L$ packets:

$$P\left[n_{\text{drop}} = k|L\right] = \begin{cases} \sum_{j=1}^{D} P_W(j) = 1 - v^D, & k = 0; \\ \sum_{j=D+(k-1)N+1}^{D+kN} P_W(j) = v^{D+(k-1)N} - v^{D+kN}, & \\ 0 < k < L; & \\ \sum_{j=D+(L-1)N+1}^{\infty} P_W(j) = v^{D+(L-1)N}, & k = L. \end{cases} \tag{5.15}$$

Based on this last equation, we can calculate the expected number of dropped packets, given $L$ packets in the talk-spurt, $E\left[n_{\text{drop}}|L\right]$:

$$E\left[n_{\text{drop}}|L\right] = \sum_{k=0}^{L} kP\left[n_{\text{drop}} = k|L\right] = v^D \frac{1 - v^{NL}}{1 - v^N} \tag{5.16}$$

[4]Typical acceptable packet delays for a PRMA system are less than 32 ms, or two frames, or $D = 40$ slots, if we have $N = 20$ slots per frame.

Furthermore, the probability that there are $L = l$ packets in a talk-spurt is simply:

$$P[L = l] = \gamma_f (1 - \gamma_f)^{l-1} \tag{5.17}$$

and the un-conditioned expected value of dropped packets is:

$$E\left[n_{\text{drop}}\right] = \sum_{l=1}^{\infty} E\left[n_{\text{drop}}|L\right] P[L = l] = \frac{v^D}{1 - (1 - \gamma_f)v^N} \tag{5.18}$$

Finally, the packet dropping probability, $P_{\text{drop}}$ is the ratio of equation (5.18) to the average number of packets generated per talk-spurt, which is just $\gamma_f$, as can be readily computed from equation (5.17):

$$P_{\text{drop}}(C, R) = \frac{\gamma_f v^D}{1 - (1 - \gamma_f)v^N} \tag{5.19}$$

In equation (5.19) above we have explicitely noted the dependence of the packet dropping probability on $C$ and $R$, through the variable $v$, see equation (5.12). If we want to un-condition the packet dropping probability from these random variables, we will have to find their distributions. To this end, note that since the probability of any slot being reserved is $r$, and slots are reserved independent from one another, the number of reserved slots in a frame $R$ is a binomial random variable, with distribution:

$$P[R = j] = \binom{N}{j} r^j (1 - r)^{N-j} \tag{5.20}$$

In order to approximate the distribution of the number of terminals in contention $C$, we think of $C$ as the number of terminals waiting in a queue for reservations. Arrivals and departures from this waiting queue are independent events and therefore the distribution of $C$ is geometric:

$$P\left[C = k|R\right] = \begin{cases} p_0(1 - p_0)^k, & k < M - R; \\ (1 - p_0)^{M-R}, & k = M - R; \\ 0, & \text{else} \end{cases} \tag{5.21}$$

where $p_0 = 1/(c+1)$. Equations (5.20) and (5.21) used in equation (5.19) allow us to compute the packet dropping probability, for any given number of terminals $M$, slots per frame $N$ and permission probability $p$. In Figure 5.4 we plot $P_{\text{drop}}$

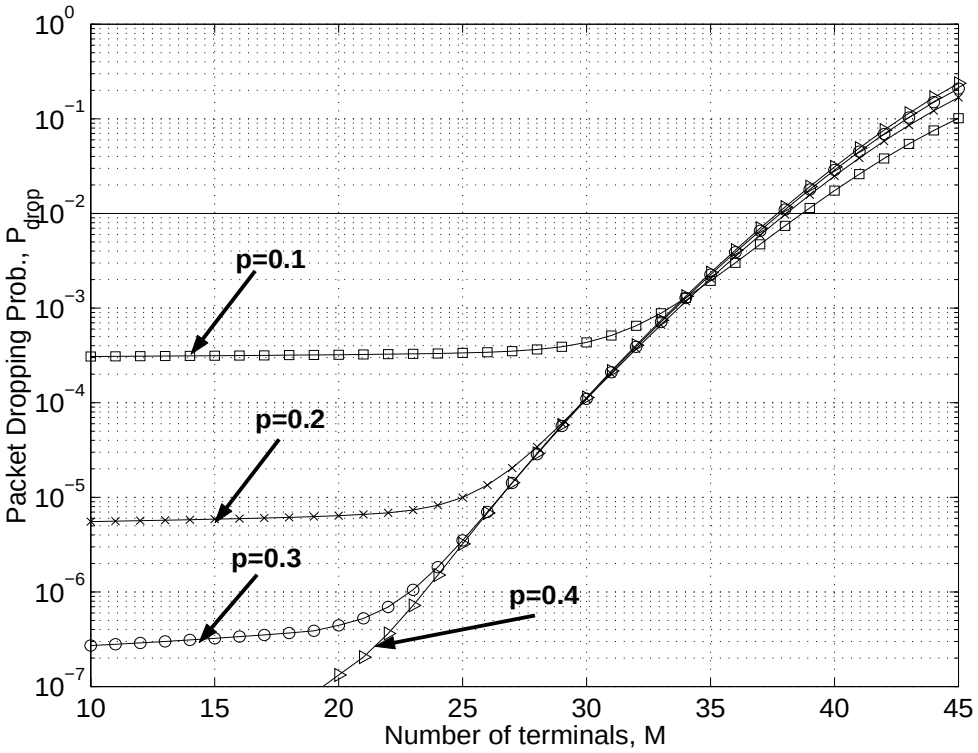



Figure 5.4: Packet dropping probability for PRMA, for several values of $p$.

versus the total number of terminals for several values of $p$. If we assume an acceptable packet dropping probability of 1%, this Figure gives us the capacity of a single cell PRMA system, and also its multiplexing capability. We note that the values of $p$ have little effect on the capacity of the system, although this is partially true. High values of $p$ will lead to more than one equilibrium points, as we show in Figure 5.3 and this may lead the system to instability. Figure 5.4 however has been created by using the smallest equilibrium point only. Second, at very small values of $p$, our simplified model of the PRMA system may not be very accurate, since now there is a non-negligible probability that a terminal will drop

all of its talk-spurt before having a chance to compete and access the channel. We see from Figure 5.4 that when the number of terminals is small and $p$ is also small, a significant percentage of packets is dropped, obviously due to complete lack of access attempts to the channel. Furthermore, by comparing Figures 5.3 and 5.4, we conclude that for values of $M$ where $P_{\text{drop}}$ is smaller than 1%, $c$ is also smaller than one. Finally, we should note that the multiplexing capability of PRMA is about 1.8 (about 37 terminals share $N = 20$ slots in a frame at capacity), while the theoretical multiplexing capability due to the intermittent voice activity is 2.35, which is the inverse of the speech duty cycle.

### 5.2.2 PRMA++

PRMA++ has evolved from PRMA and the need to integrate different services than just voice [55]. The demand for support of a multiplicity of service types, led to the requirement for a medium access mechanism that would multiplex channels that require a multiplicity of time slots per TDMA frame. Therefore, in contrast to PRMA, the time slots in PRMA++ are not considered the same, but on the up-link are divided into slots where the contention access is performed (R-slots) and slots where the physical traffic channels are supported (I-slots). In the R-slots, the mobiles transmit access bursts in a contention manner, same as for the PRMA. The access bursts contain, except the mobile's ID, the number of slots the mobile requires to satisfy its transmission requirement. The base station responds with a reservation acknowledgment in the paired acknowledgment slots of the downlink. Note that a successful contention in the R-slots does not mean that the mobile has immediate access to I-slots, since these may all be occupied. Instead, it stays in a standby mode until the base station allocates an I-slot to it, dropping packets that exceed their lifetime. The number of reservation slots

inside a frame is a design parameter that sets a trade-off between contention and data bandwidth and is decided based on simulations [54]. The similarities of PRMA++ to PRMA are obvious and their performances are also comparable. Its design has been done again without taking into account performance degradation due to co-channel interference, which as in the case of PRMA is very large. What one finds out, as we shall see, is that these protocols need rather large reuse distances to have useful packet dropping probabilities.

### 5.2.3 Centralized Packet Reservation Multiple Access (C-PRMA).

This protocol presents a significant departure in design philosophy from the previous ones [69]. Co-channel interference is taken into account implicitly both by the scheduling algorithm and by the retransmission scheme. The protocol is targeted for pico-cellular environments where round-trip delays can be considered negligible and this fact is used for immediate retransmissions of lost (interfered) packets. In this sense, it is implicitly assumed that interference will change fast enough, for a subsequent slot to be interference free. The protocol therefore takes the point of view of interference averaging: by retransmitting lost packets at random times in the future, each terminal eventually sees an average interference, an idea that has given rise to a new packet access mechanism known as *capture division packet access,* [56]–[58]. Constant slot assignment is dropped in favor of a more flexible scheme, whereby each MT packet transmission takes place in a command/response way; the BS keeps information on the state of each active mobile, that is its bandwidth requirements, the maximum tolerable delay, the current delay of the first packet in the queue and the number of retransmissions of the current packet. Based on this information, it determines which mobile has more urgent need to transmit and it commands it with a mini-slot, to which

the mobile responds immediately. Contention is done in the same way as with PRMA, that is using available slots, but now more than one mobile can content in one slot by using mini-slots.

It is worth here examining the scheduling algorithm more closely. Each BS uses a shift register, PR (called the *polling register*) whose positions are either empty, or filled with the id of a MT waiting to be polled, after its reservation has been successfully received. The positions of the PR are numbered from right to left, starting from zero, see Figure 5.5. A new packet reservation is stored in the PR in the highest available position, smaller or equal to the packet's remaining lifetime and to the right of packets with same lifetime but smaller number of retransmissions. Observe that this way, the actual lifetime of packets with many retransmissions is shortened. The total length of the PR is the maximum delay

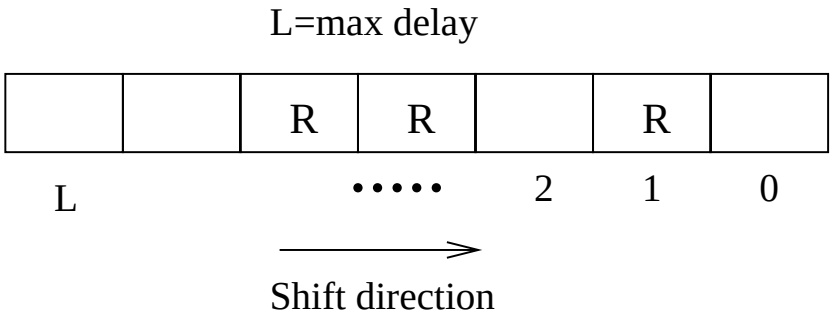



Figure 5.5: The Polling Register of C-PRMA with some positions filled with reservations of successful mobiles.

tolerable by any kind of traffic. After each time slot, the BS shifts PR one position to the right. Therefore, if a packet reservation is stored in position $k$ of the PR (with the form of the id of the corresponding MT), then that packet must be successfully transmitted by its MT within $k$ slots, otherwise it will be dropped. At each time slot the BS issues either an available time slot, where terminals can content, or a reserved one based on the following:

1. If position 0 of the PR holds a reservation, the BS always issues a reserved slot to the corresponding MT, in a last effort to receive its packet before it gets lost.

2. If PR is empty (no packets waiting for transmission) or the number of consecutive reserved slots exceeds a threshold, the BS issues an available slot. The second criterion is used to insure enough reservation bandwidth for the mobiles and is completely equivalent to the number of reservation slots in the PRMA++ protocol.

3. If none of the conditions above are satisfied, the BS starts from position 1 of the PR and finds the first occupied position with the smallest number of retransmissions and commands the corresponding MT to transmit its packet. By assigning higher priority to mobiles with fewer retransmissions it is hoped that bandwidth will not be wasted excessively for mobiles with poor channel conditions. This policy though, together with the fact that packets with many retransmissions have shorter effective lifetimes demonstrates an unfairness with mobiles that are experiencing poor channel conditions.

The scheduling algorithm of C-PRMA shows the complexity increase of current protocols in the effort to integrate multiple traffic types and reuse bandwidth as aggressively as possible. As already mentioned, co-channel interference is dealt with implicitly by retransmitting packets quickly and hoping that the channel quality changes fast. However, no particular interference statistics are calculated by the base stations, that could be used for more efficient dynamic slot allocation.

## 5.3 Summary

In this chapter we discussed the importance of MAC protocols for packet switched wireless networks. New multimedia applications as well as the expansion of the Internet into wireless have shifted attention from circuit to packet switched technologies. The extension of wireless into businesses and homes raises the problem of frequency planning, which becomes very difficult or even impossible to address in indoor environments. In that respect, new MAC protocols need to be designed that take into account the co-channel interference and adapt to it.

We briefly presented here some MAC protocols that have appeared in the literature, that try to address the medium access problem for packet switched networks. However, most of them are designed in a single cell setting and therefore the primary problem they cope with is the multiple access one. Nonetheless, the analytic methods introduced here can be extended in order to analyze MAC behavior in co-channel interference environments.

Since the original conception of PRMA, many other MAC protocols have appeared in the literature, some of which we also described here. Extensions of PRMA include, except PRMA++, other alternatives also that try to accommodate multiple types of traffic (such as non-real-time together with real-time) [61], [59], [62]. Simulative analyses of cellular PRMA systems have also appeared in the literature [63], [60]. Finally, a large volume of work has been done in non-contention protocols, or contention resolution protocols that try to avoid the instabilities inherent in a random access protocol such as PRMA, [64], [61], [65], [66].

# CHAPTER 6

# An Interference Analysis for Packet Reservation Multiple Access

## 6.1 Introduction

In the previous chapter we identified some of the most desirable properties of packet switched multiple access schemes for wireless communications. Some of these are the flexibility to incorporate applications with varying quality of service requirements, as well as the ability to statistically multiplex the usually bursty traffic, by assigning the same bandwidth on demand to different traffic flows. An additional reason for favoring packet switching versus the more traditional connection oriented access is of course the penetration of the Internet and the TCP/IP protocol suites in wireless networks [45], [46].

Obviously, research in packet switched multiple access protocols for broadband cellular systems has been intense (see [44] and references therein). The main thrust in the design of such protocols for the third generation wireless communications networks is the integration under a common bandwidth, of traffic classes requiring varying quality of service guarantees. Thus, the problem becomes one of priority based reservation and allocation of resources to multiple competing traffic types. For example, voice traffic demands very low delay and therefore should have access precedence over data traffic, while data traffic demands much

larger bandwidth, and so on. Co-channel interference, which is always present in a cellular environment is not taken into account for the design of most current packet switched protocols. The assumption is that appropriate frequency planning separates adequately common carriers/channels. However, as the number of users of wireless services increases and bandwidth becomes more scarce, interference statistics should become a major design concern for future medium access control (MAC) protocols. Note here that, the behavior of connection oriented access protocols (as those employed in the second generation wireless networks) in the presence of co-channel interference has been studied extensively. Methods of reducing such interference through power control or adaptive channel allocation have been proposed and have been shown to offer great advantages in increasing the capacity of cellular networks, through denser frequency re-use patterns [74], [82], [90], [81], [75], [78]. The situation though changes in the case of packet access protocols. Here, since no permanent bandwidth reservation is made for a specific flow, interfering sources are intermittent and variable in the same bandwidth, and over short periods of time (of the order of one packet duration). The result is that co-channel interference now changes much faster and in a much more random way, [67]. In addition to that observation, we should also consider the need for aggressive frequency re-use with little or no frequency planning for wireless personal communications services in offices, houses and other indoor as well as outdoor areas. It becomes quite clear then that, in such settings, packet access protocol performance becomes interference limited (much the same way as for CDMA systems) and therefore a good understanding of the statistics of interference that a protocol generates is crucial for its evaluation, and may also suggest ways for improving it. In this respect, analytical results when possible to be obtained, provide a deeper understanding than simulations and allow for the easy evaluation of the effect of varying different parameters of the protocol.

In this chapter, we present a co-channel interference analysis for the Packet Reservation Multiple Access (PRMA) protocol, based on a Markov modeling of the interference generation process. As we have already mentioned, PRMA is one of the first packet switching protocols proposed for wireless personal communications and its performance in cellular environments has been studied with simulations. We present an analytic approach to performance evaluation that reveals the structure of the dynamic behavior of the interference that is generated by this simple protocol.

## 6.2 PRMA Speech Terminal Model in a Multi-Cell Environment

Details for the PRMA protocol were presented in the previous chapter, where we saw how the application of equilibrium point analysis led to the determination of the probability distributions of the number of reserved slots, the number of voice terminals contending for reservations etc. We will use these single cell analytical results below, in a voice terminal model adapted to a multi-cell environment. Throughout, we assume that there are $M$ terminals in each cell and $N$ slots in the PRMA frame. Speech activity is represented by a Markov chain with two states to account for silent (state 0) and talking (state 1) periods [85]. The transition matrix is given in (6.1).

$$\mathbf{P}_{\text{speech}} = \begin{pmatrix} 1-\sigma & \sigma \\ \gamma_f & 1-\gamma_f \end{pmatrix} = \begin{pmatrix} 1-592.6\times10^{-6} & 592.6\times10^{-6} \\ 16\times10^{-3} & 1-16\times10^{-3} \end{pmatrix} \tag{6.1}$$

The modified voice terminal model for a multi-cell PRMA environment is shown in Fig. 6.1, where:

- SIL denotes that the terminal is silent.

- CON denotes that the terminal is contending for access.
- RES denotes that the terminal has a reservation for one of the $N$ slots of the PRMA frame.
- $r$ is the probability of any given slot to be reserved (same for all slots).
- $p$ is the probability with which a contending terminal accesses a vacant slot. Different contending terminals try to access vacant slots independently from each other.
- $P_c$ is the probability that a transmitted packet gets captured by the intended base station. It depends on the capture capability of the BS and also on the interference statistics. For now, we consider it known and explain later how we calculate it iteratively.

As in the previous chapter, we call $C$ the number of contending terminals in the cell (state CON), $S$ the number of terminals in silence (state SIL) and $R$ the number of terminals with reservations (state RES). Also, let $c$ and $s$ be the equilibrium values of $C$ and $S$ respectively. The differences of the MC model of Fig. 6.1 with the one of the previous chapter are only the following: all the reservation states have been collapsed to a single one, since they all correspond to the same reservation probability $r$; it has been assumed that $c < 1$ and hence $u(c) = 1$, since as it was seen in the previous chapter only these values of the equilibrium of $C$ give useful results in terms of packet dropping probability; finally, the capture probability $P_c$ has been incorporated in the necessary states.

To obtain equations for the unknown quantities $c$ and $r$ and $s$, we equate the flows in and out of states RES and SIL, and we also note that all the terminals

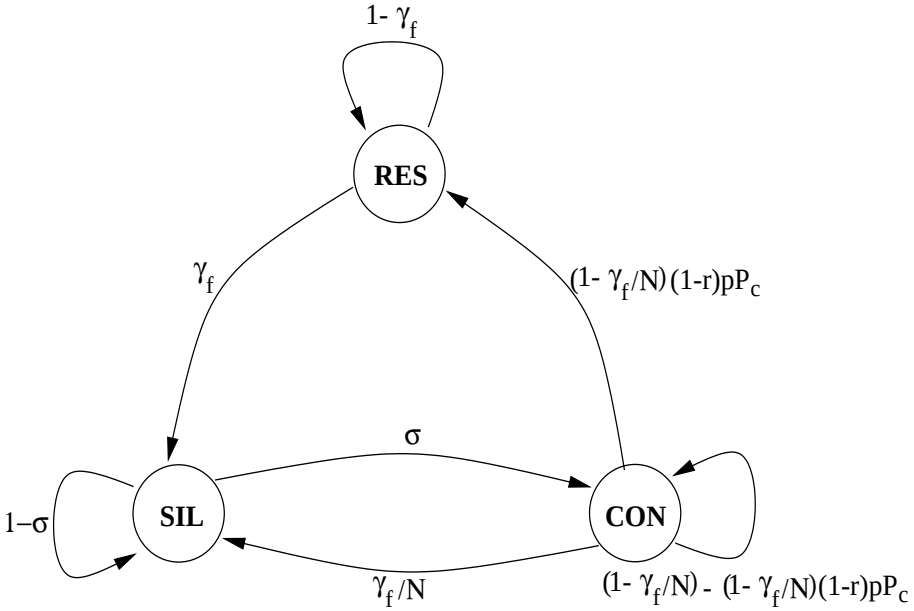



Figure 6.1: A PRMA speech terminal model, as adapted from [2].

in states SIL, RES and CON must total M:

$$r\gamma_f = c\left(1 - \frac{\gamma_f}{N}\right)(1-r)pP_c \tag{6.2}$$

$$sc = c\frac{\gamma_f}{N} + r\gamma_f \tag{6.3}$$

$$s + c + Nr = M \tag{6.4}$$

Eliminating $s$ from the above, we get for $c$ and $r$ equations (6.5) and (6.6):

$$c\left(1 + \frac{\gamma_f}{N\sigma}\right) + \left(N + \frac{\gamma_f}{\sigma}\right)\frac{cpP_c\,(1-\gamma_f/N)}{\gamma_f + cpP_c\,(1-\gamma_f/N)} - M = 0 \tag{6.5}$$

$$r = \frac{cpP_c\,(1-\gamma_f/N)}{\gamma_f + cpP_c\,(1-\gamma_f/N)} \tag{6.6}$$

Additionally, the probability distributions of $R$ and $C$ are given by (see previous chapter):

$$P(R) = \binom{K}{R} r^R (1-r)^{K-R} \tag{6.7}$$

$$P(C) = \sum_{R=0}^{K} P(C|R)P(R) \tag{6.8}$$

$$P(C|R) = \begin{cases} p_0(1-p_0)^C & \text{if } C < M-R, \\ (1-p_0)^{M-R} & \text{if } C = M-R, \\ 0 & \text{else} \end{cases} \tag{6.9}$$

where:

$$p_0 = \frac{1}{c+1}, \quad K = \min\{M, N\} \tag{6.10}$$

## 6.3 Analysis of Interference Statistics

The extension of performance evaluation for the PRMA protocol in cellular environments is of great interest. Co-channel interference impairs the protocol both during terminal contention, where packets may not be captured resulting in larger access delays, and during reservation, where additional packets are lost. The number of packets lost during a reservation period depends on the SIR dynamics, i.e. how the SIR in a single slot changes from frame to frame. In the following, we present an analysis of the SIR, based on a Markov model.

