Martin Stemick

Contributions to Resource Allocation Schemes in Multiuser OFDM-FDMA-Systems

Contributions to

Resource Allocation Schemes

in Multiuser

OFDM-FDMA-Systems

**Vom Promotionsausschuss der
Technischen Universität Hamburg-Harburg
zur Erlangung des akademischen Grades
Doktor-Ingenieur
genehmigte Dissertation**

**von
Martin Stemick**

**aus
Uelzen**

2009

Erster Gutachter:	Prof. Dr. Hermann Rohling
Zweiter Gutachter:	Prof. Dr.-Ing. habil. Udo Zölzer
Tag der mündlichen Prüfung:	23. November 2009

Acknowledgments

Table of Contents

1 Introduction and Scope

In almost all contemporary societies, mobile wireless devices have become a ubiquitous means of communication. During a time span of two decades, wireless services developed from mere telephony to broadband applications like multimedia web browsing and video data streaming. To keep pace with the increasing data rate demands of these broadband applications, the applied transmission techniques have to use a wide frequency band. In conventional *single-carrier* (SC) transmission techniques, broadband transmission goes along with very short transmit symbol durations. Additionally, in mobile wireless channels the transmit signal propagates through multiple paths with differing propagation delays. Therefore, a superposition of multiple delayed transmit signals is observed at the receiver side of an SC transmission system. This effect is known as *intersymbol interference* (ISI). In order to recover the original transmit signal, the SC receiver must spend a huge effort for signal equalization. The necessary computational complexity reaches prohibitive dimensions especially if high data rates are involved.

A much more efficient approach to cope with ISI effects in broadband transmission is the *Orthogonal Frequency Division Multiplexing* (OFDM) technique. Its basic idea is to split the channel bandwidth into multiple narrow and orthogonal frequency bands: so-called subcarriers. Due to the narrowband character of the individual subcarriers, the duration of the corresponding transmit symbols is larger by orders of magnitude than in a broadband SC system. If the symbol duration is chosen much larger than the maximum propagation delay of the channel and additionally a so-called *guard interval* (GI) is inserted after each symbol, ISI effects can be completely avoided. This dramatically reduces the equalization effort at the receiver side. Additionally, due to the utilization of multiple subcarriers in parallel, OFDM achieves very high data rates that are sufficient for all kinds of broadband applications.

Because of its excellent performance in mobile broadband radio channels, OFDM is the solely considered transmission technique of this thesis.

Another advantage of OFDM is that it allows a straightforward multiple access technique for multiuser systems: *frequency division multiple access* (FDMA). In conjunction with OFDM, this technique is referred to as OFDM-FDMA, which uses the individual subcarriers of the OFDM system as independent resources in the frequency domain. Thus, multiple wireless connections can be established simultaneously using exclusive sets of subcarriers. Also, the flexible allocation of varying numbers of subcarriers to different users inside the communication system is possible. This is of importance, since in modern communication applications the required data rates vary strongly over time and from user to user.

OFDM-FDMA also offers the possibility of *link adaption*, i. e. the adaption of the transmission mode to the current channel quality. The considered multipath propagation of signals

through the radio channel leads to strong frequency selectivity over the complete channel bandwidth. In OFDM, this translates into an individual complex fading factor for each subcarrier. In case of a single user inside the system, this effect can be utilized for *bit loading*: Each subcarrier is modulated with an individual number of bits depending on the magnitude of its fading factor.

In multiuser systems, an additional means of link adaption is available: the so-called *subcarrier selection*. Since the channel situation of each user inside the system is highly individual, two different users mostly observe a totally different fading factor on one and the same subcarrier. This so-called *multiuser diversity* (MUD) can be exploited by a channel adaptive subcarrier selection, where each user solely employs his least faded subcarriers for transmission.

The task of selecting suitable subcarriers considering channel properties and data rate demands of multiple users is referred to as *adaptive resource allocation* in the following. Due to its inherent complexity, the efficient completion of this task is an open issue in communication systems.

Therefore, the field of adaptive resource allocation in multiuser OFDM-FDMA systems is studied thoroughly in this thesis. It will be shown, that the exploitation of channel knowledge by means of adaptive resource allocation leads to large gains in system performance. In particular, optimum and heuristic allocation approaches are applied to solve selected resource allocation problems. Eventually, heuristic approaches prove to be much more computationally efficient than optimized approaches while providing an almost identical performance level.

