Diplomarbeit

Information Theoretic Aspects of Multiuser Multicarrier Systems

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines Diplom–Ingenieurs
unter der Leitung von

Ao.Univ.Prof. Dipl.-Ing. Dr.techn. Gerald Matz

Institut für Nachrichtentechnik und Hochfrequenztechnik (E 389)
eingereicht an der Technischen Universität Wien

Fakultät für Elektrotechnik und Informationstechnik

von

Michael Martl, 0026579
Kühweg 29, 9620 Hermagor

Wien, Dezember 2005

Michael Martl

Abstract

This thesis deals with multiuser multicarrier communications over unknown time and frequency selective fading channels. Since there is a growing demand for capacity, especially in mobile communications, wireless multiuser communications experiences increasing interest in recent year. The wireless channel in general is a time and frequency selective fading channel. In that case, *orthogonal frequency division multiplexing* (OFDM) promises robust high-rate transmissions. In general, no *channel state information* (CSI) is available at transmitter and receiver. This corresponds to noncoherent transmissions.

This thesis investigates noncoherent OFDM based multiuser systems, such as *Multi-Carrier-CDMA* (MC-CDMA) and *orthogonal frequency division multiple access* (OFDMA), transmitting over a time and frequency selective fading channel. The focus is on the uplink case where multiple users transmit to one receiver. An information theoretic analysis is performed with emphasis on sum system capacity and sum information rate calculation. Since in most cases no exact analytic results can be obtained, bounding techniques are applied.

First OFDM, OFDMA, and MC-CDMA are described. Then the wireless channel is discussed in terms of descriptions, statistics, and parameters, and an appropriate input-output relation will be introduced. Furthermore, an overview of known results is given. Then, upper and lower bounds on sum system capacity for MC-CDMA and OFDMA are derived. These theoretical results are supported by numerical evaluations and extensive simulation results involving different propagation scenarios.

Zusammenfassung

Diese Diplomarbeit befasst sich mit Mehrbenutzer-Mehrträger-Übertragung über unbekannte zeit- und frequenzselektive Schwundkanäle. In den letzten Jahren stieg die Nachfrage nach höherer Kanalkapazität. Speziell in der Mobilkommunikation erfährt drahtlose Mehrbenutzer-Übertragung ein steigendes Interesse. Die Freiraumübertragung erfolgt im Allgemeinen über einen zeit- und frequenzselektiven Kanal. In diesem Fall verspricht *Orthogonal Frequency Division Multiplexing* (OFDM) eine zuverlässige Übertragung mit hoher Datenrate. Im Allgemeinen ist Kanalinformation (*Channel State Information*, CSI) weder am Sender noch am Empfänger verfügbar. Man bezeichnet das als nicht-kohärente Übertragung.
Diese Arbeit untersucht nicht-kohärente OFDM-basierte Mehrbenutzersysteme wie *Multi-Carrier-CDMA* (MC-CDMA) und *Orthogonal Frequency Division Multiple Access* (OFDMA) bei Übertragung über zeit- und frequenzselektive Schwundkanäle. Der Fokus liegt auf dem Uplink-Fall, wo mehrere Benutzer zu einem Empfänger senden. Eine informationstheoretische Analyse mit Augenmerk auf Summensystemkapazität und Summeninformationsrate wird durchgeführt. Da in den meisten Fällen keine exakten analytischen Ergebnisse erreichbar sind, werden obere und untere Schranken berechnet.
Zuerst werden die Übertragungsverfahren OFDM, OFDMA und MC-CDMA beschrieben. Es folgt eine Erklärung des Mobilfunkkanals in Form von Beschreibungen, statistischen Zusammenhängen und Parametern. Dann wird eine geeignete Eingangs-Ausgangsbeziehung eingeführt. Weiters wird eine Übersicht über bekannte Ergebnisse gegeben. Danach werden obere und untere Schranken für die Summenkapazität von MC-CDMA und OFDMA abgeleitet. Diese theoretischen Ergebnisse werden erläutert durch numerische Auswertungen und ausführliche Simulationen für verschiedene Ausbreitungsszenarien.

Acknowledgment

I am indebted to my advisor Gerald Matz for helping and guiding me through my work on this diploma thesis. I would like to thank Manfred Hartmann for introducing me into the topics of my diploma thesis. I also would like to thank the Time–Frequency Signal Processing Group for their friendly integration. Furthermore I want to express my gratitude to my colleagues for helping me with various LaTeX and MatLab problems. I owe my gratitude to my family and friends for pushing me to work hard, supporting me and for being considerate of my work.

Michael Martl
Vienna, December 2005

Contents

Chapter 1

Introduction

This thesis will treat *orthogonal frequency division multiplexing* (OFDM) systems in the multiuser case. OFDM is known since the late '60s, but wasn't applied in the first years. That changed in the early ´90s and OFDM became interesting for several communication standards. For multiuser systems *Multi-Carrier-CDMA* (MC-CDMA) and *orthogonal frequency division multiple access* (OFDMA) amongst others were found. Wireless communication is one of the fields OFDM is desired to. Therefore MC-CDMA and OFDMA will be used for transmission over time and frequency selective fading channels. No *channel state information* (CSI) will be available either to the reicever nor to the transmitter. In order to characterize the whole system an information-theoretic analysis (including system capacity and information rate) will be performed. The main focus will be on the calculation of sum system capacity over all users which will allow a detailed research on the entire system capability. Since analytic calculation doesn't lead to results in some cases bounding techniques will be applied. First an approximation for a multiuser MC-CDMA system based on a single user MC-CDMA system will be presented. For this system model bounding techniques on mutual information will be applied. A second way to design a multiuser system is to introduce a fully multiuser system model. This means each user is assigned a specific input (e.g. codebook) and different a channel.
In this chapter the basic concepts of OFDM, OFDMA, and MC-CDMA will be described. Further the channel will be explained in a continuous and a discrete time point of view and a simplified input-output relation will be presented. Finally a short introduction in information theory will be given.

1.1 OFDM

OFDM first was introduced in 1966 by Chang [1]. Initially there was nearly no practical use for OFDM (for reasons of complexity) and hence only little research was done. OFDM was applied only in military communications. DFT was recognized as useful for modulation and demodulation [2]. Further the application of OFDM is proposed in [3, 4]. In the early '90s the interest increased. This went hand in hand with the increased capabilities of digital signal

processing. Today OFDM is contained in several communication standards, such as *wireless local area networks* (WLAN), like 802.11 [5] and HIPERLAN/2 [6], terrestrial digital video broadcasting (DVB-T) [7], terrestrial digital audio broadcasting (DAB-T) [8], and *asymmetric digital subscriber lines* (ADSL). Other names of OFDM are *multicarrier modulation* and *discrete multitone* (DMT). OFDM also is a candidate for the fourth generation of cellular mobile radio systems. Furthermore there is interest in the research of the combination of OFDM and *multiple-input multiple-output* (MIMO) systems [9] (termed MIMO-OFDM [10]) and use in *ultra-wideband* (UWB) systems [11].
The main principle of OFDM is to split the entire transmit band into smaller narrowband sub-channels in order to obtain frequency-flat channels. OFDM may be defined as frequency division multiplexing. By use of transmit/receive filters with overlap in time and frequency also an interpretation as code division multiplexing is possible. Since there are frequency-flat channels no channel equalization is necessary, which means less computational complexity and an easy implementation. Further OFDM is resistant to narrowband interferers and time dispersive channels. A disadvantageous aspect of OFDM is an extensive synchronization in time and frequency. Further a high peak-to-average power ratio makes high demands on the transmitter.

1.1.1 OFDM Modulator

The modulator for a pulse-shaping OFDM system [1, 2] is shown in Figure 1.1. It includes K subcarriers. The data symbols which may belong to a single high data-rate source or to multiple sources/users are denoted by $X_{n,k}$. The subscript n denotes the OFDM symbol time and k is the OFDM subcarrier index. The OFDM symbol duration is N signal samples. After upsampling by the factor of N, the transmit data is passed through the transmit filter using the transmit pulse $g[m]$ and then is modulated with the respective subcarrier center frequency. The modulated discrete-time baseband transmit signal is

$$x[m] = \sum_{n=-\infty}^{\infty} \sum_{k=0}^{K-1} X_{n,k}\, g_{n,k}[m], \tag{1.1}$$

where $g_{n,k}[m]$ are the time-frequency shifted transmit pulses $g[m]$:

$$g_{n,k}[m] \triangleq g[m - nN]\, e^{j2\pi \frac{k}{K}(m-nN)}.$$

According to (1.1), the datastream is split into K substreams, each of which is assigned to a different frequency. Assuming N OFDM symbols to be transmitted, a rectangular time-frequency lattice could be used for interpretation. This is shown in Figure 1.2. Here the data symbols are transmitted via the time-frequency-shifted pulses $g_{n,k}[m]$. In Figure 1.2 only one pulse is depicted. Note that in general there is one pulse (with particular shape) at each time-frequency lattice point and there is an overlap between neighboring pulses.
The efficient implementation of an OFDM modulator is done by a length K inverse discrete Fourier transformation (IDFT). In practical systems the DFT is implemented by a FFT.

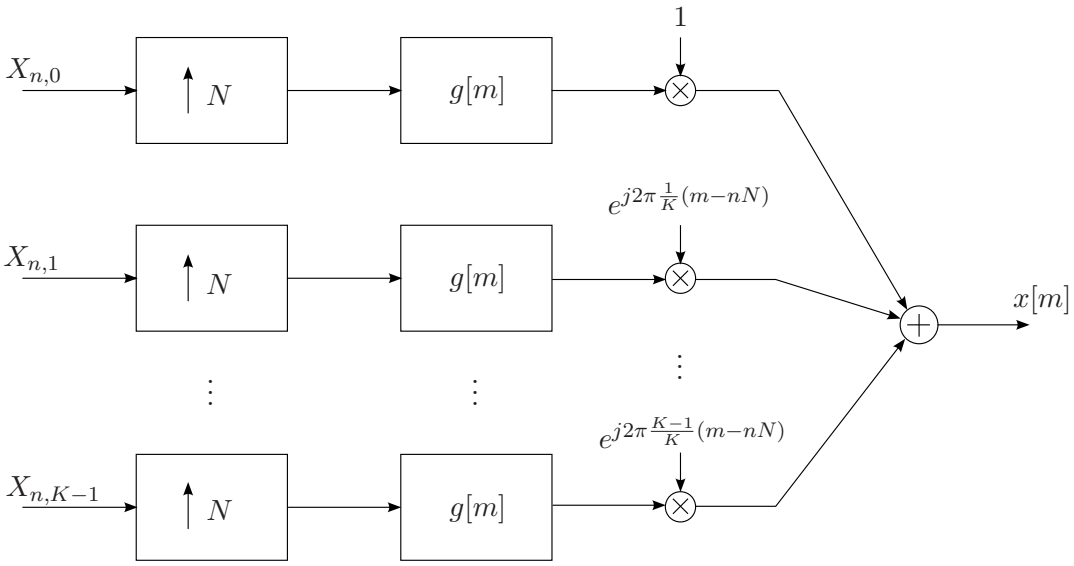

Figure 1.1: *Modulator of an OFDM system with K subcarriers, OFDM symbol duration N, and transmit-filter $g[m]$.*

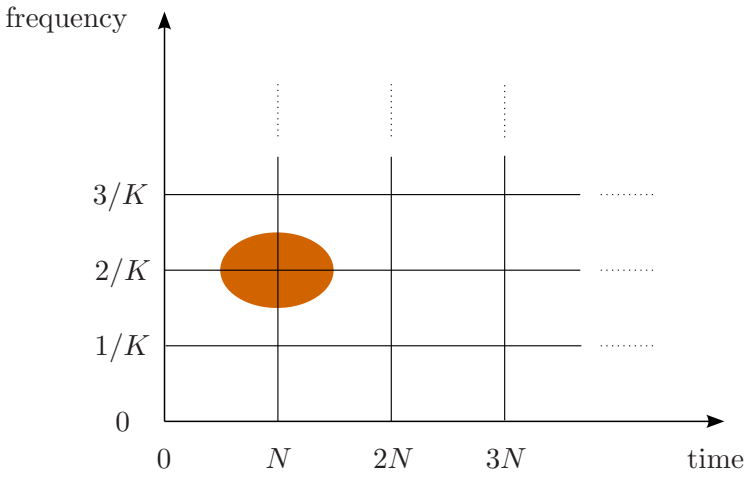

Figure 1.2: *Interpretation of OFDM modulator in terms of a time-frequency lattice. Note that in general there is one time-frequency shifted pulse at every lattice point and overlap of consecutive pulses is possible.*

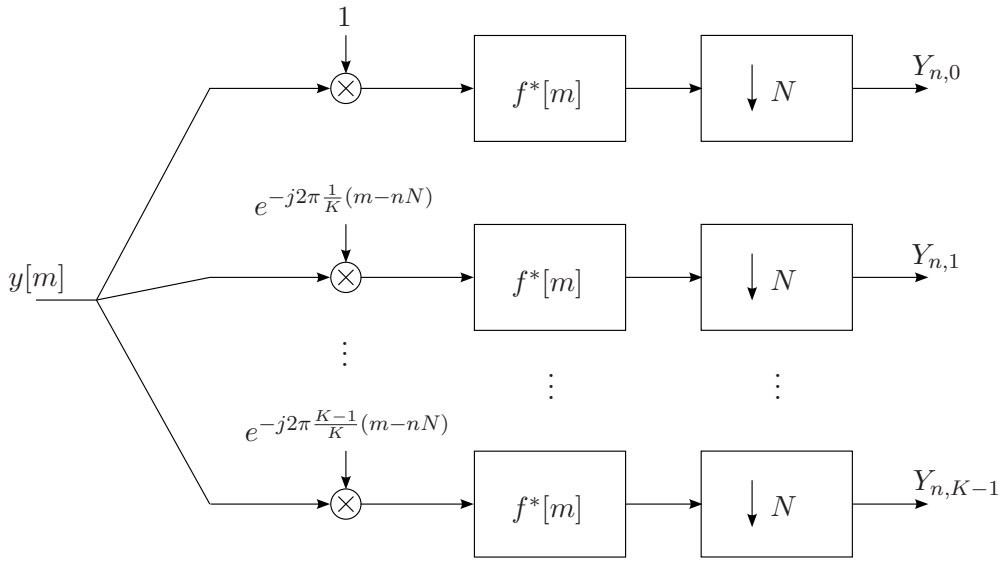

Figure 1.3: *Demodulator of an OFDM system with $f^*[m]$ denoting the receive filter.*

1.1.2 OFDM Demodulator

Equivalent to the modulator an efficient OFDM demodulator uses a DFT. Figure 1.3 shows the demodulator of a pulse-shaping OFDM system. At each OFDM symbol interval n, the demodulator derives the K sequences $Y_{n,k}$, $k = 0, \ldots, K-1$. This is done by calculation of the inner products

$$Y_{n,k} = \langle y, f_{n,k} \rangle = \sum_{m=-\infty}^{\infty} y[m] f_{n,k}^*[m],$$

where $y[m]$ and $f[m]$ denote the received signal and the receive pulse, respectively. The time-frequency shifted receive pulses are

$$f_{n,k}[m] \triangleq f[m - nN]\, e^{j2\pi\frac{k}{K}(m-nN)}.$$

Note that in the absence of noise and distortions, perfect demodulation (i.e., $Y_{n,k} = X_{n,k}$) is obtained if and only if the transmit pulse $g[m]$ and the receive pulse $f[m]$ satisfy the *biorthogonality condition* [12]

$$\langle g_{n',k'}, f_{n,k} \rangle = \delta[n - n']\, \delta[k - k']. \tag{1.2}$$

For identical transmit and receive pulses (i.e., $g[m] = f[m]$), (1.2) becomes an orthogonality condition. Note that condition (1.2) can only be fullfilled in the case $N/K \geq 1$. Typically N/K is assumed between 1.03 and 1.25.

1.1.3 Cyclic-Prefix

A cyclic-prefix OFDM (CP-OFDM) system is the most commonly applied variant of OFDM [13]. The CP can be seen as a guard interval avoiding intersymbol interference (ISI). More

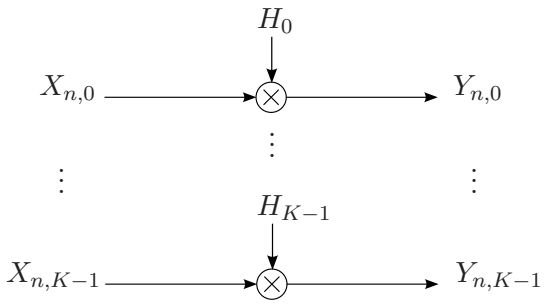

Figure 1.4: *Input-output relation for one OFDM symbol of a CP-OFDM system. The channel is time-invariant and noiseless.*

precisely, the OFDM symbol is enlarged by repeating a part of the symbol, the so called cyclic prefix. It protects consecutive OFDM symbols from ISI in the case of transmission over a multipath channel.
For time-invariant channels with an impulse response shorter than the CP, the channel input-output relation (after discarding CP) turns from linear convolution to a cyclic convolution. The output then can be expressed by a scalar multiplication,

$$Y_{n,k} = X_{n,k} H_k,$$

where n and k indicate the position in the time-frequency lattice, $Y_{n,k}$ is the output, $X_{n,k}$ is the input, and H_k the channel coefficient of the time-invariant channel (hence there is no dependence on n). This input-output relation for one OFDM symbol and a time-invariant channel is shown in Figure 1.4.
The CP also prevents intercarrier interference (ICI). Further the signal structure imposed by CP-OFDM can be exploited for time synchronization.

1.1.4 OFDMA

Orthogonal frequency division multiple access (OFDMA) [14] is a combination of OFDM and multiple access schemes as *frequency division multiple access* (FDMA) and *time division multiple access* (TDMA). In difference to OFDM where only one user can use the entire time-frequency lattice, OFDMA supports a multiuser communication. Therefore OFDMA assigns the K subcarriers to at most K single users, i.e.,

$$X_{n,k}^{(u)} = X_{n,k}, \qquad \text{for } k \in \mathcal{K}_U, \qquad \text{and } X_{n,k}^{(u)} = 0, \qquad \text{for } k \notin \mathcal{K}_U,$$

with user index u and $\mathcal{K}_U$ denoting the subset of subcarriers allocated to user u. Each user has its own distinct subcarriers. In other words there is no collision in frequency. Figure 1.5(a) shows this. By assigning the orthogonal frequency bands to the different users, multiuser interference should be avoided. For OFDMA the same drawbacks as for OFDM can be found, which includes the high peak-to-average power ratio. A multiple access interference (MAI)

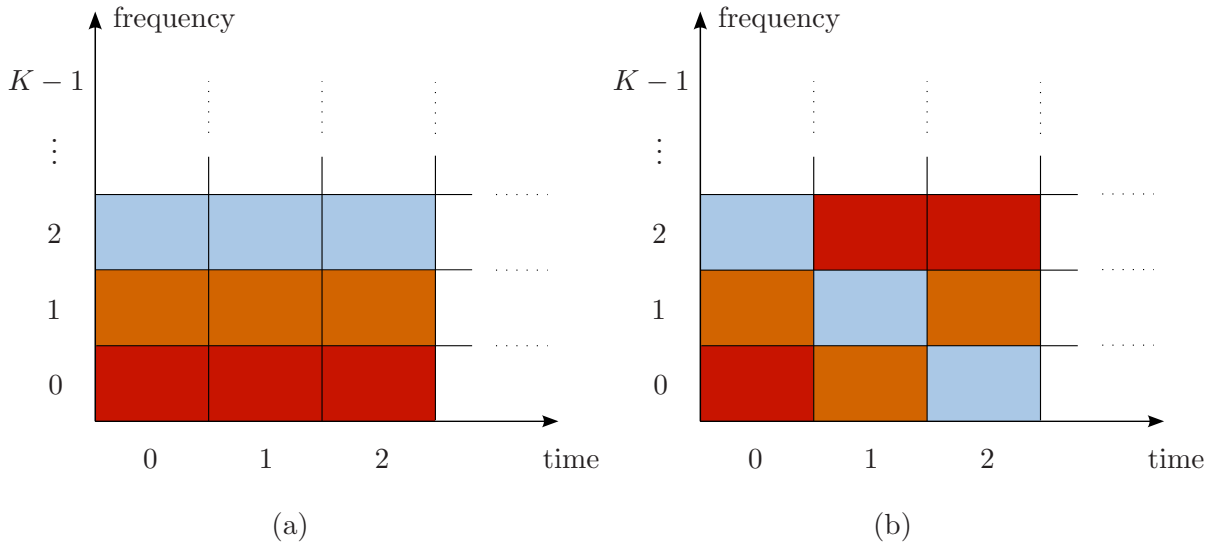

Figure 1.5: *Representation of (a) OFDMA and (b) FH-OFDMA in the time-frequency area. Note that different users are marked by different colors. The subcarriers* $0, 1, \ldots, K-1$ *and the OFDM symbols* $0, 1, \ldots, M-1$ *are shown.*

caused by symbol timing misalignments is disadvantageous since it destroys the orthogonal structure between the users. The performance of OFDMA gets poor if a single channel is adepted by deep fades.
If the assigned subcarriers are changed for every OFDM symbol, this procedure is called frequency hopping, i.e., one gets a frequency-hopped (FH)-OFDMA system. This is shown in Figure 1.5(b). By frequency hopping a frequency diversity is obtained.

1.2 Multi-Carrier CDMA

By combining *direct sequence CDMA* (DS-CDMA) and OFDM, a *Multi-Carrier-CDMA* (MC-CDMA) system is obtained. These technique has been proposed by [15–19]. Other techniques that combine DS-CDMA and OFDM are MC-DS-CDMA [18] and Multi-Tone CDMA (MT-CDMA) [19]. This overview is mainly based on [20].
In MC-CDMA, spreading sequences are used to map the datastream to the OFDM subcarriers. Simply said by MC-CDMA a serial datastream is split onto different subcarriers, each with lower data rate. By matching different spreading sequences to different users a multiuser system is implemented. Then, an overlap in frequency is realized. Hence, with MC-CDMA full collision in frequency can be obtained. The spread subcarrier sequence for a multiuser system can be written as (cf. Figure 1.6)

$$X_{n,k} = \sum_{u=0}^{U-1} X_n^{(u)} c_k^{(u)},$$

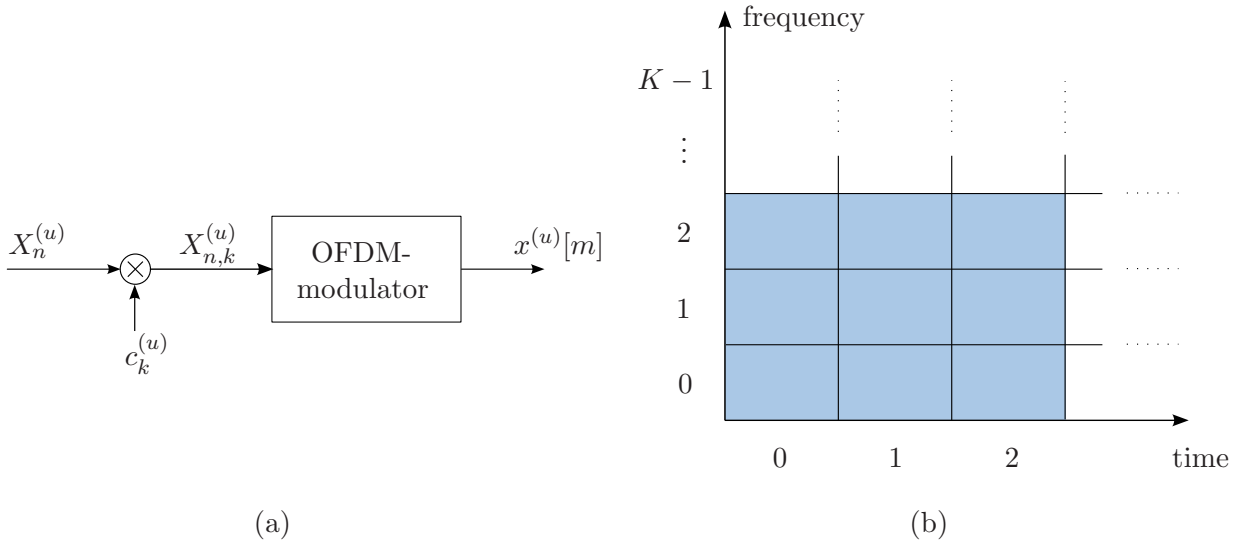

Figure 1.6: *(a) The principle of MC-CDMA modulation by concatenation of DS-CDMA ($X_{n,k}^{(u)} = X_n^{(u)} c_k^{(u)}$) and OFDM for a single user ($U = 1$); (b) The user's signal is spread over the entire time-frequency area.*

where U denotes the number of users, $X_n^{(u)}$ the input data stream of user u, and $c_k^{(u)}$ the spreading sequence of user u.
Figure 1.6 also shows the concatenation of spreader and OFDM-Modulator or DS-CDMA and OFDM, respectively. The OFDM modulation is equivalent to that in Section 1.1.1. Further Figure 1.6 shows the spread signal in the time-frequency area.

1.3 Wireless Fading Channels

In this section the main topic is the mathematical characterization of the wireless fading channel. No physical details but rather a statistical representation is used. Some of the keywords concerning wave propagation in wireless communications are free-space propagation, scattering, reflection, and diffraction. A survey of these effects is given in [21–23]. The attention will be restricted to small-scale fading effects. Large-scale fading like time variation of path loss and other statistical channel characteristics will be neglected.

1.3.1 Continuous-Time Channel Model

Channel Input-Output Relation

The continuous-time time-varying wireless channel is denoted by $\mathbb{H}_c$. The input-output relation then is [1]

$$y(t) = (\mathbb{H}_c x)(t) + u(t) = \int_\tau h(t,\tau) x(t-\tau) d\tau + u(t), \tag{1.3}$$

[1] Integrals are from $-\infty$ to ∞ unless specified otherwise.

with $x(t)$ the input, $y(t)$ the output and $h(t,\tau)$ the time-varying impulse response of the channel. Furthermore $u(t)$ denotes the additive zero-mean white Gaussian noise with power spectral density N_0. The impulse response $h(t,\tau)$ is defined in the interval $[0,\tau_{\max}]$ since $\mathbb{H}_c$ is causal with maximum delay $\tau_{\max}$ (due to physical reasons). The variation of $h(t,\tau)$ in time is limited by the maximum Doppler frequency $\nu_{\max}$.

System Descriptions

By applying Fourier transformation to the impulse response $h(t,\tau)$, the *time-dependent transfer function*

$$H_{\mathbb{H}_c}(t,f) \triangleq \int_{\tau} h(t,\tau)e^{-j2\pi f\tau}, d\tau. \tag{1.4}$$

and the *spreading function*

$$S_{\mathbb{H}_c}(\tau,\nu) \triangleq \int_{t} h(t,\tau)e^{-j2\pi\nu t}dt, \tag{1.5}$$

are obtained. The time-dependent transfer function $H_{\mathbb{H}_c}(t,f)$ describes the variation of the channel's transfer function in time, the spreading function $S_{\mathbb{H}_c}(\tau,\nu)$ describes the allocation in terms of τ and ν. A support region for $S_{\mathbb{H}_c}(\tau,\nu)$ is given through the limitations on $h(t,\tau)$ and is given by $(\tau,\nu)\in[0,\tau_{\max}]\times[-\nu_{\max}/2,\nu_{\max}/2]$. The time-dependent transfer function $H_{\mathbb{H}_c}(t,f)$ then is a lowpass function. The channel spread $\tau_{\max}\nu_{\max}$ can be used to classify channels as underspread and overspread. An underspread channel is defined by $\tau_{\max}\nu_{\max}\leq 1$ and an overspread channel by $\tau_{\max}\nu_{\max}>1$ [24]. Practical mobile radio channels are underspread.

Channel Statistics

The channel is assumed as wide-sense stationary with uncorrelated scatterers (WSSUS) and Rayleigh fading [21–23]. There is no dependence of the channel statistics on time even though the channel realizations are time-varying. Then, the correlation function of the impulse response is

$$\mathrm{E}\{h(t,\tau)h^*(t',\tau')\} = D_{\mathbb{H}_c}(t-t',\tau)\delta(\tau-\tau'),$$

with the *time-delay correlation function* $D_{\mathbb{H}_c}(\Delta t,\tau)$. The spreading function is white which gives

$$\mathrm{E}\{S_{\mathbb{H}_c}(\tau,\nu)S^*_{\mathbb{H}_c}(\tau',\nu')\} = C_{\mathbb{H}_c}(\tau,\nu)\delta(\tau-\tau')\delta(\nu-\nu'). \tag{1.6}$$

Here, $C_{\mathbb{H}_c}(\tau,\nu)>0$ is the real-valued and non-negative scattering function [25]. The support region is the same as for $S_{\mathbb{H}_c}(\tau,\nu)$, $(\tau,\nu)\in[0,\tau_{\max}]\times[-\nu_{\max}/2,\nu_{\max}/2]$. The scattering function describes the channels statistic distribution over τ and ν. By

$$\mathrm{E}\{H_{\mathbb{H}_c}(t,f)H^*_{\mathbb{H}_c}(t+\Delta t,f+\Delta f)\} = R_{\mathbb{H}_c}(\Delta t,\Delta f)$$

the *time-frequency correlation function* $R_{\mathbb{H}_c}(\Delta t,\Delta f)$ is defined. It describes the correlation of channel in terms of time- and frequency-shifts. There are relations between $C_{\mathbb{H}_c}(\tau,\nu)$,

$D_{\mathbb{H}_c}(\Delta t, \tau)$ and $R_{\mathbb{H}_c}(\Delta t, \Delta f)$ through Fourier transformation:

$$R_{\mathbb{H}_c}(\Delta t, \Delta f) = \int_\tau D_{\mathbb{H}_c}(\Delta t, \tau) e^{-j2\pi\Delta f\tau} d\tau = \int_\tau \int_\nu C_{\mathbb{H}_c}(\tau, \nu) e^{-j2\pi(\nu\Delta t - \tau\Delta f)} d\tau d\nu,$$
$$C_{\mathbb{H}_c}(\tau, \nu) = \int_\nu D_{\mathbb{H}_c}(\Delta t, \tau) e^{-j2\pi\nu\Delta t} d\Delta t.$$

For the effective support region of the time-frequency correlation function $R_{\mathbb{H}_c}(\Delta t, \Delta f)$ then $(\Delta t, \Delta f) \in [-1/\nu_{\max}, 1/\nu_{\max}] \times [-1/\tau_{\max}, 1/\tau_{\max}]$ is obtained.

Channel Parameters

The *path loss* of a WSSUS channel is given by [25]

$$\sigma^2_{\mathbb{H}_c} = R_{\mathbb{H}_c}(0,0) \triangleq \int_\tau \int_\nu C_{\mathbb{H}_c}(\tau, \nu) d\tau d\nu. \tag{1.7}$$

It describes the mean attenuation of the channel. The *delay spread* $\alpha_{\mathbb{H}_c}$ and the Doppler spread $\beta_{\mathbb{H}_c}$ are normalized second moments of the scattering function. i.e.,

$$\alpha_{\mathbb{H}_c} \triangleq \frac{1}{\sigma^2_{\mathbb{H}_c}} \int_\tau \int_\nu (\tau - \tau_0)^2 C_{\mathbb{H}_c}(\tau, \nu) d\tau d\nu,$$
$$\beta_{\mathbb{H}_c} \triangleq \frac{1}{\sigma^2_{\mathbb{H}_c}} \int_\tau \int_\nu (\nu - \nu_0)^2 C_{\mathbb{H}_c}(\tau, \nu) d\tau d\nu.$$

There, (τ_0, ν_0) defines the center of gravity of $C_{\mathbb{H}_c}(\tau, \nu)$. $\alpha_{\mathbb{H}_c}$ and $\beta_{\mathbb{H}_c}$ can be approximated by $\tau_{\max}$ and $\nu_{\max}$ as they are in the same range. The delay spread describes the range of delays through multipath propagation, the Doppler spread the range of frequency-shift caused by the time-variant channel. The reciprocals of $\alpha_{\mathbb{H}_c}$ and $\beta_{\mathbb{H}_c}$ are known as *coherence time* $T_{\mathbb{H}_c}$ and *coherence bandwidth* $B_{\mathbb{H}_c}$, respectively.

$$T_{\mathbb{H}_c} \triangleq \frac{1}{\beta_{\mathbb{H}_c}} \quad \text{and} \qquad B_{\mathbb{H}_c} \triangleq \frac{1}{\alpha_{\mathbb{H}_c}}$$

Accordingly $T_{\mathbb{H}_c}$ and $B_{\mathbb{H}_c}$ have the same range as $1/\nu_{\max}$ and $1/\tau_{\max}$, respectively. The coherence time $T_{\mathbb{H}_c}$ is a measure of the time-variance of the channel. A small $T_{\mathbb{H}_c}$ characterizes a fast fading channel and a large $T_{\mathbb{H}_c}$ a slowly fading channel. The coherence bandwidth is a measure of the bandwidth (or frequencies) which are affected by the channel response.

1.3.2 Discrete-Time Channel Model

Based on the continuous-time channel model the transition to the discrete-time channel model is done by discretization. The discrete-time channel will be denoted by $\mathbb{H}_d$. Figure 1.7 shows a discretized continuous-time channel.