### 6.3.1 The pdf of Interfering Power

Consider a terminal randomly placed in cell $j$, that transmits with unit power to its base station $\text{BS}_j$, as shown in Fig. 6.2. Then, neglecting Rayleigh fading

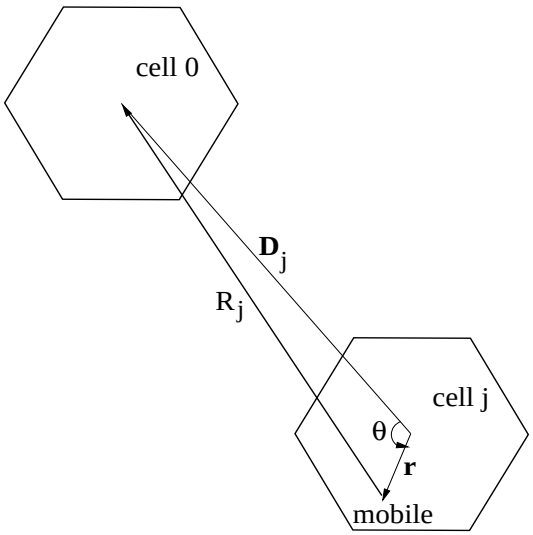



Figure 6.2: Geometry of a mobile terminal in cell $j$, interfering with the reception of base station 0.

for the moment, the received interfering power at the BS of cell 0 ($BS_0$) will be $e^{\xi_j} R_j^{-\eta}$, where $R_j$ is the distance from the terminal to $BS_0$, $\eta$ is the propagation exponent, and $\xi_j$ is a Gaussian RV with zero mean and variance $\sigma^2$, that accounts for log-normal shadowing. The pdf of this interfering power, averaged over the area of cell $j$ to account for the random position of the terminal, has been derived a number of times in the literature (for example [70]):

$$g_j(x) = P\left\{e^{\xi_j} R_j^{-\eta} \in (x, x+dx)\right\}$$
$$= \frac{1}{A} \iint \frac{\exp\left\{-\left[log(x) + \frac{\eta}{2} log\left(D_j^2 + r^2 - 2rD_j \cos(\theta)\right)\right]^2 / 2\sigma^2\right\}}{\sqrt{2\pi}\sigma x} r \, \mathrm{d}r \, \mathrm{d}\theta \quad (6.11)$$

where $A$ is the area of the hexagonal cell, $D_j$ is the distance between the center of cell $j$ and that of cell 0 and $(r, \theta)$ are the polar coordinates of the terminal, with respect to $D_j$, from the center of $BS_j$. For example, if we assume unit radius hexagonal cells, a propagation exponent of four and a log-normal shadowing component with variance 6 dB, the pdf of equation (6.11) is plotted in Fig. 6.3

for a mobile placed in cell 0 and in an (interfering) adjacent cell 1.

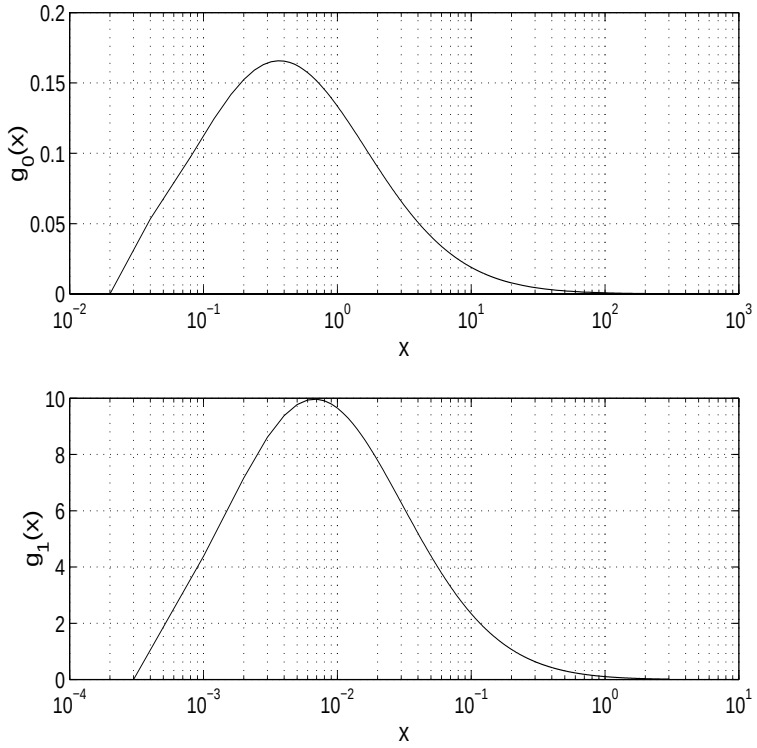



Figure 6.3: The pdf of power received at base station 0, when a mobile is placed in cell 0 or in cell 1, which can be any of the six cells adjacent to cell 0. No Rayleigh fading.

### 6.3.2 A MC Model for Interference Generation

Consider now a specific PRMA slot in cell $j$. Depending on whether some terminal in this cell transmits during the slot, interference will be generated or not in the corresponding slot in cell 0. The transmission of a packet in a slot of cell $j$ is, to a very good approximation, Markovian in structure: if the slot is reserved then packet transmission follows the source statistics which are Markovian; if the slot is not reserved, transmission of a packet depends on the number of contending

terminals and the permission probability $p$, but not on the past transmissions in that slot. Therefore, we model the slot behavior in cell $j$ as a MC, $\{S_n\}$, shown in Figure 6.4, where the states are defined as:

- EMP: the slot is empty, no terminal of cell $j$ transmits in it.
- HIT: a terminal transmits during the slot but does not get captured. Note that transitions from state HIT to itself are caused in general by different terminals trying to access the slot in two consecutive frames.
- OCC: the slot is occupied by a terminal.

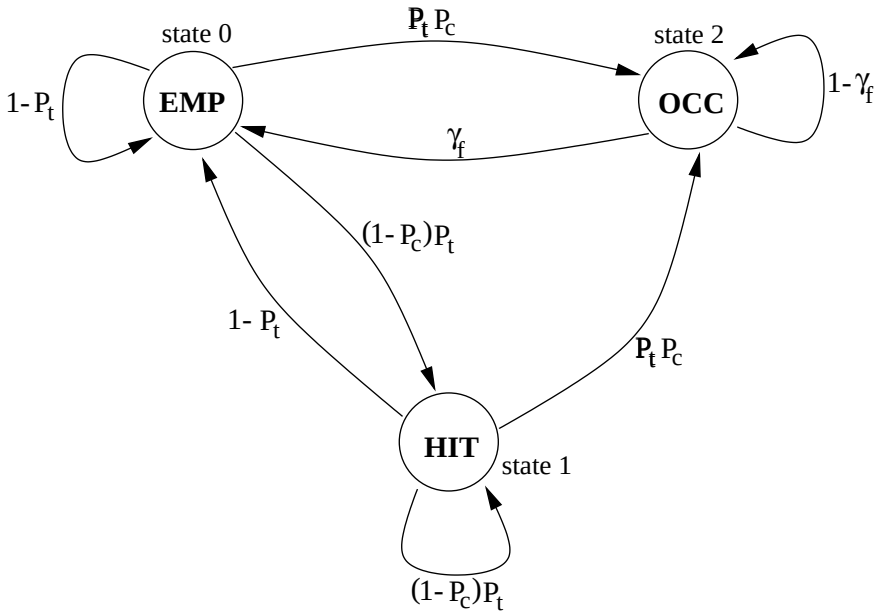



Figure 6.4: A Markov Chain Model for a PRMA slot, from interference point of view.

Jumps on the MC occur once every PRMA frame. $P_c$ is the capture probability, $P_t$ is the probability that one terminal will try to access the slot, given that it is not reserved. If $C$ terminals out of $M$ total contend in cell $j$, $P_t$ can be

calculated as:

$$P_t = \sum_{C=1}^{M} Cp(1-p)^{C-1}P(C) \tag{6.12}$$

where P(C) is given by (6.8). Here, we neglect the possibility that more than one contending terminal transmits in a slot. Analysis of the single cell PRMA protocol shows that $c$, the equilibrium value of $C$, is always smaller than one, for usefully loaded systems, i.e. for dropping probabilities of the order of $10^{-2}$. This implies that the probability of simultaneous transmission of more than one terminal in a vacant slot is quite small. Of course, this assumption does not hold at higher traffic loads where our model will not work well.

The slot generates interference $I_j$ to the corresponding slot in cell 0 in all transitions except from state EMP to itself. Jumps out of state EMP generate interference in the current frame, but not in the previous one. Jumping from state HIT to HIT, the slot generates interference both in the previous and current frames, but the interference powers are independent (conditioned on coming from the same cell $j$), since they are generated by different terminals. The same is true for transitions from HIT to OCC. Transitions from OCC to OCC generate the same interference in the previous and current frame, since the slot is occupied by the same terminal and both the shadowing and deterministic path loss are assumed constant for the span of one frame. Finally, transitions from OCC to EMP generate interference in the previous frame but not in the current one. Based on these comments, we can write the joint pdf of the interference generated by the specific slot in the previous and current frames as a sum of pdfs of mutually exclusive events. Call $s_{ik}$ the joint probability $P[S_{n-1} = i, S_n = k] \quad i, k = 0, 1, 2,$ where the numbering of states is shown in Figure 6.4. Then, the joint pdf can be

written as:

$$\begin{aligned} f(x,y) &= P[I_j(n) \in (x, x+dx), I_j(n-1) \in (y, y+dy)] \\ &= s_{00}\delta(x)\delta(y) + (s_{01}+s_{02})g_j(x)\delta(y) + (s_{11}+s_{12})g_j(x)g_j(y) \\ &\quad + (s_{10}+s_{20})g_j(y)\delta(x) + s_{22}g_j(x)\delta(x-y) \end{aligned} \tag{6.13}$$

where $g_j(x)$ is given by equation (6.11).

The knowledge of the joint pdf of the interfering power from a slot in cell $j$ in two consecutive frames allows for the calculation of the joint cdf of the SIR of a terminal transmitting in cell 0. Specifically, consider a terminal transmitting in a slot at cell 0 and call $P_0$ the power received by $\mathrm{BS}_0$, excluding Rayleigh fading. Then the joint cdf of SIR is:

$$\begin{aligned} &P\left[\mathrm{SIR}(n) \le x, \mathrm{SIR}(n-1) \le y\right] = \\ &\qquad P\left[\frac{\alpha_0^2(n)P_0}{\sum_j \alpha_j^2(n) I_j(n)} \le x,\ \frac{\alpha_0^2(n-1)P_0}{\sum_j \alpha_j^2(n-1) I_j(n-1)} \le y\right] = \\ &P\left[\alpha_0^2(n) \le x \sum_j \alpha_j^2(n)\frac{I_j(n)}{P_0},\ \alpha_0^2(n-1) \le y \sum_j \alpha_j^2(n-1)\frac{I_j(n-1)}{P_0}\right] = \\ &\qquad E\left[\left(1 - \prod_j e^{-x\alpha_j^2(n)\frac{I_j(n)}{P_0}}\right)\left(1 - \prod_j e^{-y\alpha_j^2(n-1)\frac{I_j(n-1)}{P_0}}\right)\right] = \\ &\qquad\qquad E\left[\left(1 - \prod_j \frac{1}{1+x\frac{I_j(n)}{P_0}}\right)\left(1 - \prod_j \frac{1}{1+y\frac{I_j(n-1)}{P_0}}\right)\right] \end{aligned} \tag{6.14}$$

where $\alpha_j(n), \alpha_j(n-1)$ are Rayleigh distributed random variables, taking fast fading into account. Since the length of a PRMA frame is 16 ms, a Doppler bandwidth of 11 Hz or larger would yield an autocorrelation coefficient for $\alpha_j^2(n)$ of 0.5 or smaller [72] and therefore we are justified in approximating the fading as independent. This justifies lines three and four of equation (6.14) above. The expectation of the last line of equation (6.14) is taken with respect to $I_j(n), I_j(n-1)$ whose joint pdf is given in equation (6.13) and $P_0$ whose pdf is given by equation (6.11) by setting $D_j = D_0 = 0$.

Treatment of selection diversity reception is also treated easily and the joint cdf of the SIR will be:

$$P\left[\mathrm{SIR}(n) \leq x, \mathrm{SIR}(n-1) \leq y\right] = E\left[\left(1 - \prod_j \frac{1}{1 + x\frac{I_j(n)}{P_0}}\right)^2 \left(1 - \prod_j \frac{1}{1 + y\frac{I_j(n-1)}{P_0}}\right)^2\right] \quad (6.15)$$

Knowledge of the joint cdf of the SIR allows us to model the wireless link as a Gilbert-Elliot channel, that is there are only two states, the good one (state 1) where packets get captured and the bad one (state 0) where packets are lost (i. e. the cross-over probabilities of the channel are 0 and 1 respectively). If we call $p_{ij}$ the transition probabilities, then:

$$\begin{aligned} p_{11} &= P\left[\mathrm{SIR}(n) > b \quad | \quad \mathrm{SIR}(n-1) > b\right],\ p_{10} = 1 - p_{11} \\ p_{00} &= P\left[\mathrm{SIR}(n) < b \quad | \quad \mathrm{SIR}(n-1) < b\right],\ p_{01} = 1 - p_{00} \end{aligned} \quad (6.16)$$

where $b$ is the capture ratio. Finally, the capture probability is $P_c = P[\mathrm{SIR}(n) > b]$, or the invariant probability of being in state 1.

Up to this point we have assumed that $P_c$ is known but in reality we have no means of computing it directly, since the whole system is coupled by interference and $P_c$ appears in all our equations. Instead, we use the following iterative method:

1. Choose an arbitrary value of $P_c$, usually 1, solve for $c$ and $r$ from (6.5), (6.6) and obtain $P(C)$ from (6.8).

2. Calculate the transition probabilities of the Markov chain of Figure 6.4 and from these obtain the joint pdf of equation (6.13).

3. Finally, obtain the cdf of the SIR from equations (6.14) or (6.15) by numerically evaluating the expectation. Note that, due to the delta functions in

equation (6.13), there is only need to calculate two-dimensional integrals, assuming we have the numerical values of the pdf of equation (6.11). From this cdf, $P_c$ can be calculated easily and then used as a new value in step one above.

The procedure converges to a value for $P_c$, usually within a few iterations. Although we do not have a proof of whether this is the true answer, we can argue qualitatively that there is a fixed point in the iteration in the following way: a value of $P_c$ larger than the true one, would yield a larger than the true number of reserved slots, which will lead to larger interference. Therefore, $p_{11}$ of equation (6.16) will be smaller than its real value, leading to a smaller $P_c$. By the same reasoning, a smaller value for $P_c$ would lead to smaller interference which would increase $p_{11}$, which would lead to larger $P_c$. Also, the agreement of the analysis with simulation results verify the validity of this approach.

### 6.3.3 Dropping Probability Calculation

By dropping probability, we mean the combined effect of packets dropped in contention and lost from interference. The average number of packets dropped, $\overline{N}_{\text{drop}}$ during contention is evaluated as in [52], but is usually negligible as compared to the average number of packets not captured due to interference during a reservation period. Consider now that a terminal successfully contends for a slot and that upon reserving it, has $J$ more packets left to transmit. Then, the average number of packets that will be hit by interference will be $(J - N_J)$, where $N_J$ is the average number of visits of the channel to the good state (state 1), beginning from state 1, since the first of the $J$ packets was captured. $N_J$ therefore is the

renewal function for the channel Markov chain and is given by [71]:

$$N_J = \sum_{k=0}^{J} p_{11}^{(k)} \tag{6.17}$$

where $p_{11}^{(k)}$ is the (1,1) entry of the $k$-th power of the transition matrix. Finally, the average number of packets lost will be:

$$\overline{N}_{\text{lost}} = \sum_{j=1}^{\infty} (j - N_j) P[J = j] \tag{6.18}$$

where $P[J = j]$ is the probability that there are $J = j$ packets left in a talk-spurt when a reservation begins and is calculated in appendix A. The total dropping probability is given as the ratio of the average number of dropped and lost packets over the average number of packets in a talk spurt, which is just $1/\gamma_f$:

$$P_{\text{drop}} = \gamma_f \left( \overline{N}_{\text{lost}} + \overline{N}_{\text{drop}} \right) \tag{6.19}$$

## 6.4 Numerical Results

The above analysis was used to calculate the dropping probability for a cellular PRMA system. The standard deviation of the shadowing was assumed to be 6 dB, the capture ratio was set to 8 dB and the access permission probability, $p$, was chosen as 0.3. A simulation was also done with the same parameters to verify the analysis, using Parsec, a C based discrete event parallel simulation language, developed at UCLA [73]. A regular hexagonal pattern was considered with 37 cells arranged in six rings of six cells each around a central cell. Statistics were taken at the central cell. For the simulation, a fixed number of terminals (corresponding to a certain average number of terminals per cell) were generated uniformly over the whole service area at each run and were held fixed for this run. Results were averaged over several runs. Terminals were connected to the closest BS (which for $K > 1$ is most likely the strongest one also).

Results in the form of total dropping probability versus average number of terminals per cell are presented in Figure 6.5 for cluster sizes of $K = 1, 3, 4$, and 7. Agreement between analysis and simulation is very good except at high number of terminals per cell where intra-cell collisions are not negligible. Furthermore, we should note that at low average number of terminals per cell, the total dropping probability is dominated by $\overline{N}_{\text{lost}}$, while at higher number of terminals per cell the dominating factor is $\overline{N}_{\text{drop}}$. The plots confirm the already established result in the literature [67], that every cell channel reuse ($K = 1$) is impossible for the PRMA system. Taking a value of $4 \times 10^{-2}$ as an acceptable total dropping probability, we see that a cluster size of $K = 3$ supports $M = 8$ terminals per cell, while $K = 4$ supports $M = 13$ and $K = 7$ supports $M = 33$ terminals per cell.

For the case of $K = 4$ and $M = 13$, which corresponds to $P_{\text{drop}} = 4 \times 10^{-2}$, as can be seen from Fig. 6.5, we plot the resulting Gilbert-Elliot channel in Fig. 6.6. The channel has two states, state B where SIR $< 8$ dB, and state G where SIR $> 8$ dB. The transition probabilities calculated for this example are shown on the MC.

Figure 6.7 presents only analytical results for a selection diversity of two. We can observe a great improvement for all cluster sizes. Figure 6.8 shows the system capacity (number of terminals per cell over total number of channels at $P_{\text{drop}} = 4 \times 10^{-2}$) for each cluster size and for both cases of no diversity and diversity two. Observe that the increase of system capacity for $K = 7$ is minimal, since the dropping probability in this case is limited more from congestion than from interference (at $P_{\text{drop}} = 4 \times 10^{-2}$, $M = 33$ terminals are already supported without diversity). Without diversity, capacity increases with cluster size but for diversity two the optimum cluster size is four.

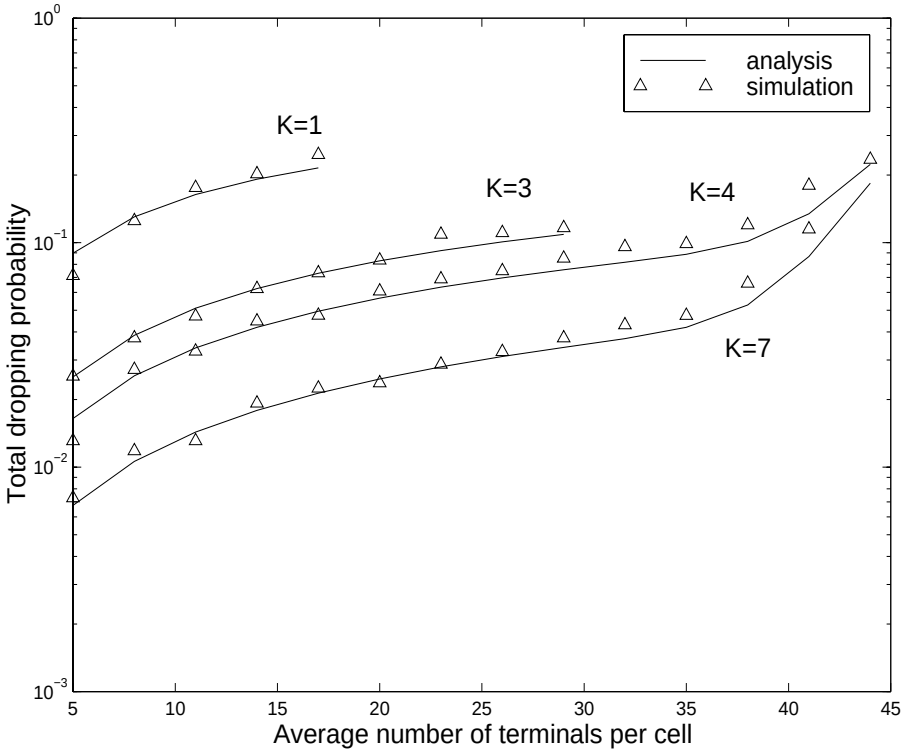



Figure 6.5: Total dropping probability vs. number of terminals per cell for no diversity and various cluster sizes.

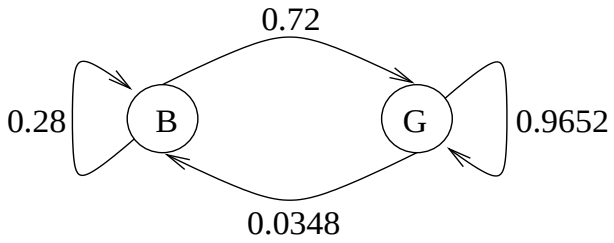



Figure 6.6: An example of the MC representing a Gilbert-Elliot channel for $K = 4$ and $M = 13$, $P_{\text{drop}} = 4 \times 10^{-2}$.

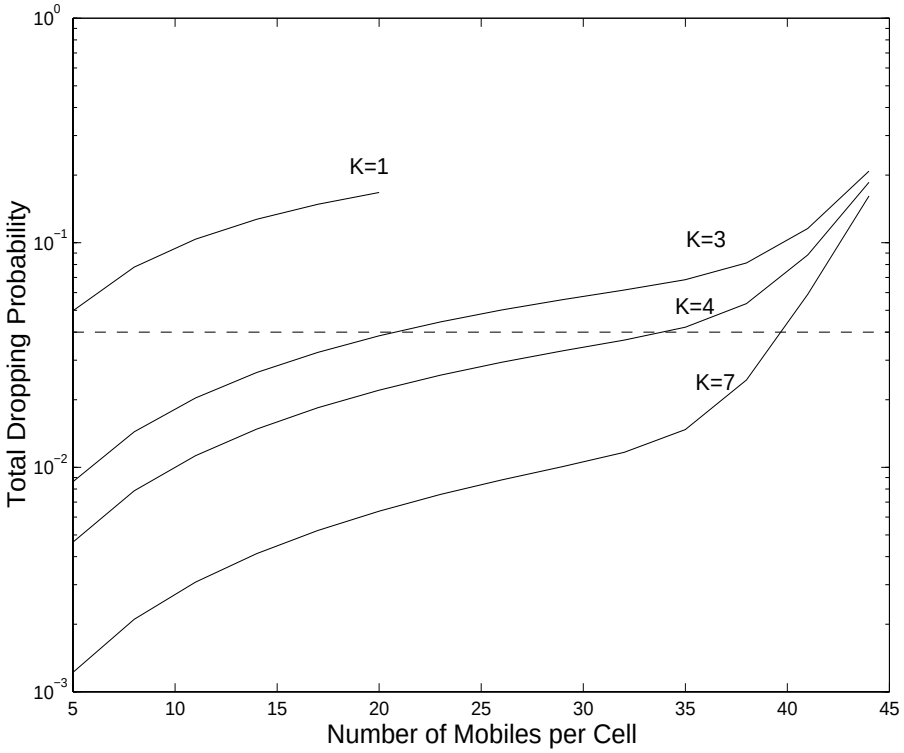



Figure 6.7: Analytical results of the total dropping probability vs. number of terminals per cell for diversity 2 and various cluster sizes.