Another issue in this respect is to provide each user inside the system with a fair share of resources, even though the individual users might observe very disproportionate channel qualities. This system aspect is referred to as *fairness* and is discussed in detail in this thesis. The above mentioned resource allocation approaches offer various possibilities to make the fairness aspect an integral part of the allocation process. These possibilities include the addition of constraints to the resource allocation or to modify whole parts of the process. The performance of the covered resource allocation approaches considering fairness is evaluated in detail identifying individual advantages and drawbacks.

This thesis is organized as follows:

Due to the vital importance of the radio channel for the design of a communication system, the basic concepts of channel influences and their modeling for system simulations are introduced in chapter 2. Based on these concepts, the fundamentals of the OFDM transmission technique are discussed in chapter 3 followed by an introduction to multiple access techniques suitable for OFDM in chapter 4. After that, chapter 5 gives an overview of the system and channel parameters used to model the communication system discussed in this thesis. Chapter 6 considers the OFDM-FDMA multiple access technique and gives basic performance figures for adaptive and non-adaptive subcarrier allocation approaches known from literature. Chapter 7 transfers these results to the more realistic system model used in this thesis, where multiple users are assumed to be uniformly distributed inside a radio cell.

The following chapters 8 to 9 discuss the research results found during this work with focuses on effective subcarrier allocation approaches for OFDM-FDMA on the one hand and fair sharing of cell capacity between users on the other hand. Regarding subcarrier allocation approaches, chapter 8 gives a performance comparison between optimized channel-adaptive subcarrier allocation approaches and heuristic approaches. Based on the results found, the issue of fairness in a multiuser system is discussed introducing various approaches to improve the fairness level and to adjust the throughput-fairness-balance inside the communication system. Novel subcarrier allocation procedures and their possible modifications are introduced in this context. Results based on these discussions are given in chapter 9 together with a thorough performance comparison between all considered approaches. Chapter 10 concludes this work with a summary.

2 Mobile Radio Channel

Communication over wireless channels is used since the beginning of the 20th century [Wea13]. The great advantage of wireless over wireline communication lies in the mobility it provides for the user. In principle, this enables message and data exchange at any location and over long distances. In order to exploit these benefits, the designer of a wireless communication system has to be well aware of the propagation properties inside the present radio channel. *To put it another way, it is the channel that actually designs the system, not the system designer.*

In wireline channels, the signals propagate along a fixed dimension and therefore experience only moderate losses and disturbances. In contrast to this, in mobile radio channels the transmitted signals propagate through space, which leads to a rapid loss of signal power and gives rise to multiple signal reflections.

In this thesis, the considered direction of transmission is the *downlink* from a fixed *Base Station* (BS) to a *Mobile Terminal* (MT), cf. Fig. 1. Due to the reciprocity of the radio channel, the following analysis of the downlink can also be applied to the *uplink* direction. The transmission area covered by the BS is referred to as a cell.
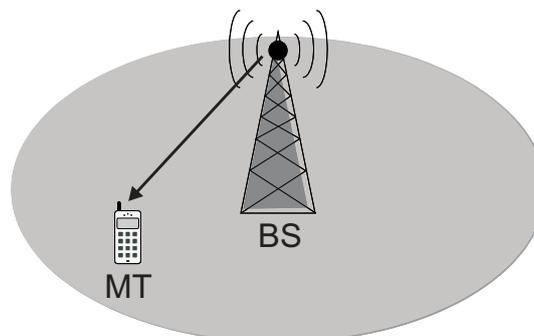
Fig. 1: Wireless transmission in downlink direction

In order to enable an error-free transmission link between an MT and its BS a detailed understanding of the signal propagation through the channel is necessary. Hence, the main properties of a mobile radio channel are introduced in this chapter.

Mobile radio channels are mostly described by division into three parts: *Large*, *Medium* and *Small Scale Effects*. First are the so-called *Large Scale Effects,* which apply to the average power of the transmit signal and describe effects that occur due to the general geometry of the propagation path. Therefore, large scale effects only change gradually with the (large scale)

movement of the MT. The large scale effects are also often referred to as *Path Loss* and are assumed to show no frequency selectivity [Tra04].

Medium scale effects describe the influence of obstacles in the local area around the MT. These effects change when the MT moves over distances, which are in the order of some tens or hundreds of meters. These effects are also known as *Shadowing* and are assumed to show no frequency selectivity.