By use of a digital-to-analog (D/A) conversion, the discrete-time input signal $x[m]$ is converted to a continuous-time signal $x(t)$. To that end, a transmit filter $\gamma(t)$ is applied, i.e.,

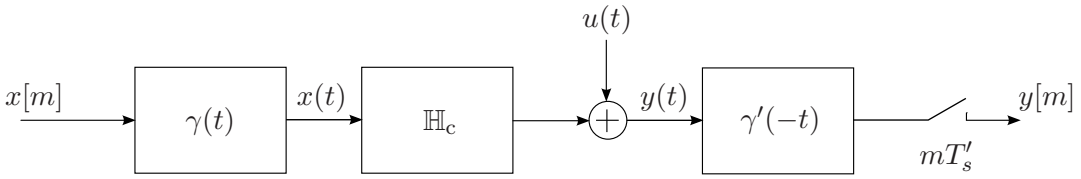

Figure 1.7: *Discretization of $\mathbb{H}_c$. Therefore the D/A conversion at the transmitter and the A/D conversion at the receiver are assumed ideal.*

$$x(t) = \sum_{m=-\infty}^{\infty} x[m]\gamma(t - mT_s).$$

Here $T_s = 1/B$ is the sample duration with B the transmission bandwidth . For $\gamma(t)$ an ideal filter is assumed (i.e., $\gamma(t) = \sqrt{B}\text{sinc}(\pi Bt)$ with $\text{sinc}(x) \triangleq \frac{\sin(x)}{x}, ||\gamma||^2 = \int_t |\gamma(t)|^2 dt = 1$). The time-varying channel is represented by $\mathbb{H}_c$ and there is additive noise $u(t)$. The received continuous-time signal $y(t)$ is passed to a receive filter $\gamma'(t)$ in order to apply noise suppression and anti-aliasing ($T'_s = 1/B'$ with the receive bandwidth $B' = B + \nu_{\max}$). Like the transmit filter it is assumed as ideal. Then

$$y[m] = \int_t y(t)\gamma'(t - mT'_s)dt \tag{1.8}$$

is the discretized output signal.

Channel Input-Output Relation

By applying the continuous-time output signal (1.3) to (1.8) $y[m]$ is obtained as

$$y[m] = \frac{1}{\sqrt{B'}} \int_\tau h(mT'_s, \tau)x(mT'_s - \tau)d\tau + u[m], \tag{1.9}$$

with

$$u[m] = \sqrt{B'} \int_t u(t)\text{sinc}(\pi B'(t - mT'_s))dt$$

the the zero-mean white Gaussian discrete-time noise with variance $\sigma_u^2 = \text{E}\{|u[m]|^2\} = N_0$. Substitution of $x(t)$ in (1.9) gives

$$y[m] = \sqrt{\frac{B}{B'}} \sum_{l=-\infty}^{\infty} x[m - l] \int_\tau h(mT'_s, \tau)\text{sinc}(\pi B[\tau - m(T'_s - T_s) - lT_s])d\tau + u[m].$$

For relative high bandwidths B' (as in wideband communication systems) an assumption on bandwidth and sampling duration, respectively can be done since the Doppler frequency $\nu_{\max}$ is small compared to B. Then, $B' = B$ and $T'_s = T_s$ and the input-output relation simplifies to

$$y[m] = (\mathbb{H}_d x)[m] + u[m] = \sum_{l=-\infty}^{\infty} h[m, l]x[m - l] + u[m], \tag{1.10}$$

with

$$h[m,l] \triangleq \int_\tau h(mT_s,\tau)\mathrm{sinc}\left(\pi B(\tau - lT_s)\right) d\tau \approx \frac{1}{B} h(mT_s, lT_s) \tag{1.11}$$

the time-varying discrete-time impulse response (approximation for large bandwidths). In general, $h[m,l]$ is nonzero for $l \in \mathbb{Z}$. Again for large bandwidths and analogous to the continuous-time channel ($h(t,\tau) = 0$ for $\tau > \tau_{\max}$) we obtain for the impulse response of discrete-time channel $\mathbb{H}_\mathrm{d}$

$$h[m,l] \approx 0 \quad \text{for} \qquad l \notin [0, L],$$

where

$$L = \left\lceil \frac{\tau_{\max}}{T_s} \right\rceil. \tag{1.12}$$

Then, $L+1$ channel delay taps are resolved and the input-output relation in (1.10) becomes

$$y[m] = (\mathbb{H}_\mathrm{d} x)[m] + u[m] = \sum_{l=0}^{L} h[m,l]x[m-l] + u[m]. \tag{1.13}$$

System Description

Analogous to the continuous-time channel the *spreading function* can be calculated from the channel impulse response by

$$S_{\mathbb{H}_\mathrm{d}}(l,\xi) \triangleq \sum_{m=-\infty}^{\infty} h[m,l] e^{-j2\pi\xi m}. \tag{1.14}$$

This has been described in [21, 26]. Insertion of (1.11) in (1.14) and using (1.5) and Poisson's sum formula gives

$$\begin{aligned} S_{\mathbb{H}_\mathrm{d}}(l,\xi) &= \int_\tau \int_\nu S_{\mathbb{H}_\mathrm{c}}(\tau,\nu)\mathrm{sinc}\left(\pi B(\tau - lT_s)\right) \sum_{m=-\infty}^{\infty} e^{-j2\pi T_s\left(\nu - \frac{\xi}{T_s}\right)m} d\tau d\nu \\ &= B \sum_{m=-\infty}^{\infty} \int_\tau S_{\mathbb{H}_\mathrm{c}}\left(\tau, (\xi+m)B\right) \mathrm{sinc}\left(\pi B(\tau - lT_s)\right) d\tau. \end{aligned}$$

Now again some simplification can be done by taking into account the support of $S_{\mathbb{H}_\mathrm{c}}(\tau,\nu)$. Since the Doppler frequencies are limited to $|\nu| \le \nu_{\max}/2$ and $\nu_{\max} \ll B/2$ there is no aliasing for $S_{\mathbb{H}_\mathrm{d}}(l,\xi)$ which gives

$$S_{\mathbb{H}_\mathrm{d}}(l,\xi) = B \int_\tau S_{\mathbb{H}_\mathrm{c}}(\tau, \xi B)\mathrm{sinc}\left(\pi B(\tau - lT_s)\right) d\tau, \qquad \xi \in [-1/2, 1/2). \tag{1.15}$$

A further approximation can be done since the bandwidth is large. Then, (1.15) becomes

$$S_{\mathbb{H}_\mathrm{d}}(l,\xi) = S_{\mathbb{H}_\mathrm{c}}(lT_s, \xi B), \qquad (l,\xi) \in [0, L] \times [-\xi_{\max}/2, \xi_{\max}/2],$$

where L is given by (1.12) and the maximum normalized Doppler frequency is

$$\xi_{\max} = T_s \nu_{\max} = \frac{\nu_{\max}}{B}.$$

If the continuous-time channel $\mathbb{H}_\mathrm{c}$ is underspread, it follows that

$$\xi_{\max} L = \frac{\nu_{\max}}{B} \left\lceil \frac{\tau_{\max}}{T_s} \right\rceil \approx \nu_{\max}\tau_{\max} < 1,$$

i.e., the support of $S_{\mathbb{H}_\mathrm{d}}(l,\xi)$ is less than one.

Channel Statistics

The correlation function of $h[m,l]$ is

$$\mathrm{E}\{h[m,l], h^*[m',l']\} = \int_\tau D_{\mathbb{H}_\mathrm{c}}\left((m-m')T_s,\tau\right) \mathrm{sinc}\left(\pi B(\tau - lT_s)\right) \mathrm{sinc}\left(\pi B(\tau - l'T_s)\right) d\tau. \tag{1.16}$$

In general different taps are correlated. But for large bandwidths this effect is negligible and the *discrete WSSUS* (DWSSUS) assumption can be applied. Uncorrelated taps in (1.16) are equivalent to $\mathrm{E}\{h[m,l], h^*[m',l']\} = 0$ for $l \neq l'$. This gives for the correlation function

$$\mathrm{E}\{h[m,l], h^*[m',l']\} = D_{\mathbb{H}_\mathrm{d}}[m-m',l]\delta[l-l'].$$

Then,

$$D_{\mathbb{H}_\mathrm{d}}[m,l] = \int_\tau D_{\mathbb{H}_\mathrm{c}}(mT_s,\tau)\, \mathrm{sinc}^2\left(\pi B(\tau - lT_s)\right) d\tau \tag{1.17}$$

is the *time-delay correlation function* of $\mathbb{H}_\mathrm{d}$ for $l = l'$. The statistics of the spreading function are given by

$$\mathrm{E}\{S_{\mathbb{H}_\mathrm{d}}(l,\xi), S_{\mathbb{H}_\mathrm{d}}(l',\xi')\} = C_{\mathbb{H}_\mathrm{d}}(l,\xi)\delta[l-l']\delta[\xi-\xi'], \qquad \xi,\xi' \in [-1/2, 1/2).$$

Hence the *scattering function* of $\mathbb{H}_\mathrm{d}$ is calculated from $D_{\mathbb{H}_\mathrm{d}}[m,l]$ by Fourier transformation:

$$C_{\mathbb{H}_\mathrm{d}}(l,\xi) = \sum_{m=-\infty}^{\infty} D_{\mathbb{H}_\mathrm{d}}[m,l] e^{-j2\pi\xi m}. \tag{1.18}$$

The relation to the continuous-time case is obtained by combining (1.18) and (1.17) and can be written as

$$C_{\mathbb{H}_\mathrm{d}}(l,\xi) = B^2 \int_\tau C_{\mathbb{H}_\mathrm{c}}(\tau,\xi B)\, \mathrm{sinc}^2\left(\pi B(\tau - lT_s)\right) d\tau. \tag{1.19}$$

For large bandwidths (1.17) and (1.19) can be approximated by

$$D_{\mathbb{H}_\mathrm{d}}[m,l] \approx \frac{1}{B^2} D_{\mathbb{H}_\mathrm{c}}(mT_s, lT_s), \tag{1.20}$$

$$C_{\mathbb{H}_\mathrm{d}}(l,\xi) \approx C_{\mathbb{H}_\mathrm{c}}(lT_s, \xi B), \qquad \xi \in [-1/2, 1/2), \tag{1.21}$$

which corresponds to a sampling of $D_{\mathbb{H}_\mathrm{c}}(\Delta t,\tau)$ and $C_{\mathbb{H}_\mathrm{c}}(\tau,\nu)$, respectively. The support region of $C_{\mathbb{H}_\mathrm{d}}(l,\xi)$ is $(l,\xi) \in [0,L] \times [-\xi_{\max}, \xi_{\max}]$.

Channel Statistics

For the *path loss* of $\mathbb{H}_{\rm d}$ we get

$$\sigma_{\mathbb{H}_{\rm d}} \triangleq \sum_{l=0}^{L} \int_{-1/2}^{1/2} C_{\mathbb{H}_{\rm d}}(l,\xi)d\xi \approx \sum_{l=0}^{L} \int_{-1/2}^{1/2} C_{\mathbb{H}_{\rm c}}(lT_s,\xi B)d\xi \approx \frac{1}{T_s}\int_{\tau}\int_{-1/2}^{1/2} C_{\mathbb{H}_{\rm c}}(\tau,\xi B)d\tau d\xi$$
$$= \int_{\tau}\int_{-B/2}^{B/2} C_{\mathbb{H}_{\rm c}}(\tau,\nu B)d\tau d\nu = \sigma_{\mathbb{H}_{\rm c}},$$

where the approximation for large bandwidth (includes small T_s)in (1.21) in order to get integration instead of summation are used. Then, the path loss $\sigma_{\mathbb{H}_{\rm d}}$ is equal to the continuous-time channel path loss $\sigma_{\mathbb{H}_{\rm c}}$ in (1.7).
The *delay spread* $\alpha^2_{\mathbb{H}_{\rm d}}$ is defined by

$$\alpha^2_{\mathbb{H}_{\rm d}} \triangleq \frac{1}{\sigma^2_{\mathbb{H}_{\rm d}}} \sum_{l=0}^{L} \int_{-\xi_{\max}}^{\xi_{\max}} (l-l_0)^2 C_{\mathbb{H}_{\rm d}}(l,\xi)d\xi,$$

and approximately equal to

$$\alpha^2_{\mathbb{H}_{\rm d}} \approx \frac{1}{\sigma^2_{\mathbb{H}_{\rm c}}} \sum_{l=0}^{L} \int_{-\xi_{\max}}^{\xi_{\max}} (l-l_0)^2 C_{\mathbb{H}_{\rm c}}(lT_s,\xi B)d\xi \approx \frac{1}{T_s^2\sigma^2_{\mathbb{H}_{\rm c}}} \int_{\tau}\int_{\nu} (\tau-\tau_0)^2 C_{\mathbb{H}_{\rm c}}(\tau,\nu) = \frac{\alpha^2_{\mathbb{H}_{\rm c}}}{T_s^2},$$

i.e.,

$$\alpha_{\mathbb{H}_{\rm d}} \approx \frac{\alpha_{\mathbb{H}_{\rm c}}}{T_s} = B\alpha_{\mathbb{H}_{\rm c}}.$$

Similarly, the *Doppler spread* $\beta^2_{\mathbb{H}_{\rm d}}$ is defined by

$$\beta^2_{\mathbb{H}_{\rm d}} \triangleq \frac{1}{\sigma^2_{\mathbb{H}_{\rm d}}} \sum_{l=0}^{L} \int_{-\xi_{\max}}^{\xi_{\max}} (\xi-\xi_0)^2 C_{\mathbb{H}_{\rm d}}(l,\xi)d\xi$$

and approximately equal to

$$\beta_{\mathbb{H}_{\rm d}} \approx T_s\beta_{\mathbb{H}_{\rm c}} = \frac{\beta_{\mathbb{H}_{\rm c}}}{B}.$$

Hence definition and approximation of *coherence time* and *coherence bandwidth* are

$$T_{\mathbb{H}_{\rm d}} \triangleq \frac{1}{\beta_{\mathbb{H}_{\rm d}}} \quad \text{and} \qquad B_{\mathbb{H}_{\rm d}} \triangleq \frac{1}{\alpha_{\mathbb{H}_{\rm d}}},$$

with the relations

$$T_{\mathbb{H}_{\rm d}} \approx \frac{T_{\mathbb{H}_{\rm c}}}{T_s} = BT_{\mathbb{H}_{\rm c}} \quad \text{and} \qquad B_{\mathbb{H}_{\rm d}} \approx T_sB_{\mathbb{H}_{\rm c}} = \frac{B_{\mathbb{H}_{\rm c}}}{B}.$$

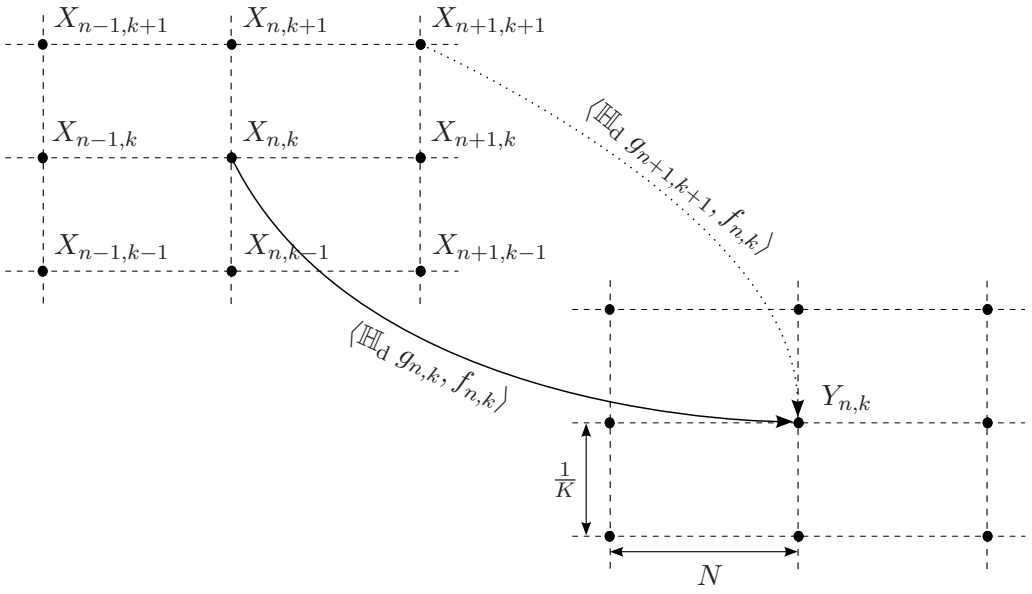

Figure 1.8: *Multiplicative input-output relation.*

1.3.3 Input-Output Relation

In this section we describe an approximate input-output relation for a single user OFDM system. This concludes OFDM modulation (1.1.1), a time-varying channel given by (1.13), and OFDM demodulation (1.1.2). Referring to [27] the system input-output relation is given by

$$Y_{n,k} = \sum_{n'=-\infty}^{\infty} \sum_{k'=0}^{K-1} \langle \mathbb{H}_\mathrm{d} g_{n',k'}, f_{n,k} \rangle X_{n',k'} + Z_{n,k}, \tag{1.22}$$

with

$$Z_{n,k} \triangleq \langle u, f_{n,k} \rangle,$$

where $\mathbb{H}_\mathrm{d}$ denotes the discrete-time channel and u is zero-mean white Gaussian discrete-time noise.

In (1.22) every $X_{n',k'}$ for $(n',k') \in \mathbb{Z} \times [0, K-1]$ contributes to each $Y_{n,k}$. Splitting (1.22) into a desired and an interference term gives

$$\langle \mathbb{H}_\mathrm{d} g_{n,k}, f_{n,k} \rangle X_{n,k} \tag{1.23}$$

for the desired term and

$$\sum_{\substack{n'=-\infty \\ n' \neq n}}^{\infty} \sum_{\substack{k'=0 \\ k' \neq k}}^{\infty} \langle \mathbb{H}_\mathrm{d} g_{n',k'}, f_{n,k} \rangle X_{n',k'} \tag{1.24}$$

for the interference term. This is shown in Figure 1.8.

The interference term (1.24) can be decomposed into *intersymbol interference* (ISI) and *inter-carrier interference* (ICI), depending on whether $n \neq n'$ or $k \neq k'$, respectively.

ICI is the interference from different subcarriers within one OFDM symbol. In most cases other than neighboring subcarriers can be neglected, e.g., by choosing appropriate transmit and receive pulses. But there is no possibility to complete avoid ICI.
Neighboring symbols much more likely contribute to ISI. Typically ISI is avoided by using a CP or a guard interval.

Approximation

By neglecting the interference terms, only the desired term in (1.23) remains (cf. [12]). Errors caused by this approximation have been investigated in [12]. For underspread channels and properly chosen pulses $g[m]$ and $f[m]$ the interference term (1.24) is approximately zero which means

$$\langle \mathbb{H}_{\mathrm{d}}\, g_{n',k'}, f_{n,k} \rangle \approx 0 \qquad \text{for} \qquad (n',k') \neq (n,k).$$

Then, (1.22) simplifies to

$$Y_{n,k} = H_{n,k} X_{n,k} + Z_{n,k}, \tag{1.25}$$

with *channel coefficients*

$$H_{n,k} \triangleq \langle \mathbb{H}_{\mathrm{d}}\, g_{n,k}, f_{n,k} \rangle.$$

This representation indeed allows to characterize a practical wireless channel using OFDM because in practice wireless channels are underspread.

1.4 Entropy, Mutual Information, and System Capacity

This short overview of basic information theoretic concepts is based on [28]. Most elementary notions and properties in information theory will be shown. The most basic concept is the calculation of the *differential entropy* of a (continuous) random variable trough its probability density function (pdf). The next step is to define *mutual information* using entropies and some of its properties. To conclude the overview, the terms of *information rate* and *system capacity* are introduced.

1.4.1 Entropy

Entropy marks the most elementary part in information theory. It's a quantity that indicates the uncertainty of a random variable. First a differentiation between discrete and random values has to be done. While the discrete random variable is described through a *probability mass function* (pmf) the continuous random variable is characterized by its pdf. Since this work is based on continuous random vectors we will use the differential entropy. For a given continuous random vector $\mathbf{X}$ with pdf $p_{\mathbf{X}}(\mathbf{x})$ the *support region* $\mathcal{S}$ is the set of outcomes $\mathbf{x}$ for which $p_{\mathbf{X}}(\mathbf{x})$ is nonzero, i.e., $p_{\mathbf{X}}(\mathbf{x}) > 0$ for $\mathbf{x} \in \mathcal{S}$ and $p_{\mathbf{X}}(\mathbf{x}) = 0$ for $\mathbf{x} \notin \mathcal{S}$. Then, the

differential entropy of $\mathbf{X}$ is calculated as

$$h(\mathbf{X}) = h(p_{\mathbf{X}}) = \mathrm{E}\left\{\log\frac{1}{p_{\mathbf{X}}(\mathbf{x})}\right\} = \int_{\mathcal{S}} p_{\mathbf{X}}(\mathbf{x})\log\frac{1}{p_{\mathbf{X}}(\mathbf{x})}d\mathbf{x} \tag{1.26}$$

$$= -\int_{\mathcal{S}} p_{\mathbf{X}}(\mathbf{x})\log\left(p_{\mathbf{X}}(\mathbf{x})\right)d\mathbf{x}. \tag{1.27}$$

Throughout the thesis the natural logarithm will be used. Furthermore there are the joint entropy $p_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y})$ and conditional entropy $p_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}|\mathbf{y})$. The joint differential entropy for the vectors $\mathbf{X}$ and $\mathbf{Y}$ with pdfs $p_{\mathbf{X}}(\mathbf{x})$ and $p_{\mathbf{Y}}(\mathbf{y})$, respectively is given by

$$h(\mathbf{X},\mathbf{Y}) = h(p_{\mathbf{X},\mathbf{Y}}) = \mathrm{E}\left\{\log\frac{1}{p_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y})}\right\} = -\int_{\mathcal{S}_{x,y}}\int p_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y})\log\left(p_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y})\right)d\mathbf{x}d\mathbf{y},$$

with $\mathcal{S}_{x,y}$ the support region of $p_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y})$. Conditional differential entropy for $\mathbf{X}$ given $\mathbf{Y}$ can be calculated as

$$h(\mathbf{X}|\mathbf{Y}) = h(p_{\mathbf{X}|\mathbf{Y}}) = \mathrm{E}\left\{\log\frac{1}{p_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}|\mathbf{y})}\right\} = -\int_{\mathcal{S}_{x,y}}\int p_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y})\log\left(p_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}|\mathbf{y})\right)d\mathbf{x}\,d\mathbf{y}$$

$$= \int_{\mathcal{S}_y} p_{\mathbf{Y}}(\mathbf{y})h(\mathbf{X}|\mathbf{Y}=\mathbf{y})\,d\mathbf{y},$$

with

$$h(\mathbf{X}|\mathbf{Y}=\mathbf{y}) = \int_{\mathcal{S}_x} p_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}|\mathbf{y})\log\left(p_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}|\mathbf{y})\right)d\mathbf{x},$$

and $\mathcal{S}_x$ and $\mathcal{S}_y$ the support regions of $\mathbf{X}$ and $\mathbf{Y}$, respectively. The effect of conditioning is a decreasing entropy since there remains less uncertainty. In system capacity calculations conditioning on e.g. input symbols or channel coefficients are common.

Gaussian random vector

One of the most common variable is the Gaussian random vector. For fixed variance, the Gaussian distribution is the one that maximizes entropy. So it always is an upper bound on the entropy of any random variable.
The differential entropy for a $M \times 1$ real-valued Gaussian random vector $\mathbf{X} \sim \mathcal{N}(\mu_{\mathbf{X}}, \mathbf{C}_{\mathbf{X}})$ with mean $\mu_{\mathbf{X}}$ and covariance matrix $\mathbf{C}_{\mathbf{X}} = \mathrm{E}\left\{(\mathbf{X}-\mu_{\mathbf{X}})(\mathbf{X}-\mu_{\mathbf{X}})^H\right\}$ is

$$h(\mathbf{X}) = \frac{1}{2}\log\left((\pi e)^M \det\mathbf{C}_{\mathbf{X}}\right). \tag{1.28}$$

For the scalar case $X \sim \mathcal{N}(\mu_X, \sigma_X)$ this gives

$$h(X) = \frac{1}{2}\log\left(\pi e\sigma_X^2\right).$$

For circularly complex Gaussian random vectors the entropy can be calculated by applying (1.28) to the real and imaginary part and summation of these results. For a $M \times 1$ complex-valued Gaussian random vector $\mathbf{X} \sim \mathcal{CN}(\mu_{\mathbf{X}}, \mathbf{C}_{\mathbf{X}})$ this gives

$$h(\mathbf{X}) = \log\left((\pi e)^M \det\mathbf{C}_{\mathbf{X}}\right). \tag{1.29}$$

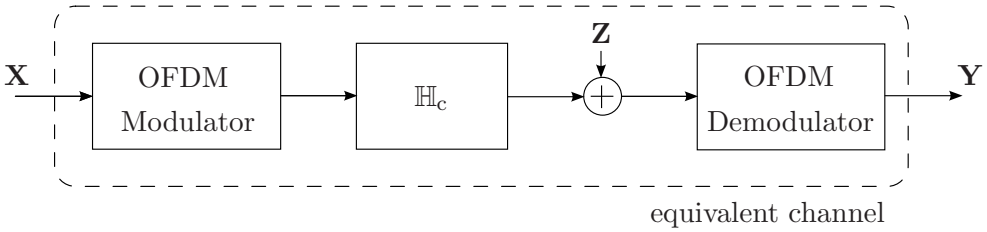

Figure 1.9: *Equivalent channel representation for system capacity calculation.*

1.4.2 Mutual Information

Mutual information is a measure of the difference of the joint distribution $p_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y})$ and the product distribution $p_{\mathbf{X}}(\mathbf{x})p_{\mathbf{Y}}(\mathbf{y})$, i.e.,

$$I(\mathbf{Y};\mathbf{X}) \triangleq \mathrm{E}\left\{\log\frac{p_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y})}{p_{\mathbf{X}}(\mathbf{x})p_{\mathbf{Y}}(\mathbf{y})}\right\} = \int\limits_{\mathcal{S}_{\mathbf{x},\mathbf{y}}}\int p_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y})\log\frac{p_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y})}{p_{\mathbf{X}}(\mathbf{x})p_{\mathbf{Y}}(\mathbf{y})}d\mathbf{x}d\mathbf{y}. \tag{1.30}$$

In other words mutual information is the amount of information that $\mathbf{Y}$ carries about $\mathbf{X}$ and conversely. Combining mutual information (1.30) and entropy (1.26) gives some further interesting properties and chain rules of mutual information. First the mutual information is lower bounded by zero, i.e., $I(\mathbf{Y};\mathbf{X}) \geq 0$, which is intuitive since $\mathbf{Y}$ could not bear less information on $\mathbf{X}$ than no information. Further it is possible to describe mutual information in terms of differential entropy in several ways. This is done by expressing (1.30) in terms of entropies and gives

$$I(\mathbf{Y};\mathbf{X}) = I(\mathbf{X};\mathbf{Y}) = h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) = h(\mathbf{X}) - h(\mathbf{X}|\mathbf{Y}), \tag{1.31}$$

$$I(\mathbf{Y};\mathbf{X}) = h(\mathbf{X}) + h(\mathbf{Y}) - h(\mathbf{X},\mathbf{Y}). \tag{1.32}$$

The term in (1.32) includes an upper limit on mutual information (i.e., $I(\mathbf{Y};\mathbf{X}) \leq h(\mathbf{X})+h(\mathbf{Y})$ which is obtained for statistically independent random variables.

1.4.3 System Capacity

System capacity denotes the capacity of a given system from input to output. Figure 1.9 shows the equivalent channel which will be used for system capacity calculation. The system input is $\mathbf{X}$, the system output $\mathbf{Y}$, and the additive noise is denoted by $\mathbf{Z}$. The continuous-time channel $\mathbb{H}_c$ isn‘t known whether to the receiver nor to the transmitter. It follows, that the equivalent channel can be represented by the conditional entropy $p_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})$. That is the pdf of the system output with given (known) system input. Then, the system capacity is defined

as

$$C_{\text{sys}} \triangleq \max_{p_{\mathbf{X}}(\mathbf{x})} I(\mathbf{Y};\mathbf{X}) = \max_{p_{\mathbf{X}}(\mathbf{x})} \left\{ \int_{\mathcal{S}_{\mathbf{x},\mathbf{y}}} \int p_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y}) \log \frac{p_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y})}{p_{\mathbf{X}}(\mathbf{x}) p_{\mathbf{Y}}(\mathbf{y})} d\mathbf{x} d\mathbf{y} \right\}$$

$$= \max_{p_{\mathbf{X}}(\mathbf{x})} \left\{ \int_{\mathcal{S}_{\mathbf{x},\mathbf{y}}} \int p_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) p_{\mathbf{X}}(\mathbf{x}) \log \frac{p_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})}{\int_{\mathcal{S}_{\mathbf{x}}} p_{\mathbf{Y}|\mathbf{X}'}(\mathbf{y}|\mathbf{x}') p_{\mathbf{X}'}(\mathbf{x}') d\mathbf{x}'} d\mathbf{x} d\mathbf{y} \right\}$$

which is the maximum of $I(\mathbf{Y};\mathbf{X})$ over all possible input pdf's $p_{\mathbf{X}}(\mathbf{x})$. An operational definition of system capacity is the highest rate at which information can be transmitted reliable within this system. For this the notion of *ergodic information rate* will be used. For a relative long transmission time the calculated mutual information will be achieved. This can be done through a sufficiently long code (compared to the fading speed of the channel [29]). Then

$$R \triangleq \lim_{T\to\infty} \frac{1}{T} I(\mathbf{Y};\mathbf{X}), \tag{1.33}$$

with total transmission time T. Hence in our case the system capacity will be the maximum of ergodic information rate over all possible inputs (all possible codebooks of $\mathbf{X}$) which is denoted by

$$S \triangleq \max\ R.$$

In contrast to the *operational channel capacity* we will focus on the *'information' channel capacity*. Hence, other information theoretic quantities like error probability and outage are not treated.

Chapter 2

Overview of Known Results

In this chapter we want to give an overview of previous results. The main topics will be multiuser OFDM systems, noncoherent channels (channels, where no channel state information (CSI) is available), and time and frequency selective fading channels. First an overview of needed properties should be pointed out. We want to deal with information theoretic quantities (such as multiuser sum system capacity) of a multiuser multicarrier system (i.e. OFDM). Although this sum system capacity isn‘t the only information quantity, we will focus our attention on it.

- **Channel/Signal Models**. In order to describe the multicarrier system, an appropriate channel and signal model has to be found. Several aspects have to be treated:
  - Single user signal model/multiuser signal model: The question of how many users there are and their mutual interference has to be declared.
  - Frequency selectivity: In general a channel is assumed frequency selective, which means different attenuation for different frequencies. Under certain circumstances a frequency-nonselective channel will be used. Frequency-nonselective channels are also known as flat fading channels.
  - Time selectivity: The channel changes in time which can be expressed by the coherence time. Depending on the duration of the transmitted symbols the channel is named a slowly fading or rapidly fading channel. Also the term block-fading is used for slowly fading channels or equivalent large coherence times.
  - Number of antennas: The more antennas there are, the more performance is within reach. In this connection *multiple-input multiple-output* (MIMO) [9, 30, 31] systems are of interest.
  - Signaling constraints: In addition to statistics of the input signal some constraints, such as average power constraint, peak-power constraint, bandwidth constraint or delay constraints should be taken into account (e.g. [32]).

- **Channel State Information (CSI)**. Here a differentiation can be done as follows:
  - CSI/no CSI: If CSI is available, no further channel statistics have to be obtained for calculating the channel output. However in many cases there is no CSI available and in this context the term *noncoherent channel* is known. In that case procedures such as channel estimation can be implemented. The goal is to achieve the most accurate CSI, which will result in a performance gain.
  - CSI at transmitter and/or receiver: For the acquisition of CSI at the receiver various approaches are possible. An overview of them is given in [12]. Making CSI available to the transmitter allows the use of adapted modulation.

Further topics to be discussed are channel estimation, synchronization or the amount of collision in frequency. Channel estimation is used to obtain CSI and thereby it allows the use of advanced signal detection techniques. So it increases performance (e.g. for MIMO systems [33]). Classical OFDM systems require perfect synchronization. CP-OFDM systems provide better synchronization but also require additional system capacity. Although neglected in many system capacity calculations it has to be mentioned. Finally different time-frequency signaling schemes (such as *frequency shift keying*(FSK) or CDMA) have to be distincted.

2.1 Single User Results

First an overview of results for single user systems will be given. An extensive survey of fading channels and their information theoretic and communication aspects is given in [32]. For the single user case, flat fading and no CSI are assumed. With CSI unavailable to transmitter and receiver it was shown that the capacity-achieving distribution of the i.i.d. input vector $\mathbf{X}$ is given by $\mathbf{X} = \mu\Phi$. Here μ is a nonnegative scalar random variable with variance equal to the average power and Φ is an isotropically distributed unit vector. This is the result of [34].