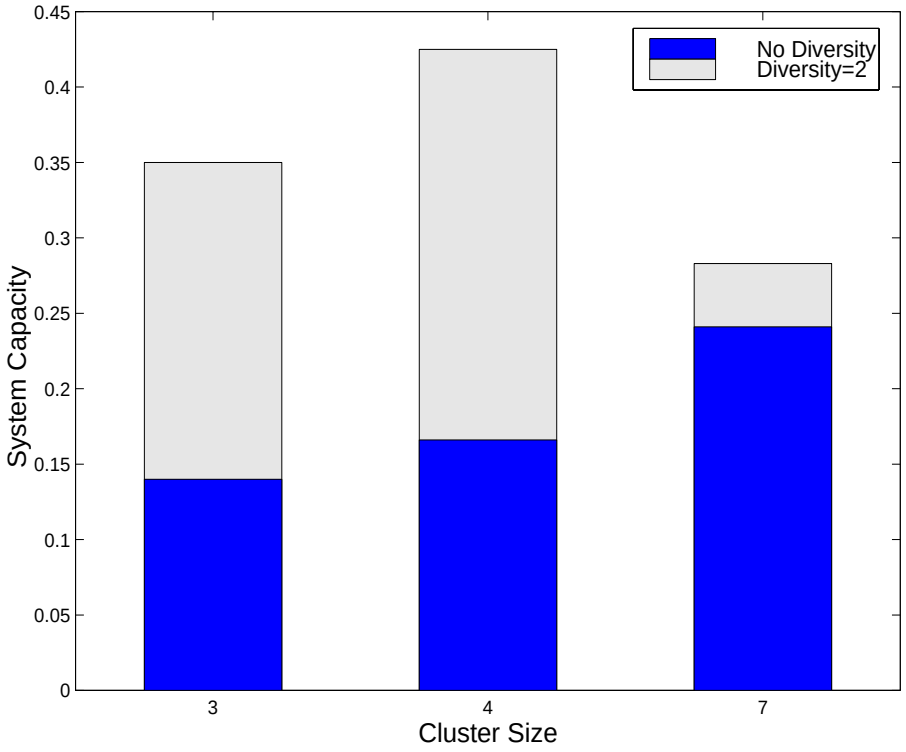



Figure 6.8: System capacity at $P_{\text{drop}} = 4 \times 10^{-2}$ vs. cluster size, for no diversity and diversity 2.

## 6.5 Summary

In this chapter, an analysis of the performance of the PRMA protocol in a cellular environment has been presented. The analysis is based on the realization that the interference generated by a slot of an interfering cell is Markovian, with the transition matrix of the Markov process depending on the source statistics of the voice terminal and the access mechanism imposed by the protocol. The joint probability distribution function (pdf) of interference coming from the same slot in two consecutive PRMA frames was evaluated, taking into account deterministic path loss, shadowing and Rayleigh fading, and used to determine the joint cumulative distribution function (cdf) of the signal to interference ratio (SIR) of a terminal in the intended base station. This result in turn is used to model the wireless link as a Gilbert-Elliot channel and the average number of packets lost due to interference is found as the expected number of visits of the link to the bad state. The major contribution of our work lies in the modeling of interference statistics and its successful application to the performance evaluation of cellular PRMA. The methodology introduced may also be helpful for the interference analysis of more complicated packet access protocols that have appeared since PRMA, and which try to integrate voice video and data [54], [69].

Simulation results of the performance of PRMA in cellular and indoor environments have been presented in [67] and [63], and a comprehensive interference analysis for circuit switched cellular systems, like GSM, is presented in [70]. An interference analysis of cellular PRMA for Rayleigh fading channels (no log-normal shadowing) was done in [68], where the number of inter-cell interfering packets received at a base station is approximated by a Poisson distribution. Finally, in [91] an analysis is presented on the capacity of PRMA in a single cell, when the base station is capable of capturing packets in the presence of multiple

intra-cell transmissions, as well as random channel errors.

# CHAPTER 7

# Algorithms for Dynamic Packet Allocation in Co-Channel Interference Limited Environments

## 7.1 Introduction

In the previous chapter we investigated analytically the co-channel interference generation process in a PRMA cellular network, where voice terminals from adjacent cells interfere each other, and performance degradation occurs (mostly) in the form of reserved packets being lost due to too low a signal to interference ratio. Of course, packet loss was directly linked to frequency reuse density and our analysis showed the infeasibility of using the same channels (carrier-slot combinations) in each cell. On the other hand, in view of the high bandwidth of new services and the potential demand growth for these services, management of the scarce and therefore expensive microwave spectrum becomes a critical issue [80]. The need for aggressive reuse of frequencies, as well as the expansion of wireless access to indoor environments such as houses or business offices [67], [86], has led to decreasing cell sizes (micro or pico-cells with radii of hundreds and tens of meters respectively). This in turn makes regular frequency reuse patterns very difficult to identify, due to the high variability of propagation conditions. A single, common frequency reuse is therefore desirable from both of the above points of view. The obvious drawback, as we saw in the previous chapter, is

the increased co-channel interference that, if not managed properly, can greatly reduce system capacity.

We discussed also in previous chapters the benefits to bandwidth efficiency, of employing packet switched access schemes in future cellular networks (through multiplexing of bursty traffic). From an interference point of view, packet access has the additional benefit of decreasing the average interference by the inverse of the duty cycle of the bursty traffic. The penalty paid here is that interference changes very fast when compared to circuit switched systems [67] and long term prediction that could be used e. g. for accurate closed loop power control [78] is impossible.

These spectrum and interference management considerations emphasize the importance of good medium access control (MAC) protocols, especially for the more challenging uplink (mobile to base) of packet switched systems. With few exceptions [70], [67], [89], MAC protocols usually deal with issues of multiple access and ignore the effect of co-channel interference. However, it is clear that in single frequency reuse environments the MAC protocol will additionally have to deal with the equally, if not more important problem of achieving a required signal to interference ratio (SIR) at the receiver, so that packets can be successfully captured.

Generally speaking, medium access control protocols try to manage co-channel interference by either averaging or avoidance. Interference averaging corresponds to CDMA systems, where each interfering transmission appears as random background noise to the intended receiver, while interference avoidance corresponds to TDMA systems where terminals can dynamically choose channels that are the least interfered. In [80], a comparison is presented on various techniques for increasing spectrum efficiency. Results show that interference averaging techniques

(like CDMA) perform better than fixed bandwidth assignment ones (like non-adaptive TDMA), while interference avoidance techniques (like adaptive TDMA or frequency hopped systems) can perform even better. Additionally, the combination of power control and dynamic channel assignment for circuit switched systems can provide two-fold or even three-fold capacity benefits over CDMA.

For packet switched systems, the use of CDMA is hindered by the difficulty of employing tight power control on a packet by packet basis, especially for bursty sources. Even if power control is feasible, interference avoidance techniques have the potential to achieve larger bandwidth efficiencies. In this respect, dynamic packet allocation (DPA) algorithms based on interference sensing and adaptive formation of good patterns of spatial channel reuse can help improve the average SIR and avoid worst case situations, and are therefore a natural complement to packet switched MAC protocols.

In this work we present two DPA algorithms that try to address the issues of reusing the same channels in every cell and managing the resulting fast changing interference, for the uplink of cellular packet switched systems. The algorithms are based on the packet reservation multiple access (PRMA) protocol and its variant, PRMA++, which we have described in earlier chapters. The first algorithm has PRMA as the underlying MAC and relies only on interference sensing and reassignment of a terminal based on the best predicted SIR channel, when the terminal's SIR falls below a threshold. The second algorithm uses the reservation slots of PRMA++ to allocate terminals to "good" channels in accordance to a form of channel segregation (CS), [87]. In addition to these algorithms a power control scheme is used that is found to greatly enhance performance by decreasing interference levels. A DPA algorithm that combined PRMA with CS and was presented in [67] is also compared here with our algorithms. Addi-

tionally, for comparison purposes we present results of a PRMA/CDMA system, where packet CDMA is used as the access protocol and the access probability is adaptively changed in an attempt to increase throughput.

We tested the algorithms with simulations for packetized voice traffic, which is a good indicator of the algorithms' performance, since it has strict delay requirements and demands relatively low dropping/interference probability (for example total packet loss rate of 1%). Three basic assumptions are made throughout this chapter:

- Round-trip delays are negligible and therefore a base station (BS) can acknowledge received packets immediately, as well as allocate slots to terminals. This assumption is valid in micro and pico-cellular environments [84].

- Each BS has knowledge of the long term attenuation (i. e. path loss and shadowing) of all of its terminals. This knowledge is easy to acquire at call-setup time and can be updated periodically (for pedestrian mobility the long term attenuation changes very slowly).

- Each BS can measure the interference in available slots in its frame (i. e. slots where no terminals are transmitting to the BS). This assumption is also justified if the slot duration is large enough (of the order of 1 ms).

## 7.2 DPA based on Reassignment/Retransmission

The short propagation delays encountered in micro and pico-cellular environments allow for timely retransmission of corrupted voice packets [84], as long as their lifetime has not expired. Here, we adopt a measurement based retransmis-

sion/reassignment scheme, whereby the BS upon receiving a corrupted packet schedules its retransmission and all further transmissions for the current talkspurt at a new available slot. The choice of the new slot depends on the remaining lifetime of the interfered packet and the latest interference measurement available for that slot. As seen in Figure 7.1, at any time $t$ each BS maintains a window

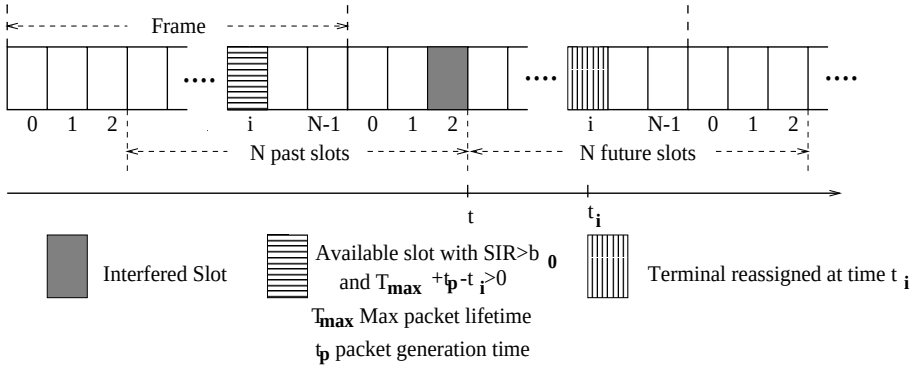



Figure 7.1: The reassignment scheme where slot $i$ satisfies the delay and SIR requirements of the retransmitted packet.

of $N$ past time slots, where $N$ is the frame size, together with a list of which of those are available, and their corresponding interference levels. Therefore, when the need for a reassignment arises, the BS uses these values to decide which of the future $N$ slots to assign. Assignment is done to that slot with the least measured interference that also satisfies the retransmitted packet's lifetime requirement. If no such slot is found, then the BS knows the packet will be dropped at the mobile and simply assigns the least interfered available slot.[1] The algorithm is described below (see also Figure 7.1):

[1]Note that for the system considered here with parameters as in Table 7.1, the remaining lifetime of the next packet after the one dropped will always be positive upon retransmission time.

**DPA1**:
At the BS upon reception of corrupted packet:
$N_{free} = 0$;
**for** i=0 **to** N-1 {
    **if**(slot i **NOT** occupied **AND** $L_{max} + t_p - t_i > 0$)
        Include slot i in list of acceptable slots;
        $++N_{free}$;
}
**if** ($N_{free} == 0$) {
    **for** i=0 **to** N-1 {
        **if**(slot i **NOT** occupied)
        Include slot i in list of acceptable slots;
        $++N_{free}$;
    }
}
**if** ($N_{free} > 0$)
    Reassign the terminal to the least interfered of the acceptable slots;
**else**
    Do not reassign the terminal to a new slot;

Reassignment upon reception of a corrupted packet is based on the following observation: when the system operates with good quality of service, for example with overall dropping/interference probability of less than 1%, terminals access and hence reserve slots with very high probability of success. Therefore, when a slot is interfered it is very likely that it will remain interfered in the future. Indeed, results reported in [67] for a PRMA system with vehicular mobility indicate a decorrelation time of about 10 frames for co-channel interference. For indoor applications where mobility will be significantly smaller, decorrelation time can be assumed to be even larger than that. Therefore, our retransmission scheme is justified, and additionally the interference measurements done in available slots, as described above, will be sufficiently reliable for slot allocation in the immediate future, as in Figure 7.1. Of course, a reallocated terminal may create excessive

interference to the new slot at an adjacent cell, triggering another reassignment, and so on. However, by using the lowest interference slots, which are the least likely to be occupied by terminals in adjacent slots and also employing power control (of which more below), this likelihood is minimized. Finally note that with this retransmission/reassignment scheme there are no extra slots allocated to retransmitted packets, i. e. each terminal still occupies only one slot per frame and packets are delivered to the BS in the correct, increasing sequence.

The DPA algorithm just described performs interference avoidance, that is, continuously attempts to assign the least interfered slots to terminals having reservations. This leads to a marked degradation of initial SIR (initial SIR is the SIR of the first packet from a new talk-spurt trying to access an available slot), since the "best" slots are allocated to terminals that already have reservations. As a result, access delay increases and a large number of packets are being dropped during contention, eventually limiting performance as traffic load increases. This observation together with the fact that the BS does not have direct control over the assigned slots from the beginning of a talk-spurt lead us to the next DPA algorithm.

## 7.3 DPA Based on Priority Ordering

Performance in single frequency reuse cellular systems can be enhanced if each BS can form a set of preferred channels that are consistently least interfered from adjacent cells. In such harsh interference environments channel occupancy per cell is much lower than what would be if a higher reuse factor was used. The idea therefore is instead of having an overall equal and low channel occupancy, to try to selectively increase the occupancy of certain slots that consistently have lower interference levels. Priority ordering algorithms, such as channel segregation [87],

attempt to adaptively form good channel reuse patterns (slot/frequency) by increasing the priority of acceptable channels and attempting to use first those with highest priority. This CS algorithm has shown remarkable performance improvement for circuit switched traffic [87] and has also been used successfully for the down-link of packet switched systems [89]. It was used in [67] for the up-link in an implicit scheme that we explain later. Inclusion of an explicit priority ordering algorithm in the up-link implies that the BS should have complete control over the channels assigned to mobile terminals. If terminals are allowed to contend for access at any slot in the frame, consistent reuse patterns cannot be formed by the BS. This naturally leads to the introduction of reservation (R) slots inside the frame in addition to regular information (I) slots, very much like the PRMA++ [54] protocol. For simplicity here, we assume that mobiles attempt to reserve a slot by transmitting the first data packet in an R-slot and not a separate reservation packet. From co-channel interference point of view the introduction of R-slots has two beneficial results:

- Assuming that all BS use the same R-slots, initial SIR as defined above is very good because interference levels in R-slots are very low. This is so, because occupancy in R-slots is low, since after initial access mobiles are reassigned by the BS to an I-slot. Of course, a sufficient number of R-slots must be allocated in a frame to avoid excessive intra-cell collisions.

- The BS has control over which slots to allocate to the terminal and hence can assign an available I-slot with highest priority, that also satisfies an SIR requirement.

A priority ordering and interference sensing algorithm is implemented as follows: each BS maintains a list of available slots, together with their priority and interference level most recently measured (same as Figure 7.1). Upon receiving a new

packet in an R-slot, the BS assigns the highest priority channel, that also exceeds an SIR threshold, and upon successful reception of the next packet in the new reserved slot, it increases the slot's priority by one. Initial priority assignment is random.

Dedicating certain slots for contention results in a reduction in available capacity for reserved terminals. On the other hand, when occupancy in a frame is temporarily very high (all of the I-slots are occupied) or many I-slots are interfered heavily, R-slots can be used as I-slots, i. e. can be allocated to terminals with reservations. Hence, the BS can assign an R-slot to a mobile terminal if there is no other slot available that satisfies the SIR requirement, and correspondingly a mobile terminal can attempt access in any slot if it finds an R-slot occupied. This is a more flexible scheme that allows better utilization of all the channels available to the BS. Note though, that this modification matters only for heavily loaded systems, since at light or moderate traffic loads there will almost always exist available slots with low interference. The algorithm is described below:

**DPA2:**
**A.**
**At the BS upon reception of a packet from new talk-spurt:**
**if** ( slot is R-slot){
$N_{free} = 0$
**for** i=0 **to** N-1 {
**if**(slot i **NOT** occupied **AND** slot i≠R-slot **AND** SIR($i$) > $b_0$)
Include slot i in acceptable slots; ++$N_{free}$;
}
**if** ($N_{free}$ == 0) {
**for** i=0 **to** N-1 {
**if**(slot i **NOT** occupied **AND** SIR($i$) > $b_0$)
Include slot i in acceptable slots;
++$N_{free}$;
}
}
**if** ($N_{free} > 0$)
Reassign the terminal to the acceptable slot having highest priority, and if new packet is received successfully in newly assigned slot, increase its priority be one.
**else**
Do not reassign the terminal to a new slot.
}
**else**
Increase priority of slot by one and do not reassign terminal;
**Upon reception of corrupted packet:**
Repeat algorithm 1, but assign slot with highest priority that also satisfies SIR> $b_0$, and if packet is received succeffully in newly assigned slot, increase priority by one.
**B.**
**At the mobile terminal upon arrival of a new talk-spurt:**
Attempt access in earliest R-slot with permission probability $p$.
**if** (R-slot occupied)
Obtain right to attempt access in any slot.

## 7.4 Implicit CS Scheme

In this DPA scheme proposed in [67] the BS allows contention only in a subset of available slots. The subset is determined according to a priority ordering CS algorithm, which is described below: If a packet is received correctly in a slot, then the priority of that slot is increased according to:

$$P_{\text{new}} = \frac{P_{\text{old}} N_{\text{old}} + 1}{N_{\text{old}} + 1} \tag{7.1}$$

$$N_{\text{new}} = N_{\text{old}} + 1 \tag{7.2}$$

where $N_{\text{old}}$ and $N_{\text{new}}$ denote the number of adaptations for the specific slot. Conversely, if a packet is not received correctly in a slot, its priority is decreased as follows:

$$P_{\text{new}} = \frac{P_{\text{old}} N_{\text{old}}}{N_{\text{old}} + 1} \tag{7.3}$$

$$N_{\text{new}} = N_{\text{old}} + 1 \tag{7.4}$$

The BS then, marks the $n$ highest priority free slots as available for access, where $n$ is a parameter for the system. In Fig. 7.2 we show a schematic of the TDMA frame and the priority list maintained by each BS.

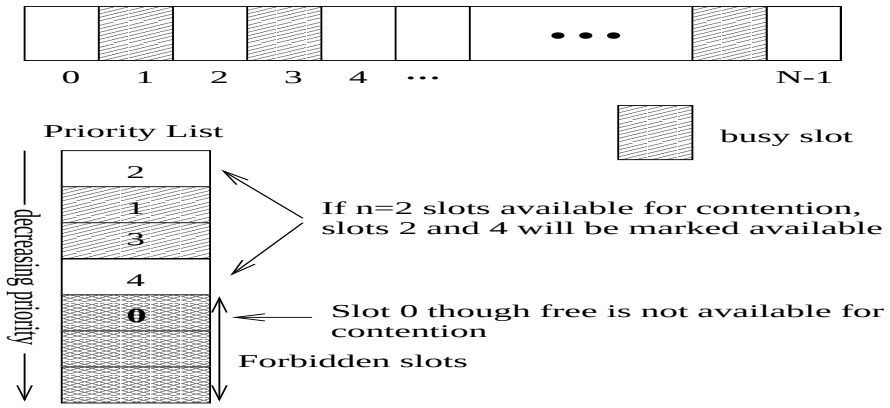



Figure 7.2: The implicit CS algorithm implemented in [67].

Since measurement based techniques always perform better than simple priority ordering schemes [89], here we use a slightly different CS algorithm, based on measurements of the interference of the slots. Thus, the BS makes available for contention a subset of the least interfered slots. For fairness of comparison we include the same retransmission and power control scheme as in our algorithms. The number of least interfered available slots where terminals are allowed to contend for access is the same as the number of R-slots in our DPA2 algorithm. We call this scheme DPA3.

## 7.5 Power Control

Signal strength based power control compensates the long-term attenuation (i. e. attenuation due to the path-loss $d^{-\eta}$ and shadowing) by controlling the transmit power of the terminals. In single frequency reuse environments it is absolutely necessary to use power control because it greatly reduces interference. In [90] it was proven that under a single dominant interferer assumption, the best power control policy is to compensate for half the long-term attenuation in dB ("square root" power control). That is, if we denote by $L$ the long term attenuation in dB, the transmit power should be $P_T = -\frac{L}{2}$ dB. In this work, in addition to square root power control we also devise a scheme for reducing transmit power when the received SIR is "too good", i. e. exceeds a second threshold above the capture ratio. This can be decided for example from the number of bits corrected by forward error correction, without the need for exact knowledge of the SIR. According to this algorithm, the power level of a packet to be transmitted will be decreased by $\Delta P$ if the SIR of the previous packet was determined to be above a threshold $b_1$ and the maximum number of adaptation iterations, $J_{\max}$ has not been reached yet. See Figure 7.3 for the flowchart. Note that when a

packet fails to be captured (SIR$< b_0$), the next packet will be transmitted again at the initial level determined by "square root" power control. In our simulations we assumed that $b_1 = b_0 + 6$ dB, $\Delta P = 1$ dB and $J_{\max} = 6$. By decreasing transmit power when SIR is more than adequate, interfering terminals interact less strongly and the DPA algorithms we described become more effective. SIR based power control that attempts to equalize SIR at the receiver (assuming the BS can estimate accurately the interference even at the occupied slots) was also tried, but was found to perform even worse than the simple square root power control, because of oscillations that resulted in power assignment. Of course, in the algorithm we propose, since power is only decreased, such instability problems do not occur.

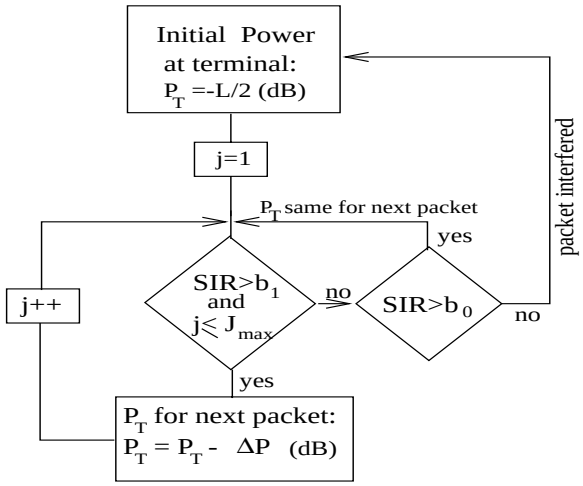



Figure 7.3: Power control algorithm.

## 7.6 A Hybrid PRMA/CDMA system

For fairness in comparison, we will assume a CDMA system with a channel rate equal to that of the PRMA system. In the CDMA system, the channel rate is

actually the chip rate, while each terminal transmits information at a rate that is $G/r$ times slower, where $G$ is the spreading gain and $r$ is the rate of the error correction code implemented. Suppose now that the channel rate is $R$ and there are $N$ slots in the PRMA system, while there are $L$ slots in the CDMA system. Then by equating the information rate of a terminal under the two systems we get:

$$\frac{R}{N} = \frac{R}{L}\frac{r}{G} \tag{7.5}$$

and solving for the spreading gain we get $G = \frac{rN}{L}$. Thus, if we use a one-half rate error correcting code, the PRMA system has $N = 20$ slots and the PRMA/CDMA system has $L = 1$ slot (i.e. the CDMA packet takes up the whole frame), we get a spreading gain $G = 10$.