The third part considers so-called *Small Scale Effects,* which stem from multiple reflections of the transmit signal at various objects. Hence, these effects are also referred to as *Multipath Propagation*. At the receiver the reflected and delayed versions of the original transmit signal combine, which leads to constructive or destructive interference. The instantaneous interference situation depends on amplitude and phase of the individual reflections and therefore will change rapidly upon movements of the MT in the order of a signal wavelength. For the same reasons, multipath fading causes frequency selectivity if the duration of a transmit symbol is in the order of the maximum path delay.

Since large, medium, and small scale effects are based on different physical phenomena, they can be described and modeled independently. The following sections will give an introduction to these classes of channel effects with a deterministic approach on the one hand and with a stochastic approach on the other hand. The analysis will take place in the baseband.

Additionally, a brief introduction to the channel models and simulation methodologies used in this thesis is given. The intention is to enable the reader to interpret the performance results of the considered transmission systems given in later chapters.

2.1 Path Loss and Shadowing

Path loss and shadowing effects describe the fading of the transmit signal when it propagates through free space, the atmosphere or absorbent materials. The attenuation thus inflicted on a signal varies when the user moves over distances which are significantly larger than the signal wavelength, therefore also the terms *large scale-* as well as *medium scale effects* are found in literature, cf. [Lin91]. These effects show only minor time-variations, especially if the user moves at low speeds. The dependency on the frequency is also very moderate. *Thus, fading caused by path loss or shadowing is assumed to be constant in the frequency- and time-domain throughout this thesis*. Hence, the influence of these effects can be described by constant factors G^{PL} and G^{SH} on the transmit power P_{TX} :

$$P_{RX} = G^{PL} G^{SH} P_{TX}. \qquad (2.1)$$

In (2.1), P_{RX} represents the receive power at the mobile terminal. Both effects will be introduced in the following, starting with path loss.

The main contribution to path loss stems from signal propagation through free space, since the transmit power P_{TX} disperses over a spherical area which is proportional to the square of the distance d as it propagates through space. If furthermore omnidirectional antennas at the MT

and the BS are assumed, the path loss solely depends on the relative distance d between MT and BS. This leads to the following definition of the path loss factor G^{PL}:

$$G^{PL}(d) = G_0 \left(\frac{d}{d_0} \right)^{-\alpha}$$ (2.2)

where the path loss exponent α has a value of $\alpha = 2$ in free space. For typical mobile channels, additional path loss contributions stem from diffraction losses at obstacles and ground-wave losses due to reflections at the earth's surface. To account for these effects, the parameter α can be adjusted in the range $\alpha \in [2,4]$. The reference factor G_0 describes the loss at the reference distance d_0, which is defined in this work to be the maximum distance d_{max} between MT and BS. The general geometry inside a cell is shown in Fig. 2.
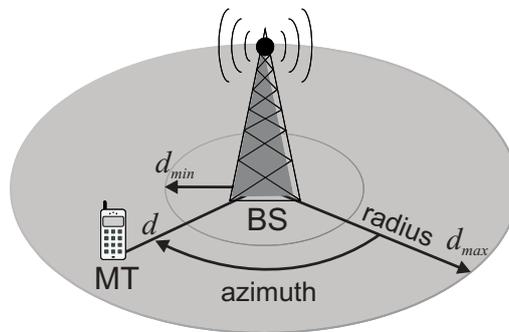
Fig. 2: Determination of cell dimension and MT position

Since multiuser communication systems are covered in this thesis, it is generally assumed that inside a cell multiple MTs are situated around the BS. To evaluate the effects of path loss on the transmission links between MTs and BS, various cellular scenarios are considered. Each cellular scenario is distinguished by the spatial distribution of the MTs inside the cell. The positions of the individual MTs can be modeled as a random variable with a characteristic distribution. Based on this assumption, the *Probability Density Function* (PDF) $p(d)$ of the distance d between BS and MT can be derived. In the following, a uniformly distributed azimuth between the connecting line from BS to each MT and the radius vector of the cell is assumed, as seen in Fig. 2.