Schafhuber in his PhD Thesis [12] has studied the information rate and system capacity of OFDM. An information-theoretic analysis with the aim of calculating the system capacity of OFDM transmission over a time and frequency selective Rayleigh fading channel is given. All these considerations are done for the wideband regime. The input is drawn from specific codebooks, such as an orthogonal codebook (i.e. concentration of transmit energy in time and in frequency; maximum peak-to-average power rate) and a constant modulus symbol alphabet (e.g. BPSK). For an orthogonal codebook the information rate is upper bounded by

$$R \leq \frac{P\sigma_{\mathbb{H}}^2}{N_0},$$

with P the transmit power, $\sigma_{\mathbb{H}}^2 \approx \sigma_{\mathbb{H}_c}^2$ the path loss, and N_0 the power spectral density of the noise. The infinite-bandwidth system capacity was calculated as

$$S = \frac{P\sigma_{\mathbb{H}}^2}{N_0}. \tag{2.1}$$

This result shows the achievability of capacity for a peaky codebook in the infinite bandwidth case and arbitrarily low error probability. Note that there is no CSI available at the receiver. The system capacity in (2.1) (with equality) is equal to the channel capacity of the time and frequency selective WSSUS channel at infinite bandwidth [24, 35, 36].
For the calculation of the information rate for constant modulus signaling the system model

$$\mathbf{Y} = \mathrm{diag}\{\mathbf{H}\}\mathbf{X} + \mathbf{Z}$$

was used. This is related to (1.25) (i.e. $Y_{n,k} = H_{n,k}X_{n,k} + Z_{n,k}$, with time index $n = 0, \ldots, M-1$, and subcarrier index $k = 0, \ldots, K-1$) and will be explained in detail in Section 3.1. The $\mathrm{diag}\{\cdot\}$ operator positions the elements of $\mathbf{H}$ on the diagonal of a matrix. The channel is assumed as underspread. Since practical wireless channels are underspread, the channel can be described by a multiplicative model.
For BPSK the entropy $h(\mathbf{Y})$ was found by calculation of the distribution of $\mathbf{Y}$ first and then calculating the entropy. A constant-modulus alphabet with $|X_{n,k}| = \sigma_X^2$, a complex Gaussian channel $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, \sigma_{\mathbb{H}}^2\mathbf{I})$, and complex Gaussian noise $\mathbf{Z} \sim \mathcal{CN}(\mathbf{0}, N_0\mathbf{I})$ were used to find the distribution of output $\mathbf{Y} \sim \mathcal{CN}(\mathbf{0}, \left(\sigma_X^2\sigma_{\mathbb{H}}^2 + N_0\right)\mathbf{I})$ which maximizes the entropy $h(\mathbf{Y})$. Then, the differential entropy is

$$h(\mathbf{Y}) \leq MK\log\left(1 + \frac{\sigma_X^2\sigma_{\mathbb{H}}^2}{N_0}\right) + MK\log(\pi e N_0).$$

Note that equality is obtained if and only if there are totally uncorrelated output symbols. The mutual information of $\mathbf{Y}$ and $\mathbf{H}$, given $\mathbf{X}$ is

$$I(\mathbf{Y};\mathbf{H}|\mathbf{X}) = \sum_{i=0}^{MK-1} \log\left(1 + \frac{\sigma_X^2}{N_0}\lambda_i\{\mathcal{R}_{\mathbb{H}}\}\right),$$

with $\mathcal{R}_{\mathbb{H}}$ a $MK \times MK$ Toeplitz correlation matrix. The first row of $\mathcal{R}_{\mathbb{H}}$ is denoted by $[\mathbf{R}_{\mathbb{H}}[0]\ \mathbf{R}_{\mathbb{H}}[1]\ \ldots\ \mathbf{R}_{\mathbb{H}}[M-1]]$, with $K \times K$ correlation matrix $\mathbf{R}_{\mathbb{H}}[m]$. The calculation of upper bound on information rate results in

$$R_{\mathrm{CM}} \leq \frac{K}{T}\log\left(1 + \frac{\sigma_{\mathbb{H}}^2\sigma_X^2}{N_0}\right) - \frac{1}{T}\sum_{l=0}^{L}\int_{-1/2}^{1/2}\log\left(1 + K\frac{\sigma_X^2}{N_0}C_{\mathbb{H}}(l,\xi)\right)d\xi,$$

where R_{CM} stands for the constant-modulus ergodic information rate. This result does not contain any CSI but the statistics of the channel. The first term in it denotes the result for the additive white Gaussian noise (AWGN) channel. The second term is a penalty term and in [12] it was shown to be a measure for the channel prediction error. For large bandwidths an approximation can be given by

$$R_{\mathrm{CM}} \lessapprox \frac{B}{TF}\log\left(1 + \frac{TF}{B}\frac{P\sigma_{\mathbb{H}}^2}{N_0}\right) - B\int_{\tau}\int_{\nu}\log\left(1 + \frac{P}{BN_0}C_{\mathbb{H}}(\tau,\nu)\right)d\tau d\nu. \tag{2.2}$$

Here the first term is the rate for the additive white Gaussian noise (AWGN) case. The second term is a penalty due to noncoherence, which results from the need of channel estimation. The

asymptotic limit for infinite bandwidth results in $R_{CM} \to 0$. Further upper and lower bounds on system capacity for an OFDM system are calculated. Then, the upper bound on C_{sys} is given through the AWGN channel capacity:

$$C_{\text{sys}} \leq \frac{B}{TF} \log\left(1 + \frac{TF}{B}\frac{P\sigma_{\mathbb{H}}^2}{N_0}\right),$$

and in the limit of infinite bandwidth

$$\lim_{B\to\infty} \frac{B}{TF} \log\left(1 + \frac{TF}{B}\frac{P\sigma_{\mathbb{H}}^2}{N_0}\right) = \frac{P\sigma_{\mathbb{H}}^2}{N_0}.$$

A reference to capacity bounding in [37] is given. Bounds on information rate for Gaussian signaling are stated but no closed-form expression.

Bölcskei and Shamai in [38] apply mutual information bounding techniques such as in [32] on different time-frequency signaling schemes such as FSK, OFDM and CDMA. The channel is a time-frequency selective fading channel. They investigate the mutual information in the wideband limit in the absence of CSI for a signal model $\mathbf{Y} = \mathbf{HX}+\mathbf{Z}$ (equivalent to that above) with $\mathbf{Y}$ the output vector, $\mathbf{H}$ the unknown channel diagonal matrix, $\mathbf{X}$ the input vector, and $\mathbf{Z}$ the circularly symmetric additive white Gaussian noise. With [1]

$$I(\mathbf{Y};\mathbf{X},\mathbf{H}) \leq \log\det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{R}_H \odot \mathbf{C}_X\right),$$
$$I(\mathbf{Y};\mathbf{H}|\mathbf{X}) = \mathrm{E}_X\left\{\log\det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{X}\mathbf{R}_H\mathbf{X}^H\right)\right\},$$

with $\mathbf{R}_H$ and $\mathbf{C}_X$ the correlations matrices of $\mathbf{H}$ and $\mathbf{X}$, respectively, the following bounds on mutual information are obtained:

$$I(\mathbf{Y};\mathbf{X}) \geq \mathrm{E}_H\left\{\log\det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{H}\mathbf{C}_X\mathbf{H}^H\right)\right\} - \mathrm{E}_X\left\{\log\det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{X}\mathbf{R}_H\mathbf{X}^H\right)\right\},$$
$$I(\mathbf{Y};\mathbf{X}) \leq \log\det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{R}_H \odot \mathbf{C}_X\right) - \mathrm{E}_X\left\{\log\det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{X}\mathbf{R}_H\mathbf{X}^H\right)\right\}. \quad (2.3)$$

Investigations on conditions for peakiness in order to achieve the perfect CSI capacity in the wideband regime for 2-PSK system lead to

$$\lim_{M\to\infty} I(\mathbf{Y};\mathbf{H}|\mathbf{X}) = \int_\tau\int_\nu \log\left(1 + \frac{1}{N_0}C_{\mathbb{H}_c}(\tau,\nu)\right) d\tau d\nu,$$

which is similar to the penalty term in R_{CM} in [12]. Hence for non-peaky time-frequency signaling schemes the system capacity will tend to zero. Further a fourthegy upper bound[2]of (2.3) for transmission over an underspread fading channel is presented as

$$I(\mathbf{Y};\mathbf{X}) \leq \frac{1}{N_0^2}\mathrm{E}_{\mathbf{X}}\left\{\sum_{i=0}^{MK-1}\lambda_i^2(\mathbf{X}\mathbf{R}_H\mathbf{X}^H)\right\}.$$

Results with fourthegy upper bound were also presented in [39, 40].

[1] $\odot$ denotes the Hadamard (pointwise) product.

[2] Fourthegy bounding uses the approximations $\log(1+x) \leq x$ and $\log(1+x) \geq x - x^2/2$.

2.2 Multiuser Results

In this section an overview of results for multiuser communication will be given. In the single user case the user has access to the entire time-frequency area. Multiuser communication has to define this allocation of time slots and subcarriers with respect to the demands of the system (e.g. high sum information rate, high amount of users). Investigations on different signaling schemes and channel statistics have to be performed.

Thomas and Cover in [28] give bounds on the capacity for the individual users. They present some results for the capacity region of Gaussian multiple user channels and the multiple access channel. The most trivial result for two users can be pointed out as

$$\begin{aligned} R_1 &< I(\mathbf{X}_1; \mathbf{Y}|\mathbf{X}_2), \\ R_2 &< I(\mathbf{X}_2; \mathbf{Y}|\mathbf{X}_1), \\ R_1 + R_2 &< I(\mathbf{X}_2, \mathbf{X}_1; \mathbf{Y}), \end{aligned}$$

with R_1, R_2 the rates of the two users, $\mathbf{X}_1, \mathbf{X}_2$ their inputs and $\mathbf{Y}$ the output of the system. The model is shown in Figure 2.1(a), where the system is described by a conditional entropy $p_{\mathbf{Y}|\mathbf{X}_1,\mathbf{X}_2}(\mathbf{y}|\mathbf{x}_1, \mathbf{x}_2)$. For the two user case and given input distribution this gives a region of achievable rate pairs depicted in Figure 2.1(b). The sum rate in Figure 2.1(b) defines the bevel, avoiding both users to transmit with highest rate. For a Gaussian multiple access system the sum of all users' information rate can be bounded as

$$\sum_{i=0}^{U-1} R^{(u)} \leq \frac{1}{2} \log \left(1 + \frac{\sum_{i=0}^{U-1} P^{(u)}}{N_0} \right),$$

with U the number of users, $R^{(u)}$ the information rate of user u, $P^{(u)}$ the power of user u, and N_0 the power spectral density of the noise.

Gallager in [41] analyzes a system with multiple users transmitting to a single receiver. A time varying multipath channel is assumed. The problem of multiuser detection is outlined, e.g. stripping. A differentiation between narrow band and wideband is done in the sense of exclusive and shared use of bandwidth by the user, respectively. There is CSI available, which is obtained by a so called coherent rake receiver (equalization at the receiver). The coherent result is

$$R_{NB}(\mathcal{U}) < \mathrm{E}\left\{ KF \ln \left[1 + \frac{\sum_{k \in \mathcal{U}} P_k |H^{(k)}(0,t)|^2}{BN_0} \right] \right\} = R_{WB}(\mathcal{U}),$$

where $R_{NB}(\mathcal{U})$ is the information rate for the narrowband case and $R_{WB}(\mathcal{U})$ that in the wideband regime (equivalent to CDMA). Note that $\mathcal{U}$ denotes the subset of active users, K the number of subcarriers, F the bandwidth of one subcarrier, B the total bandwidth, P_k/W the power spectral density with P_k the power of source k, $H^{(k)}(0,t)$ the Fourier-transformation of

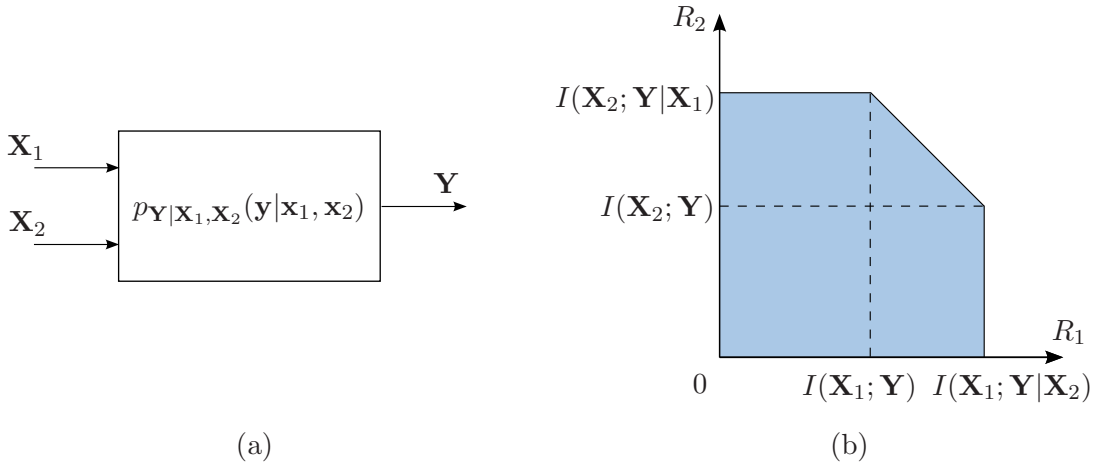

Figure 2.1: *(a) System model for two users (multiple access channel). (b) Capacity region for a multiple user system.*

the k-th user's channel impulse response at time t(it's assumed as ergodic and constant within F) and $N_0/2$ the white Gaussian noise spectral density. The narrowband case is similar to OFDMA and in other words there is no collision in frequency. The wideband case is the MC-CDMA case and there is full collision in frequency. For better comprehension the Figures 1.5 and 1.6(b) for the narrowband case and the wideband case, respectively may be useful. This result points out, that a CDMA accessing scheme is an outer bound to any other concerning the allocation of bandwidth to users. Explicit this was shown for CDMA, which is capable of higher rates than systems with slow frequency hopping. This is, because frequency hopping doesn't enhance the average mutual information. A further conclusion is, that K users (one per subcarrier), distributed in space, can obtain a higher rate than one user with equal sum power.

Multiuser aspects have also been considered in [42]. Shamai and Marzetta there used a model with U independent users transmitting over a Rayleigh block-fading channel with a coherence interval of T_{coh} symbols. There is no CSI neither at the transmitter nor at the receiver. The receiver is equipped with N receive antennas. First an upper bound for coherence interval $T_{\text{coh}} = 1$ is presented. Then, the capacity for $U > T_{\text{coh}}$ is equal to the capacity for $U = T_{\text{coh}}$ and the mutual information is maximized by choosing the input $\mathbf{X} = \mathbf{\Phi V}$ which is the result in [32] presented above. Asymptotic results for a large number of users and a long coherence interval give a bound on mutual information. The used signal model is $\mathbf{Y} = \sqrt{1/U}\mathbf{XH} + \mathbf{Z}$. Note that $\mathbf{Y}$ is the output, $\mathbf{X}$ the input, $\mathbf{X}_G$ the input with Gaussian signaling, and $\mathbf{H}$ is the channel. Then

$$E\left\{\log\det\left(\mathbf{I}+\frac{\sigma_X^2}{UN_0}\mathbf{H}^H\mathbf{H}\right)\right\}-\frac{UN}{T_{\text{coh}}}\log\left(1+\frac{\sigma_X^2\sigma_H^2T_{\text{coh}}}{UN_0}\right)\leq\frac{I(\mathbf{Y};\mathbf{X}_G)}{T_{\text{coh}}}\leq UN\log\left(1+\frac{\sigma_X^2\sigma_H^2}{N_0}\right)$$

is the result for the bounding of information rate. This bounding was done by use of the chain rules of mutual information, Jensen's inequality and the upper bound through Gaussian

assumption. With both Tcoh and U getting big with constant ratio $\beta = T_{\text{coh}}/U$ this gives the asymptotic result

$$N\log\left(1+\frac{\sigma_X^2\sigma_H^2}{N_0}\right)-\frac{N}{\beta}\log\left(1+\frac{\sigma_X^2\sigma_H^2}{N_0}\beta\right)\leq\frac{I(\mathbf{Y};\mathbf{X}_G)}{T_{\text{coh}}}\leq UN\log\left(1+\frac{\sigma_X^2\sigma_H^2}{N_0}\right).$$

Analysis for large SNR $\frac{\sigma_X^2\sigma_H^2}{N_0}$ which is appropriate for multiple users operating in an extremely high data regime leads to

$$\lim_{\text{SNR}\to\infty}\frac{I_G(\mathbf{X};\mathbf{S})}{\log\left(\frac{\sigma_X^2\sigma_H^2}{N_0}\right)}\geq T_{\text{coh}}N\left(1-\frac{N}{T_{\text{coh}}}\right).$$

Results for MIMO-systems were presented in [43, 44]. In [43] a MIMO multiple access scheme is presented, which allows to vary the amount of user collision in frequency. The channel is modeled as a frequency-selective multiple access fading channel. Perfect channel knowledge is available to the receiver but not to the transmitter. They provided results for the ergodic capacity region for joint and single user decoding. It is shown that the ergodic capacity region of a fully collision-based scheme (such as CDMA) is an outer bound on that of any other multiple access strategy. Hence for spatially well separated users, CDMA is suggested. Otherwise minimizing the amount of collision reduces the receiver complexity. For poor spatial separation orthogonal accessing schemes are beneficial as sum capacity is concerned.
In [44] also each user employs OFDM and the multiple antenna receiver has CSI of all channels. The channel is again a frequency-selective multiple access fading channel. The main target of this article is the investigation of the impact of collision in frequency on the sum capacity. The analysis is done in terms of multi-user multiplexing gain for CDMA and general tone assignments. In order to reach a high multi-user multiplexing gain, assuming good spatial separation and a large number of receive antennas, collision in frequency is stated decisive. For poor spatial separation or a small number of receive antennas no collision is needed to obtain the multiplexing gain of CMDA. An increasing number of receive antennas will require collision to maximize the multi-user multiplexing gain.

Chapter 3

System Capacity of MC-CDMA

Several works calculating the information rate and system capacity of OFDM systems were presented in Chapter 2. In this chapter, an extension of these results for MC-CDMA will be presented. Hence we will have to calculate the information rate and system capacity of a combination of DS-CDMA and OFDM (cf. Section 1.2). This will be done by using a MC-CDMA approach for multiple users. For transmission we assume a time and frequency selective Rayleigh fading channel as introduced in Section 1.3. An analysis for the wideband regime will be given.
The outline of this chapter is as follows. First a single user OFDM model equivalent to that in [12] will be introduced. Next, the MC-CDMA approach is described and applied to the single user OFDM model. The goal is to obtain the mutual information of input and output and the system capacity of MC-CDMA. Since this will be seen to be too difficult, bounds for these quantities will be derived. At last, numerical calculations and simulations are presented to interpret and confirm the theoretical results. The procedure as a whole is similar to that in [12].

3.1 Definitions and Notation

In this section we will introduce a single user OFDM model based on the approximations in Section 1.3.3. Further information theoretic notions for this case will be introduced. It will be important to separate random sequences/vectors and their realizations. Random variables will be denoted by capital letters and their realizations by lower-case letters. As vectors are denoted by boldface letters a random vector will be denoted by a boldface capital letter and its realization by a boldface lower-case letter.
We now introduce the single user OFDM model for M OFDM symbols and K subcarriers. The total transmission time then is MT (with symbol duration T) and the total bandwidth is KF (with frequency separation F). Hence we get a $MK \times 1$ channel input vector $\mathbf{X}$ and a

$MK \times 1$ channel output vector $\mathbf{Y}$ defined by

$$\mathbf{X} \triangleq [\mathbf{X}_0^T\ \mathbf{X}_1^T\ \ldots\ \mathbf{X}_{M-1}^T]^T, \qquad \text{with} \quad \mathbf{X}_n \triangleq [X_{n,0}\ X_{n,1}\ \ldots\ X_{n,K-1}]^T,$$
$$\mathbf{Y} \triangleq [\mathbf{Y}_0^T\ \mathbf{Y}_1^T\ \ldots\ \mathbf{Y}_{M-1}^T]^T, \qquad \text{with} \quad \mathbf{Y}_n \triangleq [Y_{n,0}\ Y_{n,1}\ \ldots\ Y_{n,K-1}]^T.$$

Here $X_{n,k}$ and $Y_{n,k}$ are the channel input and output, respectively for the n-th symbol and k-th subcarrier. Neglecting intersymbol and intercarrier interference the input-output relation is given by the approximation in (1.25). With the $MK \times 1$ channel vector

$$\mathbf{H} \triangleq [\mathbf{H}_0^T\ \mathbf{H}_1^T\ \ldots\ \mathbf{H}_{M-1}^T]^T, \quad \text{with} \quad \mathbf{H}_n \triangleq [H_{n,0}\ H_{n,1}\ \ldots\ H_{n,K-1}]^T,$$

the input-output relation for the total OFDM system can be expressed by

$$\mathbf{Y} = \text{diag}\{\mathbf{H}\}\mathbf{X} + \mathbf{Z} = \text{diag}\{\mathbf{X}\}\mathbf{H} + \mathbf{Z}. \tag{3.1}$$

By the diag$\{\cdot\}$ operation a diagonal matrix with the elements of the argument is built. $\mathbf{Z}$ is assumed as circularly symmetric complex Gaussian noise with covariance matrix $\mathbf{R}_Z = N_0\mathbf{I}$, i.e., $\mathbf{Z} \sim \mathcal{CN}(\mathbf{0}, N_0\mathbf{I})$. So the variance of noise σ_Z^2 will be denoted by N_0 all through the thesis. For convenience, we use the shortened notation

$$\mathbf{S} = \text{diag}\{\mathbf{H}\}\mathbf{X} = \text{diag}\{\mathbf{X}\}\mathbf{H}. \tag{3.2}$$

3.2 MC-CDMA System Capacity for Infinite Bandwidth

Using the OFDM input-output relation in (3.1) we now derive the system capacity (see Section 1.4.3) for a MC-CDMA system. The channel is a time and frequency selective WSSUS Rayleigh fading channel and neither transmitter nor receiver have CSI.

3.2.1 MC-CDMA Approximation

In order to get a multiuser system the so far described system model has to be enlarged by the MC-CDMA assumption. While in [12] the input symbols where assumed as constant-modulus we now need to adapt this to the multiuser case. The imitation of a multiuser system by a single user system is shown in Figures 3.1(a) and 3.1(b) (see also Figure 1.9 for equivalent channel). The users signals are spread over the entire bandwidth. Then, for multiple users a summation over users at the transmitter in the sense of superposition is necessary, i.e.,

$$X_{n,k} = \sum_{u=0}^{U-1} X_n^{(u)} c_k^{(u)},$$

with $X_{n,k}$ the superpositioned OFDM input symbol, $X_n^{(u)}$ the input data of user u, and $c_k^{(u)}$ the spreading sequence of user u. This superposition of all spreaded input symbols in frequency leads to a distribution of the overall input symbols $X_{n,k}$. By use of the *central limit theorem*

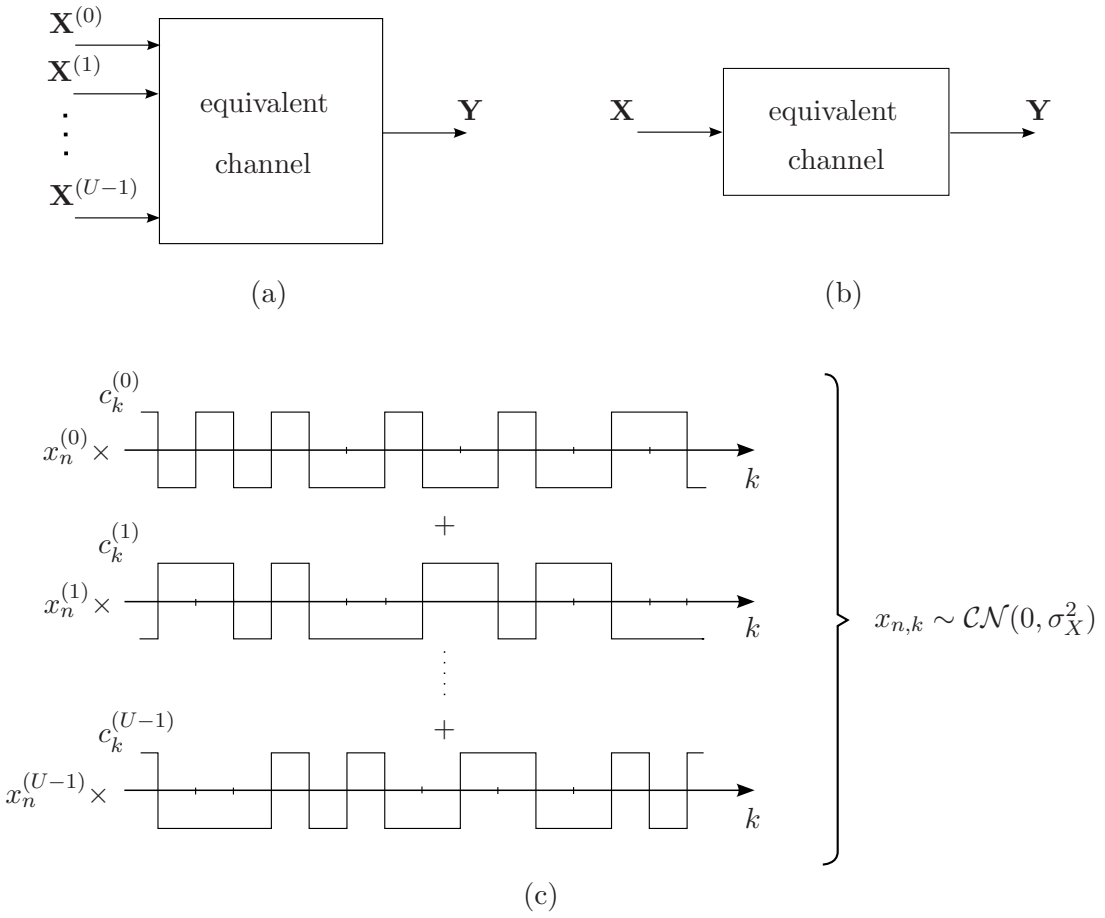

Figure 3.1: *(a) Multiuser system model with U single users applying varying codebooks transmitting over the equivalent channel. (b) Multiuser system with MC-CDMA approximation where a single user applies a circularly symmetric complex Gaussian codebook. (c) Spreading of input symbols over frequency and superposition of U users gives elements of $\mathbf{X}$.*

[45, 46] this superposition of single users applying any codebook can be described through circularly symmetric complex Gaussian input symbols $X_{n,k}$. Figure 3.1 (c) shows this. This approximation requires uncorrelated users since the individual random variables in the central limit theorem have to be independent. Further, the central limit theorem assumes a relatively large number of random variables or in our case users, respectively.

3.2.2 Information Rate Calculation for Gaussian Signaling

As pointed out in Subsection 3.2.1 the interference free OFDM system used by [12] can be used with Gaussian input symbols $X_{n,k}$ instead of a BPSK codebook. For the derivation of the mutual information

$$I(\mathbf{Y};\mathbf{X}) = I(\mathbf{Y};\mathbf{X},\mathbf{H}) - I(\mathbf{Y};\mathbf{H}|\mathbf{X}) \tag{3.3}$$

is used. With $I(\mathbf{Y};\mathbf{X},\mathbf{H}) = h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X},\mathbf{H})$ and $I(\mathbf{Y};\mathbf{X}|\mathbf{H}) = h(\mathbf{Y}|\mathbf{X}) - h(\mathbf{Y}|\mathbf{X},\mathbf{H})$ the mutual information in terms of entropy is

$$I(\mathbf{Y};\mathbf{X}) = h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}). \tag{3.4}$$

Hence the task is to find closed form expressions of $h(\mathbf{Y})$ and $h(\mathbf{Y}|\mathbf{X})$. This split of mutual information is already known from [12].

Calculation of $h(\mathbf{Y}|\mathbf{X},\mathbf{H})$

Although not needed explicitly for the calculation of $I(\mathbf{Y};\mathbf{X})$ a result for $h(\mathbf{Y}|\mathbf{X},\mathbf{H})$ is presented now. It will be used later for mutual information bounding. This calculation is equivalent to that in [12],

$$\begin{aligned} h(\mathbf{Y}|\mathbf{X},\mathbf{H}) &= \mathrm{E}_X\{\mathrm{E}_H\{h(\mathbf{Y}|\mathbf{x},\mathbf{h})\}\} = \mathrm{E}_X\{\mathrm{E}_H\{h(\mathrm{diag}\{\mathbf{h}\}\mathbf{x}+\mathbf{Z}|\mathbf{x},\mathbf{h})\}\} \\ &= \mathrm{E}_X\{\mathrm{E}_H\{h(\mathbf{Z}|,\mathbf{x},\mathbf{h})\}\} = \mathrm{E}_X\{\mathrm{E}_H\{h(\mathbf{Z})\}\} = MK\,\log(\pi e N_0). \end{aligned} \tag{3.5}$$

An interpretation therefore can be given as follows. The pdf of $\mathbf{Y}$ conditioned on $\mathbf{X}$ and $\mathbf{H}$ will only leave the uncertainty of noise, which easily can be calculated.

Calculation of $h(\mathbf{Y}|\mathbf{X})$

Again we fellow the development in [12]. Then, $h(\mathbf{Y}|\mathbf{X})$ can be developed as

$$h(\mathbf{Y}|\mathbf{X}) = \mathrm{E}_X\{\mathrm{E}_H\{h(\mathbf{Y}|\mathbf{x})\}\} = \mathrm{E}_X\left\{\log\left((\pi e)^{MK}\det\left[\mathrm{diag}\{\mathbf{x}\}\,\mathcal{R}_{\mathbb{H}}\,\mathrm{diag}\{\mathbf{x}^*\} + N_0\mathbf{I}\,\right]\right)\right\} \tag{3.6}$$

$$= \mathrm{E}_X\left\{\log\left((\pi e N_0)^{MK}\det\left[\mathbf{I} + \frac{1}{N_0}\mathrm{diag}\{\mathbf{x}\}\,\mathcal{R}_{\mathbb{H}}\,\mathrm{diag}\{\mathbf{x}^*\}\right]\right)\right\} \tag{3.7}$$

$$= MK\,\log(\pi e N_0) + \mathrm{E}_X\left\{\log\left(\det\left[\mathbf{I} + \frac{1}{N_0}\mathrm{diag}\{\mathbf{x}\}\,\mathcal{R}_{\mathbb{H}}\,\mathrm{diag}\{\mathbf{x}^*\}\right]\right)\right\} \tag{3.8}$$

with the $MK \times MK$ block-Toeplitz correlation matrix $\mathcal{R}_{\mathbb{H}} = \mathrm{E}\left\{\mathbf{H}\mathbf{H}^H\right\}$.
Then, with the results in (3.5) and (3.6) The mutual information $I(\mathbf{Y};\mathbf{H}|\mathbf{X})$ can be composed as

$$I(\mathbf{Y};\mathbf{H}|\mathbf{X}) = \mathrm{E}_X\left\{\log\left(\det\left[\mathbf{I} + \frac{1}{N_0}\,\mathrm{diag}\{\mathbf{x}\}\,\mathcal{R}_{\mathbb{H}}\,\mathrm{diag}\{\mathbf{x}^*\}\right]\right)\right\}. \tag{3.9}$$

It remains to calculate the expectation with respect to the vector $\mathbf{x}$. In this context a bound will be presented in Subsection 3.2.3 since an equality can only be obtained for a constant-modulus codebook.