We further assume that, the performance of the CDMA system is dominated by the co-channel interference that is present due to simultaneous intra and inter-cell packet transmissions. A widely used approximation in CDMA systems is that this interference is Gaussian. Employing this approximation, the bit error probability is given by the very well known equation [1]:

$$P_e = Q(\sqrt{\text{SIR}}) \tag{7.6}$$

where $Q(x)$ is the Gaussian tail function and SIR is the signal to interference ratio at the receiving BS. If tight closed loop power control is employed, all the packets transmitted by terminals within a given cell can be received by the BS at equal power $P_0$. In this case, the SIR (which is the same for all received packets) can be well approximated by [92]:

$$\text{SIR} = \frac{3GP_0}{(K-1)P_0 + P_I} \tag{7.7}$$

where $K$ is the number of intra-cell transmitted packets and $P_I$ is the collective interference coming from all other cells.

Finally, we will model the error correction code effect as a decrease in the capture ratio of the BS. This decrease corresponds to the coding gain of the code, which depends on the specific code used and the packet size. We will show a specific example in the results section below. The important thing to note here is that if the SIR is above the new lower capture threshold, all transmitted packets are received correctly by the BS.

In the hybrid PRMA/CDMA system that we just described, terminals access the channel in an uncontrolled manner, that is as long as a terminal has packets to transmit, it does so irrespective of the congestion in the cell. While this is a benefit of CDMA, and indeed at low radio occupancies all packets will be captured by the BS, at higher radio occupancies the system becomes unstable, much the same way as a regular PRMA network. What is needed is to adaptively control the probability $p$ with which terminals attempt to access the BS (the permission probability we have defined for PRMA), in such a way that the throughput is maximized. To see how this could be accomplished, call $K_{\max}$ the maximum number of packets the BS can receive with SIR above the capture ratio, at a specific interference level $P_I$. $K_{\max}$ can be readily evaluated from (7.7), where we assume unit received power:

$$K_{\max} = \frac{3G}{b_0} + 1 - P_I \tag{7.8}$$

where $b_0$ is the capture ratio, and we assume that the BS can measure the interference in a slot. In order to maximize throughput, we would like to have exactly $K_{\max}$ packets transmitted in the specific slot. Since there are already $K$ terminals reserved in the slot, only $K_{\max}$-K more terminals can be accommodated. There are a number of ways we could control the permission probability based on the above observation: for example, we could adjust $p$ such that the expected number of transmitting packets is $K_{\max}$, or we could adjust $p$ such that the prob-

ability that more than $K_{\max}$ packets are transmitted is very small. We show how to adaptively calculate $p$ for either case in appendix B. Alternative methods for adapting the permission probability for packet switched CDMA systems are presented in [93] and [94].

Parenthetically, note that the same control of the permission probability $p$ can also be applied to the regular PRMA system. However, time slots are much smaller in the PRMA system and much more control data would be transmitted by the BS. Additionally, the benefit will be much greater in the CDMA system, since many packets are transmitted in a single slot.

## 7.7 Simulation Model

A hexagonal cellular pattern was considered with 37 cells, arranged in six rings of six cells each, around a central cell. The same frequency was assumed for every cell and statistics were collected from the central cell. Rayleigh fading is ignored, assuming enough diversity at the receiver.

Simulation at the connection level was not considered, instead a fixed number of static terminals was generated uniformly over the service area at each run, and connected to the strongest BS. Each run consisted of 300000 slots. Packet loss rate results were averaged over several runs for a fixed total number of terminals. Note here that the actual number of terminals per cell is a random number.

For the PRMA/CDMA system we assumed a (1023, 578, 49) BCH code [95], that gives a coding gain of 4.5 dB at packet error probability of $10^{-2}$. See appendix C for the details of calculating this coding gain. We also assume a single slot per frame, and therefore the spreading gain is $G = 11.3$, as can be seen from equation (7.5). For the adaptation of the permission probability $p$ we considered

Table 7.1: System and simulation parameters for voice traffic.

| **P**RMA/PRMA++ PARAMETERS | |
|---|---|
| Channel rate | 722.5 Kb/s |
| Frame duration | 16 ms |
| Slots per frame | 20 |
| Data bits/packet | 512 |
| Header bits/packet | 66 |
| Maximum packet lifetime | 32 ms |
| Mean talk-spurt duration | 1 s |
| Mean silence duration | 1.35 s |
| # of R-slots (PRMA++) | 3 |
| Permission probability $p$ | 0.2 |

| **S**IMULATION PARAMETERS | |
|---|---|
| # of base stations | 37 |
| Capture ratio ($b_0$) | 8 dB |
| Capture ratio ($b_0$) (CDMA) | 3.5 dB |
| Spreading gain ($G$) (CDMA) | 11.3 |
| Path loss exponent | 4 |
| shadow fading $\sigma$ | 8 dB |
| radio occupancy | varying |

two cases: adapting $p$ such that the expected number of terminals accessing the BS is $K_{\max}$, given by (7.8), or adapting $p$ such that the probability that more than $K_{\max}$ terminals access the BS is smaller than 0.2[2]. We call these two last systems as CDMA1 and CDMA2 respectively.

A voice only system is considered, where the usual two state Markov model [52] for the voice terminals is used, with parameters as in Table 7.1. Results are reported versus the average radio occupancy, which we define as the ratio of the *e*ffective traffic load in Erlangs (average number of mobiles per cell times the duty cycle) over number of slots in the frame. Table 7.1 summarizes all the parameters of the simulation.

[2]This value resulted from experimentation as the optimum for total packet dropping probability.

## 7.8 Numerical Results

Figure 7.4 shows total packet loss probability (dropping and interference) versus radio occupancy for DPA1, DPA2, DPA3, CDMA1, and CDMA2. The DPA3 algorithm fails completely in the harsher interference environment and cannot meet the 1% total packet loss rate even for very light traffic loads. On the other hand, both of the new algorithms work remarkably well up to very high system loading. The DPA2 algorithm (slot reassignment and CS combined) performs worse than DPA1 at low radio occupancies. Due to the fact that only three contention slots are available, DPA2 has a higher rate of intra-cell collisions, which leads to more packets being dropped. However, after about 25% occupancy, the benefits of channel segregation on co-channel interference outweigh intra-cell collisions and DPA2 performs better than DPA1. On the other hand, after about 34% radio occupancy, packet loss for DPA2 increases more rapidly than for DPA1, which is again due to the limited number of contention slots. At $P_{\text{loss}} = 10^{-2}$ DPA2 achieves a radio occupancy of 35%, which is quite remarkable. Finally, the two PRMA/CDMA systems perform almost identically, and very close to DPA1.

Figure 7.5 shows SIR distributions for DPA1 and DPA2, both for packets at initial access and for packets with reservations at a radio occupancy of 34%, which corresponds to an average number of 16 terminals per cell. The difference between initial access and reserved packet SIR for DPA1 demonstrates clearly that this algorithm performs interference avoidance: by reassigning terminals when packets are interfered, reserved slots have very good SIR and those slots with high interference levels are left empty (available) for contention[3]. Both of

[3]Note that in the initial access SIR distributions we took into account only cases of a single intra-cell transmission. Obviously simultaneous intra-cell transmissions will result in collisions which will degrade initial, SIR. Here though, we are interested in studying the co-channel

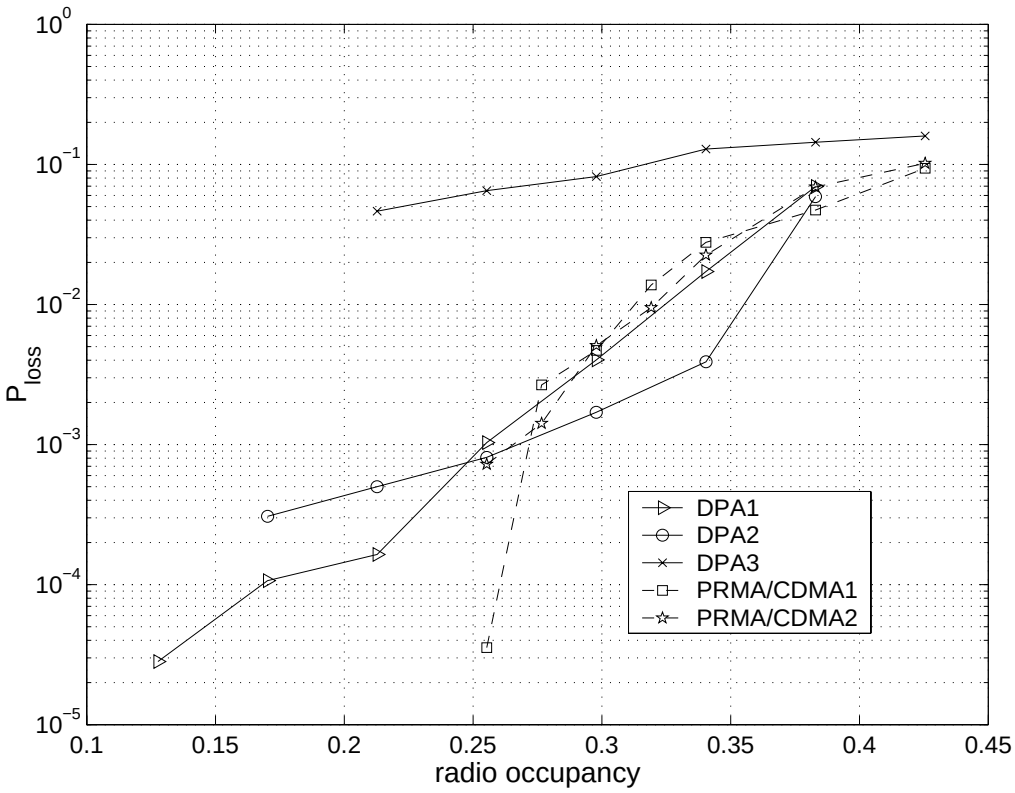



Figure 7.4: Total packet loss probability v. s. radio occupancy.

the corresponding SIR curves for the DPA2 algorithm are better at this radio occupancy which justifies the lower $P_{\text{loss}}$. Also notice that Figure 7.5 implies that the reassignment probability even at this high load is close to 1%.

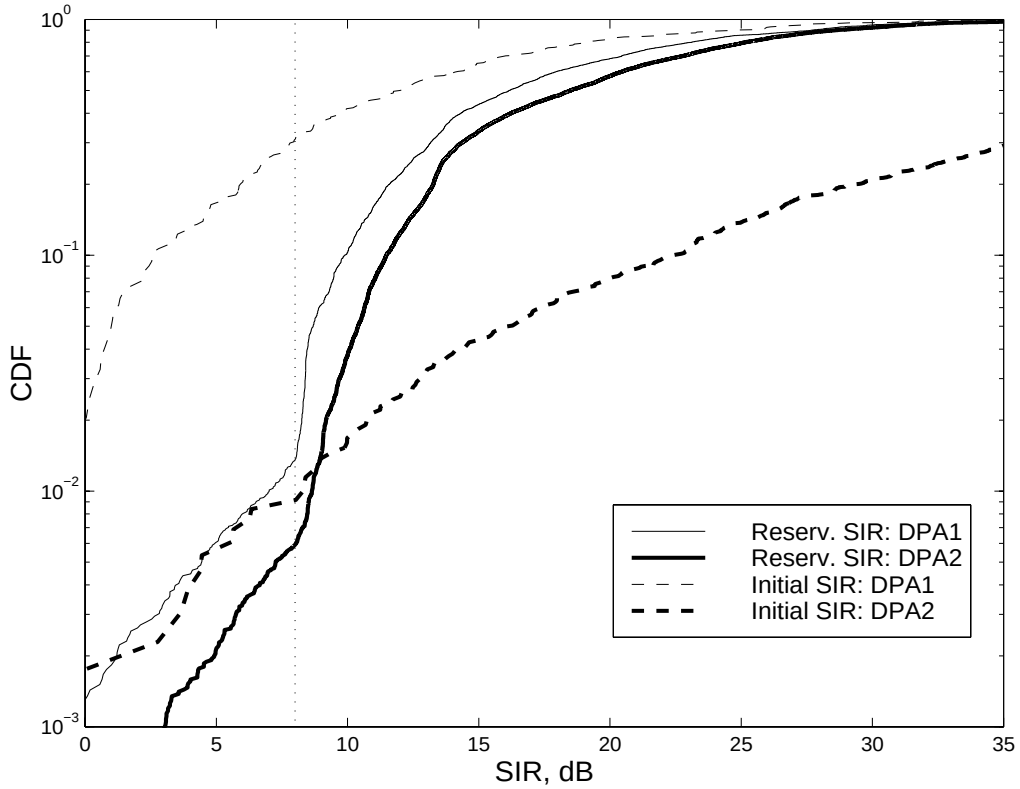



Figure 7.5: Cumulative distribution function of SIR for the two algorithms considered at 34% radio usage.

Figures 7.6 and 7.7 show slot usage probabilities at the central cell for DPA2 and for radio occupancies of 34% and 38.3%, for a single run of the simulation (i. e. the relative frequencies with which individual slots are reserved). Note that slots 0, 7 and 14 are the R-slots. It is clear from these graphs that DPA2 can successfully create a set of preferred slots for moderate traffic load, while at higher loads the I-slots become more equiprobable. Occupancy of R-slots remains

interference avoidance properties of the DPA algorithms.

small in both cases, which explains why for DPA2 $P_{\text{loss}}$ increases faster at higher radio occupancies. Thus, capacity is compromised by the use of R-slots despite the flexibility of our algorithm. Finally, Figure 7.8 shows slot usage probabilities obtained with DPA3. The implicit CS scheme is unable to create good reuse patterns and therefore slots are almost equiprobable, hence the severely degraded performance.

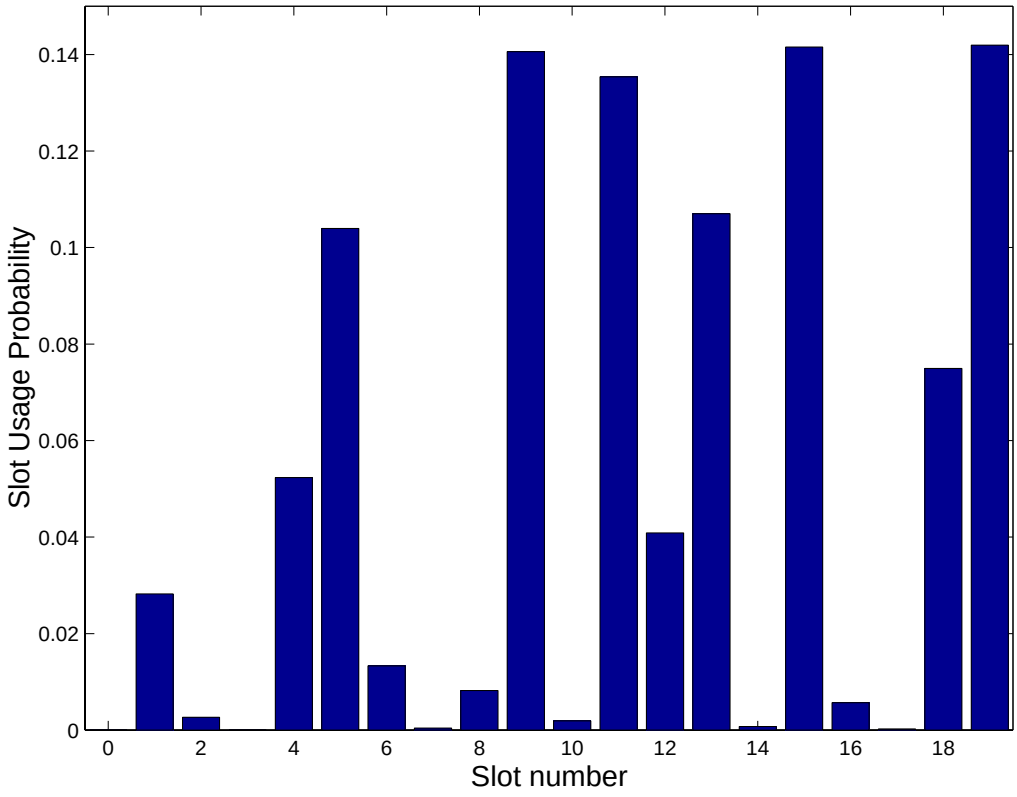



Figure 7.6: Slot usage probability of DPA2 in central cell at 34% radio occupancy.

## 7.9 Summary

In this chapter we proposed two dynamic packet allocation schemes for the uplink of cellular packet switched systems. Through simulations, we compared their

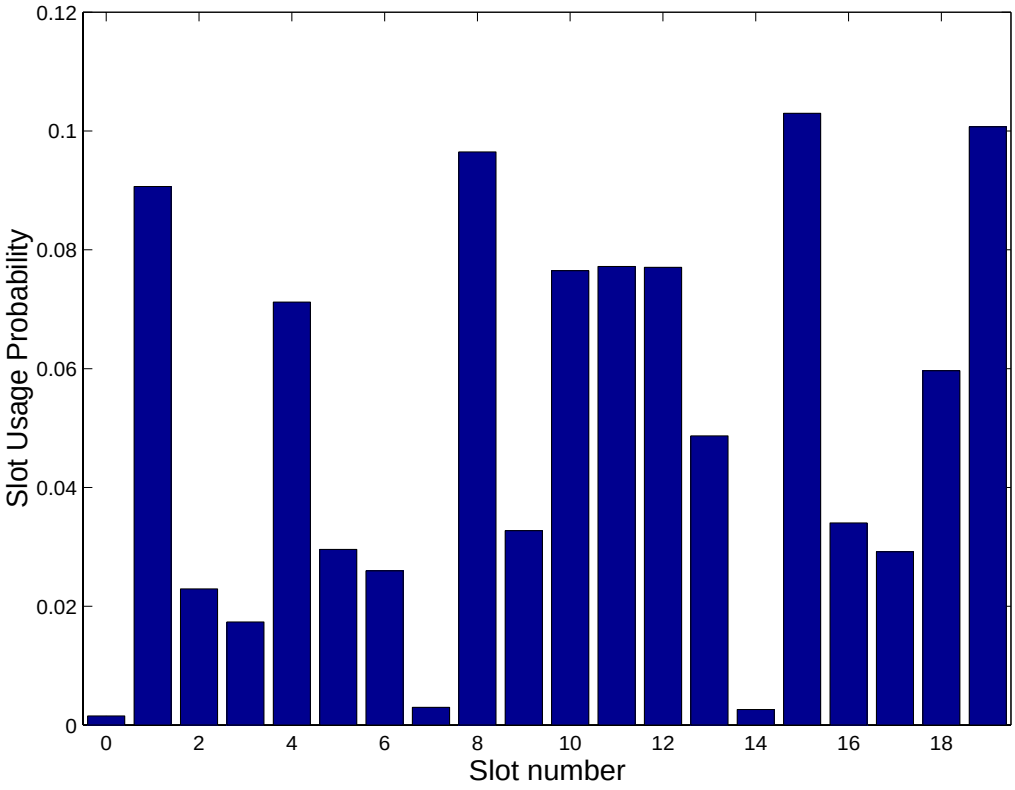



Figure 7.7: Slot usage probability of DPA2 in central cell at 38.3% radio occupancy.

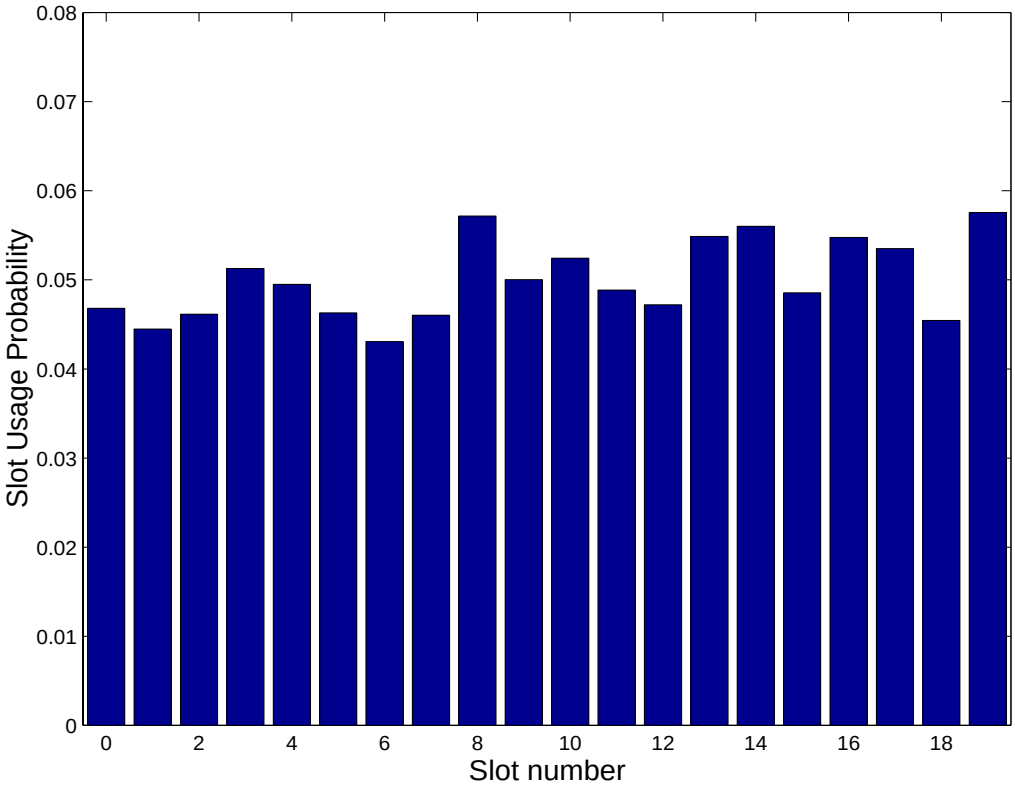



Figure 7.8: Slot usage probability of DPA3 in central cell at 34% radio occupancy.

relative performance as well as their performance against a third DPA algorithm already proposed in the literature [67] and a hybrid PRMA/CDMA scheme. Our algorithms are based on interference sensing and priority ordering as well as on power control. Both algorithms show very good performance. They achieve better than 35% radio occupancy at a total dropping/loss probability of 1%, whereas the previously reported algorithm fails completely in the harsher interference environment of our simulations. The PRMA/CDMA systems perform better at lower radio occupancies, but as radio occupancy (and hence interference) increases, there is a cross point above which the DPA algorithms outperform them. This result has actually been reported before in the literature [88]. Of course, the processing gain we assumed for the PRMA/CDMA system may seem very small, however we made that choice in order to have a fair comparison with the TDMA systems. Additionally, we should note that existing wireless products for home/office networking use comparable processing gains (e.g. the IEEE 802.11 standard specification for the physical layer has a processing gain of 10 [48]). On the other hand, if a higher processing gain is assumed then the PRMA/CDMA capacity will increase linearly with processing gain, but with a linear increase of available bandwidth the PRMA system capacity, with DCA, should also increase comparably.