In this case, the PDF (2.3) of distance d between MT and BS given below represents the first considered cellular scenario where the MTs are *distributed uniformly over the cell area* [Tra04]. In the given formula, d_{min} and d_{max} represent the minimum and maximum distance between MT and BS, respectively.

$$p(d) = \frac{2d}{d_{max}^2 - d_{min}^2} \quad d_{min} \le d \le d_{max}$$ (2.3)

Because the path loss factor G^{PL} between each MT and the BS strongly depends on the corresponding distance d between both, the spatial distribution of MTs over the cell area influences the range of path loss factors G^{PL} observed inside the cell. Thus, given the PDF $p(d)$ in conjunction with (2.2) determines the PDF $p(G^{PL})$ of the path loss factor. There, $G_{\min}$ and $G_{\max}$ represent the path loss factors at the minimum and the maximum distance ($d_{\min}$ and $d_{\max}$) to the BS, respectively.

$$p(G^{PL}) = \frac{\left(G^{PL}\right)^{-1-\frac{2}{\alpha}}}{\left(G_{\min}\right)^{-\frac{2}{\alpha}} - \left(G_{\max}\right)^{-\frac{2}{\alpha}}} \quad G_{\min} \leq G^{PL} \leq G_{\max} \tag{2.4}$$

Another possible cell scenario is to place all MTs at the same distance d to the BS, which corresponds to a *circular distribution of MTs around the BS*. In this case all MTs have an identical path loss factor G^{PL}.

Both cell scenarios introduced above - the uniform as well as the circular distribution - are used as system models in this thesis.

After this introduction to the path loss effect, the focus will now be turned to shadowing. In the following, the shadowing effect is quantified by the power factor G^{SH}. The physical cause for this effect is the power loss by propagation through obstacles like e. g. walls. Every obstacle can be modeled by an individual loss factor G_m, which leads to the overall shadowing loss factor G^{SH} between a user and a BS as shown below:

$$G^{SH} = \prod_{m=1}^{M} G_m \tag{2.5}$$

On a logarithmic scale, the loss factor $G^{SH}_{(dB)}$ is expressed by summation as in (2.6). If the number of obstacles M is sufficiently large, the shadowing $G^{SH}_{(dB)}$ can be modeled by a Gaussian random variable with the normal PDF $p\left(G^{SH}_{(dB)}\right)$, see (2.7). In equation (2.7), the parameter σ_{dB} describes the standard deviation of $G^{SH}_{(dB)}$ while μ_{dB} represents the corresponding expectation value. Both parameters σ_{dB} and μ_{dB} are also expressed on a dB-scale.

$$G^{SH}_{(dB)} = 10 \cdot \log_{10}(G^{SH}) = \sum_{m=1}^{M} 10 \cdot \log_{10}(G_m) = \sum_{m=1}^{M} G_{m(dB)} \tag{2.6}$$

$$p\left(G^{SH}_{(dB)}\right) = \frac{1}{\sqrt{2\pi}\sigma_{dB}} \exp\left(-\frac{\left(G^{SH}_{(dB)} - \mu_{dB}\right)^2}{2\sigma_{dB}^2}\right) \tag{2.7}$$

The values considered for the parameter σ_{dB} range from 4dB to 12dB. Generally, the standard deviation σ_{dB} increases, if a larger area around the MT is considered to evaluate the shadowing effects. This is due to the larger variation of obstacle positions in a wide area.

The choice of the value for μ_{dB} in (2.7) also deserves some attention: In order to avoid a constant offset to the receive power P_{RX}, the expectation value of the linear shadowing factor G^{SH} should be defined as $E\{G^{SH}\}=1$, cf. (2.1). This condition leads to the definition of the corresponding logarithmic expectation value μ_{dB} as shown in (2.8).

$$\mu_{dB} = -\sigma_{dB}^2 \frac{\ln(10)}{20} \tag{2.8}$$

The implementation of the factor G^{SH} as a random variable in computer simulations can be simplified, if $\mu_{dB}=0$ is assumed. This causes a constant power offset in (2.1), which can later be corrected by a simple subtraction.

2.2   Multipath Propagation

So far solely the channel influence due to general wave propagation was considered. The corresponding effects path loss and shadowing have no other influence on the transmission link than limiting the maximum distance in which the transmit signal can be correctly received. These effects can be mitigated by increasing the transmit power P_{TX} or using a more sensitive receiver.