Calculation of $h(\mathbf{Y})$

The first term in (3.4) is the entropy of the output vector $\mathbf{Y}$. The input-output relation in (3.1) can be written as

$$\mathbf{Y} = \mathbf{S} + \mathbf{Z}, \qquad \text{with} \quad \mathbf{S} = \mathrm{diag}\{\mathbf{H}\}\,\mathbf{X} \qquad \text{and} \quad \mathbf{Z} \sim \mathcal{CN}(\mathbf{0}, N_0\mathbf{I}). \tag{3.10}$$

For the calculation of an entropy, the distribution of $\mathbf{Y}$ has to be calculated first. Since a calculation based on the vector-product seems too complicated, we bound the entropy of the whole vector. In particular, we use the upper bound for the entropy of $\mathbf{Y}$,

$$h(\mathbf{Y}) \leq MKh(Y_{n,k}). \tag{3.11}$$

This bound corresponds to independent $Y_{n,k}$, i.e., the $Y_{n,k}$ are identically independent distributed (i.i.d). It remains to compute $h(Y_{n,k})$. We start by calculating the distribution of $Y_{n,k}$. So the first part of the basic input-output relation is

$$S_{n,k} = H_{n,k}X_{n,k} = |H_{n,k}||X_{n,k}|e^{j(\arg\{H_{n,k}\}+\arg\{X_{n,k}\})}.$$

$|H_{n,k}|$ is Rayleigh-distributed and $\arg(|H_{n,k}|)$ is uniformly distributed. The same way $|X_{n,k}|$ and $\arg(|X_{n,k}|)$ are Rayleigh-distributed and uniformly distributed, respectively. The variances are $\sigma^2_{\mathbb{H}}$ for the channel and σ^2_X for the input symbols.
The pdf of the phase of $S_{n,k}$, $\arg\{S_{n,k}\} = \arg\{H_{n,k}\} + \arg\{X_{n,k}\}$, results from a convolution of the pdf of $\arg(H_{n,k})$ and $\arg(X_{n,k})$ and is again uniform.
For the calculation of pdf of $|S_{n,k}|$ a more specialized method is needed since the product of two Rayleigh-distributed pdf's has to be obtained. By using the results in [47] the distribution $p_{|S_{n,k}|}(|s|)$ of $|S_{n,k}|$ is obtained as

$$p_{|S_{n,k}|}(|s|) = \frac{4|s|}{\sigma^2_X\sigma^2_{\mathbb{H}}}\, K_0\left(\frac{2|s|}{\sigma_X\sigma_{\mathbb{H}}}\right). \tag{3.12}$$

Here Theorem 4 on page 48 in [47] was used for two real random Gaussian vectors X_n and Y_m with mean vectors A_n and B_n and correlation coefficients ψ_0 and ψ'_0, respectively. The probability density function of the product of $r = |X_n|$ and $s = |Y_m|$, $z = rs$ is given by

$$p(z) = \frac{4}{z}\left(\frac{z^2}{4\psi_0\psi'_0}\right)^{\frac{1}{4}(n+m)} e^{-\frac{1}{2}\left(\frac{a^2}{\psi_0}+\frac{b^2}{\psi'_0}\right)} \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \frac{1}{k!j!\Gamma\left(\frac{1}{2}n+k\right)\Gamma\left(\frac{1}{2}m+j\right)}$$
$$\cdot \left(\frac{a\sqrt{z}}{2\psi_0}\right)^{2k}\left(\frac{b\sqrt{z}}{2\psi'_0}\right)^{2j}\left(\frac{\psi_0}{\psi'_0}\right)^{\frac{1}{2}(k-j)} K_{\frac{1}{2}(n-m)+k-j}\left(\frac{z}{\sqrt{\psi_0\psi'_0}}\right),$$

where $a = |A_n|$ and $b = |B_n|$, K the Bessel function of the second kind and order zero.
For the overall pdf of $S_{n,k}$ this gives

$$p_{S_{n,k}}(s) = \frac{2}{\pi\sigma^2_X\sigma^2_{\mathbb{H}}}\, K_0\left(\frac{2|s|}{\sigma_X\sigma_{\mathbb{H}}}\right), \tag{3.13}$$

where the combination of the pdfs of the magnitude and the phase of $S_{n,k}$ is done by (see [46])

$$p_{S_{n,k}}(s) = \frac{1}{|s|}\, p_{|S_{n,k}|}(|s|)\, p_{\arg\{S_{n,k}\}}(\arg\{s\}).$$

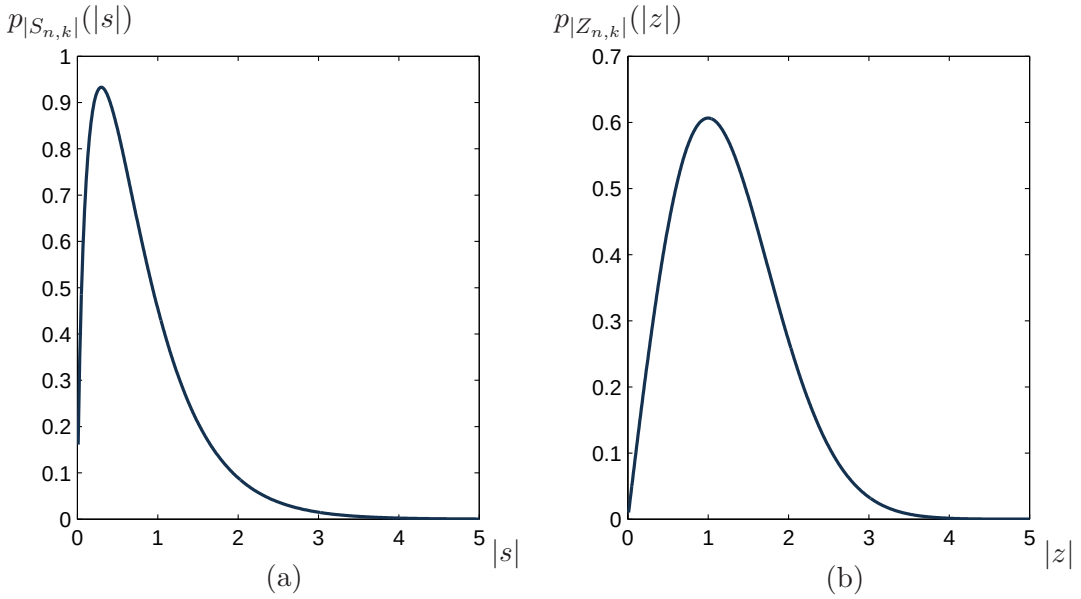

Figure 3.2: *(a) The pdf of $|S_{n,k}|$ for $\sigma_X^2 = 1$ and $\sigma_{\mathbb{H}}^2 = 1$. (b) The pdf of $|Z_{n,k}|$ is depicted for $N_0 = 1\,W/Hz$.*

According to (3.10) at next the noise has to be added which means a convolution of pdfs (in a complex manner), e.g.

$$\begin{aligned} p_{Y_{n,k}}(y_R, y_I) &= (p_{S_{n,k}} * p_{Z_{n,k}})(y_R, y_I) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_{S_{n,k}}(y'_R, y'_I)\; p_{Z_{n,k}}(y_R - y'_R, y_I - y'_I)\; dy'_R dy'_I, \end{aligned} \tag{3.14}$$

with $p_{Z_{n,k}}$ the distribution of the circularly symmetric Gaussian noise $Z_{n,k}$, given through

$$p_{Z_{n,k}}(z_r, z_I) = \frac{1}{2\pi N_0}\, e^{-\frac{z_R^2 + z_I^2}{2N_0}}, \tag{3.15}$$

or equivalently

$$p_{Z_{n,k}}(z) = \frac{1}{2\pi N_0}\, e^{-\frac{|z|^2}{2N_0}}. \tag{3.16}$$

These pdfs for the magnitude of $S_{n,k}$ and $Z_{n,k}$ are shown in Figure 3.2. In (3.14) also $p_{S_{n,k}}(s)$ is used as

$$p_{S_{n,k}}(s_R, s_I) = \frac{2}{\pi\sigma_X^2\sigma_{\mathbb{H}}^2}\, K_0\left(\frac{2\sqrt{s_R^2 + s_I^2}}{\sigma_X\sigma_{\mathbb{H}}}\right). \tag{3.17}$$

Here the mathematical complexity gets too high to reach a nice (considerable) result. Therefore a numerical calculation and a simulation will be presented in Section 3.3. On a theoretical level, bounding techniques will be applied in Subsection 3.2.3.

3.2.3 Bounds on Mutual Information for Gaussian Signaling

As pointed out above for a more complex codebook the effort in calculating mutual information and information rate accurately increases. In this section bounding techniques will be applied. For the sake of clarity the equations used are shown first,

$$I(\mathbf{Y};\mathbf{X}) = I(\mathbf{Y};\mathbf{X},\mathbf{H}) - I(\mathbf{Y};\mathbf{H}|\mathbf{X}), \tag{3.18}$$

$$I(\mathbf{Y};\mathbf{X},\mathbf{H}) = h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X},\mathbf{H}). \tag{3.19}$$

Next, the goals of bounding will be described. For the upper bound on mutual information the more detailed split will be used. Then, an upper bound on the entropy of the output symbols $h(\mathbf{Y})$ and lower bounds on $h(\mathbf{Y}|\mathbf{X},\mathbf{H})$ and $I(\mathbf{Y};\mathbf{H}|\mathbf{X})$ are required. On the other side for the lower bound on mutual information an upper bound on $I(\mathbf{Y};\mathbf{H}|\mathbf{X})$ and a lower bound on $I(\mathbf{Y};\mathbf{H},\mathbf{X})$ are necessary since the split in (3.18) will be used.

Upper Bound

For the upper bound on $I(\mathbf{Y};\mathbf{X})$ we use the more detailed term $I(\mathbf{Y};\mathbf{X}) = h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X},\mathbf{H}) - I(\mathbf{Y};\mathbf{H}|\mathbf{X})$. We first note that $h(\mathbf{Y}|\mathbf{X},\mathbf{H})$ is given by (3.5). For the upper bound on the entropy of the output vector $\mathbf{Y}$ we refer to the calculation in Section 3.2.2 where the entropy is upper bounded by (3.11). With the general fact, that Gaussian random variables maximize entropy we derive an upper bound with a Gaussian distributed $Y'_{n,k}$. This is

$$h(Y_{n,k}) \leq h(Y'_{n,k}), \quad \text{with} \qquad Y'_{n,k} \sim \mathcal{CN}(0, \sigma^2_{Y'}), \tag{3.20}$$

and hence for $Y_{n,k}$ i.i.d. this can be widened to $\mathbf{Y}$ and $h(\mathbf{Y})$. Note that Gaussian output symbols $Y_{n,k}$ are the result of a BPSK symbol alphabet on $X_{n,k}$ and not of a Gaussian codebook we supposed. Hence the result for $h(\mathbf{Y})$ is the upper bound [12]

$$h(\mathbf{Y}) \leq \mathrm{MK}\,\log\left(1 + \frac{\sigma^2_X \sigma^2_{\mathbb{H}}}{N_0}\right) + \mathrm{MK}\,\log(\pi e N_0). \tag{3.21}$$

The remaining term is a lower bound on $I(\mathbf{Y};\mathbf{H}|\mathbf{X})$. We apply $\log\,(1+x) \geq x - \frac{x^2}{2}$ to (3.9). This concept of fourthegy bounding was used amongst others by [39]. First a modification on (3.9) gives

$$\begin{aligned} I(\mathbf{Y};\mathbf{H}|\mathbf{X}) &= \mathrm{E}_X\left\{\log\left(\det\left[\mathbf{I} + \frac{1}{N_0}\,\mathrm{diag}\{\mathbf{x}\}\,\mathcal{R}_{\mathbb{H}}\,\mathrm{diag}\{\mathbf{x}^*\}\right]\right)\right\} \\ &= \mathrm{E}_X\left\{\sum_{i=0}^{MK-1}\log\left(1 + \frac{1}{N_0}\,\lambda_i\left\{\mathrm{diag}\{\mathbf{x}\}\,\mathcal{R}_{\mathbb{H}}\,\mathrm{diag}\{\mathbf{x}^*\}\right\}\right)\right\}, \end{aligned}$$

where the eigenvalues are denoted by $\lambda_i\{\cdot\}$. We introduce the abbreviation $\mathcal{R}'_{\mathbb{H}} = \mathrm{diag}\{\mathbf{x}\}\,\mathcal{R}_{\mathbb{H}}\,\mathrm{diag}\{\mathbf{x}^*\}$

and apply $\log(1+x) \geq x - \frac{x^2}{2}$, which gives

$$I(\mathbf{Y};\mathbf{H}|\mathbf{X}) \geq \mathrm{E}_X \left\{ \sum_{i=0}^{MK-1} \left(\frac{1}{N_0}\, \lambda_i\{\mathcal{R}'_{\mathbb{H}}\} - \frac{1}{2N_0^2}\, \lambda_i^2\{\mathcal{R}'_{\mathbb{H}}\} \right) \right\}$$
$$= \frac{1}{N_0}\, \mathrm{E}_X \left\{ \sum_{i=0}^{MK-1} \lambda_i\{\mathcal{R}'_{\mathbb{H}}\} \right\} - \frac{1}{2N_0^2}\, \mathrm{E}_X \left\{ \sum_{i=0}^{MK-1} \lambda_i^2\{\mathcal{R}'_{\mathbb{H}}\} \right\}. \tag{3.22}$$

Noting that $\sum_i \lambda_i\{\mathcal{R}'_{\mathbb{H}}\}$ equals the trace $\mathrm{Tr}(\mathcal{R}'_{\mathbb{H}})$ and $\sum_i \lambda_i^2\{\mathcal{R}'_{\mathbb{H}}\}$ the Frobenius norm (Schur-norm) $||\mathcal{R}'_{\mathbb{H}}||_F^2 = \mathrm{Tr}(\mathcal{R}'_{\mathbb{H}}\mathcal{R}'^H_{\mathbb{H}})$ we can write

$$I(\mathbf{Y};\mathbf{H}|\mathbf{X}) \geq \frac{1}{N_0}\mathrm{Tr}\left\{\mathrm{E}_X\{\mathcal{R}'_{\mathbb{H}}\}\right\} - \frac{1}{2N_0^2}\mathrm{Tr}\left\{\mathrm{E}_X\left\{\mathcal{R}'_{\mathbb{H}}\mathcal{R}'^H_{\mathbb{H}}\right\}\right\}.$$

Resubstitution of $\mathcal{R}'_{\mathbb{H}}$ gives

$$I(\mathbf{Y};\mathbf{H}|\mathbf{X}) \geq \frac{1}{N_0}\, \mathrm{Tr}\left\{\mathrm{E}_X\{\mathrm{diag}\{\mathbf{x}^*\}\, \mathrm{diag}\{\mathbf{x}\}\}\, \mathcal{R}_{\mathbb{H}}\right\}$$
$$- \frac{1}{2N_0^2}\, \mathrm{Tr}\left\{\mathrm{E}_X\left\{\mathrm{diag}\{\mathbf{x}^*\}\, \mathrm{diag}\{\mathbf{x}\}\, \mathcal{R}_{\mathbb{H}}\, \mathcal{R}_{\mathbb{H}}^H\, \mathrm{diag}\{\mathbf{x}^*\}\, \mathrm{diag}\{\mathbf{x}\}\right\}\right\}$$
$$= \frac{\sigma_X^2}{N_0}\, \mathrm{Tr}\left\{\mathcal{R}_{\mathbb{H}}\right\} - \frac{3\sigma_X^4}{N_0^2}\, \mathrm{Tr}\left\{\mathcal{R}_{\mathbb{H}}\, \mathcal{R}_{\mathbb{H}}^H\right\}.$$

where we used the shift-property of $\mathrm{Tr}(\cdot)$[1] and

$$\mathrm{E}_X\left\{\mathrm{diag}\{\mathbf{x}^*\}\, \mathrm{diag}\{\mathbf{x}\}\right\} = \sigma_X^2\mathbf{I},$$
$$\mathrm{E}_X\{\mathrm{diag}\{\mathbf{x}^*\}\, \mathrm{diag}\{\mathbf{x}\}\, \mathrm{diag}\{\mathbf{x}^*\}\, \mathrm{diag}\{\mathbf{x}\}\} = \mathrm{E}_X\{|x|^4\}\, \mathbf{I} = 6\, \sigma_X^4\mathbf{I},$$

where we assumed an i.i.d. codebook. Finally, $\mathrm{Tr}(\mathcal{R}_{\mathbb{H}})$ equals the path loss $\sigma_{\mathbb{H}}^2$ and hence the lower bound on $I(\mathbf{Y};\mathbf{H}|\mathbf{X})$ can be written as

$$I(\mathbf{Y};\mathbf{H}|\mathbf{X}) \geq \frac{\sigma_X^2}{N_0}\, \sigma_{\mathbb{H}}^2 - \frac{3\sigma_X^4}{N_0^2}\, \|\mathcal{R}_{\mathbb{H}}\|_F^2. \tag{3.23}$$

Lower Bound

Using the first term in (3.18) we first search for a lower bound on $I(\mathbf{Y};\mathbf{X},\mathbf{H})$. In order to get a tighter bound a result of [38] is used,

$$I(\mathbf{Y};\mathbf{X},\mathbf{H}) = \mathrm{E}_H\left\{\log\det\left[\mathbf{I} + \mathbf{H}\mathbf{C}_X\mathbf{H}^H\right]\right\}. \tag{3.24}$$

The second term for the lower bound is the upper bound on $I(\mathbf{Y};\mathbf{H}|\mathbf{X})$. This derivation is based on that in [12].

[1] $\mathrm{Tr}(\mathbf{ABC}) = \mathrm{Tr}(\mathbf{CAB}) = \mathrm{Tr}(\mathbf{BCA})$

Starting with the equation in (3.9) by applying Jensen's inequality[2], gives

$$\begin{aligned} I(\mathbf{Y};\mathbf{H}|\mathbf{X}) &= \mathrm{E}_X\left\{\log\left(\det\left[\mathbf{I}+\frac{1}{N_0}\,\mathrm{diag}\{\mathbf{x}^*\}\,\mathrm{diag}\{\mathbf{x}\}\,\mathcal{R}_{\mathbb{H}}\right]\right)\right\} \\ &\leq \log\left(\det\left[\mathbf{I}+\frac{1}{N_0}\,\mathrm{E}_X\left\{\mathrm{diag}\{\mathbf{x}^*\}\,\mathrm{diag}\{\mathbf{x}\}\right\}\,\mathcal{R}_{\mathbb{H}}\right]\right) \\ &= \log\left(\det\left[\mathbf{I}+\frac{\sigma_X^2}{N_0}\,\mathcal{R}_{\mathbb{H}}\right]\right), \end{aligned}$$

where we used $\mathrm{E}_X\left\{\mathrm{diag}\{\mathbf{x}^*\}\,\mathrm{diag}\{\mathbf{x}\}\right\} = \sigma_X^2\,\mathbf{I}$. With $\lambda(\mathcal{R}_{\mathbb{H}})$ the eigenvalues of $\mathcal{R}_{\mathbb{H}}$ we get

$$I(\mathbf{Y};\mathbf{H}|\mathbf{X}) \leq \sum_{i=0}^{MK-1}\log\left(1+\frac{\sigma_X^2}{N_0}\lambda_i\{\mathcal{R}_{\mathbb{H}}\}\right). \tag{3.25}$$

Combination of Bounds

With the composition in (3.19) and the corresponding results (3.21), (3.5) and (3.23) the upper bound is

$$I(\mathbf{Y};\mathbf{X}) \leq \mathrm{MK}\,\log\left(1+\frac{\sigma_X^2\sigma_{\mathbb{H}}^2}{N_0}\right)-\frac{\sigma_X^2}{N_0}\,\sigma_{\mathbb{H}}^2+\frac{3\sigma_X^4}{N_0^2}\,\|\mathcal{R}_{\mathbb{H}}\|_F^2. \tag{3.26}$$

The the upper bound on information rate can be calculated. This calculation follows (1.33) with total transmission time MT and gives

$$R=\lim_{M\to\infty}\frac{1}{MT}\,I(\mathbf{Y};\mathbf{X}) \leq \frac{B}{TF}\log\left(1+\frac{TF}{B}\frac{P\sigma_{\mathbb{H}}^2}{N_0}\right)-\frac{P\sigma_{\mathbb{H}}^2}{N_0}+\frac{3P^2}{BN_0^2}\int_\tau\int_\nu C_{\mathbb{H}}^2(\tau,\nu)\,d\tau d\nu. \tag{3.27}$$

By Combination of the Results in (3.24) and (3.25) and with (3.18) the lower bound on mutual information is obtained as

$$I(\mathbf{Y};\mathbf{X}) \geq \mathrm{E}_H\{\log\det(\mathbf{I}+\mathbf{H}\mathbf{C}_{\mathbf{x}}\mathbf{H}^{\mathbf{H}})\}-\sum_{i=0}^{MK-1}\log\left(1+\frac{\sigma_X^2}{N_0}\,\lambda_i\{\mathcal{R}_{\mathbb{H}}\}\right).$$

Here the first term is equivalent to the additive white Gaussian noise case and the second term can be regarded as a noncoherence penalty term. This penalty is caused by the necessity of estimating the channel.

3.3 Simulation Results and Numerical Evaluation

In order to assess the difference of our results to previous ones and to obtain further insights regarding the distributions, entropy and others, simulations and numerical evaluations were done.

[2] Jensen's inequality says, that $f\left(\sum_{i=1}^n\lambda_i x_i\right)\geq\sum_{i=1}^n\lambda_k f(x_k)$ for a concave function f and $f\left(\sum_{i=1}^n\lambda_i x_i\right)\leq\sum_{i=1}^n\lambda_k f(x_k)$ for a convex function. For probability calculations the summation may be exchanged by the expectation operator $\mathrm{E}(\cdot)$, then $f(\mathrm{E}X)\leq\mathrm{E}f(X)$ for a concave function [28].

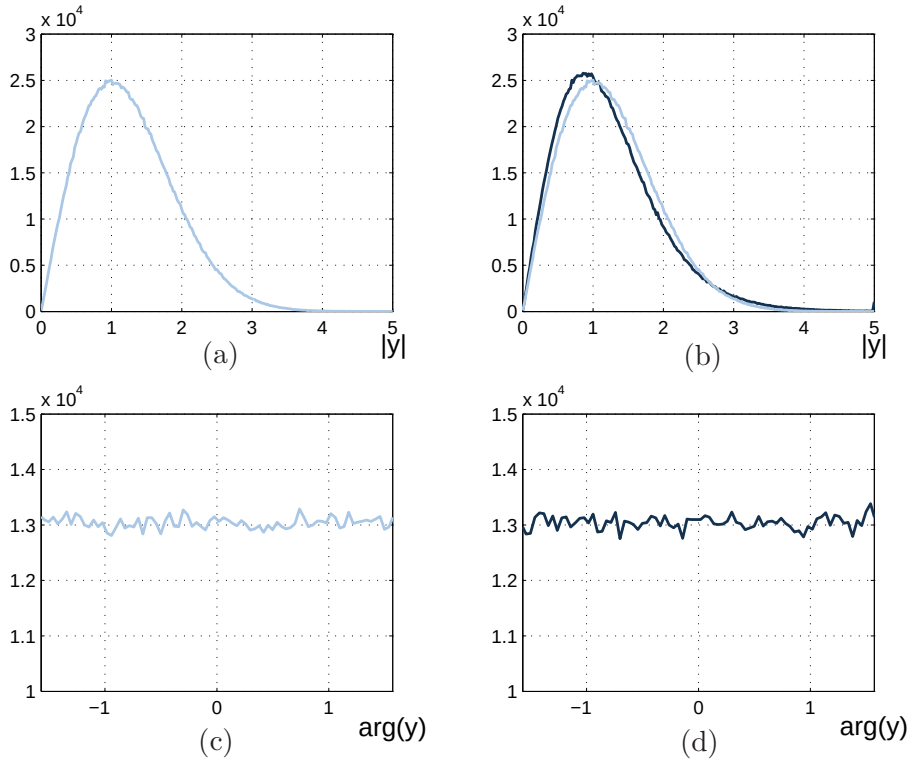

Figure 3.3: *Histograms of $|Y_{n,k}|$ for (a) BPSK input (b) Gaussian input; histograms of $arg(Y_{n,k})$ for (c) BPSK input and (d) Gaussian input. Parameters used were $\sigma_X^2 = 1$, $\sigma_{\mathbb{H}}^2 = 1$ and $N_0 = 1\,W/Hz$. The number of realizations was* 2048000.

3.3.1 Monte Carlo Simulation Results

The simulations were done for a constant-modulus input $\mathbf{X}$ and a Gaussian distributed input $\mathbf{X}$. The channel coefficients $H_{n,k}$ are assumed i.i.d. complex Gaussian with path loss $\sigma_{\mathbb{H}}^2$. This assumption is because the correlations are not relevant for the bound on $h(Y_{n,k}$.

Figure 3.3 shows the simulation of the pdf of the receive symbol $Y_{n,k}$ of an OFDM system. The simulation is based on the input-output relation in (3.1) with BPSK and Gaussian input symbols with $\sigma_X^2 = 1$. The path loss is $\sigma_{\mathbb{H}}^2 = 1$, and the power spectral density of noise-vector $\mathbf{Z}$ is $N_0 = 1$W/Hz. This unrealistic assumptions (extremely low SNR) were made to emphasize the differences of the various input symbol constellations. With independent symbols and channel coefficients the simulation simply can be done using (1.25). For Gaussian input (Figure 3.3(b)) the distribution of $Y_{n,k}$ gets narrower than that for BPSK input. For comparison the distribution of the magnitude of $Y_{n,k}$ for BPSK input constellation which is Rayleigh-distributed with variance $\sigma_Y = 2$ is shown in Figure 3.3(b). This can also (and better) be seen in Figure 3.4 for a different noise variance. Even though the distribution of the output for Gaussian input symbols seems to be Rayleigh it isn't (cf. (3.12)), but it is for BPSK. More precisely, BPSK input causes complex Gaussian distributed output $Y_{n,k}$, which

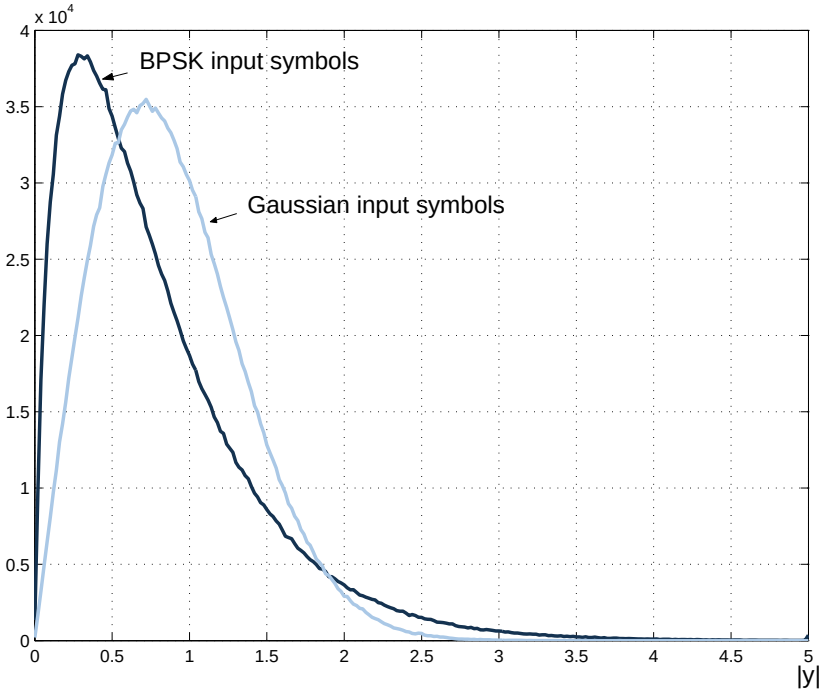

Figure 3.4: *Histograms of $|Y_{n,k}|$ for BPSK and $|Y_{n,k}|$ Gaussian input symbols with $\sigma_X^2 = 1$, $\sigma_{\mathbb{H}}^2 = 1$ and $N_0 = 10^{-3}\,W/Hz$. The number of realizations was* 2048000.

means maximization of entropy. The distribution of the phases for Gaussian and BPSK input can be regarded as identically distributed. From this diagrams, the fact that $Y_{n,k}$ has lower entropy for a Gaussian codebook than for a constant-modulus codebook can be confirmed. Then, the entropy $h(Y_{n,k})$ for BPSK input denotes an upper bound on the entropy $h(Y_{n,k})$ for Gaussian input. In another argumentation the entropy of a less widespread pdf (Gaussian input) will be lower since there is less uncertainty contained.

3.3.2 Numerical Evaluation

For numerical evaluation a convolution of the pdf in (3.12) divided by $|s|$ (because of pdf-transformation) and the absolute-part-pdf of (3.16),

$$p_{|Z_{n,k}|}(|z|) = \frac{|z|}{2\pi N_0}\, e^{-\frac{|z|^2}{2N_0}},$$

divided by $|z|$, has to be done by using (3.14). This convolution was done numerically. Figure 3.5 shows the distributions of $S_{n,k}$ and $Z_{n,k}$ in the complex plane. The distribution of $S_{n,k}$ (cf. (3.13)) contains the Bessel function of second kind and order 0 which describes the maximum point at the origin tending to infinity. Hence these numerical calculations will suffer certain inaccuracy. The distribution of $Z_{n,k}$ is circular complex Gaussian. Then, the distribution of output symbols is obtained by convolution $p_{Y_{n,k}}(y_R, y_I) = (p_{S_{n,k}} * p_{Z_{n,k}})(y_R, y_I)$ (cf. (3.14)).

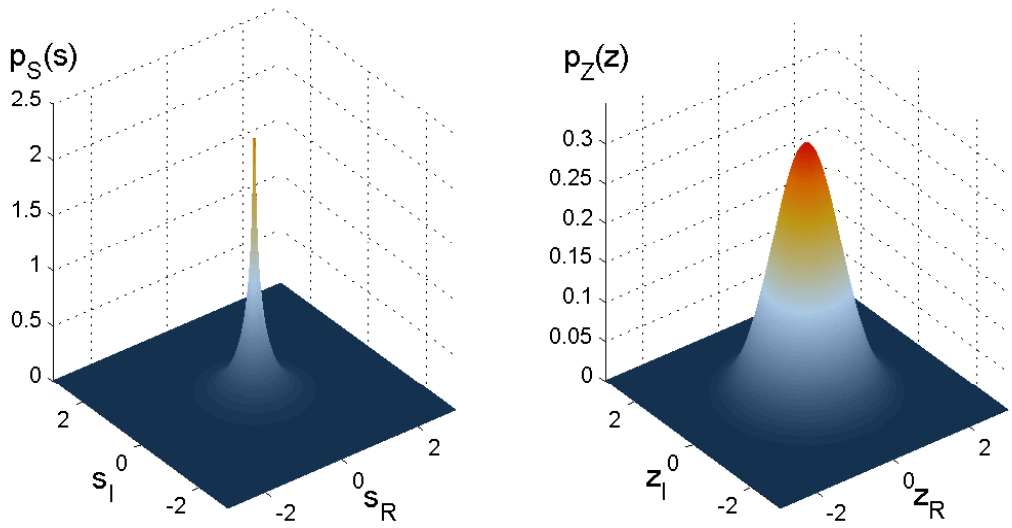

Figure 3.5: *Numerically evaluated pdfs of (a) $S_{n,k}$ and (b) $Z_{n,k}$ for $\sigma_X^2 = \sigma_{\mathbb{H}}^2 = 1$ and $N_0 = 1\,W/Hz$.*

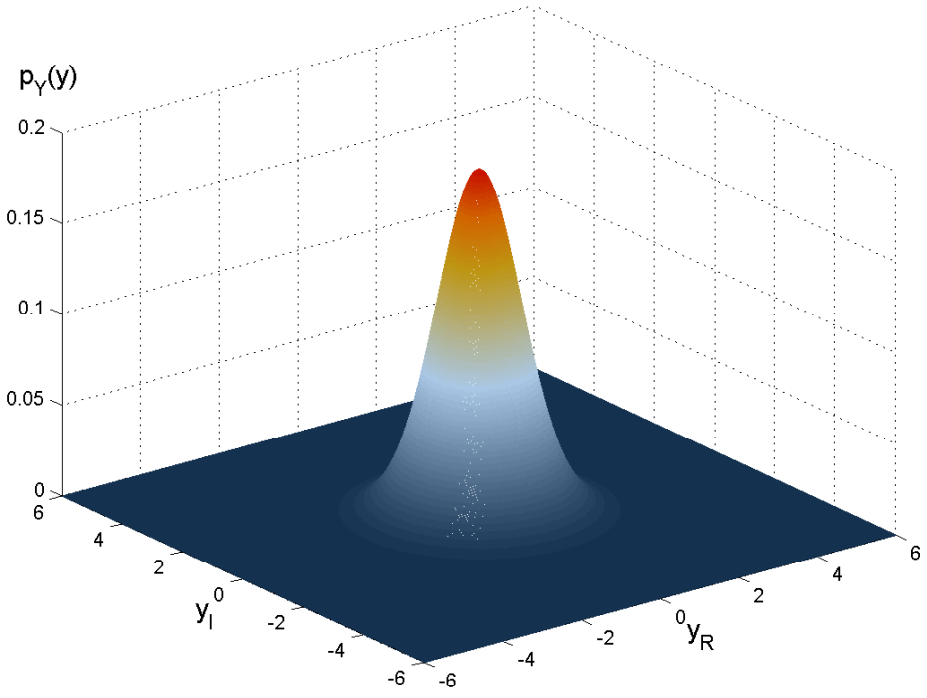

Figure 3.6: *Numerically evaluated pdf of the output $Y_{n,k}$ for $\sigma_X^2 = \sigma_{\mathbb{H}}^2 = 1$ and $N_0 = 1\,W/Hz$.*

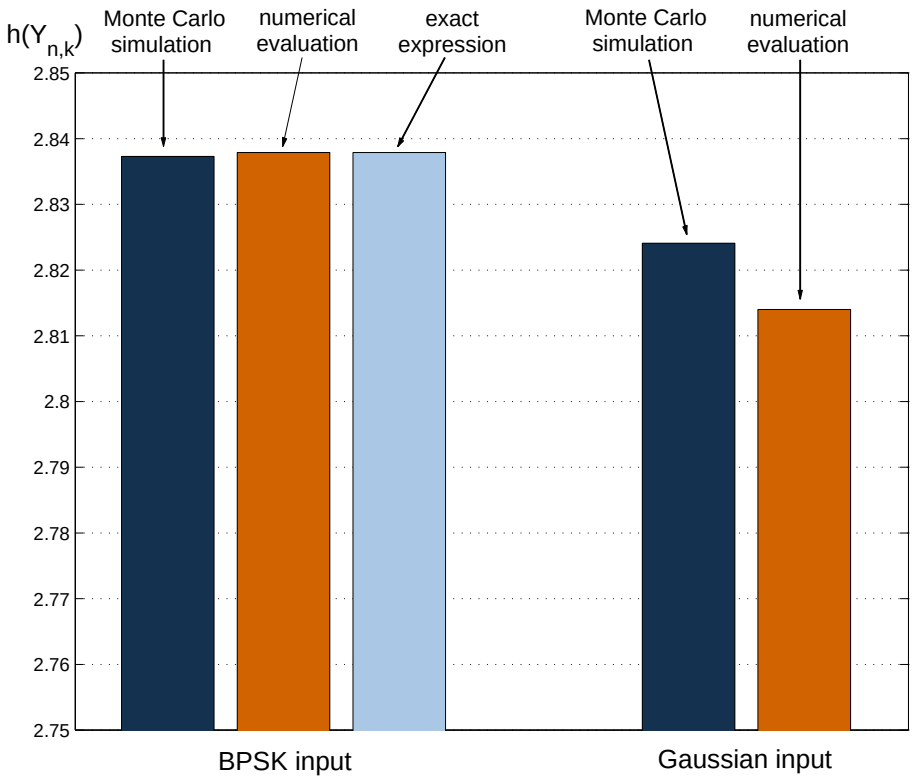

Figure 3.7: *Differential entropy of output $Y_{n,k}$ for BPSK input (left) and Gaussian input (right) obtained with Monte Carlo simulation, numerical evaluation and exact expression. All results are (as above) for $\sigma_X^2 = \sigma_{\mathbb{H}}^2 = 1$ and $N_0 = 1\,W/Hz$.*

This is shown in (Figure 3.6). It can be seen, that the pdf is similar to a complex Gaussian pdf, however it is slightly more concentrated at the origin.