Our algorithms attest to the fact that relatively simple interference measurement and power control techniques are able to achieve the highly desirable goal of single frequency reuse in third generation wireless services. Our algorithms were tested for voice traffic. Of course, it can be argued that future wireless traffic will be far more data-centric than voice-centric. However, voice is a good benchmark for testing DCA algorithms, because it has very strict delay requirements and is also rather long-hold, therefore it should provide a worst cast bound. Video traffic will present very similar qualitative performance with voice, since it has

similar characteristics. However, very bursty data traffic has very different characteristics. Fortunately, very bursty traffic will naturally randomize interference and therefore simple retransmission schemes will work very well, without any interference measurements. This fact has been noticed and verified in [84].

# CHAPTER 8

# Multicasting Real Time Traffic Over Ad-Hoc Networks: A Routing and Medium Access Control Algorithm

## 8.1 Introduction

In the previous chapters we dealt with medium access control (MAC) protocols for packet switched cellular networks. The main characteristic of such networks is the presence of an infrastructure of fixed base stations, where each can control the access mechanism of the mobile terminals that lie within its area (cell). An implicit result of the presence of base stations is that the wireless access is limited only to a single hop: from the mobile to the base station and vice-versa. Thus, from a connectivity point of view, cellular networks have a *star topology*, with the base station at the center and the mobile terminals stemming from it.

This chapter will deal with a different kind of wireless networks, namely with *ad-hoc* networks. Ad-hoc, or multi-hop wireless networks are *infrastructure-less* networks with no fixed routers, or base stations. Two main features distinguish them from the conventional cellular networks: the lack of centralized (base station) control and the fact that mobile node interconnection is achieved in a peer to peer, multi-hop fashion. That is, if a mobile terminal wants to communicate with another mobile terminal that is out of its range, it must use intermediate

terminals as relays to get its traffic forwarded to the final destination. The intermediate terminals used for forwarding data are usually assumed to be of the same type as the source and the destination, and they are also roaming about in the same area as the source and destination. Thus, from a connectivity point of view, ad-hoc networks have a more general and random graph topology than the star topology of cellular networks. Example applications of ad-hoc networks are emergency search-and-rescue operations, business or conference meetings in which the participants want to quickly and seamlessly inter-network, military operations, and sensor networks.

Two main problems need to be solved efficiently for relaying traffic in ad-hoc networks: the first one is *routing* of data from the source through multiple hops to the destination. Since the wireless access is not single hop any more, as in cellular networks, a source needs to know a path towards each potential destination. This path information can be either explicit, containing all the intermediate relays to the destination, or implicit, containing only the neighbor node that is the first relay towards the destination. Since the member nodes of ad-hoc networks are usually mobile, the routing protocol needs to adapt fast to changing topologies by finding new routes to destinations and deleting stale ones. The second problem is that of medium access control. Since ad-hoc networks do not have a star topology, there are additional impairments in accessing the wireless medium, apart from simple contention at the receiver. Two of the most prominent ones are the *hidden terminal* and the *exposed terminal*. The hidden terminal problem can be explained with the help of Figure 8.1, where four terminals are shown, all having the same communication range (only C's range is shown for clarity). Say that terminal A has initiated communication with terminal B, transmitting packets to it. Terminal C, being out of range of terminal A's transmission, is unaware of it and therefore can also attempt to initiate communication with

terminal D. However, at this point a collision will take place at terminal B, which will not be able to receive A's transmissions any more. We say that terminal C is a hidden terminal for B. The exposed terminal problem can be considered the dual of the hidden terminal and is explained with the help of Figure 8.2. Say that terminal B has now initiated transmission towards terminal A. Also, assume that terminal C wants to initiate transmission towards terminal D. However, terminal C is sensing the wireless medium busy due to B's transmission and therefore will defer from transmitting to D, in case it would cause interference to some neighboring node. It is clear though that D's transmission in this case would not cause any interference (since terminal A is out of range of terminal D) and that D is unnecessarily deferring. We say that C is an exposed terminal to terminal B.

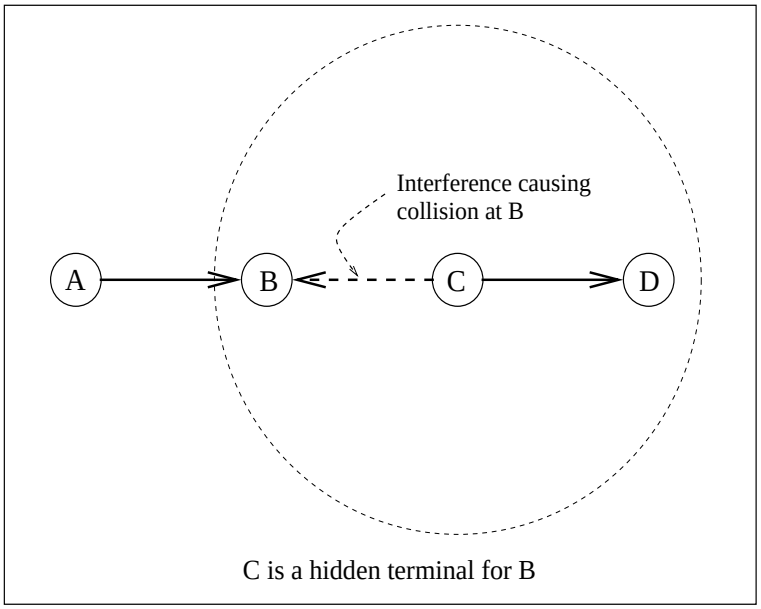



Figure 8.1: An illustration of the hidden terminal problem.

In this chapter we will tackle both the routing and medium access control problems for a specific application of ad-hoc networks, namely multicasting real-

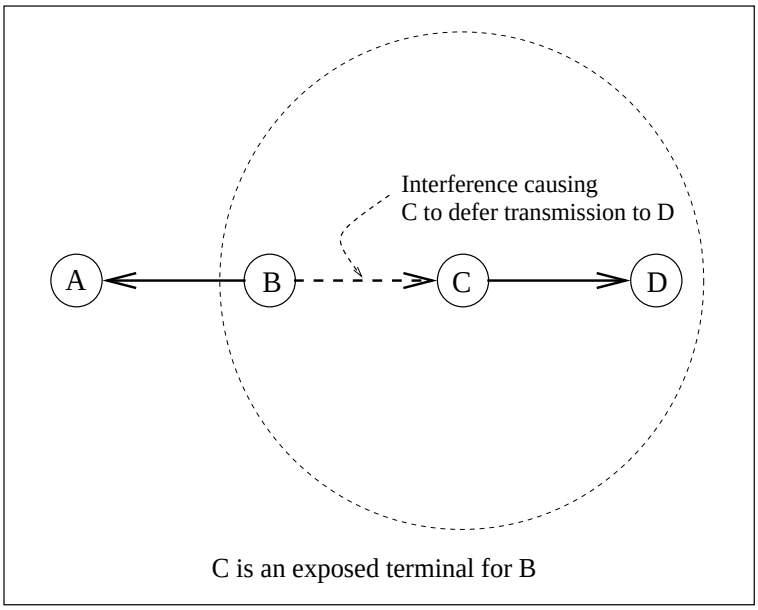



Figure 8.2: An illustration of the exposed terminal problem.

time traffic. Multicasting, in its simplest form that we will study here, consists of a single source transmitting the same information to many receivers, which we call the members of the multicast session. Consider Figure 8.3, where we show an ad-hoc network with one terminal as the source of the multicast data and all the other terminals as the receivers. Here, we can identify both problems we discussed above: we need a protocol that identifies routes from the source to all the destinations of the ad-hoc network (routing) and additionally allows efficient sharing of the wireless medium so that collisions are minimized[1] and the hidden and exposed terminal problems are not present. Specifically, we will present a protocol that simultaneously performs affiliation of terminals to the multicast session, solving the routing problem, and resource reservation, solving the problem of efficient medium access. The multicast application is specifically geared towards real-time traffic, that being either constant bit rate (CBR) or

[1]The protocol we will present actually eliminates collisions in the absence of mobility.

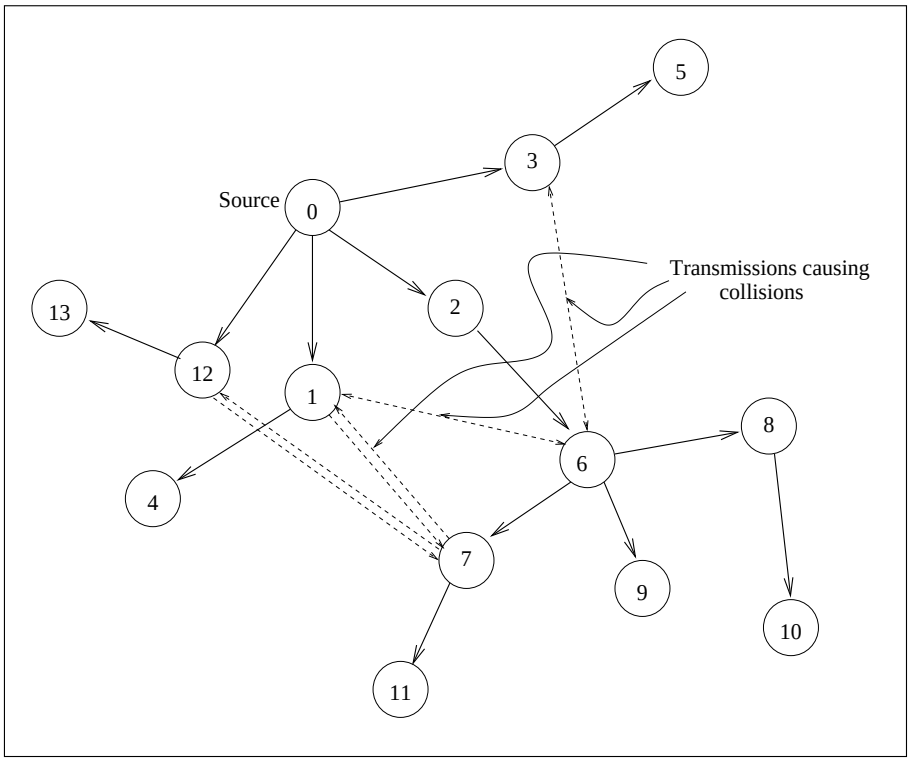



Figure 8.3: An ad-hoc network where a single source (terminal 0) multicasts data to all other terminals.

variable bit rate (VBR). In CBR traffic, the source generates periodically packets, with a specified and fixed period. VBR traffic on the other hand is semi-periodic: the source generates packets with a specific mean period (and hence a specific mean rate) but can occasionally generate more or less packets within a period. Usually applications that generate such kind of traffic are delay sensitive, that is the packets generated can only tolerate a very limited delay[2]. Therefore, the MAC protocol must be able to provide explicit reservations to such packets (in much the same way PRMA provides reservations for voice packets in cellular systems), in order to guarantee their prompt delivery to the destination. Since we are dealing with an ad-hoc network here, these reservations must actually take place on multiple hops, so that an end-to-end bandwidth reservation exists.

In the following sections we will explain in detail our affiliation/reservation protocol and show performance results, in terms primarily of throughput under terminal mobility.

## 8.2 Problem Definition

The problem, as we discussed earlier is the following: we have an ad-hoc network with a single source that needs to multicast real-time traffic (either CBR or VBR) to a set of receivers. We need to design an affiliation and resource reservation protocol such that all the receivers of the multicast data can receive the real-time traffic collision free. Additionally, new members of the multicast session should be able to join by negotiating bandwidth reservations with already existing neighboring members of the session. The protocol should be able to adapt to mobility, by re-negotiating affiliations and bandwidth reservations as links break and collisions start to happen. Finally, the affiliation mechanism should be such

[2]Such applications are usually voice or video.

that the number of relays is minimized, as is the number of hops to reach the most distant receivers.

Since we need bandwidth reservations, bandwidth will be reserved in terms of time division multiple access (TDMA) slots for a multicast session and the multicast structure will be built so as to avoid collisions and jamming of transmissions of multicast group members. CBR multicast data will be streamed through the reserved channels from the source to the destination periodically, thereby ensuring a bound on delay jitter, while for VBR traffic we will study the trade-off between reserved and random-access bandwidth in order to achieve a certain packet loss rate. Our approach is that of concurrent TDMA *collision-free slot scheduling* and *route building,* so that a multicast mesh is created with reserved bandwidth along the different routes.

A lot of work has been previously done on broadcast scheduling in ad-hoc networks, which is akin to our problem [100], [101], [102]. In [100], broadcast scheduling was shown to be a NP-complete problem. The broadcast scheduling problem can be stated as follows:

- In an ad-hoc network, how can one schedule transmissions among nodes, such that, when a node transmits a packet, every neighboring node receives it? Note that the scheduling protocol needs to ensure that when two nodes transmit, their packets do not collide at a third node (hidden terminal). It should also ensure that two neighboring nodes do not transmit at the same time, since they will be unable to listen to each other's transmissions.

Any of the algorithms developed for solving the above problem could be used to multicast data from a single source in an ad-hoc network. However, multicasting is different and has properties that can be exploited to improve the spatial re-use of bandwidth (slots):

- In multicasting, not all nodes need to transmit: only a subset of nodes relay packets to all other nodes.

- Neighboring nodes that relay multicast packets from a single session to different "children" can simultaneously transmit packets if they do not jam the reception of any other neighboring node. That is, nodes that relay at the same time the same packet, do not need to hear each other's transmissions, because these transmissions are not intended for each other.

In our protocol, it is up to the receivers to join the multicast session and bandwidth is reserved by means of signaling packets exchanged between one hop neighbors. In order to satisfy its bandwidth requirement, a node can connect to different "parent" nodes in different slots. Multicasting is therefore done on more robust *mesh* [99], instead of a tree. Some links will eventually fail due to mobility, but by using the same signaling mechanism, receivers will change parent nodes, bandwidth will be reassigned and the multicast mesh will be re-configured for the new topology. Through simulations we find that at pedestrian speeds, throughputs above 90% are achieved, when our protocol is used.

## 8.3 TDMA Frame Structure

We assume half-duplex transceivers and hence, nodes cannot transmit and receive simultaneously. Therefore, since we are dealing with periodic traffic (either CBR or VBR), we assume that the slotted time is grouped into super-frames, consisting of two frames each, frame 0 and frame 1. In one of these frames a node receives data and in the other it transmits, if it is a relay. We denote the transmit frame of node $N_i$ by $F_i$; $F_i$ takes on a value of 0 or 1 depending on whether $N_i$ can transmit in frame 0 or in frame 1 respectively. Figure 8.4 depicts an example, showing a

super-frame and its constituent frames, the first one of which is a transmit frame while the second, a receive frame. The reserved and random-access portions of each frame are also shown.

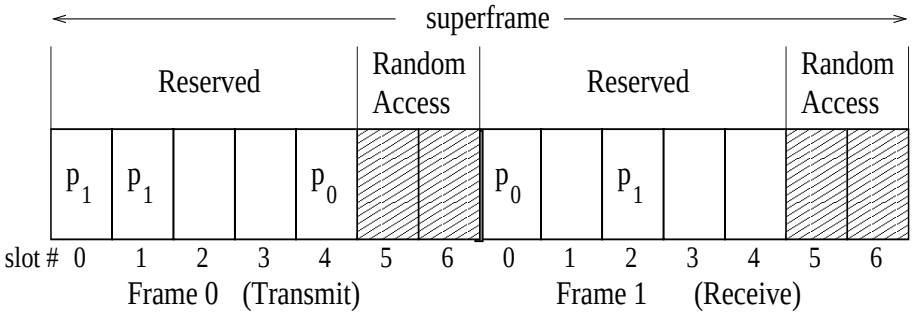



Figure 8.4: TDMA frame structure. The node receives in frame 1 and transmits in frame 0 (hence $F = 0$).

Multiple slots can be reserved for a session by a node, and used in each frame. Packets transmitted by the source in each frame are numbered sequentially for that frame and each node marks its receive (transmit) slots with the *f*rame-sequence number of the packet to be received (transmitted) in that slot. For example, Figure 8.4 shows that a node receives packets with frame-sequence numbers 0 and 1 in slots 0 and 2 and transmits packet 0 in slot 4 and packet 1 in slots 0 and 1 (multiple transmissions are for supporting different children that cannot receive in the same slot).

Packets in a frame that are in excess of the reserved number of slots are transmitted/received in the random-access portion of the frame (this would happen in the case of VBR traffic for packets in excess of the current mean source rate). The parent node notifies its children of the specific random-access slots they have to listen to by appending the relevant information to packets transmitted during the reserved portion of the bandwidth.

Transmit scheduling for signaling (i.e. control information) between nodes is

done in a round-robin fashion on a separate channel. That is, we assume that there is a second slotted channel, with much slower rate than the data channel, where there are $N$ slots per frame, with $N$ being the number of terminals in the ad-hoc network. Each terminal has a unique slot where it transmits its signaling information. These transmissions are obviously done collision free and therefore all one hop neighbors of the transmitter receive the signaling packet. The type of control information exchanged will be described in detail in a later section. Note here that we have assumed a very simple signaling scheme, which is also not scalable in the number of nodes of the ad-hoc network. The implications of a more realistic scheme which would allow for collisions in signaling packets have not been researched yet, but will be left as future study. We should note though that the effect of collisions in signaling packets is that certain nodes may have stale information about their neighbors, due to missing their signaling information. This could lead to more unsuccessful attempts to reserve bandwidth, thus deteriorating performance. An obvious remedy would be to increase the signaling rate, in hope of updating more often the state of each node to its neighbors' data-bases. Unfortunately, increasing signaling rate beyond a point will lead to more collisions. In any case, such trade-offs should be studied with simulations in order to deduce useful conclusions.

## 8.4 Channel Model and Connectivity

We assume that all nodes transmit with the same power, $P_T$. The power received by a node $N_0$ due to the transmission of a node $N_j$ is given by $P_{j0} = G_{j0}P_T$, where $G_{j0}$ denotes the power attenuation on the path between nodes $N_0$ and $N_j$. Here, we assume that $G_{ij}$ follows a simple propagation decay law, i.e. $G_{ij} = d_{ij}^{-\eta}$, where $d_{ij}$ is the distance between nodes $N_i$ and $N_j$ and $\eta$ is the power loss exponent.

The signal to interference ratio (SIR) at node $N_0$, when node $N_j$ transmits, is given by:

$$\text{SIR} = \frac{G_{j0}}{\sum_{i \neq j} G_{i0}} \tag{8.1}$$

where nodes $N_i$ have frame numbers $F_i = F_j$, and transmit in the receive slot of node $N_0$. A packet is received correctly if its SIR is above a certain threshold $\gamma$. Finally, we assume that, in the absence of co-channel interference a node can communicate with nodes up to a distance $d_{\max}$ away. This distance typically depends on the transmit power $P_T$, the noise floor and the sensitivity of the receivers. Note here that, the larger the transmit power, the fewer will be the number of relays that will exist in the multicast mesh and thus, the required bandwidth for that multicast session will be smaller. On the other hand, with smaller transmit powers, more nodes will need to relay packets, and bandwidth requirements for that session will also increase.

In the following sections it will be assumed that nodes have knowledge of the path losses to all of their one-hop neighbors. Indeed, the slow variations of the channel, due to path loss and shadow fading are easily tracked, in contrast to fast fading variations due to multipath which, usually, cannot be estimated easily [80].

Since nodes might have more than one reserved slot in a TDMA frame, they can rely on different parents to receive packets in the different slots. This adds a degree a flexibility to the reservation process, since a single parent node may not be able to relay certain packets in a frame, either because it does not have them, or because it does not have enough slots to transmit all the packets, or because some of its children may experience too high an interference in the slots where the parent can transmit the packets. Consider for example Figure 8.5. Let all the nodes shown in the figure have three slots reserved for receiving packets $p_0$, $p_1$, and $p_2$. Initially, node $N_0$ relays these packets to node $N_2$ in slots 2, 3, and

4. However, node $N_1$ that is relaying the same packets to node $N_3$ in slots 0, 1, and 2 strays into the neighborhood of node $N_2$, and thereby causes interference to node $N_2$'s transmission in slot 2. Thus, node $N_2$ will no longer receive packet $p_0$. If node $N_0$ cannot relay $p_0$ to $N_2$ in some other slot (for any of the reasons mentioned above), $N_2$ can receive that packet from node $N_1$, and therefore should now attempt to become a child of that node also, with the intent to receive a packet $p_0$ from that node in slot 0. Thus, in this scenario nodes receive packets from multiple parents.

It should be noted here that, in the case of VBR traffic, there is a drawback in having multiple parents relay packets to a single child: if there are packets to be transmitted in random-access slots, these will be transmitted by all parents, although only one actually needs to transmit, for the packets to be received by the child. This will increase the possibility of collisions in the random-access slots. Considering Figure 8.5 for example, if nodes $N_0$ and $N_1$ are parents of node $N_2$, and they have packets that must be relayed to $N_2$ in random-access, both $N_0$ and $N_1$ will transmit these packets, whereas only $N_0$'s transmission would be enough for the packets to be relayed to $N_2$.

## 8.5 Overview of the Protocol

Our protocol deals primarily with the transmission scheduling problem for a multicast session in a wireless ad-hoc network. The mechanism will be presented, by which a node reserves TDMA slots and attaches itself to the multicast mesh, when at least one of its neighbors is a multicast member. We do not consider the case in which none of the neighbors is a multicast member. In such a case, a search might be initiated to find a route from the member to the existing multicast mesh and all the nodes along the route from that member to the mesh

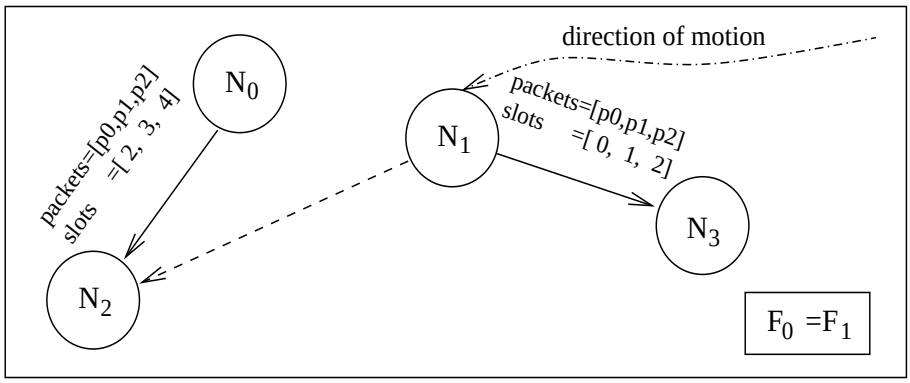



Figure 8.5: Node $N_1$'s transmissions interfere the reception at node $N_2$ in slot 2. Node $N_2$ can connect to node $N_1$ to receive packet $p_0$ in slot 0, if it cannot receive $p_0$ in another slot from $N_0$.

will be forced to join the multicast session as relays [99]. Our protocol is receiver initiated, i.e. it is up to the receiver to attempt to connect to the multicast mesh. No special routing information is needed at relay nodes; they would only need to maintain the IDs of their children, the slots with which they transmit to them, and the frame-sequence numbers of the packets they transmit in each slot. Of course, each node will try to attach itself to relays that are as close to the source as possible, but minimum hop routing is a secondary consideration, when a node cannot reserve enough bandwidth from minimum hop relays. Finally, maintenance information, which will be detailed below, is only exchanged among one hop neighbors and is not propagated to more distant nodes.

### 8.5.1 Data Structures Maintained at Each Node

A node is characterized by a set $\mathcal{Q}$ (for a particular multicast session), defined as:[3]

[3]Here and throughout this work, bold typeface denotes a vector quantity.

$$\mathcal{Q} = \{ID, HC, F, RS, RS_{\text{des}}, RS_{\text{min}}, \mathbf{PID},$$
$$\mathbf{RxSlot}, \mathbf{RxSeq}, \mathbf{SIR}, \mathbf{G_p}, \mathbf{TxSlot}, \mathbf{TxSeq}, \mathbf{US}\} \quad (8.2)$$

where:

- $ID$: the node's unique ID;
- $HC$: the node's hop-count, i.e. the number of hops from the source;
- $F$: a bit which indicates the transmit frame of the super-frame (Figure 8.4);
- $RS$: the number of receive slots that are *c*urrently reserved by the node;
- $RS_{\text{des}}$: the desired number of reserved slots, as estimated by the node, based on the current traffic load;
- $RS_{\text{min}}$: the minimum acceptable number of reserved slots. This depends on the acceptable packet loss rate[4]. Note that, $RS$ and the available number of random-access slots determine the *c*urrent packet loss rate at the node;
- **PID**: a vector which represents the parents of the node, $PID[i]$, $i = 0, 1, \ldots, RS - 1$. Note that all parents must have identical hop counts, and that some or all of the elements of $PID[i]$ might be identical (i.e. the same parent transmits multiple packets in each frame to its child);
- **RxSlot**: a vector which lists the receive slots of the node, $RxSlot[i]$, $i = 0, 1, \ldots, RS - 1$;

[4]Note that packets that are transmitted in random-access slots may not be received correctly if the SIR at the receiver node is below a certain threshold. These packets will be considered lost.

- **RxSeq**: a vector indicating the frame-sequence in which packets are received in the receive slots, $RxSeq[i]$, $i = 0, 1, \ldots, RS-1$. The node receives packet with sequence number $RxSeq[i]$ in slot $RxSlot[i]$. This is required for sorting packets according to their frame-sequence numbers;

- **SIR**: a vector containing the SIR of the node in its receive slots **RxSlot**, $SIR[i]$, $i = 0, 1, \ldots, RS - 1$;

- $\mathbf{G_p}$: a vector representing the path losses between the node and each of its parents, $G_p[i]$, $i = 0, 1, \ldots, RS - 1$;

- **TxSlot**: a vector which lists the slots in which the node transmits (empty, if the node is not a relay);

- **TxSeq**: a vector indicating the sequence in which packets are transmitted in the transmit slots. That is, the node transmits packet with frame-sequence number $TxSeq[i]$ in slot $TxSlot[i]$;

- **US**: a vector which lists the slot numbers of the transmit slots unusable by the node. These are slots in which the node cannot transmit since such a transmission would cause excessive interference to some neighboring node that receives in the same slot.

As long as a node has at least $RS_{\min}$ slots reserved and it successfully receives packets from its parents in these slots, it will consider itself to be connected to the multicast mesh. If the node receives less than $RS_{\min}$ packets in the reserved portion of the bandwidth, it becomes disconnected, and will set its $RS = 0$ and will attempt to reconnect to the multicast session.

### 8.5.2 Signaling Information Exchanged

Nodes participating in the multicast session take turns to transmit their status information on a signaling channel in a round-robin fashion[5]. Specifically, a node $N_i$ will transmit the following set in its own signaling slot:

$$\mathcal{Q}_T = \{ID, HC, F, RS, \mathbf{RxSlot_1}, \mathbf{RxSeq_1}, \mathbf{PID_1}, \mathbf{TxSlot}, \mathbf{TxSeq}, \mathbf{US}\} \quad (8.3)$$

where $\mathbf{RxSlot_1}$, $\mathbf{RxSeq_1}$ and $\mathbf{PID_1}$ contain vectors **RxSlot**, **RxSeq** and **PID** respectively, and potentially some additional terms, as explained below.

This information, collected by neighboring nodes, has a dual purpose:

a) Node $N_i$, through its broadcast signaling message, can attempt to reserve more receive slots, if $RS < RS_{\text{des}}$. This includes the case when $RS = 0$, i.e. when the node is disconnected. The node then, would append the required extra receive slots to vector **RxSlot** of equation (8.2). It will also append the potential parents from whom it can receive in these slots and the corresponding frame-sequence numbers of the packets it is missing, to vectors **PID** and **RxSeq** respectively (if $RS = 0$, then these vectors are all empty). Thus, vectors $\mathbf{RxSlot_1}$, $\mathbf{PID_1}$ and $\mathbf{RxSeq_1}$ of (8.3) are created.

b) Neighboring nodes will update their neighborhood data-bases by means of the data broadcasted by $N_i$. The entry, $NB_i$, of the neighborhood data-base contains the fields of $\mathcal{Q}_T$, the estimated path loss between the two nodes ($NB_i.G$) and the time stamp which represents the last time that node $N_i$ sent an update message ($NB_i.t$). In addition, neighboring nodes will update their corresponding packet schedules (**TxSlot** and **TxSeq**), and the field representing their unusable slots (**US**), as will be explained

[5]Since signaling information is transmitted in round-robin fashion, it is interference free and is received by nodes up to a distance $d_{\max}$ away.

below. To this end, nodes maintain a *c*hildren data-base and an *u*nusable slots data-base. An entry, $CH_i$, of the children data-base has the fields:

- $CH_i.ID$: the child's ID;
- $CH_i$.**RxSlot**: the vector of *r*eceive slots of the child, i. e. slots in which the child receives from the specific node;
- $CH$.**t**: each component of this vector, $CH.t[i]$ denotes the last time at which the node transmitted to the child in the corresponding receive slot $CH.RxSlot[i]$.

An entry $US_i$ of the unusable slots data-base has the fields:

- $US_i.s$: the position of the unusable slot in the frame;
- $US_i.t$: the time at which this entry was last refreshed.

The node can render a slot usable again, if the slot is not refreshed in the *US* before a timeout period.

### 8.5.3 Scheduling Rules

The goal of the multicast scheduler is to create and maintain a multicast mesh such that:

- each node successfully receives all desired packets transmitted in the reserved portion of the bandwidth;
- the reserved bandwidth adaptively changes in order to accommodate the dynamics of the mean source rate, and satisfies the minimum required bandwidth for a specific source rate.

- the maximum distance in terms of number of hops of any node from the source, in the multicast mesh, is minimized.

It is obvious from the definitions above that a node $N_0$ that is a relay and has transmit frame $F_0$ in the super-frame, will cause co-channel interference to the reception of all nodes (except its children of course) that have a receive frame $F_0$ and receive in the transmit slots of $N_0$. Nodes with receive frame $F_0$, whose SIR drops below the threshold $\gamma$ because of node $N_0$'s transmission, will not able to receive correctly, packets destined for them in the interfered slots. Hence, node $N_0$ should not transmit in slot $s$, if the SIR of any receiving neighbor in that slot drops below $\gamma$ as a result of the transmission. Moreover, note that two or more nodes (not necessarily neighbors) could at the same time, and independently, decide that slot $s$ can be used for transmissions, if such transmissions do not bring the SIR of any simultaneously receiving neighboring node below $\gamma$. However, by transmitting simultaneously (without knowing of each other's transmission), they can eventually jam some receiving node. To minimize the possibility of such occurrences, the SIR threshold for rendering a slot unusable could be set higher than $\gamma$. The first scheduling rule can be stated as follows:

**Rule 1** *A node $N_0$ with transmit frame $F_0$ cannot transmit in slot $s$ if there is at least one Node $N_j$ in its neighborhood, having receive frame $F_0$, such that $N_j$ is not a multicast child of $N_0$, receives in slot $s$ and*

$$SIR(N_j) = \frac{G_{ij}}{\sum_{k \neq i} G_{kj} + G_{0j}} < \gamma_1, \tag{8.4}$$

*for some preset $\gamma_1$, such that $\gamma_1 > \gamma$, and where $N_i$ is a parent of $N_j$ and nodes $N_k$ have transmit frame $F_k = F_0$ and transmit in slot $s$.*

Note that in order for node $N_0$ to decide whether it can use slot $s$, it needs to know the resulting SIR of all its neighbors that would be interfered within slot

$s$. Furthermore, to estimate the resulting SIR, $N_0$ must have knowledge of the factors $G_{ij}$ and $\sum_{k \neq i} G_{kj}$ of equation (8.4) for all neighbors $N_j$. In the event that such information is deemed too expensive to be relayed, rule 1 can be simplified to the following:

**Rule 1 (a)** *[No SIR information] A node $N_0$ with transmit frame $F_0$ cannot transmit in slot $s$ if there is at least one node $N_j$ in its neighborhood with receive frame $F_0$, such that $N_j$ is not a multicast child of $N_0$ and receives in slot $s$.*

Obviously, this is a more restrictive rule than the first one and will lead to somewhat higher bandwidth requirements. Furthermore, note that since nodes maintain information only about one-hop neighbors, a node $N_0$ conforming with rule 1 can still jam another node that is more than one-hop away and receives in the same frame and slot as those in which $N_0$ transmits. Without information from nodes more than one hop away, such a possibility cannot be excluded, but if the transmit power levels are small such effects are negligible, as simulations have also confirmed.

A node $N_0$ wishing to reserve a slot $s$ for receiving from a potential parent node $N_j$ can do so if $N_j$ is able to transmit in slot $s$ and if the SIR of $N_0$ in $s$ is above a certain threshold (which will be taken to be higher than the threshold necessary for correct reception of a packet). Therefore, the second scheduling rule is stated as follows:

**Rule 2** *A node $N_0$ can reserve slot $s$ for receiving from a neighboring node $N_j$, if:*

- *$N_0$ has receive frame $F_j$;*
- *$N_j$ can transmit in slot $s$, as determined from rule 1;*

- *the SIR of $N_0$ in slot $s$ satisfies:*

$$SIR(N_0) = \frac{G_{j0}}{\sum_{k \neq j} G_{k0}} > \gamma_2, \tag{8.5}$$

*for some preset $\gamma_2$, such that $\gamma_2 > \gamma$, and where nodes $N_k$ have transmit frame $F_k = F_j$ and transmit in slot $s$.*

Note here that, node $N_0$ can easily estimate the SIR in slot $s$ by just knowing the path loss $G_{j0}$ to its potential parent node and the interference in slot $s$. Both the path loss and the interference can be easily estimated [80].

### 8.5.4 Connection Procedure

In order to explain the connection procedure of the protocol, assume that node $N_0$ has a smaller number of reserved slots than what is desired, but the node is not disconnected from the multicast mesh, that is $RS_{\min} \leq RS < RS_{\text{des}}$ for $N_0$. First, $N_0$ will determine the frame-sequence numbers of the packets from the multicast session that it is not receiving in the reserved portion of the bandwidth.[6] Then, using its neighborhood data-base $NB$, the node will look for neighbors that are already transmitting these packets (for supporting other nodes) whose hop-count is less than its own by one. This will be done by inspecting the fields $NB_j.\mathbf{T}xSeq$ in the database, for all neighbors $N_j$ such that $NB_j.HC = HC - 1$. For each such neighbor $N_j$, that transmits one of the missing packets, say packet with frame-sequence number $NB_j.TxSeq[k]$ in slot $NB_j.TxSlot[k]$, node $N_0$ will estimate the SIR for $N_j$'s transmission in that slot. If $N_0$ senses a receive power $P$ in that slot, then an estimate of the SIR, for $N_j$'s transmission would be:

$$\text{SIR} = \frac{NB_j.G \cdot P_T}{P - NB_j.G \cdot P_T}. \tag{8.6}$$

[6]Note that packets with frame-sequence numbers from 0 to $RS_{\text{des}} - 1$ will be scheduled to be received in the reserved portion of the bandwidth, while packets with larger frame-sequence numbers will be received in the random-access portion of the bandwidth.

If the estimated SIR is larger than the threshold $\gamma$, node $N_0$ can receive the corresponding missing packet from node $N_j$. Therefore, it will append $NB_j.ID$, $NB_j.TxSlot[k]$ and $NB_j.TxSeq[k]$ to **PID**, **RxSlot** and **RxSeq**, respectively, and include these fields in the signaling packet. This process continues until either node $N_0$ meets its desired reservation bandwidth, $RS_{\text{des}}$, or there is no other neighbor that is already relaying the missing packet, and from which $N_0$ could receive that packet.

If after this procedure, node $N_0$ is still missing multicast packets, it will identify neighboring nodes with hop-count one less than its own that can add transmit slots to their $\mathbf{T}xSlot$, in order to relay the missing packets. To this end, for each neighbor $N_j$, $N_0$ will examine the set of slots left after subtracting the slots in $NB_j.\mathbf{T}xSlot$ as well the slots in $NB_j.\mathbf{U}S$. For each one of the remaining slots (if there are any), $N_0$ will estimate the SIR, if $N_j$ were to relay a packet to $N_0$ in that slot:

$$\text{SIR} = \frac{NB_j.G \cdot P_T}{P}. \tag{8.7}$$

where $P$ is the power sensed in the slot. If the $SIR > \gamma$, $N_0$ can receive packets from $N_j$ in the specific slot and will proceed to make $N_j$ its parent. It will add $N_j$'s ID, the specific slot and the frame-sequence number of the missing packet to vectors **PID**, **RxSlot** and **RxSeq** respectively.

If node $N_0$ is disconnected from the multicast mesh ($RS = 0$), it also follows the above procedure, but it first tries to use as parents neighbors with the minimum hop count. If $N_0$ cannot find $RS_{\min}$ slots, to receive packets from these neighbors, it will try to obtain the packets from neighbors with a hop count greater than the minimum by one, and so on. It is worth noting that, since all parents of a node have the same hop count, routing cycles are precluded.

As mentioned earlier, node $N_0$ will broadcast a maintenance packet in its

signaling slot. This packet will contain the fields shown in equation (8.3). When node $N_0$'s neighbors receive this broadcast packet, they will first examine the field which lists the parent ID's of $N_0$, $\mathbf{P}ID_1$. Each neighbor node $N_j$ will find elements $PID_1[k]$ that contain its own ID. For each such element, $N_j$ will check whether $N_0$ is in its children data-base, and is already receiving packets with frame-sequence number $RxSeq_1[k]$ in slot $RxSlot_1[k]$ from $N_j$. If so, $N_j$ will simply refresh the pertinent time entry in its children data-base, i.e. $CH.\mathbf{t}$. If $N_0$ is not in its children data-base, node $N_j$ will check if it already relays packets with frame-sequence number $RxSeq_1[k]$ in slot $RxSlot_1[k]$ to some other children. If this is the case, then $N_j$ will add node $N_0$ to its children data-base as a child receiving the above mentioned packets in the above mentioned slot. Finally, if $N_j$ is not transmitting in slot $RxSlot_1[k]$, and this slot is not unusable (i.e. it is not in the $\mathbf{U}S$ of $N_j$), $N_j$ will add $N_0$ as its child as in the above case. It will also add slot $RxSlot_1[k]$ to its transmit slots vector $\mathbf{T}xSlot$ and $RxSeq_1[k]$ to the transmit sequence vector $\mathbf{T}xSeq$.

Next, each neighbor node $N_j$ will consider the elements $PID_1[k]$ of $\mathbf{P}ID_1$, that do not contain its own ID. This would imply that node $N_0$ will be receiving packets from different neighbors in the corresponding slots. For each such element, for which $ID \neq PID_1[k]$, $N_j$ will add the slot (of the transmit frame), with number $RxSlot_1[k]$ to its unusable slot vector $\mathbf{U}S$, but only if it is not already using this slot to transmit. By adding slot $RxSlot_1[k]$ to $\mathbf{U}S$, node $N_j$ will preclude itself from using this slot in the future to relay packets, and hence potentially interfere with $N_0$'s reception in the same slot. Note that, $N_j$ might be able to use slot $RxSlot_1[k]$ for transmission, if $N_0$'s SIR in that slot was sufficiently high in spite of such a transmission. In that case, $N_j$ would have to be able to estimate $N_0$'s SIR and that would entail $N_0$ broadcasting SIR and interference information for each one of its receive slots. While this variation is possible, it is not considered

here.

It is to be noted that, if the network were to be static, the protocol is designed so as to ensure that nodes that are already connected to the multicast mesh are not jammed by new members that try to connect. Consider for example, Figure 8.6. Node $N_0$ broadcasts its maintenance packet at time $t$. Suppose that node

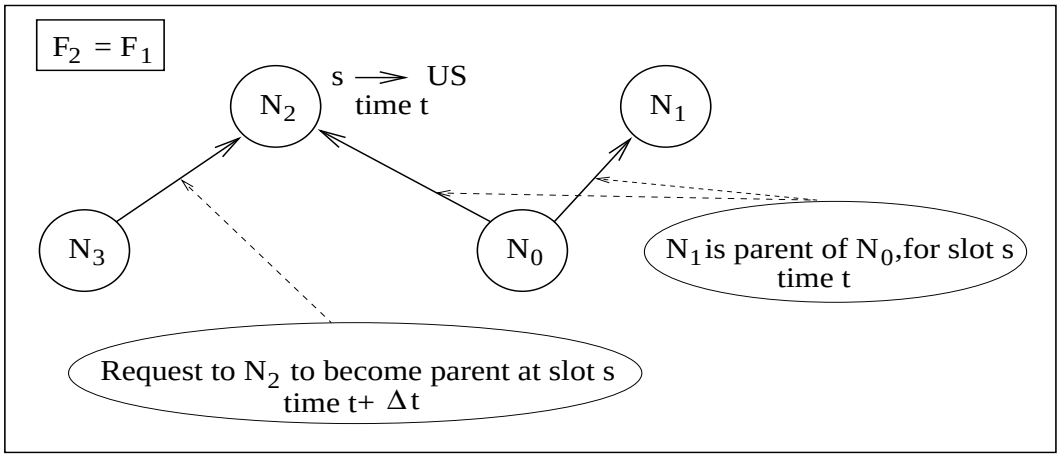



Figure 8.6: The reception of node $N_0$ is protected from jamming since slot $s$ is in the unusable slot vector of node $N_2$.

$N_1$'s ID is contained in $\mathbf{P}ID_1$ for slot $s$, either because $N_0$ is already connected to $N_1$ and receives from $N_1$ in slot $s$, or because it requests for the first time to receive packets from $N_1$ in slot $s$. At that time, neighbor $N_2$, upon receiving $N_0$'s message, will update (or add) slot $s$ in its unusable slots data-base. Assume also that Node $N_2$ has not yet broadcasted its maintenance packet, when node $N_3$ transmits a maintenance packet in which $N_2$ is identified as a parent which would transmit in the same slot $s$ (this implies that $N_2$'s signaling slot comes after that of $N_3$ in the round robin signaling period). Node $N_2$ will simply ignore this message since slot $s$ is already in its unusable slot vector, and hence cannot be used for transmission. Thus, if node $N_0$ has successfully connected to node

$N_1$, its connection is protected from being jammed. Node $N_3$ will learn of the updated unusable slots of node $N_2$ later in the same signaling period, and in the next signaling period may try to connect to $N_2$ by means of another slot, or may attempt to connect to another node. On the other hand, if $N_0$ was attempting to connect to node $N_1$ and this attempt failed for some reason (for example $N_0$ also had "stale" information with regard to slot $s$, and this slot is already in the unusable slot vector of $N_1$) slot $s$ will not be updated in the unusable slots data-base $US$ of $N_2$ for some time and will subsequently be rendered usable again.

Of course, under mobility, as nodes move into the vicinity of each other, excessive interference will be caused in reserved slots and nodes will lose connectivity with some of their parents. Then, following the procedures outlined earlier, they will try to re-affiliate to the multicast mesh.

## 8.6 Protocol Performance

Performance of the protocol we presented was evaluated using Parsec, a C based, discrete event parallel simulation language, developed at UCLA [73]. Fifty nodes were considered and were dispersed in an area of one square kilometer. In all cases, the minimum SIR required for a packet to be successfully received was set at $\gamma = 10$ dB and the path loss exponent $\eta$ was set to four. We consider CBR traffic to begin with, and each node needs to reserve only one slot in the superframe, for receiving packets ($RS_{\text{des}} = RS_{\text{min}} = 1$). The first simulation we performed assumed that all nodes were static. Our aim was to compute the number of slots, as well as the number of relays required in order to connect all the nodes to the multicast mesh. We computed this number for various values of the maximum transmission range $d_{\text{max}}$. Results from one hundred different

Table 8.1: Number of slots per superframe and number of relays.

| $d_{\max}(m)$ | Slots | Relays |
|---|---|---|
| 250 | 8 | 19 |
| 300 | 8 | 14 |
| 350 | 6 | 9 |
| 400 | 6 | 9 |
| 500 | 4 | 6 |

node positions were found and averaged, and are presented in Table 8.1. The number of slots reported is per superframe. This table shows that, as expected (since for a given node the number of neighbors increases with $d_{\max}$), fewer relays are required for larger values of $d_{\max}$, in order to have complete connectivity, and hence fewer slots are sufficient to support the multicast session. As $d_{\max}$ decreases, the average degree of nodes decreases. Thus, more relays are needed to support the multicast session, and one expects that the number of slots required for complete connectivity increases. However, spatial re-use of TDMA slots increases with decreasing $d_{\max}$, since the co-channel interference also decreases in this case; for $d_{\max} = 250$ meters, 19 relays are needed and 8 slots are required, while for $d_{\max} = 500$ meters, 6 relays and 4 slots are enough.

Next, we evaluate the performance of the protocol under mobility. We fix the values of $d_{\max}$ to be 250 meters. We also assume that there are ten slots per superframe assigned for the session, although from table 8.1 we see that only eight are needed on the average. The extra slots will improve performance in terms of received packets, as will be seen below. We assume that, out of the 50 nodes, 10 nodes are mobile and all mobile nodes move with the same speed. Nodes randomly select a direction (an angle is chosen uniformly between 0 and

360 degrees) and move along a straight line. They choose a new random direction after a random interval which is exponentially distributed with a mean value of one minute. Finally, nodes that reach the boundary of the roaming area simply bounce back into the area by choosing a new random direction. A signaling period is assumed to follow every 1000 data slots. Results were computed for 7 different combinations of mobile nodes (beginning from the same initial network) and an average of the combinations is presented in Figure 8.7 in the form of percentage of multicast data packets received (throughput) versus the speed of the mobiles. Two cases are considered, wherein, the source node is either static or mobile. Observe that, if the source node is static, the throughput remains very high (90%) even at high speeds. However, when the source node is mobile, the throughput suffers as speed increases.

Throughput improves with the number of slots available for reservation. Figure 8.8 presents throughput versus number of slots in the superframe assigned for the multicast session, for the case in which the source is mobile and the speed considered is 15 Km/h. As it can be seen from Table 8.1 and Figure 8.8, although only 8 slots are needed on the average for full connectivity of the static network, throughput under mobility would be poor when only 8 slots are assigned. However, by assigning 10 slots for the multicast session, we observe a 10% improvement in throughput.