3.3.3 Comparison of Simulation and Numerical Evaluation

In this Subsection the Monte Carlo simulated and numerically calculated pdfs of the output vector $\mathbf{Y}$ are used for analysing its statistical properties. The most important value here is be the entropy of the output. In Figure 3.7 the entropy of $\mathbf{Y_{n,k}}$ obtained with different input symbols and Monte Carlo simulation and numerical evaluation is presented. Here, the above mentioned fact that the entropy for BPSK input is higher than that for a Gaussian input is confirmed. There are also negligible differences (caused by too small resolution at the sampling of the pdfs) between Monte Carlo simulation, numerical calculation and accurate calculation according to (1.29). In Figure 3.8 similar results are shown for $N_0 = 10^{-3}$W/Hz. Here also the difference between Monte Carlo simulation and numerical calculation for Gaussian input constellation can be seen. In Figure 3.8 also an entropy for a numerical result with a resolution (sampling of pdf) of 401×401 bins is shown. Since the pdf of $S_{n,k}$ gets infinite at the origin, the quantization underlying the empirical histogram causes an error. Therefore a more accurate

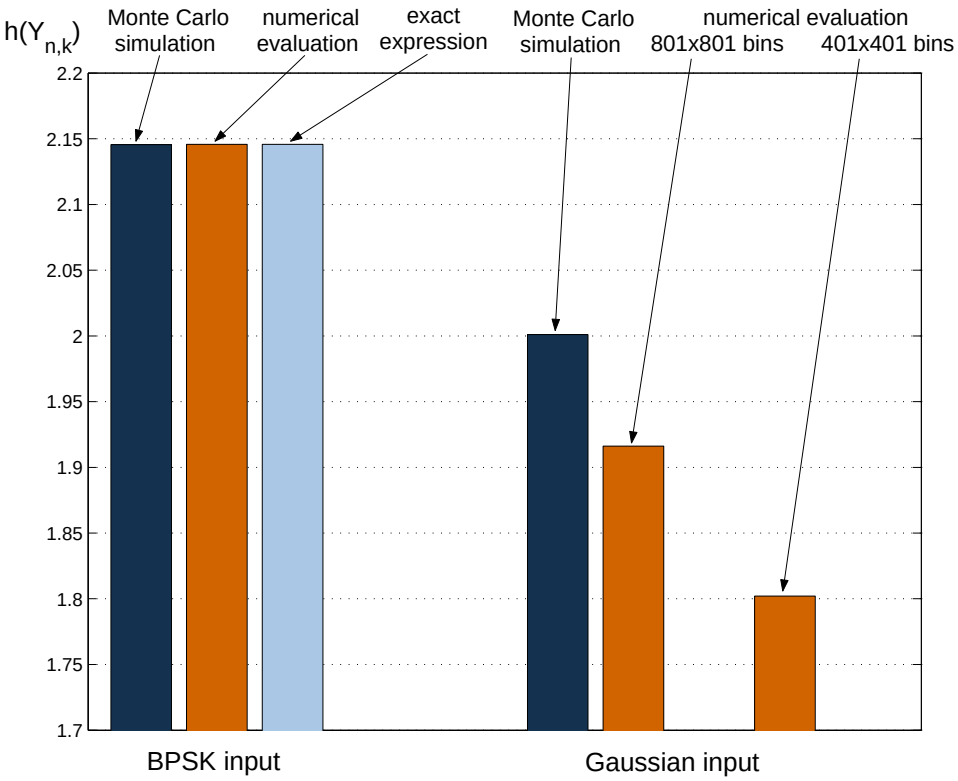

Figure 3.8: *Differential entropy of output $Y_{n,k}$ for BPSK input (left) and Gaussian input (right) obtained with Monte Carlo simulation, numerical evaluation (varying resolution) and exact expression. All results are for $\sigma_X^2 = \sigma_{\mathbb{H}}^2 = 1$ and $N_0 = 10^{-3}\,W/Hz$.*

result can be obtained for a higher resolution of 801×801bins which is also shown in 3.8. For increasing resolution the result obtained with Monte Carlo simulation should be reached.

3.3.4 Simulation of Upper Bound on Information Rate

The simulation of the upper bound on information rate is done according to (3.27). Figure 3.9 shows the result for this fourthegy upper bound for two different channels with flat scattering functions and channel spreads $\tau_{\max}\nu_{\max} = 10^{-2}$ and $\tau_{\max}\nu_{\max} = 10^{-3}$. The simulation setup was taken from [12], where an IEEE 802.11a related system was simulated. The subcarrier spacing was $F = 312.5$kHz and $TF = N/K = 1.25$. The transmit power was $P = 1$mW, the path loss for both channels $\sigma_{\mathbb{H}} = 90$dB and $N_0 = k_0 \cdot 400K = 4.1421 \cdot 10^{-21}$W with the Boltzmann constant $k_0 = 1.3807 \cdot 10^{-23}$Ws/Hz. For comparison the result of [12] in (2.2) and the additive white Gaussian noise (AWGN) information rate are depicted as dashed curves. The dark solid curves depicts the fourthegy upper bound for $\tau_{\max}\nu_{\max} = 10^{-3}$ the bright solid curve the fourthegy upper bound for $\tau_{\max}\nu_{\max} = 10^{-2}$. It can be seen that the fourthegy upper bound isn‘t very tight for low to medium frequencies. This can be described

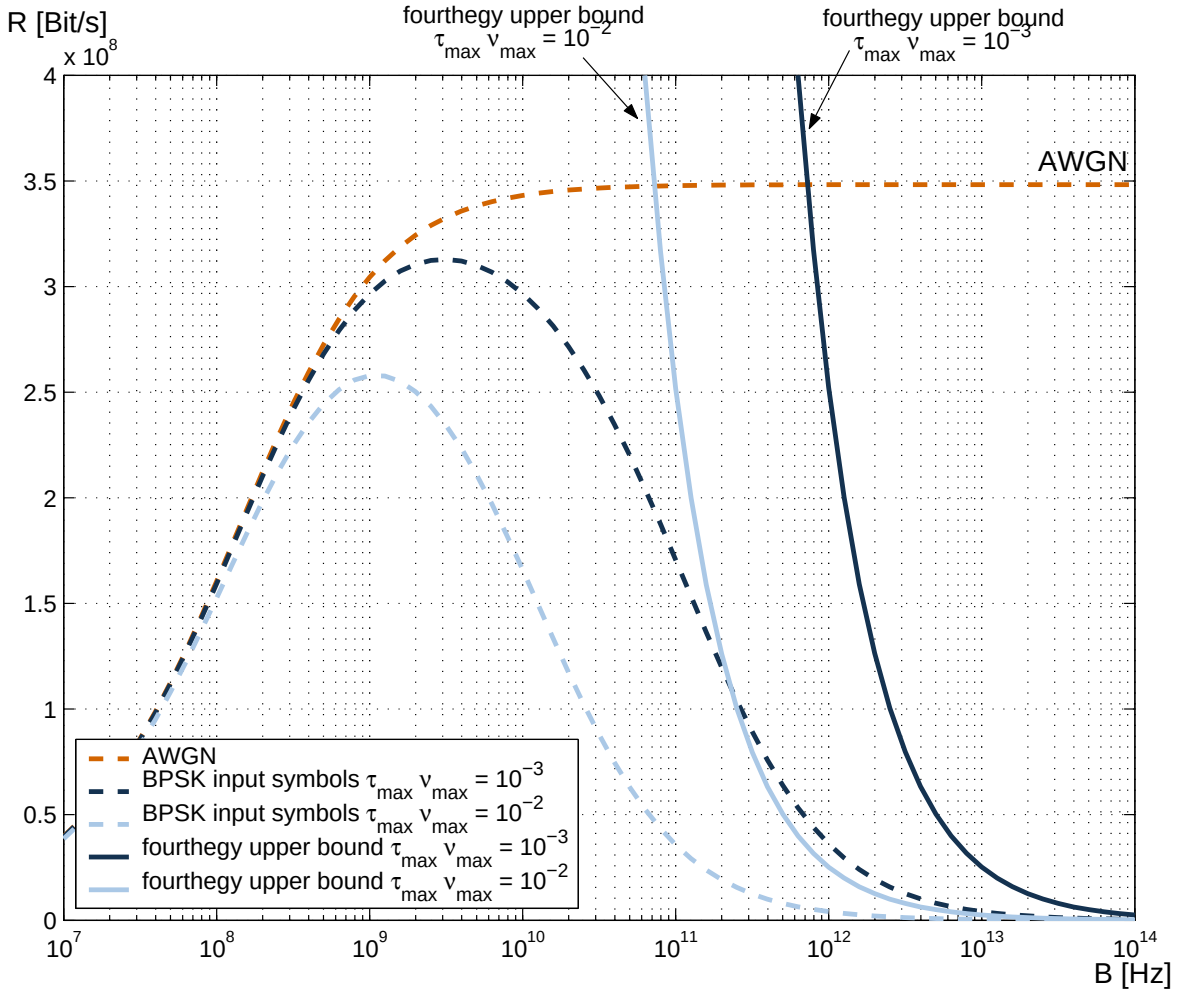

Figure 3.9: *Fourthegy upper bound on information rate for channel spreads* $\tau_{\max}\nu_{\max} = 10^{-3}$ *(dark solid curve) and* $\tau_{\max}\nu_{\max} = 10^{-2}$ *(bright solid curve).*

through the tightness of fourthegy bounding ($\log(1+x) \geq x - \frac{x^2}{2}$) that is only given for small arguments and hence SNR. However, it correctly predicts that in the infinite bandwidth case the information rate will tend to zero. This effect can be explained with an increasing effort in channel estimation (since the accuracy of estimation is inverse proportional to bandwidth B). The second (intuitive) conclusion is that a low channel spread is desirable in order to enable higher rates.

Chapter 4

System Capacity for a Multiuser OFDM System

In Chapter 3 we investigated the system capacity of MC-CDMA, where we used a single user OFDM model combined with CDMA. Results for sum system capacity were presented for a high number of users. In this chapter we will use a multiuser OFDM model. Hence each user can be distinguished through its own codebook (i.e., MC-CDMA or OFDMA) and channel. Further the number of users doesn't underlie any constraint. This multiuser system will be subjected to mutual information bounding techniques used in [38, 42]. We will apply these bounding techniques on mutual information and information rate, respectively. All through the derivations the focus will be on the sum information rate. This is done according to a total power constraint over all users. Then, mainly numerical evaluations for the upper bound on information rate will be presented. Using these results the impact of number of users, input codebook, channel statistics, and power allocation will be shown. First, lower an introduction to the multiuser OFDM model will be given.

4.1 Definitions and Notation

The definitions of one users vectors and matrices are similar to those in Section 3.1. Under the assumption of independent users we now extend the model to a multiuser OFDM model. There, U users transmit to a single receiver (uplink case). All users apply an OFDM modulator and the receiver employs an OFDM demodulator. Each user is transmitting over its independent channel $\mathbb{H}_U$. At the receiver there is a superposition of all users output vectors and circularly symmetric complex Gaussian noise is added. This multiuser OFDM model for an uplink scenario is shown in Figure 4.1. The received signal will be given by the $MK \times 1$ vector $\mathbf{Y}$ as before but the user specific transmit signal and channel will get a user index u. First a

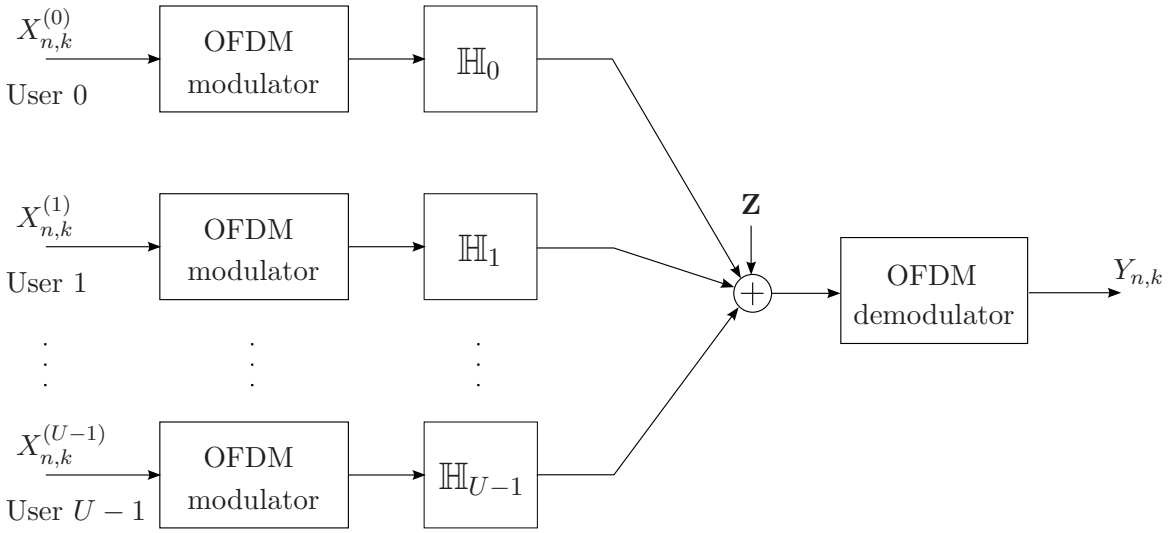

Figure 4.1: *Multiuser OFDM model for an uplink scenario.*

summation over all user signals at the receiver gives

$$\mathbf{Y} = \sum_{u=0}^{U-1} \operatorname{diag}\left\{\mathbf{H}^{(u)}\right\} \mathbf{X}^{(u)} + \mathbf{Z} = \sum_{u=0}^{U-1} \operatorname{diag}\left\{\mathbf{X}^{(u)}\right\} \mathbf{H}^{(u)} + \mathbf{Z}, \tag{4.1}$$

with U the number of users transmitting to the receiver. In matrix-vector notation this is

$$\mathbf{Y} = \mathcal{H}_U \mathbf{X}_U + \mathbf{Z}, \tag{4.2}$$

with the $MKU \times 1$ input vector

$$\mathbf{X}_U = \left[\mathbf{X}^{(0)^T} \mathbf{X}^{(1)^T} \dots \mathbf{X}^{(U-1)^T}\right]^T,$$

and the overall channel matrix of size $MK \times MKU$

$$\mathcal{H}_U = \left[\operatorname{diag}\left\{\mathbf{H}^{(0)}\right\} \operatorname{diag}\left\{\mathbf{H}^{(1)}\right\} \dots \operatorname{diag}\left\{\mathbf{H}^{(U-1)}\right\}\right].$$

The $MK \times 1$ matrix $\mathbf{Z}$ is circularly symmetric complex Gaussian noise ($\mathbf{Z} \in \mathcal{N}(\mathbf{0}, N_0\mathbf{I})$). Different from the input vector in Chapter 3 the input vector $\mathbf{X}_U$ now combines all user input vectors. A total power constraint on the input symbols, i.e., $\mathrm{E}\{\mathbf{X}_U^H \mathbf{X}_U\} = MKU\sigma_X^2$ is assumed. Note that the diagonal form of the channel matrix for user u results from the multiplicative model in (1.25). To take into account correlations on one users channel (different users will be assumed as independent) $\mathcal{H}_U$ gets a matrix with entries that are not only positioned in a diagonal way like above, but on every element of $\mathcal{H}_U$. Equivalently the multiuser input-output relation can be written as

$$\mathbf{Y} = \mathcal{X}_U \mathbf{H}_U + \mathbf{Z}, \tag{4.3}$$

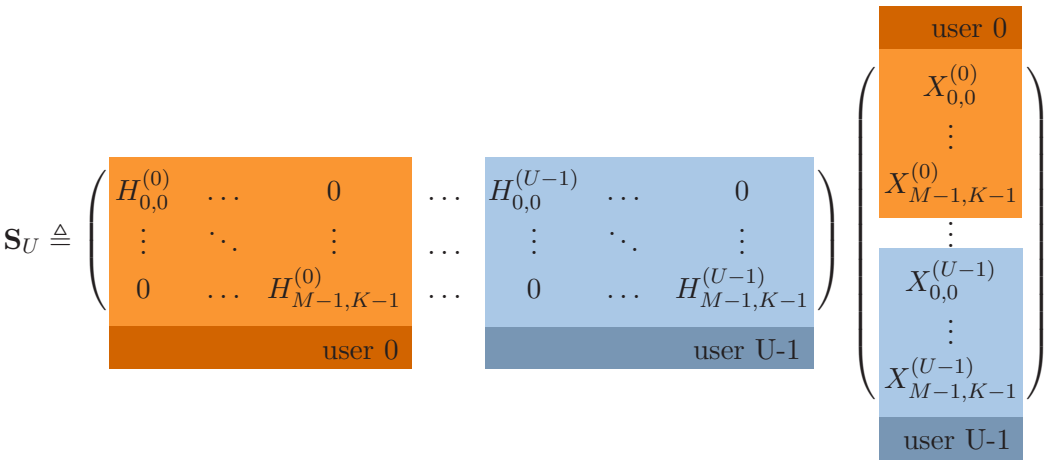

Figure 4.2: *Interpretation of the multiple user OFDM model in* (4.2) *through* $\mathbf{S}_U$ (4.5).

with the channel vector of size $MKU \times 1$

$$\mathbf{H}_U = \left[\mathbf{H}^{(0)^T}\mathbf{H}^{(1)^T} \dots \mathbf{H}^{(U-1)^T}\right]^T, \tag{4.4}$$

and the $MK \times MKU$ input matrix

$$\mathcal{X}_U = \left[\operatorname{diag}\left\{\mathbf{X}^{(0)}\right\} \operatorname{diag}\left\{\mathbf{X}^{(1)}\right\} \dots \operatorname{diag}\left\{\mathbf{X}^{(U-1)}\right\}\right].$$

For convenience, we use the shortened notation

$$\mathbf{S}_U = \mathcal{H}_U\mathbf{X}_U = \mathcal{X}_U\mathbf{H}_U. \tag{4.5}$$

To ease the understanding $\mathbf{S}_U = \mathcal{H}_U\mathbf{X}_U$ is depicted in Figure 4.2.

4.2 System Capacity Bounding I

In this section the bounding techniques of Bölcskei and Shamai in [38] will be applied to the multiuser model introduced above. We assume no channel knowledge neither at transmitter nor at receiver. The channel itself is assumed to be a time and frequency selective Rayleigh fading underspread WSSUS channel. In order to preserve generality, the derivations are done for a general codebook of $\mathbf{X}^{(u)}$ as far as possible. An exception will be made for bounding purposes.

First the signal model is adapted to the multiuser case. We are using the input-output relation in (4.2) or equivalently (4.3). We start splitting the sum mutual information[1] $I(\mathbf{Y};\mathbf{X})$ into

$$I(\mathbf{Y};\mathbf{X}) = I(\mathbf{Y};\mathbf{X},\mathbf{H}) - I(\mathbf{Y};\mathbf{H}|\mathbf{X}). \tag{4.6}$$

Note that all considerations are for the sum rate of all users. Using (4.6) we calculate upper and lower bounds for $I(\mathbf{Y};\mathbf{X})$.

[1] For convenience, we omit the subscript U and write e.g., $\mathbf{X} \equiv \mathbf{X}_U$.

4.2.1 Upper Bound on Mutual Information

In order to obtain an upper bound the terms in (4.6) now will be bounded. First the mutual information $I(\mathbf{Y};\mathbf{X},\mathbf{H})$ is upper bounded by that of an additive Gaussian noise channel where the transmitter has total control of the input, in that case $\mathbf{S} = \mathcal{H}\mathbf{X}$. If $\mathbf{S}$ is made circularly symmetric jointly complex Gaussian with correlation equal to that of $\mathcal{H}\mathbf{X}$, one obtains

$$I(\mathbf{Y};\mathbf{X},\mathbf{H}) \leq \log\ \det\ \left(\mathbf{I} + \frac{1}{N_0}\mathbf{R}_H \odot \mathbf{C}_X\right), \tag{4.7}$$

with the covariance matrices $\mathbf{C}_X$ of the transmit signal and $\mathbf{R}_H$ of the channel. The $MKU \times MKU$ matrices $\mathbf{C}_X$ and $\mathbf{R}_H$ are given by

$$\mathbf{C}_X = \mathrm{E}\left\{\mathbf{X}\mathbf{X}^H\right\} = \begin{pmatrix} \mathbf{C}_X^{(0)} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_X^{(1)} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{C}_X^{(U-1)} \end{pmatrix}, \tag{4.8}$$

with users input covariance matrices $\mathbf{C}_X^{(u)} = \mathrm{E}\left\{\mathbf{X}^{(u)}\mathbf{X}^{(u)\,H}\right\}$ and

$$\mathbf{R}_H = \mathrm{E}\left\{\mathbf{H}\mathbf{H}^H\right\} = \begin{pmatrix} \mathbf{R}_H^{(0)} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_H^{(1)} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{R}_H^{(U-1)} \end{pmatrix},$$

with users channel covariance matrices $\mathbf{R}_H^{(u)} = \mathrm{E}\left\{\mathbf{H}^{(u)}\mathbf{H}^{(u)\,H}\right\}$. Note that $\mathbf{0}$ denotes the $MK \times MK$ all-zero matrix. Noting this, (4.7) gets

$$I(\mathbf{Y};\mathbf{X},\mathbf{H}) \leq \log \prod_{u=0}^{U-1} \det\ \left(\mathbf{I} + \frac{1}{N_0}\mathbf{R}_H^{(u)} \odot \mathbf{C}_X^{(u)}\right) = \sum_{u=0}^{U-1} \log\ \det\ \left(\mathbf{I} + \frac{1}{N_0}\mathbf{C}_S^{(u)}\right). \tag{4.9}$$

In (4.9), $\mathbf{R}_H^{(u)} \odot \mathbf{C}_X^{(u)}$ corresponds to the covariance matrix of $\mathbf{S}^{(u)}$, i.e., $\mathbf{C}_S^{(u)} = \mathbf{R}_H^{(u)} \odot \mathbf{C}_X^{(u)}$. The goal for the second part of the upper bound is a lower bound or an exact expression for $I(\mathbf{Y};\mathbf{H}|\mathbf{X})$. The input output relation in (4.3) is used. For $I(\mathbf{Y};\mathbf{H}|\mathbf{X})$ the channel $\mathbf{H}$ acts as transmit signal and the 'channel' $\mathcal{X}$ is known. In [38] an equality is given, which here reads

$$I(\mathbf{Y};\mathbf{H}|\mathbf{X}) = \mathrm{E}_X\left\{\log\ \det\ \left(\mathbf{I} + \frac{1}{N_0}\,\mathcal{X}\,\mathbf{R}_H\mathcal{X}^H\right)\right\}. \tag{4.10}$$

Defining $\mathbf{A} = \frac{1}{N_0}\,\mathcal{X}\,\mathbf{R}_H\mathcal{X}^H$ and using the eigenvalue decomposition of $\mathbf{A}$ leads to

$$I(\mathbf{Y};\mathbf{H}|\mathbf{X}) = \mathrm{E}_X\left\{\log \prod_{i=1}^{r(\mathbf{A})} (1+\lambda_i(\mathbf{A}))\right\} = \sum_{i=1}^{r(\mathbf{A})} \mathrm{E}_X\left\{\log\ (1+\lambda_i(\mathbf{A}))\right\},$$

with $r(\mathbf{A})$ and $\lambda_i(\mathbf{A})$ denoting the rank and i-th eigenvalue of matrix $\mathbf{A}$, respectively. To obtain an upper bound on $I(\mathbf{Y};\mathbf{H}|\mathbf{X})$, a fourthegy lower bound (like in (3.22)) is used. Then

$$\begin{aligned} I(\mathbf{Y};\mathbf{H}|\mathbf{X}) &\geq \sum_{i=1}^{r(\mathbf{A})} \mathrm{E}_X\left\{\lambda_i(\mathbf{A}) - \frac{\lambda_i^2(\mathbf{A})}{2}\right\} = \sum_{i=1}^{r(\mathbf{A})} \mathrm{E}_X\left\{\lambda_i(\mathbf{A})\right\} - \frac{1}{2}\sum_{i=1}^{r(\mathbf{A})} \mathrm{E}_X\left\{\lambda_i^2(\mathbf{A})\right\} \\ &= \mathrm{E}_X\left\{\sum_{i=1}^{r(\mathbf{A})} \lambda_i(\mathbf{A})\right\} - \frac{1}{2}\mathrm{E}_X\left\{\sum_{i=1}^{r(\mathbf{A})} \lambda_i^2(\mathbf{A})\right\} = \mathrm{E}_X\left\{\mathrm{Tr}(\mathbf{A})\right\} - \frac{1}{2}\,\mathrm{Tr}(\mathbf{A}\mathbf{A}^H), \end{aligned}$$

with $\mathrm{Tr}(\mathbf{A}) = \sum_{i=1}^{r(\mathbf{A})} \lambda_i(\mathbf{A})$ and $\mathrm{Tr}(\mathbf{A}\mathbf{A}^H) = \sum_{i=1}^{r(\mathbf{A})} \lambda_i^2(\mathbf{A})$. Resubstitution of $\mathbf{A}$ gives

$$I(\mathbf{Y};\mathbf{H}|\mathbf{X}) \geq \frac{1}{N_0}\,\mathrm{E}_X\left\{\mathrm{Tr}\left(\mathcal{X}\,\mathbf{R}_H\mathcal{X}^H\right)\right\} - \frac{1}{2N_0^2}\,\mathrm{E}_X\left\{\mathrm{Tr}\left(\mathcal{X}\,\mathbf{R}_H\mathcal{X}^H\,\mathcal{X}\,\mathbf{R}_H^H\,\mathcal{X}^H\right)\right\}, \tag{4.11}$$

with

$$\mathrm{E}_X\left\{\mathrm{Tr}\left(\mathcal{X}\,\mathbf{R}_H\mathcal{X}^H\right)\right\} = \mathrm{Tr}\left(\mathrm{E}_X\left\{\mathcal{X}^H\mathcal{X}\right\}\mathbf{R}_H\right) = \mathrm{Tr}\left(\mathbf{C}_X \odot \mathbf{R}_H\right) = MKU\sigma_X^2\sigma_{\mathbb{H}}^2, \tag{4.12}$$

which holds for any input codebook (i.e., MC-CDMA and OFDMA), and

$$\begin{aligned} &\mathrm{E}_X\left\{\mathrm{Tr}\left(\mathcal{X}\,\mathbf{R}_H\mathcal{X}^H\,\mathcal{X}\,\mathbf{R}_H^H\mathcal{X}^H\right)\right\} \\ &= \sum_{i_1,i4=0}^{MK-1}\sum_{i_2,i_3,i_5,i_6=0}^{MKU-1} (\mathcal{X})_{i_1,i_2}(\mathbf{R}_X)_{i_2,i_3}\left(\mathcal{X}^H\right)_{i_3,i_4}(\mathcal{X})_{i_4,i_5}\left(\mathbf{R}_X^H\right)_{i_5,i_6}\left(\mathcal{X}^H\right)_{i_6,i_1} \\ &= \sum_{u=0}^{U-1}\sum_{u'=0}^{U-1}\sum_{i=0}^{MK-1}\sum_{j=0}^{MK-1} \mathrm{E}_X\left\{r_{i,j}^{(u)}\; x_i^{(u)}x_j^{(u)*}r_{i,j}^{(u')*}x_i^{(u')*}x_j^{(u')}\right\} \\ &= \sum_{k=0}^{MK-1}\sum_{l=0}^{MK-1}\sum_{u=0}^{U-1} \mathrm{E}_X\left\{r_{k,l}^{(u)}\; x_k^{(u)}x_l^{(u)*}\sum_{u'=0}^{U-1}\left(r_{k,l}^{(u')*}x_k^{(u')*}x_l^{(u')}\right)\right\}. \end{aligned} \tag{4.13}$$

Here $r_{i,j}^{(u)}$ denotes the correlation of the MK channel coefficients of user u, which is equal to the channels correlation matrix entry $\left(\mathbf{R}_X^{(u)}\right)_{i,j}$. The input symbols $x_i^{(u)}$ are taken as the i-th element of users input matrix $\mathbf{X}^{(u)}$. For input correlations of the same user ($u = u'$) and with $\mathrm{E}_X\left\{x_k^{(u)}x_l^{(u)*}x_k^{(u)*}x_l^{(u)}\right\} = \mathrm{E}_X\left\{|x_k^{(u)}|^2\,|x_l^{(u)}|^2\right\}$ we introduce one users $MK \times MK$ fourth moment matrix

$$\mathbf{K}_X^{(u)} = \begin{pmatrix} \mathrm{E}_X\left\{|x_{0,0}^{(u)}|^4\right\} & \mathrm{E}_X\left\{|x_{0,0}^{(u)}|^2|x_{0,1}^{(u)}|^2\right\} & \cdots & \mathrm{E}_X\left\{|x_{0,0}^{(u)}|^2|x_{M-1,K-1}^{(u)}|^2\right\} \\ \mathrm{E}_X\left\{|x_{0,1}^{(u)}|^2|x_{0,0}^{(u)}|^2\right\} & \mathrm{E}_X\left\{|x_{0,1}^{(u)}|^4\right\} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{E}_X\left\{|x_{M-1,K-1}^{(u)}|^2|x_{0,0}^{(u)}|^2\right\} & \cdots & \cdots & \mathrm{E}_X\left\{|x_{M-1,K-1}^{(u)}|^4\right\} \end{pmatrix}, \tag{4.14}$$

where the input symbols are $x_{0,0}^{(u)}, x_{0,1}^{(u)}, \ldots, x_{M-1,K-1}^{(u)}$.

Then, with the result in (4.13) and the fourth moment matrix $\mathbf{K}_X^{(u)}$ the remaining expectation term in (4.11) can be written as

$$\mathrm{E}_X\left\{\mathrm{Tr}\left(\mathcal{X}\,\mathbf{R}_H\mathcal{X}^H\,\mathcal{X}\,\mathbf{R}_H^H\mathcal{X}^H\right)\right\} = \mathrm{vec}^T\left\{\sum_{u=0}^{U-1}\mathbf{R}_H^{(u)}\odot\mathbf{R}_H^{(u)}\odot\left[\mathbf{K}_X^{(u)}-\mathbf{C}_X^{(u)}\odot\mathbf{C}_X^{(u)}\right]\right\}\mathbf{1} + \mathrm{vec}^T\left\{\sum_{u=0}^{U-1}\sum_{u'=0}^{U-1}\mathbf{R}_H^{(u)}\odot\mathbf{R}_H^{(u')}\odot\mathbf{C}_X^{(u)}\odot\mathbf{C}_X^{(u')}\right\}\mathbf{1}, \quad (4.15)$$

where $\mathrm{vec}^T\{\cdot\}\mathbf{1}$ (with $\mathbf{1}$ denoting the size $(MK)^2\times 1$ all ones vector) was used to add up all matrix entries. Here the hermitian conjugation of $\mathbf{R}_H^{(u)}$ and $\mathbf{C}_X^{(u)}$ were suppressed due to hermitian nature of correlation matrices ${\mathbf{R}_H^{(u)}}^H=\mathbf{R}_H^{(u)}$ and ${\mathbf{C}_X^{(u)}}^H=\mathbf{C}_X^{(u)}$, respectively. Further the expectation terms in (4.13) which contain more than one users correlation coefficients simplify to

$$\mathrm{E}_X\left\{x_k^{(u)}x_l^{(u)^*}x_k^{(u')^*}x_l^{(u')}\right\} = \mathrm{E}_X\left\{x_k^{(u)}x_l^{(u)^*}\right\}\mathrm{E}_X\left\{x_k^{(u')^*}x_l^{(u')}\right\},$$

due to the assumption of independent users. Then, these terms can be expressed through the users correlation matrices $\mathbf{C}_X^{(u)}$.

An interpretation of the first term in (4.15) leads to the impact of peakiness. By the subtraction $\mathbf{K}_X^{(u)}-\mathbf{C}_X^{(u)}\odot\mathbf{C}_X^{(u)}$ there only remain terms $\mathrm{E}_X\left\{|x_k^{(u)}|^2|x_l^{(u)}|^2\right\}-\mathrm{E}_X^2\left\{x_k^{(u)}x_l^{(u)}\right\}$ (where again $x_i^{(u)}$ denotes the i-th element of users input matrix $\mathbf{X}^{(u)}$) describing statistical relations of different symbols, i.e., for PSK input symbols this gives

$$\mathbf{K}_X^{(u)}-\mathbf{C}_X^{(u)}\odot\mathbf{C}_X^{(u)} = \begin{pmatrix} 0 & \sigma_X^4 & \dots & \sigma_X^4 \\ \sigma_X^4 & 0 & \ddots & \sigma_X^4 \\ \vdots & \ddots & \ddots & \vdots \\ \sigma_X^4 & \dots & \sigma_X^4 & 0 \end{pmatrix}.$$

For Gaussian distributed input symbols one would obtain $\mathbf{K}_X^{(u)}-\mathbf{C}_X^{(u)}\odot\mathbf{C}_X^{(u)}=\sigma_X^4\mathbf{1}\mathbf{1}^T$. More general, this can be explained using input symbols

$$X_{n,k} = \begin{cases} A, & \text{with probability } p, \\ 0, & \text{with probability } 1-p, \end{cases}$$

where p denotes the probability of $X_{n,k}$ to have magnitude A. Therefore p also is a measure of peakiness, e.g. $p\approx 0$ characterises a peaky codebook. Then, we obtain

$$\mathbf{K}_X^{(u)}-\mathbf{C}_X^{(u)}\odot\mathbf{C}_X^{(u)} = \sigma_X^4\begin{pmatrix} \frac{1-p}{p} & 1 & \dots & 1 \\ 1 & \frac{1-p}{p} & \ddots & 1 \\ \vdots & \ddots & \ddots & \vdots \\ 1 & \dots & 1 & \frac{1-p}{p} \end{pmatrix}.$$

Hence for a peaky codebook (e.g.,$(1-p)/p \gg 0$) , as it is supposed to achieve system capacity in [12] low correlation will be obtained and the first term in (4.15) is getting large.
Another interpretation follows the following conversion, i.e.,

$$\begin{aligned}
\mathrm{vec}^T\left\{\sum_{u=0}^{U-1}\mathbf{R}_H^{(u)}\odot\mathbf{R}_H^{(u)}\odot\mathbf{K}_X^{(u)}\right\}\mathbf{1} &= \sum_{u=0}^{U-1}\sum_{i=0}^{MK-1}\sum_{j=0}^{MK-1}\left(\mathbf{R}_H^{(u)}\odot\mathbf{R}_H^{(u)}\odot\mathbf{K}_X^{(u)}\right)_{i,j}\\
&= \sum_{u=0}^{U-1}\sum_{i=0}^{MK-1}\sum_{j=0}^{MK-1}\left(\mathbf{R}_H^{(u)}\right)_{i,j}\left(\mathbf{K}_X^{(u)}\right)_{i,j}\left(\mathbf{R}_H^{(u)}\right)^*_{j,i}\\
&= \sum_{u=0}^{U-1}\mathrm{Tr}\left\{\left(\mathbf{R}_H^{(u)}\odot\mathbf{K}_X^{(u)}\right)\mathbf{R}_H^{(u)H}\right\}\\
&= \sum_{u=0}^{U-1}\left\langle\mathbf{R}_H^{(u)}\odot\mathbf{K}_X^{(u)},\mathbf{R}_H^{(u)H}\right\rangle.
\end{aligned}$$

Here by use of the inner product, the impact of input symbols (expressed through the fourth moment) on the channel is investigated. With the correlation matrix of $\mathbf{S}$, i.e., $\mathbf{C}_S^{(u)}=\mathbf{R}_H^{(u)}\odot\mathbf{C}_X^{(u)}$ the second part of the first term in (4.15) can be calculated as follows:

$$\begin{aligned}
\mathrm{vec}^T\left\{\sum_{u=0}^{U-1}\mathbf{R}_H^{(u)}\odot\mathbf{R}_H^{(u)}\odot\mathbf{C}_X^{(u)}\odot\mathbf{C}_X^{(u)}\right\}\mathbf{1} &= \mathrm{vec}^T\left\{\sum_{u=0}^{U-1}\mathbf{C}_S^{(u)}\odot\mathbf{C}_S^{(u)}\right\}\mathbf{1}\\
&= \sum_{u=0}^{U-1}\sum_{i=0}^{MK-1}\sum_{j=0}^{MK-1}\left(\mathbf{C}_S^{(u)}\right)_{i,j}\left(\mathbf{C}_S^{(u)}\right)_{j,i}\\
&= \sum_{u=0}^{U-1}\mathrm{Tr}\left(\mathbf{C}_S^{(u)}\odot\mathbf{C}_S^{(u)}\right)\\
&= \sum_{u=0}^{U-1}\left\langle\mathbf{C}_S^{(u)},\mathbf{C}_S^{(u)}\right\rangle. \qquad (4.16)
\end{aligned}$$

It is a sum over the fourth moments of all users at the receiver.
So far only the effects of single users input and channel coefficients were treated, but no interactions between users. This can be described through the second term in (4.15). By use of a similar formulation as in (4.16) it can be rewritten as

$$\begin{aligned}
\mathrm{vec}^T\left\{\sum_{u=0}^{U-1}\sum_{u'=0}^{U-1}\mathbf{R}_H^{(u)}\odot\mathbf{R}_H^{(u')}\odot\mathbf{C}_X^{(u)}\odot\mathbf{C}_X^{(u')}\right\}\mathbf{1} &= \mathrm{vec}^T\left\{\sum_{u=0}^{U-1}\sum_{u'=0}^{U-1}\mathbf{C}_S^{(u)}\odot\mathbf{C}_S^{(u')}\right\}\mathbf{1}\\
&= \sum_{u=0}^{U-1}\sum_{u'=0}^{U-1}\sum_{i=0}^{MK-1}\sum_{j=0}^{MK-1}\left(\mathbf{C}_S^{(u)}\right)_{i,j}\left(\mathbf{C}_S^{(u')}\right)_{j,i}\\
&= \sum_{u=0}^{U-1}\sum_{u'=0}^{U-1}\mathrm{Tr}\left(\mathbf{C}_S^{(u)}\odot\mathbf{C}_S^{(u')}\right)\\
&= \sum_{u=0}^{U-1}\sum_{u'=0}^{U-1}\left\langle\mathbf{C}_S^{(u)},\mathbf{C}_S^{(u')}\right\rangle. \qquad (4.17)
\end{aligned}$$

Here the inner product gives information on the similarity of two different users correlation at the receiver. Hence a maximization of (4.17) is achieved through equivalent input-channel combinations of all users. Then, the lower bound on $I(\mathbf{Y};\mathbf{H}|\mathbf{X})$ can be combined to

$$I(\mathbf{Y};\mathbf{H}|\mathbf{X}) \geq \frac{MKU}{N_0}\sigma_X^2\sigma_{\mathbb{H}}^2 - \frac{1}{2N_0^2}\sum_{u=0}^{U-1}\left\langle \mathbf{R}_H^{(u)} \odot \mathbf{K}_X^{(u)}, \mathbf{R}_H^{(u)^H}\right\rangle + \frac{1}{2N_0^2}\sum_{u=0}^{U-1}\left\langle \mathbf{C}_S^{(u)}, \mathbf{C}_S^{(u)}\right\rangle - \frac{1}{2N_0^2}\sum_{u=0}^{U-1}\sum_{u'=0}^{U-1}\left\langle \mathbf{C}_S^{(u)}, \mathbf{C}_S^{(u')}\right\rangle. \tag{4.18}$$

By combination of (4.9), (4.12), the first term in (4.15) and (4.17) the total upper bound on mutual information is obtained as

$$I(\mathbf{Y};\mathbf{X}) \leq \sum_{u=0}^{U-1}\log\ \det\ \left(\mathbf{I} + \frac{1}{N_0}\mathbf{R}_H^{(u)} \odot \mathbf{C}_X^{(u)}\right) - \frac{MKU}{N_0}\sigma_X^2\sigma_{\mathbb{H}}^2 + \frac{1}{2N_0^2}\sum_{u=0}^{U-1}\left\langle \mathbf{R}_H^{(u)} \odot \mathbf{K}_X^{(u)}, \mathbf{R}_H^{(u)^H}\right\rangle + \frac{1}{2N_0^2}\sum_{\substack{u=0\\u\neq u'}}^{U-1}\sum_{u'=0}^{U-1}\left\langle \mathbf{C}_S^{(u)}, \mathbf{C}_S^{(u')}\right\rangle. \tag{4.19}$$

The first term in (4.19) (which is similar to a parallel AWGN channel) can further be bounded by

$$\begin{aligned}\sum_{u=0}^{U-1}\log\ \det\ \left(\mathbf{I} + \frac{1}{N_0}\mathbf{R}_H^{(u)} \odot \mathbf{C}_X^{(u)}\right) &= \sum_{u=0}^{U-1}\sum_{i=0}^{r\left(\frac{1}{N_0}\mathbf{R}_H^{(u)}\odot\mathbf{C}_X^{(u)}\right)}\log\left(1+\lambda_i\left(\frac{1}{N_0}\mathbf{R}_H^{(u)} \odot \mathbf{C}_X^{(u)}\right)\right)\\ &\leq \sum_{u=0}^{U-1}\sum_{i=0}^{r\left(\frac{1}{N_0}\mathbf{R}_H^{(u)}\odot\mathbf{C}_X^{(u)}\right)}\lambda_i\left(\frac{1}{N_0}\mathbf{R}_H^{(u)} \odot \mathbf{C}_X^{(u)}\right)\\ &= \sum_{u=0}^{U-1}\mathrm{Tr}\left(\frac{1}{N_0}\mathbf{R}_H^{(u)} \odot \mathbf{C}_X^{(u)}\right) = \frac{MKU}{N_0}\sigma_X^2\sigma_{\mathbb{H}}^2,\end{aligned}$$

where the bound will be tight for low SNR. Then, the sum mutual information for low SNR is obtained as

$$I(\mathbf{Y};\mathbf{X}) \leq \frac{1}{2N_0^2}\sum_{u=0}^{U-1}\left\langle \mathbf{R}_H^{(u)} \odot \mathbf{K}_X^{(u)}, \mathbf{R}_H^{(u)^H}\right\rangle + \frac{1}{2N_0^2}\sum_{\substack{u=0\\u\neq u'}}^{U-1}\sum_{u'=0}^{U-1}\left\langle \mathbf{C}_S^{(u)}, \mathbf{C}_S^{(u')}\right\rangle.$$

In order to maximize the upper bound two different aspects correlated to the two terms in (4.20) can be discussed. First the impact of the channel input (in terms of fourth moment) on the channels correlation structure should be small. Hence an appropriate selection of codebook contributes to the maximization of sum mutual information. It was shown, that a peaky codebook will maximize this term.

The second term seems to be much more important because there are more terms to summarize and more interesting it allows insights on the effects of a multiuser system, since the inner

product expresses similarity of the user's input-channel combinations. Then, for maximization of (4.20) equivalent input-channel combinations are beneficial (i.e.,$\mathbf{C}_S^{(u)} \approx c\mathbf{C}_S^{(u')}$). This includes the fact, that different channel statistics may be balanced by different codebooks.

4.2.2 Lower Bound on Mutual Information

For the derivation of the lower bound on the system capacity again (4.6) is used. A further splitting and bounding of mutual information $I(\mathbf{Y};\mathbf{X})$ gives

$$I(\mathbf{Y};\mathbf{X}) = I(\mathbf{Y};\mathbf{X}|\mathbf{H}) + I(\mathbf{Y};\mathbf{H}) - I(\mathbf{Y};\mathbf{H}|\mathbf{X}) \geq I(\mathbf{Y};\mathbf{X}|\mathbf{H}) - I(\mathbf{Y};\mathbf{H}|\mathbf{X}), \tag{4.20}$$

where the trivial upper bound $I(\mathbf{Y};\mathbf{H}) \geq 0$ was used. By assuming circularly symmetric jointly complex Gaussian transmit signals the first term is lower bounded by

$$I(\mathbf{Y};\mathbf{X}|\mathbf{H}) \geq \mathrm{E}_H \left\{ \log\ \det\ \left(\mathbf{I} + \frac{1}{N_0}\, \mathcal{H}\mathbf{C}_X\mathcal{H}^H \right) \right\}. \tag{4.21}$$

The term in (4.21) now is equivalent to (4.10) with exchange of $\mathcal{X}$ by $\mathcal{H}$ and $\mathbf{R}_H$ by $\mathbf{C}_X$. Applying the fourthegy bounding technique we obtain a result dual to (4.18):

$$\begin{aligned} I(\mathbf{Y};\mathbf{X}|\mathbf{H}) \geq \frac{MKU}{N_0}\sigma_X^2\sigma_{\mathbb{H}}^2 - \frac{1}{2N_0^2}\sum_{u=0}^{U-1}\left\langle \mathbf{C}_X^{(u)} \odot \mathbf{K}_H^{(u)}, \mathbf{C}_X^{(u)H} \right\rangle + \frac{1}{2N_0^2}\sum_{u=0}^{U-1}\left\langle \mathbf{C}_S^{(u)}, \mathbf{C}_S^{(u)} \right\rangle \\ - \frac{1}{2N_0^2}\sum_{u=0}^{U-1}\sum_{u'=0}^{U-1}\left\langle \mathbf{C}_S^{(u)}, \mathbf{C}_S^{(u')} \right\rangle. \end{aligned} \tag{4.22}$$

with the fourth moment matrix of the channel $\mathbf{H}^{(u)}$,

$$\mathbf{K}_H^{(u)} = \begin{pmatrix} \mathrm{E}_H\left\{|h_{0,0}^{(u)}|^4\right\} & \mathrm{E}_H\left\{|h_{0,0}^{(u)}|^2|h_{0,1}^{(u)}|^2\right\} & \dots & \mathrm{E}_H\left\{|h_{0,0}^{(u)}|^2|h_{M-1,K-1}^{(u)}|^2\right\} \\ \mathrm{E}_H\left\{|h_{0,1}^{(u)}|^2|h_{0,0}^{(u)}|^2\right\} & \mathrm{E}_H\left\{|h_{0,1}^{(u)}|^4\right\} & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{E}_H\left\{|h_{M-1,K-1}^{(u)}|^2|h_{0,0}^{(u)}|^2\right\} & \dots & \dots & \mathrm{E}_H\left\{|h_{M-1,K-1}^{(u)}|^4\right\} \end{pmatrix}. \tag{4.23}$$

Next, simplifications for the complex Gaussian channel coefficients can be applied. For the fourth moment of channel symbols $\mathrm{E}_H\left\{|h_{k,l}^{(u)}|^4\right\}$ (same user, same symbol) in (4.23) we get

$$\mathrm{E}_H\left\{|h_{k,l}^{(u)}|^4\right\} = 2\,\sigma_{\mathbb{H}}^4,$$

and for the same user and different symbols $(k,l) \neq (k',l')$

$$\mathrm{E}_H\left\{|h_{k,l}^{(u)}|^2|h_{k',l'}^{(u)}|^2\right\} = \mathrm{E}_H\left\{h_{k,l}^{(u)}h_{k,l}^{(u)*}\right\}\mathrm{E}_H\left\{h_{k',l'}^{(u)}h_{k',l'}^{(u)\ *}\right\} + \mathrm{E}_H\left\{h_{k,l}^{(u)}h_{k',l'}^{(u)\ *}\right\}\mathrm{E}_H\left\{h_{k,l}^{(u)*}h_{k',l'}^{(u)}\right\},$$

by use of Isserlis' theorem[2]. The matrix $\mathbf{K}_H^{(u)}$ in (4.23) then can be expressed by the path loss $\sigma_{\mathbb{H}}^2$ and the elements of the channel's covariance matrix $\mathbf{R}_H$:

$$\mathbf{K}_H^{(u)} = \sigma_{\mathbb{H}}^2 \left[\mathbf{I} + \mathbf{1}\mathbf{1}^T\right] + \mathrm{vec}^T\{\mathbf{R}_H^{(u)}\}\mathrm{vec}\{\mathbf{R}_H^{(u)}\}.$$

Note that this bounding on the first term started with a Gaussian assumption on the input symbols. The second term in (4.20) is upper bounded by applying Jensen's inequality as

$$\begin{aligned} I(\mathbf{Y};\mathbf{H}|\mathbf{X}) &\leq \log\ \det\ \left(\mathbf{I} + \frac{1}{N_0}\mathrm{E}_X\left\{\sum_{u=0}^{U-1}\mathrm{diag}\{\mathbf{X}^{(u)}\}^H\mathrm{diag}\{\mathbf{X}^{(u)}\}\mathbf{R}_H^{(u)}\right\}\right) \\ &= \log\ \det\ \left(\mathbf{I} + \frac{1}{N_0}\sum_{u=0}^{U-1}\mathrm{E}_X\left\{\mathrm{diag}\{\mathbf{X}^{(u)}\}^H\mathrm{diag}\{\mathbf{X}^{(u)}\}\right\}\mathbf{R}_H^{(u)}\right) \\ &= \log\ \det\ \left(\mathbf{I} + \frac{\sigma_X^2}{N_0}\sum_{u=0}^{U-1}\mathbf{P}^{(u)}\mathbf{R}_H^{(u)}\right), \end{aligned} \tag{4.24}$$

with $MK \times MK$ matrix

$$\mathbf{P}^{(u)} = \begin{pmatrix} p_{0,0}^{(u)} & 0 & \dots & 0 \\ 0 & p_{0,1}^{(u)} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & p_{M-1,K-1}^{(u)} \end{pmatrix},$$

describing the power allocation of one user to different subcarriers. Hence it is used to define the input codebook, e.g., for OFDMA the coefficients of $\mathbf{P}^{(u)}$ are

$$p_{n,k}^{(u)} = \begin{cases} 1, & k \in \mathcal{K}_U, \\ 0, & k \notin \mathcal{K}_U, \end{cases}$$

with $\mathcal{K}_U$ denoting the subset of subcarriers allocated to user u. In contrast, $p_{n,k}^{(u)} = 1$ for all n, k with MC-CDMA. The total lower bound then is combined to

$$\begin{aligned} I(\mathbf{Y};\mathbf{X}) \geq\ & \frac{MKU}{N_0}\sigma_X^2\sigma_{\mathbb{H}}^2 - \frac{1}{2N_0^2}\sum_{u=0}^{U-1}\left\langle \mathbf{C}_X^{(u)} \odot \mathbf{K}_H^{(u)}, \mathbf{C}_X^{(u)H}\right\rangle \\ & - \frac{1}{2N_0^2}\sum_{u=0}^{U-1}\sum_{\substack{u'=0 \\ u\neq u'}}^{U-1}\left\langle \mathbf{C}_S^{(u)}, \mathbf{C}_S^{(u')}\right\rangle - \log\ \det\ \left(\mathbf{I} + \frac{\sigma_X^2}{N_0}\sum_{u=0}^{U-1}\mathbf{P}^{(u)}\mathbf{R}_H^{(u)}\right). \end{aligned} \tag{4.25}$$

4.2.3 System Capacity Bounding

The bounds for system capacity are obtained by combining the formulas for system capacity and ergodic information rate in Section 1.4.3, i.e.,

$$R \triangleq \lim_{M\to\infty}\ \frac{1}{MT}\ I(\mathbf{Y};\mathbf{X}) \tag{4.26}$$

[2]The fourth moment of circular complex Gaussian random variables is given through $\mathrm{E}\{x_1x_2^*x_3x_4^*\} = \mathrm{E}\{x_1x_2^*\}\ \mathrm{E}\{x_3x_4^*\} + \mathrm{E}\{x_1x_4^*\}\ \mathrm{E}\{x_2^*x_3\}$ [48, 49].

with the results in (4.16) and (4.25). Since a closed form expression for general assumptions on the channel's statistics can not be found here only the result for an uncorrelated channel (i.e. $\mathbf{R}_H^{(u)} = \sigma_{\mathbb{H}}^2\mathbf{I}$), uncorrelated input symbols (i.e. $\mathbf{C}_X^{(u)} = \sigma_X^2\mathbf{I}$ and $\mathbf{K}_X^{(u)} = \sigma_X^4\left[\mathbf{I} + 2\mathbf{1}\mathbf{1}^T\right]$) for every user is derived as

$$R \leq R_{\mathrm{UB}} = U\frac{B}{TF}\log\left(1 + \frac{TF}{B}\frac{P\sigma_{\mathbb{H}}^2}{UN_0}\right) - \frac{P\sigma_{\mathbb{H}}^2}{N_0} + \frac{TF}{B}\frac{3}{2}\frac{P^2\sigma_{\mathbb{H}}^4}{U^2N_0^2} + \frac{TF}{B}\frac{U(U-1)P^2\sigma_{\mathbb{H}}^4}{2\,U^2N_0^2}, \quad (4.27)$$

where $\sigma_X^2 = PT/K$ and $K = B/F$ with P the constant transmit power (constant in the sum over all users) and B the entire bandwidth.

4.3 System Capacity Bounding II

As above, a system capacity bounding will be applied to the multiuser OFDM model defined in Section 4.1. Further an extension to statistical characterizations like the scattering function will be done. This derivation is related to system capacity bounding of Shamai and Marzetta in [42]. We start with a reformulation of the signal model. The time-dependent transfer function of the channel is defined in (1.4). From the spreading function the time-dependent transfer function can be calculated by combination of (1.4) and (1.5). The channel coefficients in the multicarrier system then are

$$H_{n,k}^{(u)} = \sum_{\tau=0}^{L_\tau}\sum_{\nu=-L_\nu/2}^{L_\nu/2} S_{\mathbb{H}}(\tau,\nu)e^{-j2\pi\left(\frac{\tau k}{K} - \frac{\nu n}{M}\right)}d\tau d\nu.$$

Note that a single user channel is treated first, with the extension to the multiuser case following afterwards. The non-zero support region of the discretized spreading function $S_{\mathbb{H}}(kL_\tau/K, nL_\nu/M)$ is defined in $[0, L_\tau] \times [-L_\nu/2, L_\nu/2]$ and $L_\tau L_\nu < MK$ with the maximum multipath delay L_τ and the maximum Doppler frequency L_ν. Therefore, the nonzero elements of the spreading function are arranged into a length $L_\tau L_\nu$ vector as follows:

$$\mathbf{h}_r^{(u)} = \begin{pmatrix} S_{\mathbb{H}_c}(0,0) \\ S_{\mathbb{H}_c}(0,1) \\ \vdots \\ S_{\mathbb{H}_c}(L_\tau, L_\nu) \end{pmatrix}.$$

Then, the $MK \times 1$ stacked channel vector for user u is defined as[3]

$$\mathbf{H}^{(u)} = \left(\mathbf{F}_{K\times L_\tau} \otimes \mathbf{F}_{M\times L_\nu}\right)\mathbf{h}_r{}^{(u)},$$

with $\mathbf{F}$ the DFT matrix of size $K \times L_\tau$ and $M \times L_\nu$, respectively. As we will apply the signal

[3]Here $\otimes$ denotes the Kronecker product [50].

model in (4.3) with the $MKU \times 1$ channel vector in (4.4) we get

$$\mathbf{H} = \begin{pmatrix} (\mathbf{F}_{K\times L_\tau} \otimes \mathbf{F}_{M\times L_\nu})\, \mathbf{h}_r{}^{(0)} \\ (\mathbf{F}_{K\times L_\tau} \otimes \mathbf{F}_{M\times L_\nu})\, \mathbf{h}_r{}^{(1)} \\ \vdots \\ (\mathbf{F}_{K\times L_\tau} \otimes \mathbf{F}_{M\times L_\nu})\, \mathbf{h}_r{}^{(U-1)} \end{pmatrix} = \mathbf{I} \otimes (\mathbf{F}_{K\times L_\tau} \otimes \mathbf{F}_{M\times L_\nu})\, \mathbf{h}_r,$$

with $\mathbf{h}_r = \left[\mathbf{h}_r^{(0)^T} \mathbf{h}_r^{(1)^T} \ldots\ \mathbf{h}_r^{(U-1)^T}\right]^T$,. Next, we write $\mathbf{F}$ instead of $\mathbf{I} \otimes (\mathbf{F}_{K\times L_\tau} \otimes \mathbf{F}_{M\times L_\nu})$. The signal model in (4.3) then gets

$$\begin{aligned} \mathbf{S} &= \left[\mathrm{diag}\left\{\mathbf{X}^{(0)}\right\} \mathrm{diag}\left\{\mathbf{X}^{(1)}\right\}\ \ldots\ \mathrm{diag}\left\{\mathbf{X}^{(U-1)}\right\}\right] \mathbf{H} \\ &= \left[\mathrm{diag}\left\{\mathbf{X}^{(0)}\right\} \mathrm{diag}\left\{\mathbf{X}^{(1)}\right\}\ \ldots\ \mathrm{diag}\left\{\mathbf{X}^{(U-1)}\right\}\right] \mathbf{F}\ \mathbf{h}_r = \mathcal{X}_U\ \mathbf{h}_r, \end{aligned}$$

with

$$\mathcal{X}_U = \left[\mathrm{diag}\left\{\mathbf{X}^{(0)}\right\} \mathrm{diag}\left\{\mathbf{X}^{(1)}\right\}\ \ldots\ \mathrm{diag}\left\{\mathbf{X}^{(U-1)}\right\}\right] \mathbf{F}.$$

For one users input signal this is equal to

$$\mathcal{X}_U^{(u)} = \mathrm{diag}\left\{\mathbf{X}^{(u)}\right\} (\mathbf{F}_{K\times L_\tau} \otimes \mathbf{F}_{M\times L_\nu})\,.$$

Note that this signal model assumes independent users. The whole signal model then equals

$$\mathbf{Y} = \mathcal{X}_U\ \mathbf{h}_r + \mathbf{Z},$$

where $\mathbf{Z}$ is circularly symmetric complex Gaussian noise ($\mathbf{Z} \in \mathcal{N}(\mathbf{0}, N_0\mathbf{I})$). In this chapter the shorter notation $\mathbf{S}$ will be used for $\mathbf{S} = \mathcal{X}_U\ \mathbf{h}_r$. Throughout the whole derivation of upper and lower bound the input power constraint is

$$\mathrm{Tr}\left(\mathrm{E}_X\left\{\mathcal{X}_U\ \mathcal{X}_U{}^H\right\}\right) = MKU\sigma_X^2.$$

4.3.1 Upper Bound on Mutual Information

First an upper bound on mutual information will be derived. We are using again the chain rule of mutual information in (4.6),

$$I(\mathbf{Y}; \mathcal{X}_U) = I(\mathbf{Y}; \mathcal{X}_U, \mathbf{H}) - I(\mathbf{Y}; \mathbf{H}|\mathcal{X}_U). \tag{4.28}$$

Note that the shift through Fourier transformation resulting in $\mathcal{X}_U$ changes mutual information since there is loss or gain of information through an non-invertible DFT. This DFT is non-invertible because the transformation matrix $\mathbf{F}$ isn‘t unitary. Out of the chain rule in (4.28) a simple upper bound on mutual information can be found by dropping the second term $I(\mathbf{Y}; \mathbf{H}|\mathcal{X}_U) > 0$ which gives

$$I(\mathbf{Y}; \mathcal{X}_U) \leq I(\mathbf{Y}; \mathcal{X}_U, \mathbf{H}). \tag{4.29}$$

Then, a maximization of $I(\mathbf{Y}; \mathcal{X}_U, \mathbf{H})$ has to be found. The term is maximized by making $\mathbf{S} = \mathcal{X}_U\ \mathbf{h}_r$ zero-mean complex Gaussian with covariance $\mathbf{C}_S = \mathrm{E}_X\{\mathcal{X}_U\ \mathbf{R}_{h_r}\ \mathcal{X}_U{}^H\}$. This gives

$$I(\mathbf{Y}; \mathcal{X}_U, \mathbf{H}) \leq \log\ \det\ \left(\mathbf{I} + \frac{1}{N_0}\ \mathrm{E}_X\left\{\mathcal{X}_U\ \mathbf{R}_{h_r}\ \mathcal{X}_U{}^H\right\}\right), \tag{4.30}$$

with

$$\mathbf{R}_{h_r} = \mathrm{E}_H\left\{\mathbf{h}_r\ \mathbf{h}_r^H\right\} = \begin{pmatrix} \mathbf{R}_{h_r}^{(0)} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \mathbf{R}_{h_r}^{(U-1)} \end{pmatrix}, \tag{4.31}$$

the covariance matrix of the channel. Referring to (1.6) for one user u this can be denoted by

$$\mathbf{R}_{h_r}^{(u)} = \mathrm{diag}\left\{C_{\mathbb{H}_\mathrm{c}}^{(u)}(\tau, \nu)\right\} = \begin{pmatrix} C_{\mathbb{H}_\mathrm{c}}^{(u)}(0,0) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & C_{\mathbb{H}_\mathrm{c}}^{(u)}(L_\tau, L_\nu) \end{pmatrix}. \tag{4.32}$$

Then, the correlation term in (4.30) can be rewritten to

$$\begin{aligned} &\mathrm{E}_X\left\{\mathcal{X}_U\ \mathbf{R}_{h_r}\ \mathcal{X}_U{}^H\right\} \\ &\quad= \mathrm{E}_X\left\{\mathcal{X}_U^{(0)}\mathbf{R}_{h_r}^{(0)}\left[\mathcal{X}_U^{(0)}\right]^H + \mathcal{X}_U^{(1)}\mathbf{R}_{h_r}^{(1)}\left[\mathcal{X}_U^{(1)}\right]^H + \dots + \mathcal{X}_U^{(U-1)}\mathbf{R}_{h_r}^{(U-1)}\left[\mathcal{X}_U^{(U-1)}\right]^H\right\} \\ &\quad= \mathrm{E}_X\left\{\sum_{u=0}^{U-1}\mathcal{X}_U^{(u)}\mathbf{R}_{h_r}^{(u)}\left[\mathcal{X}_U^{(u)}\right]^H\right\} = \sum_{u=0}^{U-1}\mathrm{E}_X\left\{\mathcal{X}_U^{(u)}\mathbf{R}_{h_r}^{(u)}\left[\mathcal{X}_U^{(u)}\right]^H\right\} = \sum_{u=0}^{U-1}\mathbf{C}_S^{(u)}. \end{aligned}$$

With this result the upper bound can be written as

$$I(\mathbf{Y}; \mathcal{X}_U) \leq \log\ \det\ \left(\mathbf{I} + \frac{1}{N_0}\sum_{u=0}^{U-1}\mathbf{C}_S^{(u)}\right). \tag{4.33}$$

Then, it easily can be shown that the additive white Gaussian case marks an upper bound on mutual information by assuming $\mathbf{C}_S^{(u)} \approx \mathcal{CN}\left(\mathbf{0}, \sigma_{\mathbb{H}}^2\sigma_X^2\mathbf{I}\right)$. The trivial Gaussian upper bound then is

$$I(\mathbf{Y}; \mathcal{X}_U) \leq UMK \log\ \left(1 + \frac{\sigma_{\mathbb{H}}^2\sigma_X^2}{N_0}\right). \tag{4.34}$$

4.3.2 Lower Bound on Mutual Information

The lower bound is found through the splitting of mutual information in (4.28) and the bounding

$$I(\mathbf{Y}; \mathcal{X}_U, \mathbf{H}) \geq I(\mathbf{Y}; \mathcal{X}_U|\mathbf{H}).$$

Then, two terms have to be considered for the lower bound, e.g.

$$I(\mathbf{Y}; \mathcal{X}_U) \geq I(\mathbf{Y}; \mathcal{X}_U|\mathbf{H}) - I(\mathbf{Y}; \mathbf{H}|\mathcal{X}_U). \tag{4.35}$$

For the first term in (4.35) we start with

$$I(\mathbf{Y};\mathcal{X}_U|\mathbf{H}) \geq MK\ \mathrm{E}_H\left\{\log\ \det\ \left(1+\frac{\sigma_X^2}{N_0}\ \mathbf{h}_r\mathbf{h}_r^H\right)\right\},$$

where we assumed Gaussian distributed and white input $\mathcal{X}_U$ in order to maximize $I(\mathbf{Y};\mathcal{X}_U|\mathbf{H})$. We will use this assumption for the derivation of the entire lower bound. By applying a fourthegy lower bound ($\log(1+x) \geq x - \frac{x^2}{2}$) we obtain

$$\begin{aligned} I(\mathbf{Y};\mathcal{X}_U|\mathbf{H}) &\geq MK\ \frac{\sigma_X^2}{N_0}\ \mathrm{Tr}\left\{\mathrm{E}_H\left\{\mathbf{h}_r\ \mathbf{h}_r^H\right\}\right\} - MK\ \frac{\sigma_X^4}{N_0^2}\ \mathrm{Tr}\left\{\mathrm{E}_H\left\{\mathbf{h}_r^H\mathbf{h}_\mathrm{r}\ \mathbf{h}_\mathrm{r}^H\mathbf{h}_\mathrm{r}\right\}\right\} \\ &= MK\ \frac{\sigma_X^2}{N_0}\ \mathrm{Tr}\left\{\mathbf{R}_{h_\mathrm{r}}\right\} - MK\ \frac{\sigma_X^4}{N_0^2}\left(\mathrm{Tr}^2\left\{\mathbf{R}_{h_\mathrm{r}}\right\} + \mathrm{Tr}\left\{\mathbf{R}_{h_\mathrm{r}}^2\right\}\right), \end{aligned} \tag{4.36}$$

were we used

$$\mathrm{Tr}\left\{\mathrm{E}_H\left\{\mathbf{h}_\mathrm{r}^H\mathbf{h}_\mathrm{r}\ \mathbf{h}_\mathrm{r}^H\mathbf{h}_\mathrm{r}\right\}\right\} = \mathrm{Tr}^2\left\{\mathbf{R}_{h_\mathrm{r}}\right\} + \mathrm{Tr}\left(\mathbf{R}_{h_\mathrm{r}}^2\right).$$

Note, that $\mathcal{X}_U$ already was assumed to be Gaussian for the derivation of the first term in (4.35). The maximization of the second term in (4.35), $I(\mathbf{Y};\mathbf{H}|\mathcal{X}_U)$ is again done by making $\mathbf{H}$ zero-mean complex Gaussian. This constellation is equal to the (however virtual) case of the user sending $\mathbf{H}$ over a channel $\mathcal{X}_U$. With $\mathbf{S} = \mathcal{X}_U\ \mathbf{h}_r$ this gives

$$\begin{aligned} I(\mathbf{Y};\mathbf{H}|\mathcal{X}_U) &\leq \mathrm{E}_X\left\{\log\ \det\ \left(\mathbf{I}+\frac{1}{N_0}\ \mathbf{S}\mathbf{S}^H\right)\right\} \\ &= \mathrm{E}_X\left\{\log\ \det\ \left(\mathbf{I}+\frac{1}{N_0}\ \mathcal{X}_U\ \mathbf{R}_{h_r}{\mathcal{X}_U}^H\right)\right\} \\ &\leq \log\ \det\ \left(\mathbf{I}+\frac{1}{N_0}\ \mathrm{E}_X\left\{\mathcal{X}_U\ \mathbf{R}_{h_r}{\mathcal{X}_U}^H\right\}\right), \end{aligned}$$

where we used Jensen's inequality for the second inequation. Then, the result is

$$I(\mathbf{Y};\mathbf{H}|\mathcal{X}_U) \leq \log\ \det\ \left(\mathbf{I}+\frac{1}{N_0}\ \mathrm{E}_X\left\{\mathcal{X}_U\ \mathbf{R}_{h_r}{\mathcal{X}_U}^H\right\}\right), \tag{4.37}$$

which is equivalent to the result in (4.24).