Next, we consider the case of VBR traffic. Here, we assume that nodes can estimate accurately the mean traffic rate [103], [104][7], and focus on the problem of partitioning a total number of slots into reservable and random-access portions, such that a specific packet loss rate is achieved. Obviously, by assigning reserved

[7]It must be noted that estimating the mean traffic rate of a VBR traffic source is not an easy problem. Also, during abrupt increases of the mean, packets will be queued and will be relayed mostly through random-access slots, until more bandwidth is reserved, thus increasing the packet dropping rate. We do not consider these problems in our simplified model.

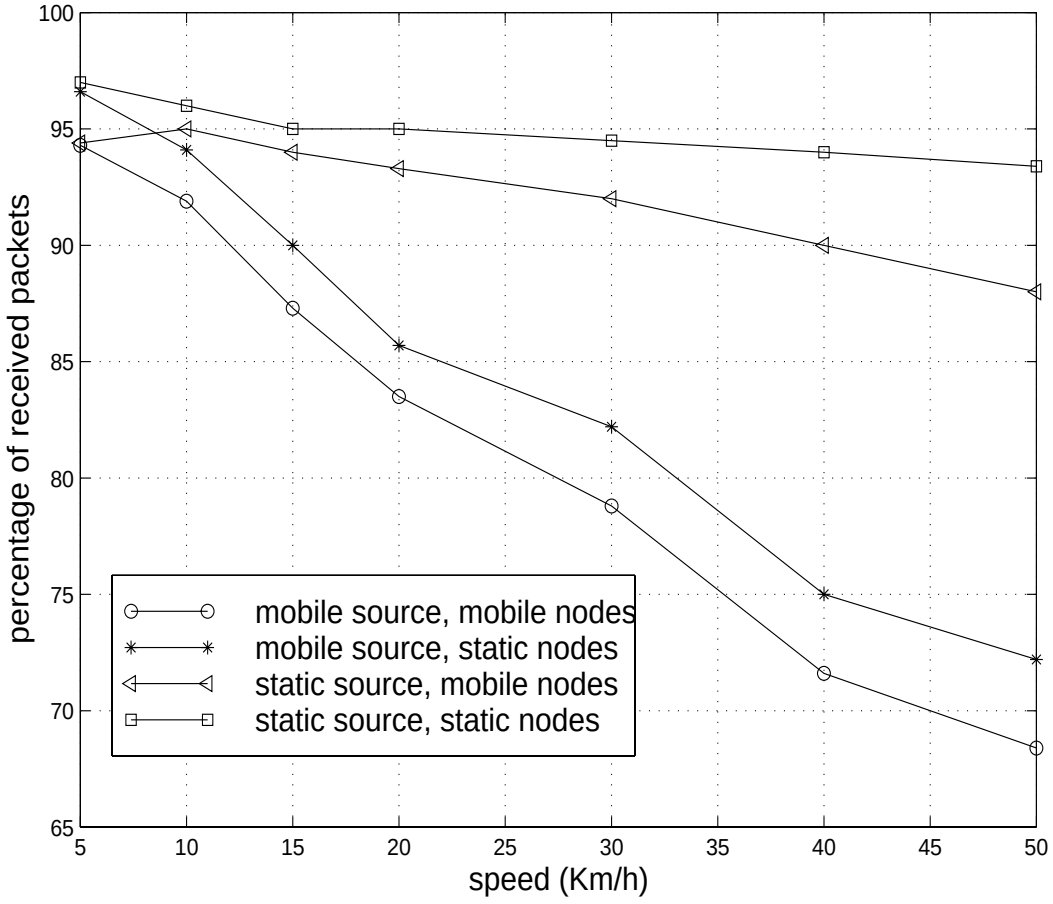



Figure 8.7: Percentage of received packets versus speed, for 10 mobile nodes and the cases of the source being static and mobile.

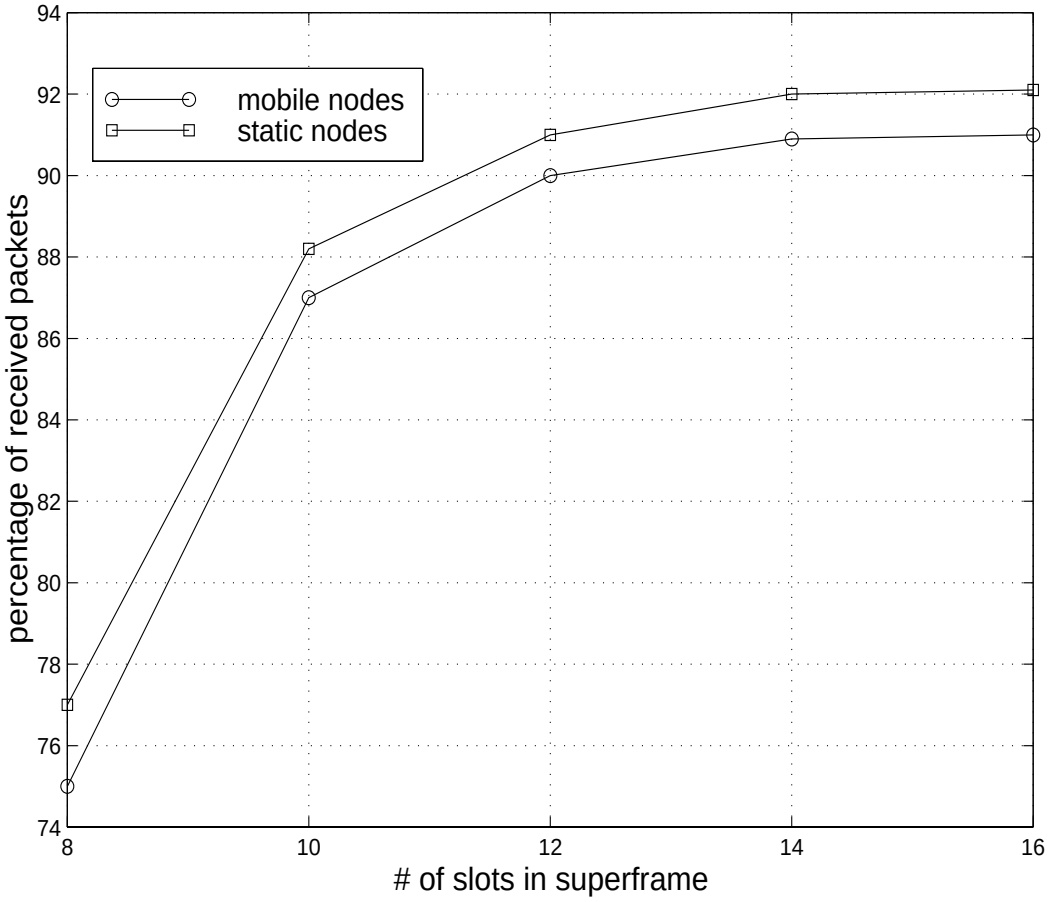



Figure 8.8: Percentage of received packets versus # of slots in superframe, for mobile source and speed of 15 Km/h.

slots to support the peak source rate, the packet loss rate will be zero (in a static network) but bandwidth utilization will be very poor. The question is, how much bandwidth above the mean do we have to assign to a multicast session in order to have an acceptable packet loss rate? In our simplified model, the packet arrival process at the source is assumed to be Poisson distributed, with mean rate of one packet per superframe (with the assumptions made above we can scale the mean to be one). Packets that are in excess of the reserved bandwidth are transmitted in the random-access portion of the transmit frame. There is no queuing and therefore delay jitter is bounded by the duration of one superframe (10 ms). We have set $RS_{\min} = RS_{\text{des}}$ and the total number of slots to be allocated for the multicast session to enable contention free scheduling is obtained by multiplying $RS_{\text{des}}$ by 8 (which is the number required for $RS_{\text{des}} = 1$, as seen from Table 8.1). Figure 8.9 presents the packet loss probability versus total available number of slots in a superframe, for different $RS_{\text{des}}$, for a static network of 50 nodes and with $d_{\max} = 250$. Statistics were gathered from nodes with the maximum hop-count, which are expected to have the worst packet loss rate. From Figure 8.9, note that the abscissa of the left-most point of each curve corresponds to the total number of reserved slots, which is a multiple of eight. The corresponding packet loss rate is the fraction of packets per frame that are in excess of $RS_{\text{des}}$, since there are no random-access slots in this case and these packets are dropped at the source. From Figure 8.9, it is obvious that assigning only one reserved slot ($RS_{\text{des}} = 1$) for each node, and receiving all packets in excess of the first one in the random-access portion, results in very severe packet loss rate. However, for $RS_{\text{des}} = 2$ (in which case, as can be seen from Figure 8.9, there are 16 reserved slots and about 10% of packets are transmitted using random-access), using 24 random-access slots, for a total of 40 slots in the superframe, the packet loss rate is slightly above 1%. Note here that $RS_{\text{des}} = 2$ corresponds to the mean source

rate plus one standard deviation.

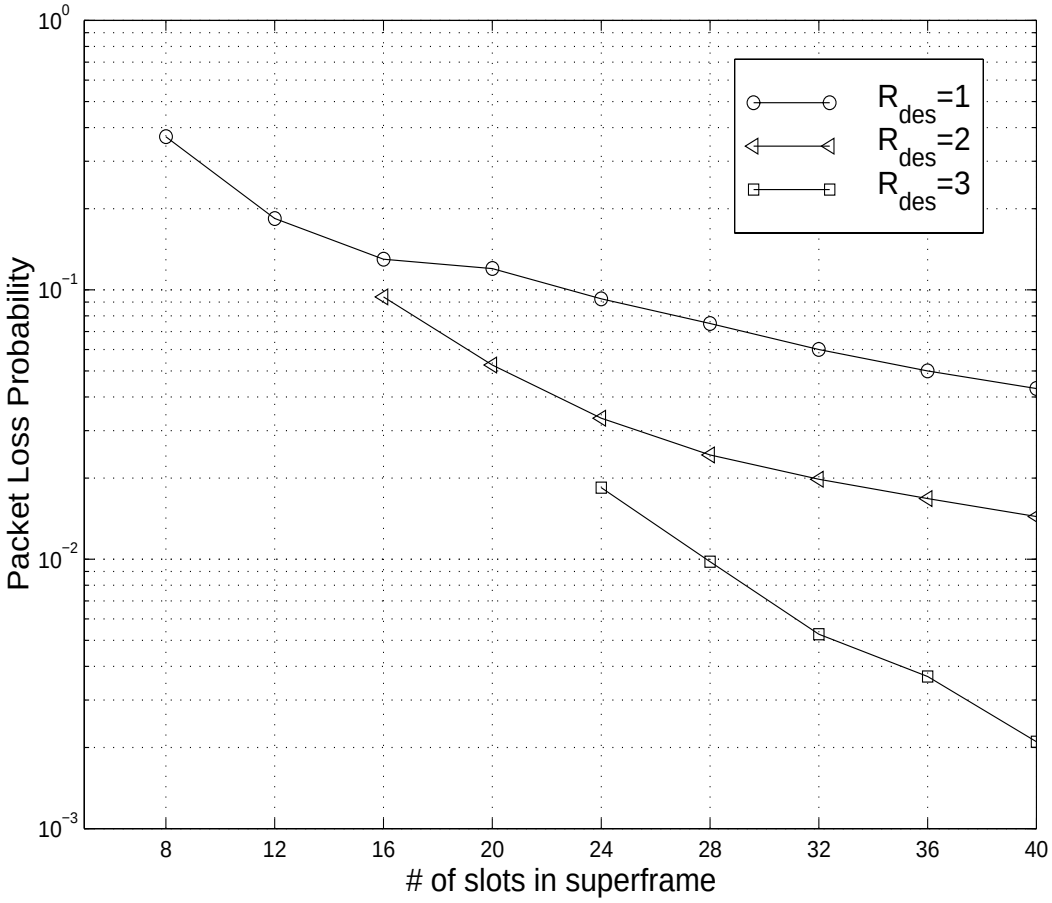



Figure 8.9: Packet loss probability for a Poisson arrival rate with mean one packet per superframe vs. total number of slots in the superframe and for different number of reserved slots, $RS_{\text{des}}$.

## 8.7 Summary

In this chapter we proposed a novel protocol to stream real-time multicast data to nodes in a wireless ad-hoc network. The multicast structure is built with receiver initiated affiliations and is such that collisions are avoided when data

is being streamed. The protocol is distributed and the information is being exchanged only between one hop neighbors. Real time connections are guaranteed bandwidth, while bandwidth is also re-used spatially. The multicast structure which is built has self-healing features.

Simulation results show that throughput above the 90%th percentile for pedestrian speeds ($\sim$ 10 Km/h) is achieved. The proportion of reserved and random-access bandwidth needed for VBR sources was also studied with a simplified model. Results quantify the trade-off between bandwidth reservation and packet loss rate for the multicast session.

A study of the exact signaling bandwidth needed for the exchange of maintenance packets is under way. Additionally, we are looking for ways to adapt the random-access strategies (like RTS/CTS or CSMA and its variants [105]) for the signaling channel, which would make our protocol completely scalable and more robust to mobility. We would identify this as the main drawback of our protocol. Apart from the scalability issue, which is very important, another problem with the current design is that signaling overhead is independent of the mobility of the network. In other words, the signaling is the same whether nodes move fast or slow. In case of low mobility this means that we waste a lot of bandwidth unnecessarily, since changes in the network state maybe very slow. On the other hand, under high mobility the signaling rate may not be sufficient to update fast enough the network state. In other words, right now we have more of a centralized signaling scheme, while we need a distributed one that will be reactive rather than proactive.

Many MAC protocols have been designed for ad-hoc wireless networks, though almost all of them are asynchronous, that is transmissions take place in an unslotted and un-coordinated manner. For such protocols see [106], [48], [107], [108],

[109] and references therein. For routing protocols designed for ad-hoc networks, see [97], [110], [111], [112], [113] and references therein.

# CHAPTER 9

# Conclusion

In this dissertation we addressed two important issues of wireless transmission: The first was propagation in indoor environments. This is an area much less researched than the area of outdoor propagation. The indoor environment has its own peculiar features that makes it interesting, and additionally it is becoming increasingly important for new home–office applications that have started to appear. We used the FDTD method from microwave engineering to predict the channel inside a residence and compared results with measurements. We demonstrated that FDTD works quite well in predicting the channel, and although it requires a very large amount of computer memory, it gives a substantial amount of data to do very detailed statistical analyses of all of the field components. The statistical analysis of the electromagnetic field revealed a number of things: the field power drops roughly as the third power of distance, at least for the relatively small distances that we measured in the residence (around fifteen meters from the transmitter). The delay spread of the channel drops exponentially and is quite small, around 20 ns for the transmit power we used. The individual multipath components are well approximated as lognormal, while if we approximate the channel by a single complex tap, the distribution of channel power is heavily location dependent. When there is line of sight, the channel is as expected Rician, while if there is no line of sight, the channel distribution is best described by a Nakagami distribution, but with largely varying parameters. Finally, the channel

autocorrelation differs drastically from the uniform scattering Jakes model. An interesting problem that we did not address is the analysis of field components other than the perpendicular one. Such analysis needs to be done in order to find how much energy is distributed to the horizontal directions, and results of such an analysis can have significant implications in antenna design for indoor applications.

The second issue we studied was that of medium access control protocols for co-channel interference environments. Specifically, we considered PRMA and analyzed its behavior in a cellular environment with small frequency re-use factors. We saw that for the relatively long hold voice traffic that we considered, interference can be accurately modeled as Markovian. Through an equilibrium point analysis for the resulting Markov chain we were able to find the total number of packets dropped both due to contention and due to excessive interference. The conclusion of this analysis is that co-channel interference becomes a very severe impairment for small re-use factors, to the point that the capacity of the system is practically zero. Motivated by that fact, we subsequently studied dynamic channel allocation algorithms that use actual measurements of the power present in the TDMA slots and allocate slots according to best signal to interference ratio. Additionally, base stations maintain a list of "preferred" slots, increasing the preference of a slot when packets are received correctly in it. The combination of interference measurements and slot prioritization improves dramatically the capacity of the system, even when all the slots are re-used in every micro-cell. We find that very significant role in this improvement is also played by power control. The basic power control is the known square root one, that compensates for half the power loss in dB. However, a new scheme that reduces power as long as the SIR of the received packets is too high is also seen to give improved results. Comparisons with a hybrid PRMA/CDMA system show that, under the

same bandwidth constraint DCA outperforms PRMA/CDMA. This result verifies the argument that interference avoidance performs better than interference averaging. Interesting extensions obviously include the consideration of traffic types other than voice only, and the integration of such types of traffic. Finally, we studied a MAC protocol for wireless multicast in an ad-hoc network. Ad-hoc networks are fundamentally different from the micro-cellular networks we studied before, because now there are no base stations to perform centralized scheduling. We developed a distributed algorithm that uses only one-hop neighborhood information to do bandwidth reservation and affiliation of terminals to the multicast session. We studied performance for mobile networks and saw that for pedestrian speeds our algorithm achieves throughput above 0.9. One major drawback of the algorithm is that it uses a round-robin signaling scheme, which is not scalable in the number of nodes, and also not adaptive to network mobility.

# APPENDIX A

## Calculation of $P[J = j]$ of equation (6.18)

In order to calculate the probability distribution of the number of packets remaining in a talk-spurt when a terminal obtains a reservation, first we condition on the number of contending terminals $C$ as well as the number of terminals already having a reservation, $R$. We will consider these as known for the following discussion. As we discussed in the previous chapter, for a terminal in state CON to obtain a reservation, the following conditions should simultaneously hold:

a) The slot is not reserved, probability $(1 - R/N)$.

b) The terminal is the only one in its cell that obtains permission to transmit, probability $p(1-p)^{C-1}$.

c) The packet transmitted by the terminal is captured, probability $P_c$.

Therefore, the probability that the terminal does not reserve the current empty slot is:

$$v = 1 - \left(1 - \frac{R}{N}\right) p(1-p)^{C-1} P_c \tag{A.1}$$

Now, we write $P[J = j|C, R]$ as follows:

$$\begin{aligned} P[J = j|C, R] &= \sum_{l=j}^{\infty} P[N_D = l - j|L = l, C, R] P[L = l] \\ &= \sum_{k=0}^{\infty} P[N_D = k|L = k + j, C, R] P[L = k + j] \end{aligned} \tag{A.2}$$

where $N_D$ is the number of packets dropped in a talk-spurt and $L$ is the total number of packets in the talk-spurt. The first probability factor in the sum has been given in equation (5.15) as a function of $v$, and the second in equation (5.17). By direct substitution in equation (A.2) and calculation of the sums we obtain:

$$P[J=j|C,R] = \begin{cases} \gamma_f(1-\gamma_f)^{j-1}\left[1-\dfrac{\gamma_f v^D}{1-v^N(1-\gamma_f)}\right], & j \geq 1 \\ \dfrac{\gamma_f v^D}{1-v^N(1-\gamma_f)}, & j=0 \end{cases} \tag{A.3}$$

Note here that, the probability distribution given in (A.3) is conditioned on $C$ and $R$. We un-condition by averaging over these random variables:

$$P[J=j] = \sum_{n=0}^{M}\sum_{m=0}^{M} P[J=j|C=n,R=m]P[C=n|R=m]P[R=m] \tag{A.4}$$

where the last two probability distributions are given by equations (6.9) and (6.7) respectively.

# APPENDIX B

# Calculation of Permission Probability $p$ for a PRMA/CDMA System

We want to adaptively change the permission probability $p$ of a hybrid PRMA/CDMA system such that we maximize radio occupancy, or equivalently the number of terminals per cell that the system can support at packet drop probability of $10^{-2}$. Note here that the BS has knowledge of the total number of terminals in its cell, as well as the number of terminals that are currently active. Furthermore, the process by which terminals turn active-inactive evolves independently of the process by which terminals obtain reservations. To see this fact, which was first realized in [91], we will find the Markov process that models the speech-silence transitions of the terminals in a cell. Call $N_S(t)$ the number of terminals that are silent at time $t$, while $M$ is the total number of terminals in the cell. Also, call $t_1$ and $t_2$ the mean talk-spurt and silence times respectively. The actual talk-spurt and silence times are distributed exponentially, and therefore $N_S(t)$ is a Markov process. We can write for the transition probabilities:

$$P\left[N_S(t+\tau)=i|N_S(t)=i-1\right]=\binom{M-i+1}{1}\left(1-e^{-\tau/t_1}\right)e^{-(M-i)\tau/t_1} \quad \text{(B.1)}$$

$$P\left[N_S(t+\tau)=i-1|N_S(t)=i\right]=\binom{i}{1}\left(1-e^{-\tau/t_2}\right)e^{-(i-1)\tau/t_2} \quad \text{(B.2)}$$

Equation (B.1) has the following interpretation: there are $M-i+1$ talking terminals and from these one has to become idle, probability $\left(1-e^{-\tau/t_1}\right)$, while

$M-i$ have to keep talking, probability $e^{-(M-i)\tau/t_1}$. Equation (B.2) has a similar interpretation. Having these transition probabilities, we can draw the transition diagram of the speech-silence process, as shown in Fig. B.1. If we define $\rho = \frac{t_2}{t_1}$ and solve for the stationary probability of the Markov chain, we obtain:

$$P\left[N_S = i\right] = \binom{M}{i} \frac{\rho^i}{(1+\rho)^M} \tag{B.3}$$

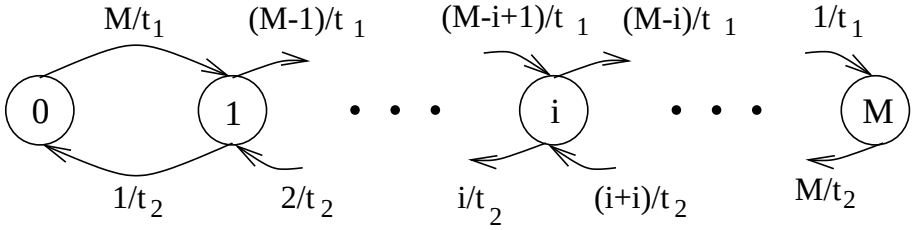



Figure B.1: The transition diagram of the speech – silence process for the PRMA/CDMA system.

As we already noted, the BS has knowledge of the number of terminals, say $R$ that are already active. Therefore, if we call $N_C$ the number of terminals in contention, we can find the conditional probability that $N_C = c$ given $R = K$ as follows:

$$\begin{aligned} P\left[N_C = c | R = K\right] &= P\left[N_S = M - K - c | R = K\right] \\ &= \binom{M-K}{c} \left(\frac{\rho}{1+\rho}\right)^{M-K} \rho^{-c} \end{aligned} \tag{B.4}$$

Now, given $N_C$ terminals in contention, and each terminal accessing the channel independently with probability $p$, the probability that $N_{\text{acc}}$ terminals attempt to transmit is given by:

$$P\left[N_{\text{acc}} = l | N_C = c\right] = \binom{c}{l} p^l (1-p)^{c-l} \tag{B.5}$$

Say that we want to adjust the permission probability $p$ such that the probability that more than $J$ terminals access the channel is below a threshold value $\alpha$. We have, from (B.4) and (B.5):

$$P\left[N_{\text{acc}} > J | R = K\right] = \sum_{c=J+1}^{M-K} P\left[N_{\text{acc}} > J | N_C = c\right] P\left[N_C = c | R = K\right]$$
$$= \left(\frac{\rho}{1+\rho}\right)^{M-K} \sum_{c=J+1}^{M-K} \binom{M-K}{c} \rho^{-c} \sum_{l=J+1}^{c} \binom{c}{l} p^l (1-p)^{c-l} \quad \text{(B.6)}$$

For any value of $\alpha$, we have to solve the above non-linear equation for $p$. Note that (B.6) depends on the number of terminals that currently hold reservations.