The entire lower bound is given through the combination of (4.37) and (4.36) as

$$\begin{aligned} I(\mathbf{Y};\mathcal{X}_U) \geq\ & MK\ \frac{\sigma_X^2\sigma_\mathbb{H}^2}{N_0} - MK\ \frac{\sigma_X^4}{N_0^2}\left[\sigma_\mathbb{H}^4 + \mathrm{Tr}\left\{\mathbf{R}_{h_\mathrm{r}}^2\right\}\right] \\ & - \log\ \det\ \left(\mathbf{I}+\frac{1}{N_0}\ \mathrm{E}_X\left\{\mathcal{X}_U\ \mathbf{R}_{h_\mathrm{r}}{\mathcal{X}_U}^H\right\}\right). \end{aligned}$$

Finally the result is

$$\begin{aligned} I(\mathbf{Y};\mathcal{X}_U) \geq\ & MK\ \frac{\sigma_X^2\sigma_\mathbb{H}^2}{N_0} - MK\ \frac{\sigma_X^4}{N_0^2}\left[\sigma_\mathbb{H}^4 + \mathrm{Tr}\left\{\mathbf{R}_{h_\mathrm{r}}^2\right\}\right] \\ & - \log\ \det\ \left(\mathbf{I}+\frac{1}{N_0}\ \mathrm{E}_X\left\{\mathcal{X}_U\ \mathbf{R}_{h_\mathrm{r}}{\mathcal{X}_U}^H\right\}\right), \end{aligned}$$

where $\mathrm{E}_X\left\{\mathcal{X}_U\ \mathbf{R}_{h_\mathrm{r}}{\mathcal{X}_U}^H\right\}$ also may be written as $\mathrm{E}_X\left\{\mathcal{X}_U\ \mathbf{R}_H{\mathcal{X}_U}^H\right\}$. The dependence of the lower bound on sum information rate on the scattering function is given through the relations in (4.31) (4.32). Note, that Gaussian distributed and white input $\mathcal{X}_U$ was assumed for the derivation of the lower bound.

F	312.5 kHz
TF	1.25
P	1 mW
$\sigma^2_{\mathbb{H}}$	90 dB
N_0	$4.1421 \cdot 10^{-21}$ W

Table 4.1: *Simulation parameters for an IEEE 802.11a related system.*

4.4 Simulation Results

In order to get an evaluation of the bounds derived above, some simulations are presented next. First results for exact expressions will be presented. Afterwards numerical evaluation results are depicted.

4.4.1 Upper bound on sum information rate for uncorrelated channel

The simulations are based on the upper bound on information rate derived in (4.27) (using bounding techniques presented in [38]) and will be done for an IEEE 802.11a related system as in Chapter 3. Table 4.1 subsumes these values that are applied for the simulations unless specified otherwise.

Upper bound on sum information rate for different number of users

Figure 4.3 shows the sum information rate over bandwidth. This numerical evaluation is based on the upper bound derived in (4.27). The AWGN case which characterises an upper bound on information rate is depicted as a dashed line. Then, the two-user case and a result for $U = 1000$ are depicted through a dark solid line and a bright solid line, respectively. The achievable capacity region (or equivalently the achievable rates) can be found below the bounded area. For low bandwidths the AWGN is more tight and hence will be the limit. At higher bandwidths (beginning at intersection of AWGN rate and R_{UB}) the fourthegy upper bound is more tight and indicates the upper bound. For an increasing number of users the fourthegy upper bound on sum information rate decreases. So for uncorrelated channels an increasing number of users is disadvantageous in terms of sum information rate. Further the upper bound on sum information rate in the infinite bandwidth case tends to zero.

Upper bound on sum information rate for user-specific path losses

In order to describe the influence of the path losses on the sum information rate first an exact expression has to be found. The path losses of different users are denoted by $\sigma^2_{\mathbb{H}_U}$. In the two-user case this gives an upper bound on sum information rate (derived from (4.27)) which

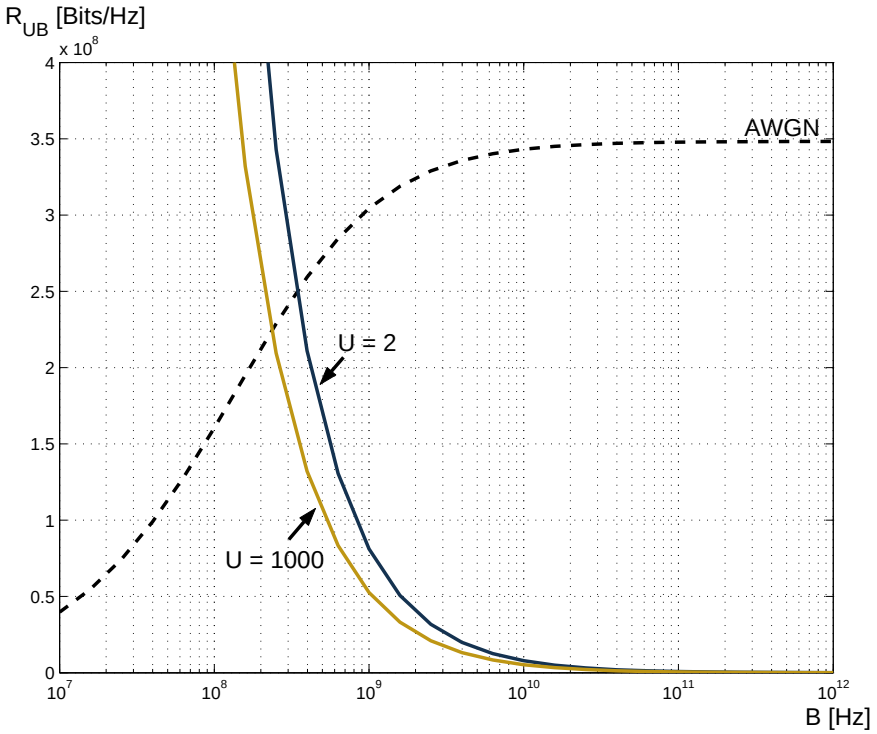

Figure 4.3: *Fourthegy upper bound on sum information rate over bandwidth for independent channel coefficients and different number of users.*

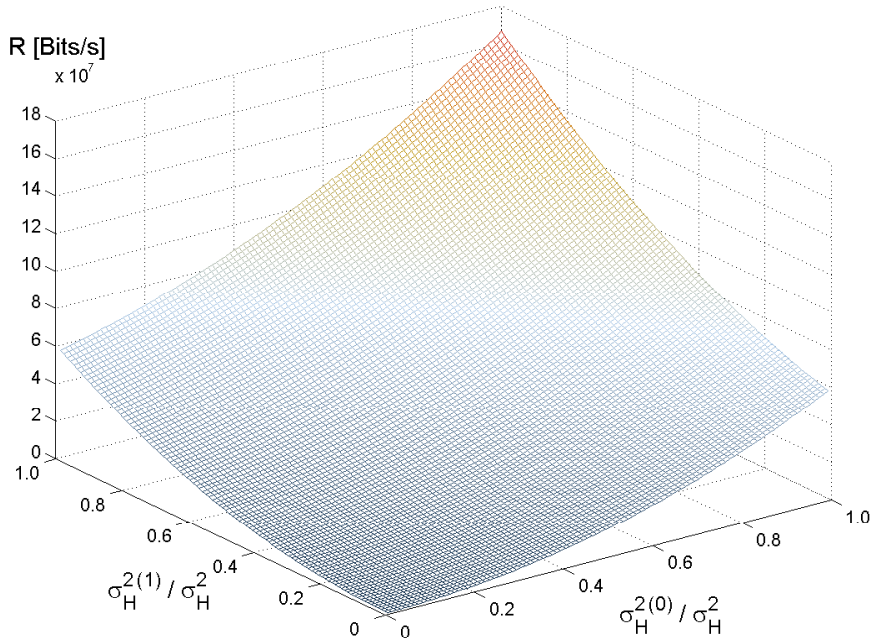

Figure 4.4: *Upper bound on sum information rate for varying path loss* $\sigma_{\mathbb{H}}^{2\,(u)}/\sigma_{\mathbb{H}}^2 = 0.1 \ldots 1$ *at a bandwidth of* $B = 500MHz$.

is

$$R_{\mathrm{UB}} = \frac{B}{TF}\left[\log\left(1+\frac{TF}{B}\frac{P\sigma^2_{\mathbb{H}_0}}{2N_0}\right)+\log\left(1+\frac{TF}{B}\frac{P\sigma^2_{\mathbb{H}_1}}{2N_0}\right)\right] - \frac{P\left(\sigma^2_{\mathbb{H}_0}+\sigma^2_{\mathbb{H}_1}\right)}{2N_0} + \frac{TF}{B}\frac{3\,P^2\left(\sigma^4_{\mathbb{H}_0}+\sigma^4_{\mathbb{H}_1}\right)}{8\,N_0^2} + \frac{TF}{B}\frac{P^2\sigma^2_{\mathbb{H}_0}\sigma^2_{\mathbb{H}_1}}{2\,U^2N_0^2}.$$

This is shown in Figure 4.4 for a bandwidth of $B = 500$MHz. Here the upper bound on information rate is shown in dependence of the users path losses which are depicted normalized to $\sigma^2_{\mathbb{H}}$ (defined in Table 4.1). For increasing path losses, e.g. $\sigma^{2\,(u)}_{\mathbb{H}}/\sigma^2_{\mathbb{H}}$ is getting small, the sum information rate is decreasing. Hence, it's intuitive to desire low path losses in order to enable higher sum rates. No trade-offs between the different user's path losses can be noticed and therefore a low path loss is preferable for all users. Note that the maximum rate (i.e. $\sigma^2_{\mathbb{H}_0} = \sigma^2_{\mathbb{H}_1} = \sigma^2_{\mathbb{H}}$) can be found in Figure 4.3.

Upper bound on sum information rate for different user power allocation

Here an overall power constraint $\sum_{u=0}^{U-1} P^{(u)} = UP$ is applied, with P defined in Table 4.1. Then, the upper bound on sum information rate in the two-user case is derived as

$$R_{\mathrm{UB}} = \frac{B}{TF}\left[\log\left(1+\frac{TF}{B}\frac{P^{(0)}\sigma^2_{\mathbb{H}}}{2N_0}\right)+\log\left(1+\frac{TF}{B}\frac{P^{(1)}\sigma^2_{\mathbb{H}}}{2N_0}\right)\right] - \frac{P\sigma^2_{\mathbb{H}}}{N_0} + \frac{TF}{B}\frac{3\left(P^{(0)^2}+P^{(1)^2}\right)\sigma^4_{\mathbb{H}}}{8\,N_0^2} + \frac{TF}{B}\frac{P^{(0)}P^{(1)}\sigma^4_{\mathbb{H}}}{4\,N_0^2},$$

where $P^{(0)}$ denotes the power allocated to user 0. Figure 4.5 shows this for a bandwidth $B = 500$MHz (cf. Figure 4.3 and Figure 4.4). For uniformly distributed power, i.e. $P^{(0)} = P^{(1)}$, the upper bound on sum information rate is minimized. A maximum of R_{UB} is obtained by allocation of the total power to one user (single user case). Hence in the case of an uncorrelated channel the single user case is desirable in order to achieve higher data rates. Again a disadvantageous effect of increasing number of users for transmission over uncorrelated channels can be approved.

Investigations on the tightness of fourthegy lower bound on $I(\mathbf{Y};\mathbf{H}|\mathbf{X})$

Since the fourthegy bounding technique can be quite loose under certain conditions, an investigation on this is presented here. To do this a Monte Carlo simulation of (4.10) and a numerical evaluation for (4.18) at $B = 10$GHz were done and compared. The Monte Carlo simulation was done for Gaussian distributed input symbols. The numerical evaluation then has to use $\mathrm{E}_X\left\{|x^{(u)}_{k,l}|^4\right\} = 2\,\sigma^4_X$.

First the influence of the number of users U on $I(\mathbf{Y};\mathbf{H}|\mathbf{X})$ was investigated for $M = 5$ and $K = 12$. In Figure 4.6(a) the dark solid curve shows the numerically evaluated fourthegy

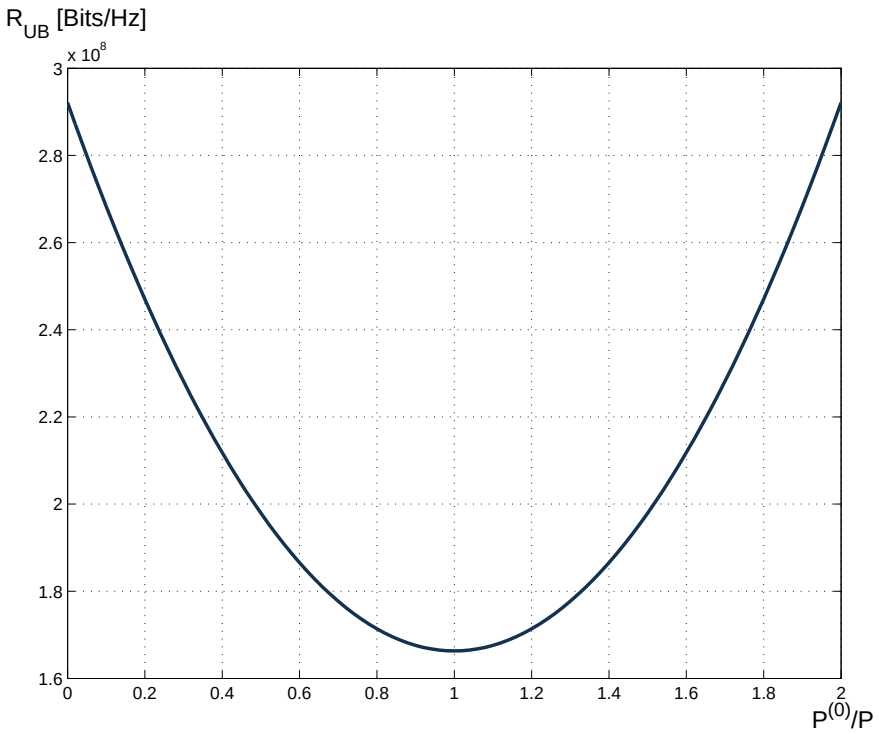

Figure 4.5: *Fourthegy upper bound on sum information rate for varying user power allocation for $U = 2$ with $P^{(0)} + P^{(1)} = 2P$ for $B = 500MHz$.*

lower bound and the bright solid curve with markers depicts the Monte Carlo simulation result. There is nearly no gap between the two curves. So the effect of the amount of users on the tightness can be neglected.

In Figure 4.6(b) the effect of increasing frame length on $I(\mathbf{Y};\mathbf{H}|\mathbf{X})$ is depicted for $K = 12$. Again the dark solid curve shows the numerically evaluated fourthegy lower bound and the bright solid curve with markers is the Monte Carlo simulation result. As there is no dependence of thigthness on U there is even no dependence on M .

Finally the effect of varying path loss on the tightness of the fourthegy lower bound is shown in 4.7. The curves are depicted in the same way as above for $M = 5$ and $K = 12$. The difference between the curves increases with increasing path loss (i.e., the SNR is getting better). Figure 4.7 shows that for a large SNR the fourthegy lower bound is less tight.

The investigation on the tightness of fourthegy lower bound on $I(\mathbf{Y};\mathbf{X}|\mathbf{H})$ will be equivalent. It is needed to develop the lower bound on sum mutual information $I(\mathbf{Y};\mathbf{X})$ and it has to be shown the same as before with change of $\mathbf{X}$ and $\mathbf{H}$. Then, the results are similar to that above.

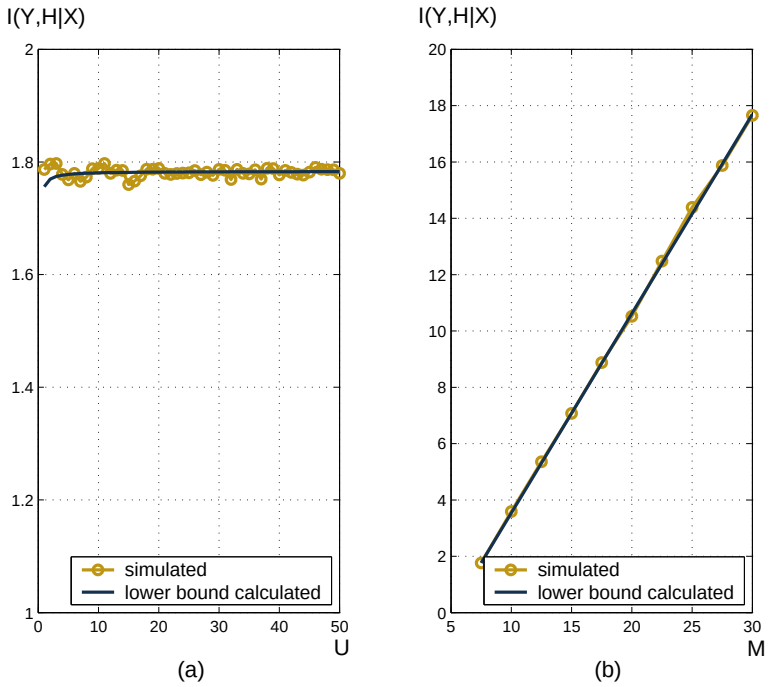

Figure 4.6: *Comparison of simulated* $I(\mathbf{Y};\mathbf{H}|\mathbf{X})$ *in* (4.10) *and fourthegy lower bound in* (4.18) *for different (a) amount of users* U *and (b) frame length* M*.*

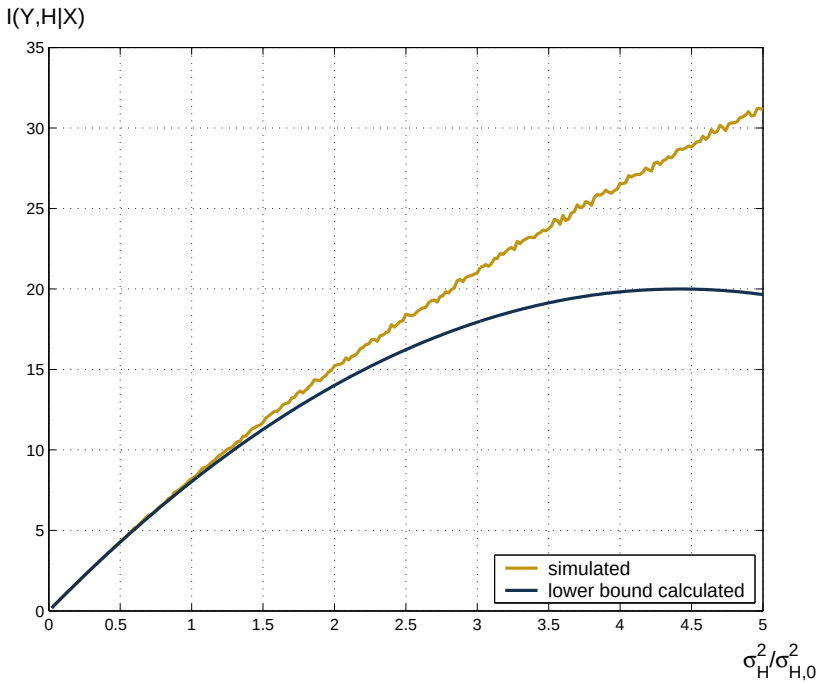

Figure 4.7: *Comparison of simulation for* $I(\mathbf{Y};\mathbf{H}|\mathbf{X})$ *in* (4.10) *and fourthegy lower bound in* (4.18) *for varying path losses (equivalently, SNR).*

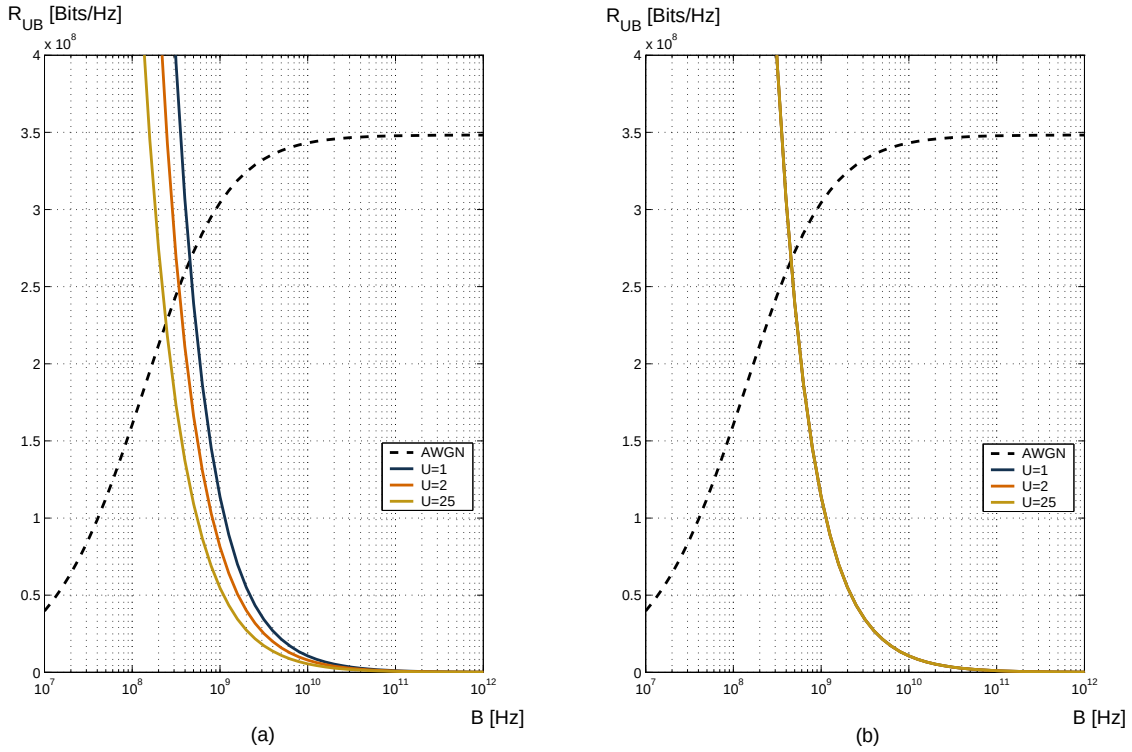

Figure 4.8: *Sum information rate for (a) MC-CDMA and (b) OFDMA over bandwidth for different amount of users (U).*

4.4.2 Upper bound on sum information rate for general assumptions on the codebook and the channel

In this subsection simulations for different channel statistics, i.e. scattering functions, channel spreads and different codebooks, i.e. MC-CDMA and OFDMA as well as spreading of symbols are presented. All simulations are based on the fourthegy upper bound on sum information rate in (4.16) (bounding technique of Bölcskei/Shamai). While in the case of uncorrelated channel coefficients the results for MC-CDMA and OFDMA are equal, for more specific channel statistics there are slight differences. For simulations again the parameters in Table 4.1 will be used. Exceptions will be made for the subcarrier spacing F because of complexity reduction reasons. The simulation of OFDMA requires much more effort than that of MC-CDMA. Hence some simulations will only be done for MC-CDMA and a reference to OFDMA will be given.

MC-CDMA and OFDMA for varying amount of users in the case of uncorrelated channel

Although more general channel statistics could be applied to (4.16) this simulation was done for an uncorrelated channel, i.e. $\mathbf{R}_H^{(u)} = \sigma_{\mathbb{H}}^2\mathbf{I}$ in order to extract the effect of MC-CDMA and OFMDA. Further a comparison with the result shown in Figure 4.3 can be done. The

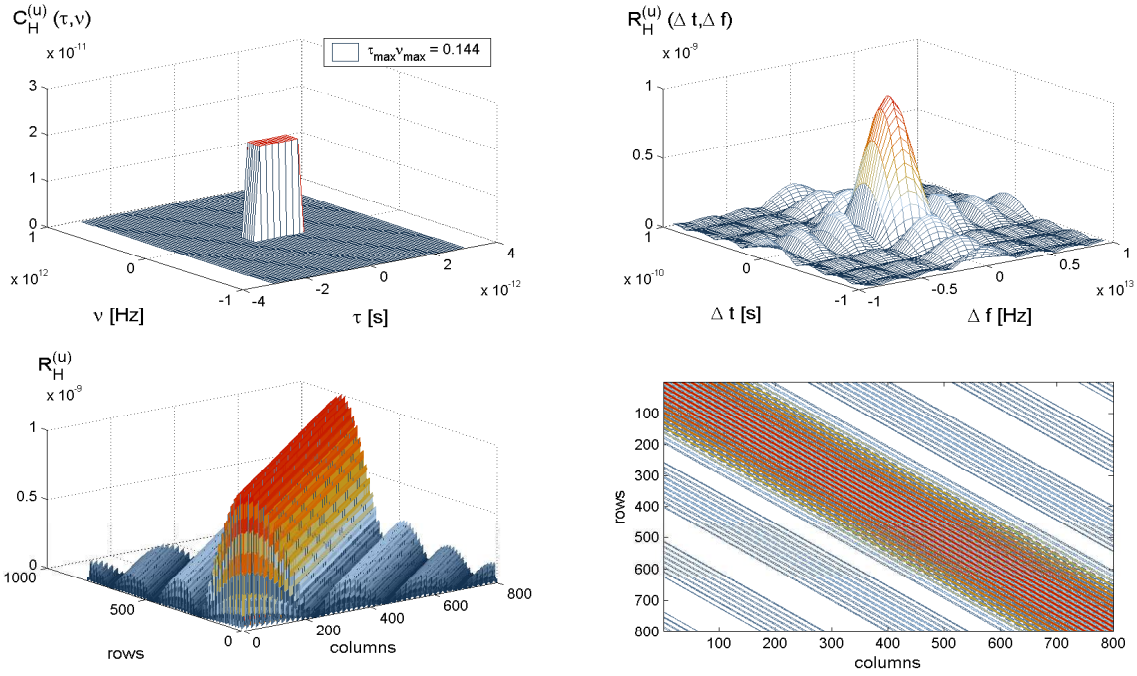

Figure 4.9: *Calculation of* $\mathbf{R}_H^{(u)}$ *from* $\mathbf{C}_{\mathbb{H}}^{(u)}(\tau, \nu)$.

parameters in Table 4.1 are used and the number of users is varying $U = 1, 2, 25$. Figure 4.8 shows the upper bound on sum information rate R_{UB} over bandwidth B for MC-CDMA and OFDMA, respectively. The dashed curve is assigned to the AWGN case, the darkest solid curve denotes the single user case, and the brightest shows the result for $U = 25$. The achievable rate region then can be found below the AWGN curve and the fourthegy upper bound R_{UB} for low bandwidths and high bandwidths, respectively. In the single user case ($U = 1$) the fourthegy upper bound will be equal for both, MC-CDMA and OFDMA. With increasing number of users the OFDMA case is beneficial since no loss of rate occurs. This is due to the separation of subchannels. For MC-CDMA an increasing number of users is equal to an increasing uncertainty on channel statistics at the receiver. In general, the upper bound in the infinite bandwidth limit tends to zero.

Investigations on a varying channel spread $\tau_{\max}\nu_{\max}$

First a short overview of the calculation of $\mathbf{R}_H^{(u)}$ from scattering function is depicted in Figure 4.9 (a). It shows the scattering function in the upper left-hand corner and the channel correlation coefficients out of a two-dimensional Fourier transformation in the upper right-hand corner. Then, the $\mathbf{R}_H^{(u)}$ matrix is composed out of this correlation coefficients (cf. 4.9 lower right-hand corner).

Now some investigations on the effect of different channel spreads on sum information rate will be performed. The channel spread $\tau_{\max}\nu_{\max}$ will be defined as the non-zero support region

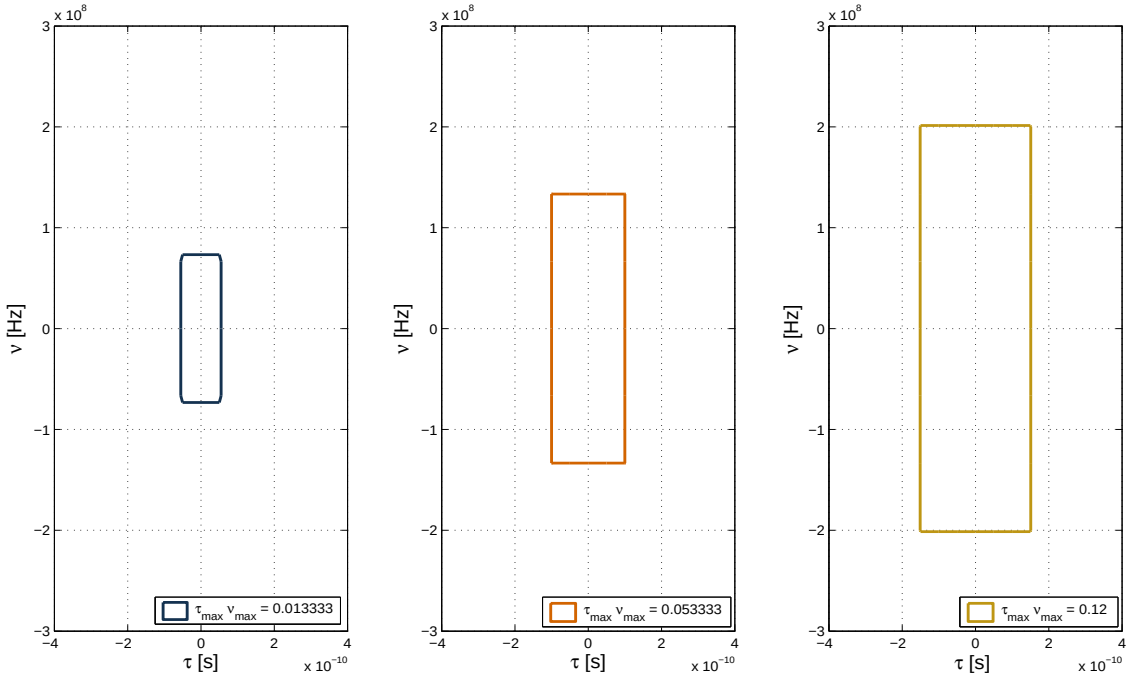

Figure 4.10: *2-dimensional map of the users scattering functions with varying channel spreads.*

of the scattering function $C_{\mathbb{H}}(\tau, \nu)$. Within this support region the scattering function is assumed flat (cf. 4.9). In Figure 4.10 different scattering functions in terms of channel spread are depicted. For these scattering functions Figure 4.11 shows the dependence of a MC-CDMA system on channel spread. The channel spread was assumed equal for all users channel (i.e., $\tau_{\max}^{(0)}\nu_{\max}^{(0)} = \tau_{\max}^{(1)}\nu_{\max}^{(1)}$, where $\tau_{\max}^{(u)}\nu_{\max}^{(u)}$ is the channel spread of user u). In Figure 4.11 the AWGN case is depicted by the dashed line, the smallest channel spread by the darkest solid line, and the biggest channel spread by the brightest solid line. Hence it can be seen, that an increasing channel spread implies lower rates. The limiting case is given by the uncorrelated channel considered previously. On the other hand a scattering function equal to a Dirac impulse at the origin will lead to an upper bound that in the ergodic limit drifts to infinity. This could be explained as a trend to the AWGN case for infinite bandwidth. Therefore for quite low channel spreads (highly correlated channels) it‘s hard to find a representative convergent result. Because of this, meaningful OFDMA results (which are situated above the MC-CDMA results) are more difficult to obtain. In Figure 4.11 the OFDMA result (not shown) would lie slightly above that of MC-CDMA.

In Figure 4.12 the upper bound on sum information rate in the two user case is depicted over different channel spreads (but equal for all users: $\tau_{\max}^{(0)}\nu_{\max}^{(0)} = \tau_{\max}^{(1)}\nu_{\max}^{(1)}$). The numerical evaluation was done for bandwidth $B = 1\text{GHz}$, $M = 100$, and $K = 8$. The bright line with the markers denotes the OFDMA result which lies above the MC-CDMA result. As above a small channel spread is necessary to enable higher rates. For low channel spreads there is an

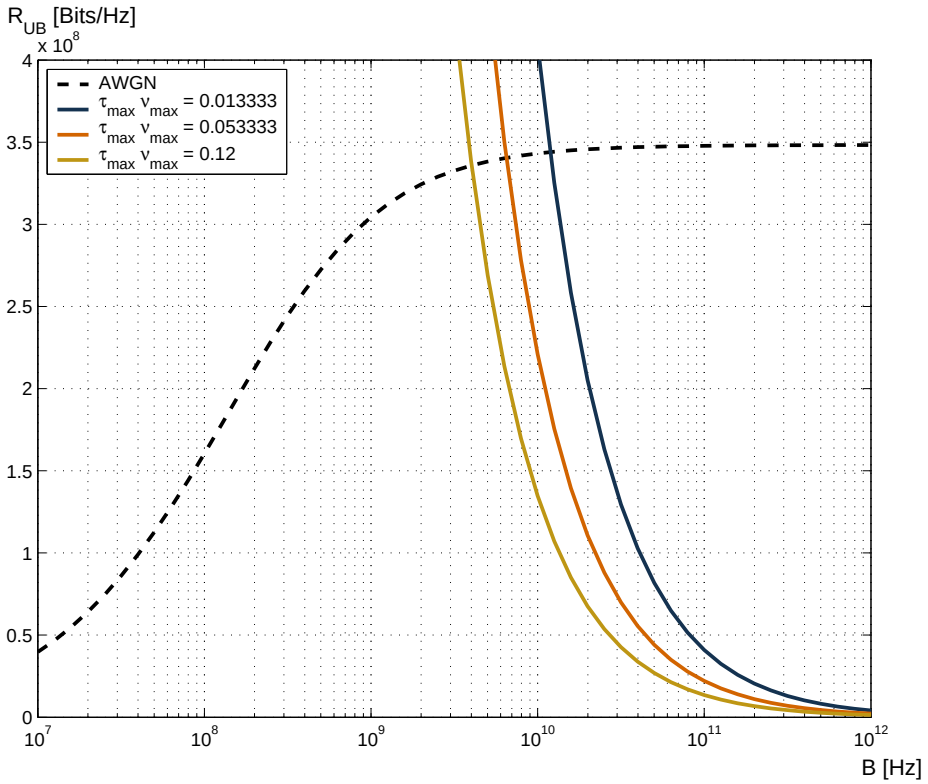

Figure 4.11: *Impact of varying channel spreads* ($\tau_{\max}^{(0)}\nu_{\max}^{(0)} = \tau_{\max}^{(1)}\nu_{\max}^{(1)}$) *on the upper bound on sum information rate for a MC-CDMA system with* $M = 60$ *and* $K = 16$.

exponential rising of rate which could be described through the beneficial effects in channel state estimation. As low channel spreads imply high correlation at the channel less channel coefficients have to be estimated. This often used equivalence to channel estimation may be useful in order to understand noncoherent channels.