If we want to adjust the permission probability $p$ such that the expected number of terminals in contention that attempt to access the channel is $J$, we have to find $E\left[N_{\text{acc}} | R\right]$. This is easily done using (B.4) and (B.5):

$$E\left[N_{\text{acc}} | R\right] = E_{N_C}\left\{E\left[N_{\text{yacc}} | R, N_C\right]\right\}$$
$$= \left(\frac{\rho}{1+\rho}\right)^{M-k} \sum_{c=0}^{M-k} \binom{M-K}{c} \rho^{-c} \sum_{l=0}^{c} l \binom{c}{l} p^l (1-p)^{c-l} = p\frac{M-K}{1+\rho} \quad \text{(B.7)}$$

Therefore, the permission probability $p$ that assures $E\left[N_{\text{acc}} | R = K\right] = J$ is given by:

$$p = J\frac{1+\rho}{M-K} \quad \text{(B.8)}$$

# APPENDIX C

# Calculation of Coding Gain for a PRMA/CDMA System

Assume that we are using a (L, K, t) binary block error correcting code[1]. Further assume that the crossover probability of the equivalent binary symmetric channel is $P_e$. Then, the probability that a codeword is received in error is simply:

$$P_{\text{cerr}} = \sum_{i=t+1}^{L} \binom{L}{i} P_e^i (1-P_e)^{L-i} \approx \binom{L}{t+1} P_e^{t+1} (1-P_e)^{L-t-1} \tag{C.1}$$

The crossover probability under a Gaussian interference assumption for the interference limited channel is given by:

$$P_e = Q\left(\sqrt{\text{SIR}}\right) \tag{C.2}$$

Therefore, if we want a packet error probability of $10^{-2}$ and we assume that a packet comprises a single codeword, we find from (C.1) the corresponding crossover probability and from (C.2) the necessary SIR that achieves that probability. The ratio of this SIR over the one that achieves a crossover probability of $10^{-2}$ in an uncoded system gives us the coding gain.

[1]The notation (L, K, t) means that a codeword has L bits of which K are information carrying and the code can correct t bits.

# References


[1] J. G. Proakis, *Digital Communications,* third edition, McGraw Hill, New York, NY, 1995.

[2] W. C. Jakes, editor, *Wireless Mobile Communications,* IEEE Press, New York, NY, 1994.

[3] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: information-theoretic and communications aspects," *IEEE Trans. Inform. Theory,* pp. 2619-92, Oct. 1998.

[4] E. Biglieri, D. Divsalar, P. McLane, and M. Simon, *Introduction to Trellis-Coded Modulation with Applications,* Macmillan, New York, NY, 1991.

[5] M. Abramowitz, and I. Stegun, editors, *Handbook of Mathematical Functions,* Dover, New York, NY, 1965.

[6] T. S. Rappaport, *Wireless Communications,* Prentice Hall, New Jersey, 1996.

[7] H. Hashemi, "The indoor radio propagation channel," *IEEE Proc.,* pp. 943-967, July 1993.

[8] J. Anderson, T. Rappaport, and S. Yoshida, "Propagation measurements and models for wireless communications channels," *IEEE Commun. Mag.,* Nov. 1994.

[9] R. Valenzuela, "A ray tracing approach to predicting indoor wireless transmission," in *Proc. IEEE Vehicular Technology Conference, VTC 1993,* pp. 214-218, 18-20 May, 1993.

[10] S. Seidel, and T. Rappaport, "Site-specific propagation prediction for wireless in-building personal communication system design," *IEEE Trans. Veh. Technol.,* pp. 879-91, Nov. 1994.

[11] T. Russel, C. Bostian, and T. Rappaport, "A deterministic approach to predicting microwave diffraction by buildings for microcellular systems," *IEEE Trans. Antennas Propagat.,* pp. 1640-49, Dec. 1993.

[12] C. A. Balanis, *Advanced Engineering Electromagnetics,* Wiley, New York, 1989.

[13] D. Cox, "Delay Doppler characteristics of multipath delay spread and average excess delay for 910 MHz urban mobile radio paths," *IEEE Trans. Antennas Propagat.,* pp. 625-35, Sept. 1972.

[14] T. Rappaport, "Characterization of UHF multipath radio channels in factory buildings," *IEEE Trans. Antennas Propagat.,* pp. 1058-69, Aug. 1989.

[15] K. S. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas Propagat.,* pp. 302-7, May 1966.

[16] W. Lee, *Mobile Cellular Telecommunications Systems,* McGraw Hill, New York, NY, 1989.

[17] D. Pozar, *Microwave Engineering,* Adison–Wesley, Reading, MA, 1990.

[18] D. M. Sheen, S. M. Ali, M. D. Abouzahra, and J. A. Kong, "Application of the three dimensional finite-difference time-domain method to the analysis of planar microstrip circuits," *IEEE Trans. Microwave Theory Tech.,* pp, 849-57, July 1990.

[19] A. Taflove, *Computational Electromagnetics: The Finite-Difference Time-Domain Method*, Artech House, Boston, 1995.

[20] G. Mur, "Absorbing boundary conditions for the finite difference approximation of the time domain electromagnetic field equations," *IEEE Trans. Electromagn. Compat.,* pp. 377-82, Nov. 1981.

[21] K. Mei, and J. Fang, "Superabsorption-a method to improve absorbing boundary conditions," *IEEE Trans. Antennas Propagat.,* pp. 1001-10, Sept. 1992.

[22] Z. Yusheng, and W. Wenbing, "The studies of the stability of FDTD with Mur's absorbing boundary condition of second order in 3-D scattering problems," *IEEE Microwave Guided Let.,* pp. 120-22, March 1996.

[23] J. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *J. Computational Phys.,* pp. 185-200, Oct. 1994.

[24] K. Pahlavan, and A. Levesque, *Wireless Information Networks,* Wiley, New York, NY, 1995.

[25] H. Zaghloul, G. Morrison, and M. Fattouche, "Comparison of indoor propagation channel characteristics at different frequencies," *Electronics Let.,* pp. 2077-79, Oct. 1991.

[26] G. Turin, "Introduction to spread-spectrum antimultipath techniques and their application to urban digital radio," *IEEE Proc.,* March 1980.

[27] J. Nail, "Designing Discone Antennas," *Electronics,* August 1953.

[28] S. Howard, K. Pahlavan, "Measurement and analysis of the indoor radio channel in the frequency domain," *IEEE Trans. Instrumentation Meas.,* pp. 751-55, Oct. 1990.

[29] K. Pahlavan, S. Howard, and T. Sexton, "Decision feedback equalization of the indoor radio channel," *IEEE Trans. Commun.,* pp. 164-70, Jan. 1993.

[30] M. Nakagami, "The m-distribution–a general formula of intesity distribution of rapid fading," in W. Hoffman, editor, *Statistical Methods of Radio Wave Propagation*, Pergamon, New York, 1960.

[31] H. Wang, and P. Chang, "On verifying the first-order markovian assumption for a rayleigh fading channel model," *IEEE Trans. Veh. Technol.,* pp. 353-57, May 1996.

[32] K. Luebbers, *The Finite Difference Time Domain Method for Electromagnetics,* CRC Press, Boca Raton, FL, 1993.

[33] H. Bertoni, *Radio Propagation for Modern Wireless Systems,* Prentice Hall, New Jersey, 1999.

[34] G. Pelosi, R. Coccioli, and S. Selleri, *Quick Finite Elements for Electromagnetic waves,* Artech House, Boston, MA, 1998.

[35] P. Silvester, and R. Ferrari, *Finite Elements for Electrical Engineers,* third edition, Cambridge University Press, 1996.

[36] M. Tsai, F. De Flaviis, O. Frodham, and N. Alexopoulos, "Modeling planar arbitrarily-shaped microstrip elements in multi-layered media," *IEEE Trans. Microwave Theory Tech.,* pp. 330-37, March 1997.

[37] F. De Flaviis, M. Noro, R. Diaz, G. Franceschetti, and N. Alexopoulos, "A vector potential formulation for the solution of electromagnetic problems," *IEEE Microwave Letters,* pp. 310-12, Sept. 1998.

[38] F. De Flaviis, M. Noro, N. Alexopoulos, R. Diaz, and G. Franceschetti, "Extensions to complex materials of the Fitzgerald model for the solution of electromagnetic problems," *Electromagnetics,* pp. 35-65, 1998.

[39] P. Aoyagi, J. Lee, and R. Mittra, "A hybrid Yee algorithm/scalar-wave equation approach," *IEEE Trans. Microwave Theory Tech.,* pp. 1593-1600, Sept. 1993.

[40] L. Lapidus, and G. Pinder, *Numerical Solutions of Partial Differential Equations in Science and Engineering,* Wiley, New York, NY, 1982.

[41] D. S. Jones, *The Theory of Electromagnetics,* Macmillan, New York, NY, 1964.

[42] M. Sucher, J. Fox, editors, *Handbook of Microwave Measurements,* Polytechnic Press, 1963.

[43] X. Zhang, and K. Mei, "Time-domain finite difference approach to the calculation of the frequency-dependent characteristics of microstrip discontinuities," in *Proc. IEEE International Microwave Symposium, MTT-S 1988,* pp. 1775-88, Anaheim, CA, May 25-27, 1988.

[44] I. Akyildiz, J. McNair, L. Martorell, R. Puigjaner, and Y. Yesha, "Medium access control protocols for multimedia traffic in wireless networks," *IEEE Network,* pp. 39-47, July 1999.

[45] C. Perkins, "Mobile IP," *IEEE Commun. Mag.,* pp. 84-99, May 1997.

[46] P. Bhagwat, C. Perkins, and S. Tripathi, "Network layer mobility: an architecture and survey," *IEEE Personal Commun.,* pp. 54-64, June 1996.

[47] J. Haartsen, "The bluetooth radio system,", *IEEE Personal Commun.,* pp. 28-36, Feb. 2000.

[48] *IEEE P802.11 Draft Standard for Wireless LAN: Medium Access Control (MAC) and Physical Layer (PHY) Specifications,* IEEE, July 1996.

[49] V. Bhagavath, "Emerging high-speed xDSL access services: architectures, issues, insights, and implications," *IEEE Commun. Mag.,* pp. 106-14, Nov. 1999.

[50] A. Nordbotten, "LMDS systems and their application," *IEEE Commun. Mag.,* pp. 150-4, June 2000.

[51] D. Goodman, R. Velenzuela, K. Gayliard, and B. Ramamurthi, "Packet reservation multiple access for local wireless communications," *IEEE Trans. Commun.,* pp. 885-90, Aug. 1989.

[52] S. Nanda, D. Goodman, and U. Timor, "Performance of PRMA: a packet reservation protocol for cellular systems," *IEEE Trans. Veh. Technol.,* pp. 584-98, Aug. 1991.

[53] D. Goodman, and S. Wei, "Efficiency of packet reservation multiple access," *IEEE Trans. Veh. Technol.,* pp. 170-6, Feb. 1991.

[54] J. Dunlop, D. Robertson, P. Cosimini, and J. De Vile, "Development and optimization of a statistical multiplexing mechanism for ATDMA," in *Proc. IEEE Vehicular Technology Conference, VTC 1994,* pp. 1040-44, Stockholm, Sweden, May 1994.

[55] J. DeVile, "A reservation multiple access scheme for an adaptive TDMA air-interface," in *Proc. Fourth WINLAB Workshop on Third Generation Wireless Information Networks*, New Jersey, USA, Oct. 1993.

[56] F. Borgonovo, L. Fratta, M. Zorzi, and A. Acampora, "Capture division packet access: a new cellular access architecture for future PCNs," *IEEE Commun. Mag.,* pp. 154-62, Sept. 1996.

[57] F. Borgonovo, M. Zorzi, L. Fratta, V. Trecordi, and G. Bianchi, "Capture division packet access for wireless personal communications," *IEEE J. Select. Areas Commun.,* pp. 609-22, May 1996.

[58] M. Zorzi, and F. Borgonovo, "Performance of capture division packet access with slow shadowing and power control," *IEEE Trans. Veh. Technol.,* pp. 687-96, Aug. 1997.

[59] S. Nanda, "Stability evaluation and design of the PRMA joint voice data system," *IEEE Trans. Commun.,* pp. 2092-104, May 1994.

[60] L. Jalloul, S. Nanda, and D. Goodman, "Packet reservation multiple access over slow and fast fading channels,", in *Proc. IEEE,* pp. 354-59.

[61] R. Fantacci, and L. Zoppi, "A combined reservation random access polling protocol for voice-data transmissions in wireless packet network," *IEEE Trans. Veh. Technol.,* pp. 652-62, March 1999.

[62] G. Anastasi, D. Grillo, and L. Lenzini, "An access protocol for speech/data/video integration in TDMA-based advanced mobile systems," *IEEE J. Select. Areas Commun.,* pp. 1498-1510, Oct. 1997.

[63] M. Orange, K. Sowerby, A. Coulson, and K. Butterworth, "Performance of a speech-data PRMA system in an in-building environment," in *Proc. IEEE Vehicular Technology Conference, VTC 1997,* pp. 1518-1522, Phoenix AZ, May 1997.

[64] N. Amitay, and S. Nanda, "Resource auction multiple access (RAMA) for statistical multiplexing of speech in wireless PCS," *IEEE Trans. Veh. Technol.,* pp. 584-96, Aug. 1994.

[65] M. Paterakis, and A. Clearly, "On the voice-data integration in third generation wireless access communication networks," *Europ. Trans. Telecommun.,* pp. 11-18, Jan.-Feb. 1994.

[66] A. Clearly, and M. Paterakis, "An investigation of stack based algorithms for voice packet transmission in microcellular wireless environments," in *Proc. IEEE International Conference on Communications, ICC 1995,* pp. 1375-81, Seattle, WA, June 1995.

[67] M. Frullone, G. Riva, P. Grazioso, and C. Carciofi, "PRMA performance in cellular environments with self-adaptive channel allocation strategies," *IEEE Trans. Veh. Technol.,* pp. 657-65, Nov. 1996.

[68] M. Orange, and K. Sowerby, "The effect of receiver capture and co-channel interference on PRMA," in *Proc. IEEE Vehicular Technology Conference, VTC 1996,* pp. 1741-45, Atlanda, GA, April 1996.

[69] G. Bianchi, F. Borgonovo, L. Fratta, L. Musumeci, and M. Zorzi, "C-PRMA: a centralized packet reservation multiple access for local wireless communications," *IEEE Trans. Veh. Technol.,* pp. 422-36, May 1997.

[70] M. Zorzi, "On the analytical computation of the interference statistics with applications to the performance evaluation of mobile radio systems," *IEEE Trans. Commun.,*, pp. 103-09, Jan. 1997.

[71] E. Cinlar, *Introduction to Stochastic Processes,* Prentice Hall, New Jersey, 1975.

[72] J. Parsons, *The Mobile Radio Propagation Channel,* Halsted Press, New York, 1992.

[73] http://pcl.cs.ucla.edu/projects/parsec/

[74] G. Pottie, "A highly adaptive radio transceiver for personal communications applications," (invited) *ISSSE '95* San Francisco, October 1995 (invited).

[75] N. Bambos, and G. Pottie, "Power control based admission policies in cellular radio networks," in *Proc. IEEE GLOBECOM 1992,* pp. 863-67, Orlando, FL, Dec. 6-9, 1992.

[76] N. Bambos, and S. Kandukuri, "Power control multiple access (PCMA) for wireless communication networks," in *Proc. IEEE INFOCOM 2000.*

[77] K. Leung, "A Kalman filter method for power control in broadband wireless networks," in *Proc. IEEE INFOCOM 1999,* pp. 948-56.

[78] S. Chen, N. Bambos, and G. Pottie, "Radio link admission algorithms for wireless networks with power control and active link protection," *technical report,* UCLA-ENG-94-25.

[79] S. Tasaka, "Stability and performance of the R-ALOHA packet broadcast system," *IEEE Trans. Commun.,* pp. 717-726, Aug. 1983.

[80] G. Pottie, "System design choices for personal communications," *IEEE Personal Commun.,* (invited) pp. 50-67, Oct. 1995.

[81] C. Wang, and G. Pottie, "Interference avoidance and power control strategies for coded frequency-hopped cellular systems," in *Proc. IEEE International Conference on Communications, ICC 1995,* Seattle, WA, June 1995.

[82] C. Hansen, and G. Pottie, "Line probing for interference coupled communications systems," in *Proc. IEEE International Information Theory Symposium, ISIT 1995,* Whistler British Columbia, Canada, Sept. 1995.

[83] N. Mitrou, T. Orinos, and E. Protonotarios, "A reservation multiple access protocol for microcellular mobile-communication systems," *IEEE Trans. Veh. Technol.,* pp. 340-51, Nov. 1990.

[84] G. Bianchi, F. Borgonovo, M. De Marco, and V. Trecordi, "A simulation study of cellular systems based on the capture division packetized access (CDPA) technique," in *Proc. IEEE International Conference on Communications, ICC 1995*, pp. 1399-1403, Seattle, WA, June 1995.

[85] P. Brady, "A model for generating on-off speech patterns in two way conversations," *Bell Syst. Tech. J.,* vol. 48, no. 7, pp. 2445-2472, Sept. 1969.

[86] T. Fong, P. Henry, K. Leung, X. Qiu, and N. Shankaranarayanan, "Radio resource allocation in fixed broadband wireless networks," *IEEE Trans. Commun.,*, pp. 806-18, June 1998.

[87] Y. Akaiwa, and H. Andoh, "Channel segregation-a self organized dynamic channel allocation method: application to TDMA/FDMA microcellular system," *IEEE J. Select. Areas Commun.,* pp. 949-54, Aug. 1993.

[88] S. Krishnamurthy, A. Acampora, and M. Zorzi, "On the capacity of TDMA and CDMA for broadband wireless packet access," in *Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC 1998,* pp. 167-72, Boston, MA, Sept. 8-11 1998.

[89] J. Chuang, and N. Sollenberger, "Spectrum resource allocation for wireless packet access with application to advanced cellular internet service," *IEEE J. Select. Areas Commun.,* pp. 820-29, Aug. 1998.

[90] Z. Haas, J. Winters, and D. Jonson, "Simulation results of the capacity of cellular systems," *IEEE Trans. Veh. Technol.,* pp. 805-17, Nov. 1997.

[91] X. Qiu, and V. Li, "On the capacity of packet reservation multiple access with capture in personal communication systems," *IEEE Trans. Veh. Technol.,* pp. 666-75, Nov. 1996.

[92] M. Pursley, "Performance evaluation for phase-coded spread spectrum multiple access communication–part I: system analysis," *IEEE Trans. Commun.,* pp. 795-799, Aug. 1977.

[93] A. Brand, and A. Aghvami, "Performance of a joint CDMA/PRMA protocol for mixed voice/data transmission for third generation mobile communication," *IEEE J. Select. Areas Commun.,* pp. 1698-707, Dec. 1996.

[94] B. Brand, and A. Aghvami, "Multidimensional PRMA with prioritized Bayesian broadcast-a MAC strategy for multiservice traffic over UMTS," *IEEE Trans. on Veh. Technol.,* pp. 1148-61, Nov. 1998.

[95] S. Lin, and D. Costello, *Error Control Coding,* Prentice Hall, New Jersey, 1983.

[96] A. Ephremides, and B. Hajek, "Information theory and communication networks: an uncosummated union," *IEEE Trans. Inform. Theory,* pp. 2416-34, Oct. 1998.

[97] C. Perkins, and P. Bhagwat, "Highly dynamic destination-sequenced distance vector routing (DSDV) for mobile computers," in *Proc. ACM SIGCOMM 1994,* pp. 234-44, London, UK, Aug. 1994.

[98] J. Broch et al., "A performance comparison of multi-hop wireless ad hoc network routing protocols," in *Proc. ACM MOBICOM 1998,* pp. 87-97, Dallas TX.

[99] J. Garcia-Luna-Aceves, and E. Madruga, "The core-assisted mesh protocol," *IEEE J. Select. Areas Commun.,* pp. 1380-94, Aug. 1999.

[100] A. Ephremides, and T. Truong, "Scheduling broadcasts in multihop radio networks," *IEEE Trans. Commun.,* pp. 456-60, April 1990.

[101] S. Ramanathan, and E. Lloyd, "Scheduling algorithms for multihop radio networks," *IEEE/ACM Trans. on Networking,* pp. 166-77, 1993.

[102] I. Chlamtac, and A. Farago, "Making transmission schedules immune to topology changes in multi-hop packet radio networks," *IEEE/ACM Trans. on Networking,* pp. 23-29, Feb. 1994.

[103] S. Chong, and J. Gosh, "Predictive dynamic bandwidth allocation for efficient transport of real-time VBR video over ATM," *IEEE J. Select. Areas Commun.,* pp. 12-23, Jan. 1995.

[104] D. Connors, G. Pottie, B. Ryu, and S. Dao, "A quality of service based medium access control protocol for real-time sources," *Baltzer/ACM J. on Spec. Topics in Mobile Netw. and Appl. (MONET),* to appear Fall 1999.

[105] D. Bertsekas, and R. Gallager, *Data Networks,* second edition, Prentice Hall, New Jersey, 1992.

[106] V. Bharghavan, A. Demers, S. Shenker, and L. Zhang, "MACAW: A media access protocol for wireless LAN's," in *Proc. ACM SIGCOMM 1994,* pp. 212-25, London, UK, Aug. 1994.

[107] J. Garcia-Luna-Aceves, and C. Fullmer, "Floor acquisition multiple access (FAMA) in single-channel wireless networks," *Baltzer/ACM Mobile Networks and Applications,* pp. 157-74, 1999.

[108] Z. Tang, and J. Garcia-Luna-Aceves, "Hop reservation multiple access for multichannel packet radio networks," *Computer Communications,* pp. 877-86, Elsevier, May 2000.

[109] Z. Haas, and J. Deng, "Dual busy tone multiple access (DBTMA)-performance results," in *IEEE Wireless Commun. and Networking Conference, WCNC 1999,* pp. 1328-32, New Orleans, LA, Sept. 1999.

[110] D. Maltz, J. Broch, J. Jetcheva, and D. Johnson, "The effects of on-demand behavior in routing protocols for multihop wireless ad hoc networks," *IEEE J. Select. Areas Commun.,* pp. 1439-53, Aug. 1999.

[111] C. Perkins, E. Royer, and S. Das, "Ad hoc on-demand distance vector (AODV) routing," *http://www.ietf.org/internet-drafts/draft-ietf-manet-aodv-06.txt.*

[112] D. Park, and S. Corson, "A highly adaptive distributed routing algorithm for mobile wireless networks," in *Proc. IEEE INFOCOM 1997,* pp. 1405-13, Kobe, Japan, 1997.

[113] R. Sivakumar, P. Sinha, and V. Bharghavan, "CEDAR: a core-extraction distributed ad hoc routing algorithm," *IEEE J. Select. Areas Commun.,* pp. 1454-65, Aug. 1999.