Figure 4.13 again shows the upper bound on information rate for bandwidth $B = 10$GHz, $M = 80$ and $K = 4$. Here, the two users are adepted by different channel spreads but the sum will remain constant, e.g. $\tau_{\max}^{(0)}\nu_{\max}^{(0)} + \tau_{\max}^{(1)}\nu_{\max}^{(1)} = 0.1$. Again the bright line with the markers denotes the OFDMA result which lies above the MC-CDMA result (dark line). In the case of equal channel spread for both users (i.e., $\tau_{\max}^{(0)}\nu_{\max}^{(0)} = \tau_{\max}^{(1)}\nu_{\max}^{(1)} = 0.05$) the upper bound has its minimum. In the case where one user uses a channel with large channel spread while the second transmits over a channel with small channel spread there is a considerable increase of rate. Hence at least one channel with a small channel spread can lead to a higher sum information rate than two with the same sum of channel spreads. An explanation of this effect can again be obtained in terms of channel estimation. For channel estimation one highly correlated channel has more impact than even more spread channels. The comparison of MC-CDMA and OFDMA shows an advantage of OFDMA.

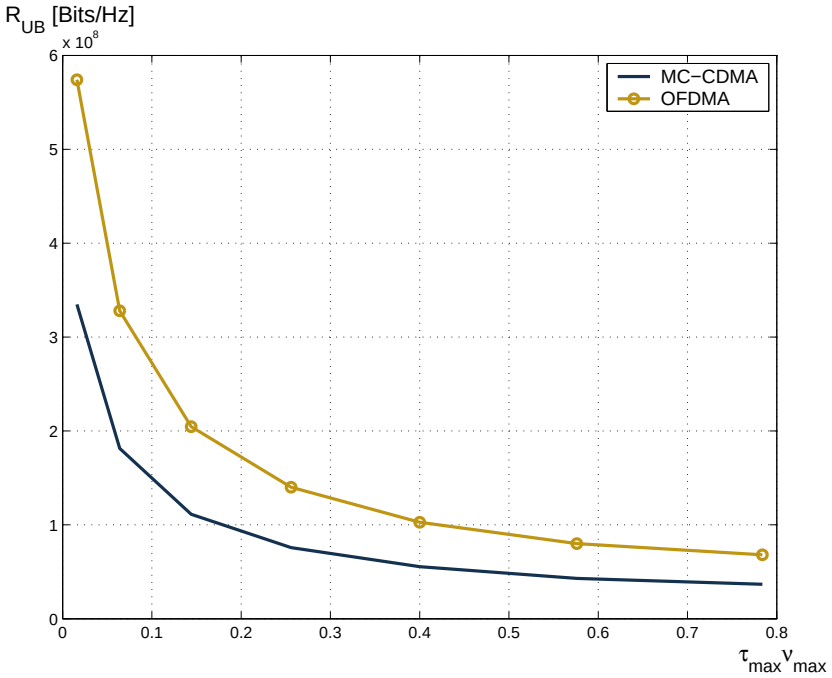

Figure 4.12: *Sum information rate over varying channel spreads for MC-CDMA and OFDMA with bandwidth $B = 1GHz$, $M = 100$ and $K = 8$. The two users have the same channel spread ($\tau_{\max}^{(0)}\nu_{\max}^{(0)} = \tau_{\max}^{(1)}\nu_{\max}^{(1)}$).*

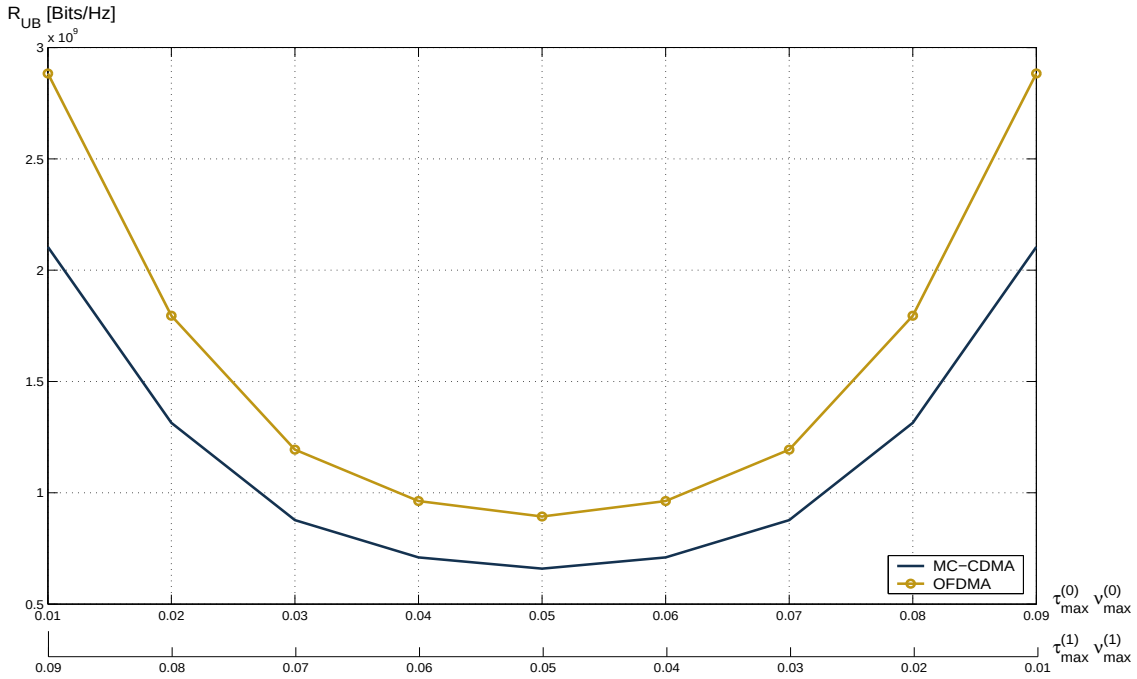

Figure 4.13: *Sum information rate over varying channel spreads for MC-CDMA and OFDMA with bandwidth $B = 10GHz$, $M = 80$ and $K = 4$. The sum of the user's channel spreads is constant, i.e. $\tau_{\max}^{(0)}\nu_{\max}^{(0)} + \tau_{\max}^{(1)}\nu_{\max}^{(1)} = 0.1$.*

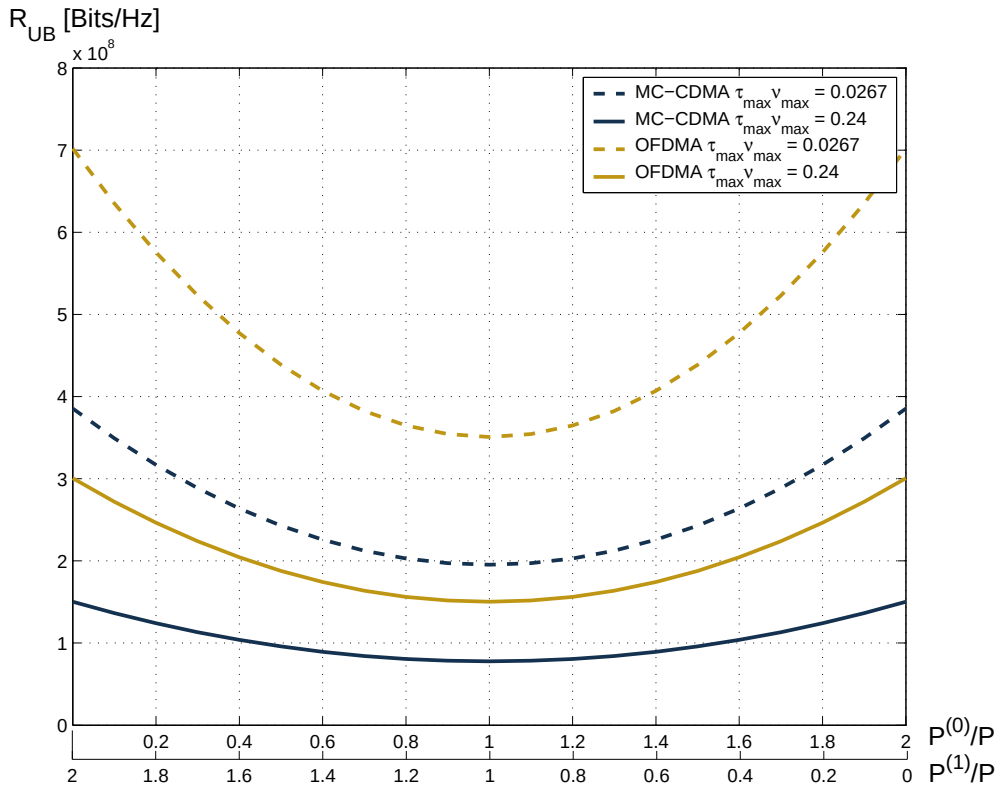

Figure 4.14: *Sum information rate for MC-CDMA and OFDMA for a varying power allocation, using* $P^{(0)} + P^{(1)} = 2P$ *with* $P = 1mW$, $M = 120$, $K = 4$ *and a bandwidth of* $B = 10GHz$.

Investigations on a different power allocation in the two-user case

For the two-user case Figure 4.14 shows the impact of varying power allocation on sum information rate. This result is similar to that shown in Figure 4.5 with additional dependence on channel spread and codebook. The results are for bandwidth $B = 10$GHz, $K = 4$, and $M = 120$ (for convergence reasons). The sum of the two users powers remains constant, i.e. $P^{(0)} + P^{(1)} = 2P$ and the channel for both is assumed equal with the same channel spread $\tau_{\max}\nu_{\max}$. In Figure 4.14 the bright curves show the result for OFDMA and the dark curves are the results for MC-CDMA. The different channel spreads are marked through dashed ($\tau_{\max}\nu_{\max} = 0.0267$) and solid ($\tau_{\max}\nu_{\max} = 0.24$) curves. The transmit powers are normalized to the transmit power P defined in Table 4.1. Then, the sum information rate maximizing case is given for a low channel spread and the single user case. It's beneficial in comparison to uniformly distributed transmit power. This again can be explained with improved conditions for channel estimation since it's easier to estimate the channel for a higher SNR. In comparison of codebooks again an advantage of OFDMA over MC-MCDMA can be recognized. In an estimation point of view the reason for this is the better time-frequency allocation at OFDM, in other words an improved peakiness.

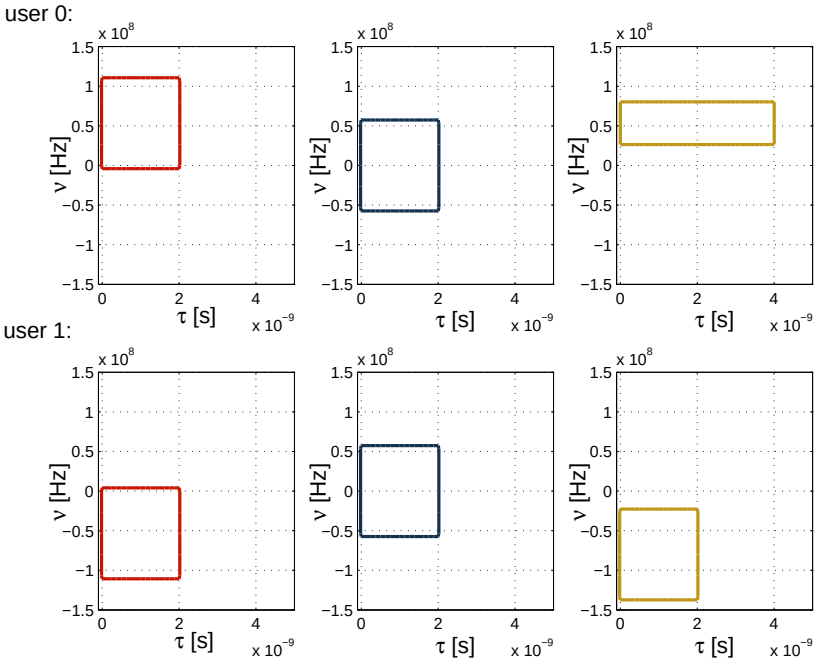

Figure 4.15: *Scenarios with different scattering functions for $U = 2$. Left: disjoint regions; center: overlapping regions; right: a more general case with disjoint regions. The channel spread is fixed at a value of $\tau_{\max}^{(0)}\nu_{\max}^{(0)} = \tau_{\max}^{(1)}\nu_{\max}^{(1)} = 0.21333$.*

Simulations for different channels

So far we treated simulations on different channel spreads or number of users. Now more general setups will be used to search for the impact of different scattering functions on the upper bound on sum information rate. These investigations will be done for the two-user case. The first three scenarios are shown in Figure 4.15. Here the support regions of the scattering functions $\mathbf{C}_{\mathbb{H}}^{(u)}(\tau, \nu)$ for user 0 and user 1 are shown above and below, respectively. The first scenario on the left hand side shows disjoint support regions for the two users. The second scenario depicts equal support regions and the third scenario on the right hand side a more general case of non-overlapping scattering functions with more multipath delay for user 0. Note that the channel spread remains constant for every user, i.e. $\tau_{\max}^{(0)}\nu_{\max}^{(0)} = \tau_{\max}^{(1)}\nu_{\max}^{(1)} = 0.21333$. Then, the upper bound on information rate over bandwidth is shown in Figure 4.16. The dashed line shows the AWGN upper bound which will be tight for small bandwidths. The dark line shows the result for scenario two where full overlapping was assumed. The brighter line shows the result for the disjoint support regions of scenario one and the brightest line depicts the more general disjoint case of scenario three. All solid curves are for MC-CDMA. Equivalently, the OFDMA results are shown through markers of the same color. Then, it can be seen that identical scattering regions are beneficial in a maximization of sum information rate sense. The more separation between the scattering regions of two user, the smaller the upper bound and hence the sum rate will be. An explanation can be given through the entire

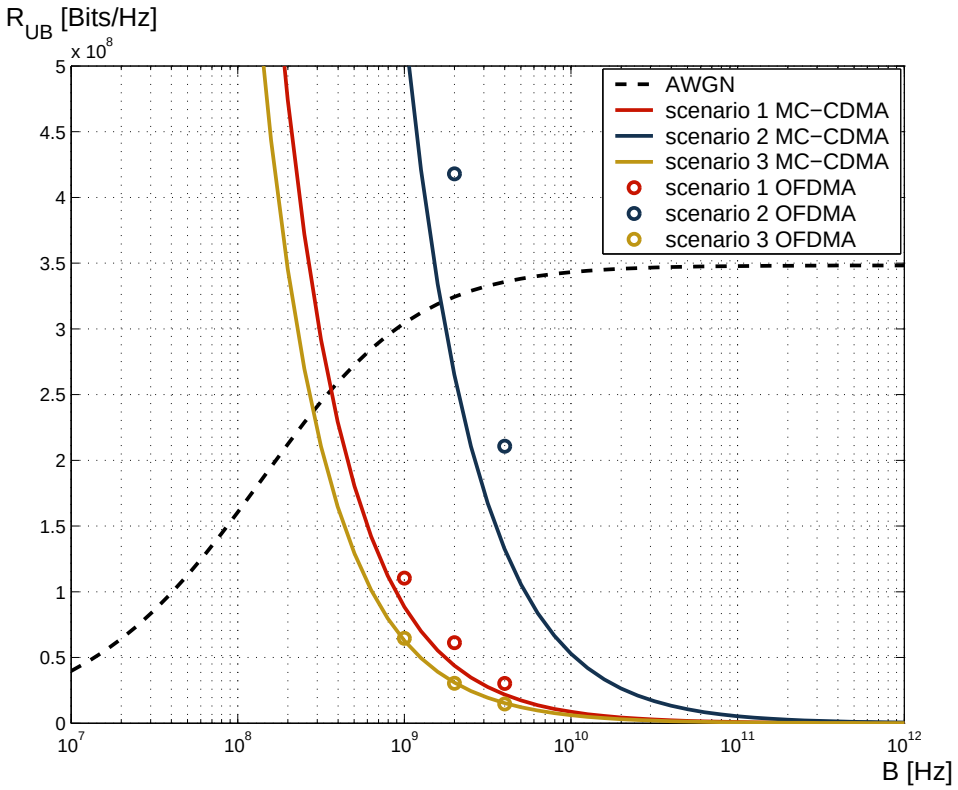

Figure 4.16: *Sum information rates for MC-CDMA and OFDMA for the simulation scenarios shown in Figure 4.15.*

time-delay region used by the users. With increasing joint support regions the upper bound on information rate decreases. In Figure 4.16 also the advantageous effect of an OFDMA system on the upper bound is shown. For a higher upper bound the difference between the MC-CDMA and the OFDMA result increases. Another information gained from this simulation concerns the location of scattering function. Hence a scattering function positioned near the origin ($\tau_{\max}\nu_{\max} = 0$) should result in a higher rate.

A second variation now is presented through the scenarios shown in Figure 4.17. Here in the two-user case the scattering function of user 1 remains constant while that of user 0 varies. Both are centered at the origin and therefore overlapped. The dark line characterises scenario one with different small channel spreads, the brighter line scenario two with equal channel spreads, and the brightest line characterises different large channel spreads. The MC-CDMA result for R_{UB} over bandwidth is shown in Figure 4.18. Here again the dashed curve depicts the upper bound through AWGN. It can be seen that smaller channel spreads for user 0 (since the channel spread of user 1 remains constant) enable higher rates. As above, one channel with low channel spread is desirable in order to achieve higher sum rate.

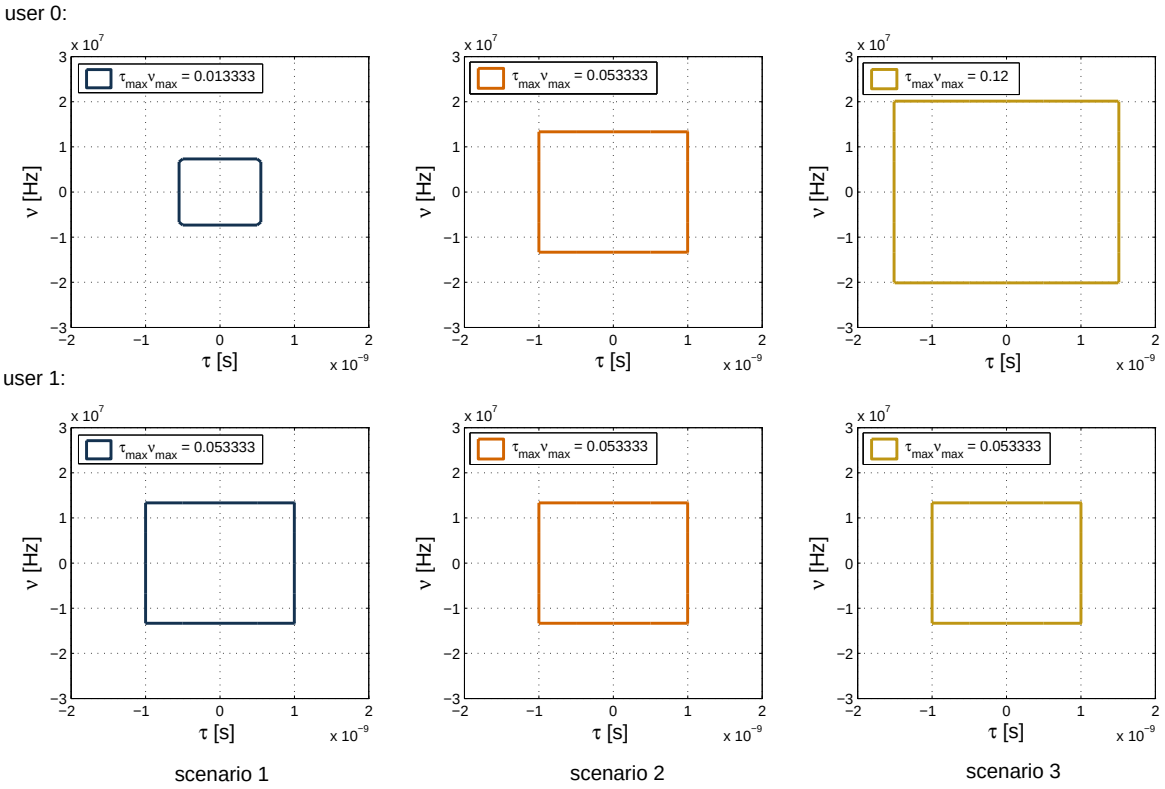

Figure 4.17: *Scenarios for 2D-scattering functions for different channels of user* 0 *(top) and user* 1 *(bottom).*

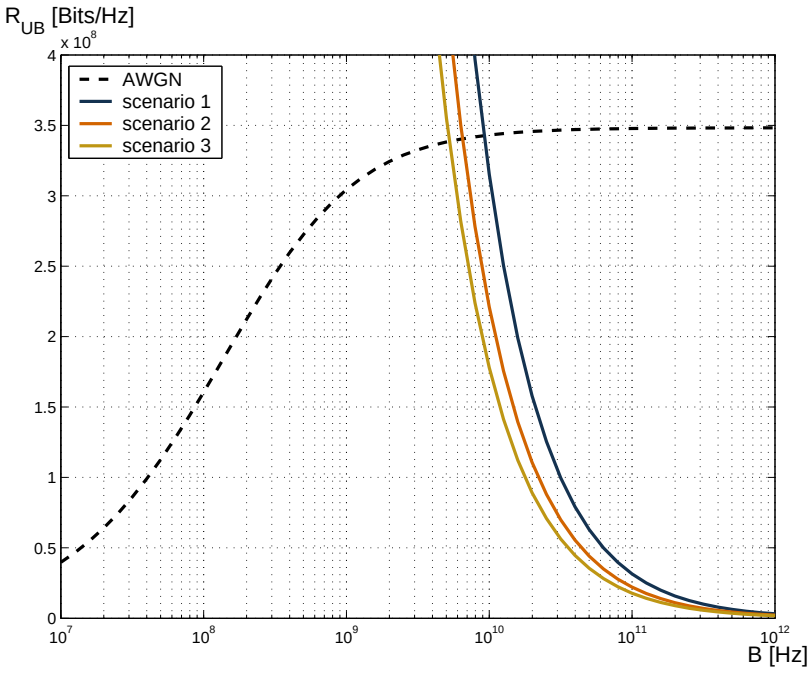

Figure 4.18: *Sum information rate for MC-CDMA for simulation scenarios in Figure 4.17.*

Chapter 5

Conclusions

In this chapter a summary of the results presented above and an outlook to further topics will be given. OFDM based multiuser systems, such as Multi-Carrier-CDMA (MC-CDMA) and orthogonal frequency division multiple access (OFDMA) transmitting over a time and frequency selective fading channel were investigated. The channel was assumed as noncoherent (no CSI available). The focus was on the uplink case where multiple users transmit to one receiver. An analysis was performed in an information theoretic context with emphasis on sum system capacity and sum information rate calculation.

- In Chapter 1 an introduction to OFDM systems was given. For a pulse-shaping OFDM system a modulator and demodulator were shown. The OFDM based multiuser communication systems MC-CDMA and OFDMA were explained. Wireless fading channels were presented in terms of input-output relation, channel statistics, and channel parameters for the continuous-time and discrete-time case. An approximate multiplicative input-output relation for small delays and Doppler spreads was shown. Further an introduction to information theoretic parameters like entropy, mutual information, and system capacity was given.

- Since the search for previous results was a main topic of this diploma thesis, an overview of known results was presented in Chapter 2. First an overview of relevant channel and signal properties was given. Next, single user results for time-frequency selective noncoherent channels were presented. Further multiuser results (including achievable rate regions) were presented.

- In Chapter 3 a multiuser system was viewed as a single user system with modified input statistics. First a known single user model was presented. Then an approximation for MC-CDMA was introduced and partial results for the calculation of sum system capacity were found. Since a closed form expression for sum system capacity couldn't be found, upper and lower bounds on sum system capacity were derived. Simulation results and numerical evaluations were presented. The simulation of the upper bound on

information rate shows that the information rate tends to zero in the infinite bandwidth limit. Further, a low channel spread is desirable in order to enable higher rates.

- In Chapter 4, bounds on sum information rate for MC-CDMA and OFDMA were presented. A multiuser model was introduced and applied to known bounding techniques combined with a fourthegy bounding technique. For the noncoherent case, closed form expressions for upper and lower bounds on sum information rate were found. Then, numerical evaluations for this fourthegy upper bound were done. The tightness of the upper bound was confirmed for low SNR. For OFDMA it was shown that for uncorrelated channels there is no dependence on the amount of users while for MC-CDMA an increasing number of users is disadvantageous. Simulations for different user power allocations show that the single user case is maximizing the sum information rate. Investigations for general channel statistics show that a low channel spread is preferable. Further, more overlapping scattering functions for different users are preferable over disjoint support regions. All results show the advantage of OFDMA over MC-CDMA in terms of the upper bound on sum rate.

Finally an outlook to further research on this topic will be given.

- This thesis was focused on sum information rate. Further investigations should address the achievable rate region in order to show trade-offs between the users (the sum information rate only characterises one boundery point of this achievable rate region).

- All models in this work are based on the uplink scenario. This results may be used to investigate the downlink scenario, e.g. one basestation is transmitting to multiple receivers (users). Downlink specific scenarios for wireless channels may be included.

- Regarding the recent interest in multiple antenna systems an extension of the results to the MIMO case could be performed.

- For the bounds presented in Chapter 4, further numerical evaluations in order to investigate the influence of peakiness could be done. The same could be done for the similarity of users input and channel statistics. Further the influence of different terms in the bounds could be extracted for various propagation scenarios.

Bibliography

[1] R. W. Chang, "Synthesis of band-limited orthogonal signals for multi-channel data transmission," *Bell Syst. Tech. J.*, vol. 45, pp. 1775–1796, Dec. 1966.

[2] S. B. Weinstein and P. M. Ebert, "Data transmission by frequency division multiplexing using the discrete Fourier transform," *IEEE Trans. Comm. Technol.*, vol. 19, pp. 628–634, Oct. 1971.

[3] L. J. Cimini, "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *IEEE Trans. Comm.*, vol. 33, pp. 665–675, July 1985.

[4] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Comm. Mag.*, vol. 28, pp. 5–14, May 1990.

[5] IEEE P802 LAN/MAN Committee, "The working group for wireless local area networks (WLANs)." http://grouper.ieee.org/groups/802/11/index.html.

[6] M. Radimirsch and V. Vollmer, "HIPERLAN type 2 standardisation — An overview," in *Proc. European Wireless '99, ITG-Fachtagung*, (Munich, Germany), pp. 139–144, Oct. 1999.

[7] ETSI, "Digital video broadcasting (DVB); framing structure, channel coding and modulation for digital terrestrial television." Draft EN 300 744, V1.2.1, 1999. http://www.etsi.org.

[8] ETSI, "Digital audio broadcasting (DAB) to mobile, portable and fixed receivers." ETS 300 401, 1995. http://www.etsi.org.

[9] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecomm.*, vol. 10, pp. 585–596, Nov. 1999.

[10] H. Bölcskei, "Principles of MIMO-OFDM wireless systems," in *Handbook on Signal Processing for Communications* (M. Ibnkahla, ed.), Boca Raton, FL: CRC Press, to appear.

[11] G. R. Aiello and G. D. Rogerson, "Ultra-wideband wireless systems," *IEEE Microwave Magazine*, vol. 4, pp. 36–47, June 2003.

[12] D. Schafhuber, *Wireless OFDM Systems: Channel Prediction and System Capacity*. PhD thesis, Vienna University of Technology, March 2004, (online at www.nt.tuwien.ac.at/dspgroup/dschafhu/public.htm).

[13] B. Muquet, Z. Wang, G. B. Giannakis, M. de Courville, and P. Duhamel, "Cyclic prefixing or zero padding for wireless multicarrier transmissions?," *IEEE Trans. Comm.*, vol. 50, pp. 2136–2148, Dec. 2002.

[14] IEEE P802 LAN/MAN Committee, "The working group on broadband wireless access standards." http://grouper.ieee.org/groups/802/16/index.html.

[15] J. L. N. Yee and G. Fettweis, "Multicarrier CDMA in indoor wireless radio networks," in *PIMRC93*, (Yokohama, Japan), pp. 109–113, Sept 1993.

[16] A. B. A. Chouly and S. Jourdan, "Orthogonal multicarrier techniques applied to direct sequence spread spectrum CDMA systems," *Proceedings of the IEEE Global Telecommunications Conference 1993*, pp. 1723–1728, 1993.

[17] K. Fazel and L. Papke, "On the performance of convolutionally-coded CDMA/OFDM for mobile communication system," in *PIMRC93*, (Yokohama, Japan), pp. 468–472, Sept. 1993.

[18] V. M. DaSilva and E. S. Sousa, "Performance of orthogonal CDMA codes for quasisynchronous communication systems," in *Proceedings of IEEE ICUPC 1993*, (Ottawa, Canada), pp. 995–999, Oct. 1993.

[19] L. Vandendorpe, "Multitone direct sequence CDMA system in an indoor wireless environment," in *Proceedings of IEEE SCVT 1993*, (Delft, The Netherlands), pp. 4.1:1–8, Oct. 1993.

[20] L. Hanzo, M. Münster, B. J. Choi, and T. Keller, *OFDM and MC-CDMA for Broadband Multi-User Communications, WLANs and Broadcasting*. IEEE Press and Wiley, 2003.

[21] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 3rd ed., 1995.

[22] J. D. Parsons, *The Mobile Radio Propagation Channel*. London: Pentech Press, 1992.

[23] W. C. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.

[24] R. S. Kennedy, *Fading Dispersive Communication Channels*. New York: Wiley, 1969.

[25] G. Matz and F. Hlawatsch, "Time-frequency characterization of random time-varying channels," in *Time-Frequency Signal Analysis and Processing: A Comprehensive Reference* (B. Boashash, ed.), ch. 9.5, pp. 410–419, Oxford (UK): Elsevier, 2003.

[26] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Comm. Syst.*, vol. 11, pp. 360–393, 1963.

[27] W. Kozek and A. F. Molisch, "Nonorthogonal pulseshapes for multicarrier communications in doubly dispersive channels," *IEEE J. Sel. Areas Comm.*, vol. 16, pp. 1579–1589, Oct. 1998.

[28] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[29] W. Rhee and J. M. Cioffi, "On the capacity of multiuser wireless channels with multiple antennas," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2580–2595, Oct. 2003.

[30] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.

[31] H. Sampath, S. Talwar, J. Tellado, V. Erceg, and A. Paulraj, "A fourth-generation MIMO-OFDM broadband wireless system: Design, performance, and field trial results," *IEEE Comm. Mag.*, vol. 40, pp. 143–149, Sept. 2002.

[32] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inf. Theory*, vol. 44, pp. 2619–2692, Oct. 1998.

[33] A. Paulraj, R. U. Nabar, and D. Gore, *Introduction to space-time wireless communications*. Cambridge (UK): Cambridge Univ. Press, 2003.

[34] T. L. Marzetta and B. M. Hochwald, "Fundamental limitations on multiple-antenna wireless links in Rayleigh fading," in *Proc. IEEE Int. Symp. Information Theory (ISIT 98)*, (Cambridge, MA), p. 310, Aug. 1998.

[35] I. Jacobs, "The asymptotic behavior of incoherent m-ary communication systems," *Proc. IEEE*, vol. 51, pp. 251–252, Jan. 1963.

[36] J. R. Pierce, "Ultimate performance of m-ary transmissions on fading channels," *IEEE Trans. Inf. Theory*, vol. 12, pp. 2–5, Jan. 1966.

[37] I. Telatar and D. Tse, "Capacity and mutual information of wideband multipath fading channels," *IEEE Trans. Inf. Theory*, vol. 46, pp. 1384–1400, July 2000.

[38] H. Bölcskei and S. Shamai (Shitz), "Peaky time-frequency signaling over underspread fading channels." in preparation.

[39] M. Médard and R. G. Gallager, "Bandwidth scaling for fading multipath channels," *IEEE Trans. Inf. Theory*, vol. 48, pp. 840–852, April 2002.

[40] V. G. Subramanian and B. Hajek, "Broad-band fading channels: Signal burstiness and capacity," *IEEE Trans. Inf. Theory*, vol. 48, pp. 809–827, April 2002.

[41] R. G. Gallager, "An inequality on the capacity region of multiaccess fading channels," in *Communication and Cryptography—Two Sides of One Tapestry*, pp. 129–139, Boston, MA: Kluwer, 1994.

[42] S. Shamai (Shitz) and T. L. Marzetta, "Multiuser capacity in block fading with no channel state information," *IEEE Trans. Inf. Theory*, vol. 48, pp. 938–942, April 2002.

[43] S. Visuri and H. Bölcskei, "MIMO-OFDM multiple access with variable amount of collision," in *Proc. IEEE International Conf. on Communications (ICC), Paris, France*, pp. 286–291, June 2004.

[44] S. Visuri and H. Bölcskei, "On multiple accessing for frequency-selective MIMO channels," in *European Signal Processing Conference (EUSIPCO)*, pp. 523–527, Sept. 2004.

[45] H. Weinrichter and F. Hlawatsch, *Stochastische Grundlagen nachrichtentechnischer Signale*. Wien: Springer, 1991.

[46] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill, 1984.

[47] K. S. Miller, *Multidimensional Gaussian Distributions*. New York, US: John Wiley and Sons, Inc., 1964.

[48] L. Isserlis, "On a formula for the product-moment coefficient in any number of variables," *Biometrika*, vol. 12, pp. 134–139, 1918.

[49] I. S. Reed, "On a moment theorem for complex Gaussian processes," *IEEE Trans. Inf. Theory*, vol. 8, no. 3, pp. 194–195, 1962.

[50] J. W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits and Systems*, vol. CAS-25, pp. 772–781, Sept. 1978.