But there is another channel influence, which has a much more disturbing effect on the transmission link and thus will be described in extensive detail. This channel influence is known as *multipath propagation*. It is caused by the fact that a transmit signal not necessarily reaches the receiver over a straight path. In contrast, the transmit signal can be reflected at conducting surfaces of stationary or moving obstacles and thus reaches the receiver multiple times via different paths. Such a situation is depicted in Fig. 3. The reflecting objects are also referred to as *scatterers*. Each reflection leads to an individual attenuation and phase shift of the transmit signal, depending on the material and shape of the reflecting scatterer. Additionally, the length of each path causes the signal to be delayed. Thus, multiple attenuated and delayed versions of the original transmit signal superimpose at the receiver. The multipath channel can therefore be described as a superposition of signal paths with individual delay, attenuation and phase. In addition to these indirect propagation paths, a direct *line-of-sight* (LOS) path may also exist.
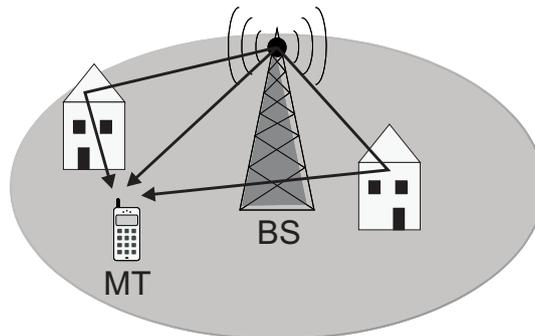
Fig. 3: Signal reflections causing multipath propagation

One disturbing influence of multipath propagation on the transmit signal is that the superposition of delayed signals at the receiver leads to either destructive or constructive interference. It will be shown in the following that this interference not only may change rapidly over time but additionally shows frequency selectivity.

Since the transmit signal consists of a sequence of data symbols, their delayed superposition at the receiver also causes *intersymbol interference* (ISI), i.e. the overlapping of adjacent symbols. Especially this effect can cause a serious performance drawback for broadband transmission systems: Since their symbol duration is short compared to the channel delays, this causes each symbol to be interfered by a multitude of preceding symbols.

In contrast to previously discussed channel effects, multipath propagation is not described by a mere factor, but by an impulse response $h(t)$ and a transfer function $H(f)$ respectively. In the following sections an analytical model for multipath propagation is introduced. It is shown, that the multipath channel can be described by a linear time-variant (LTV) or a linear-time-invariant (LTI) system, respectively. The latter analysis is much simpler and therefore introduced first. It is based on the assumption that scatterers and the MTs are stationary for a time duration considerably longer than the duration of a transmit symbol. Subsequently, the case of moving MTs and objects is analyzed. This leads to a time-variant channel with rapidly changing fading behavior. Since the complexity of the analytical description increases for channel scenarios with many rapidly moving objects, also methods of stochastic channel description are introduced. These methods simplify channel analysis by introducing stochastic models for the multipath channel behavior.

2.2.1 Time-invariant Multipath Channel Model

In the previous sections, the radio channel was solely characterized by its power transfer characteristics. In this section we will also consider distortions of the transmit signal caused by multipath propagation.

The behavior of a multipath channel as depicted in Fig. 3 can be completely analyzed using a LTI system model [Bel63]. This model is sufficient if the MT and all reflecting objects assume fixed positions inside the cell. A complete description of the multipath channel is then given by the channel impulse response $h(t)$, cf. (2.9).

$$h(t) = \sum_{p=1}^{N_P} \alpha_p \delta\left(t - \tau_p\right) \tag{2.9}$$

Each propagation path is represented by a complex attenuation factor α_p and a path delay τ_p. The receive signal $y(t)$ is determined by a convolution of the transmit signal $x(t)$ with the channel impulse response $h(t)$:

$$y(t) = x(t) * h(t). \tag{2.10}$$

An equivalent description of the channel is obtained using the Fourier transform of $h(t)$ given as

$$H(f) = \sum_{p=1}^{N_P} \alpha_p \cdot \exp(-j2\pi f \tau_p). \qquad (2.11)$$

Equation (2.11) shows the channel transfer function $H(f)$, which generally exhibits a frequency selective behaviour if the observed frequency band is larger than the reciprocal of the longest path delay τ_p. Fig. 4 shows an example of such a transfer function considering a bandwidth of 20MHz and a multipath channel with a maximum delay of $\tau = 3.2\mu s$. Since in this case a time-invariant channel is considered, the representation in the frequency domain is sufficient.
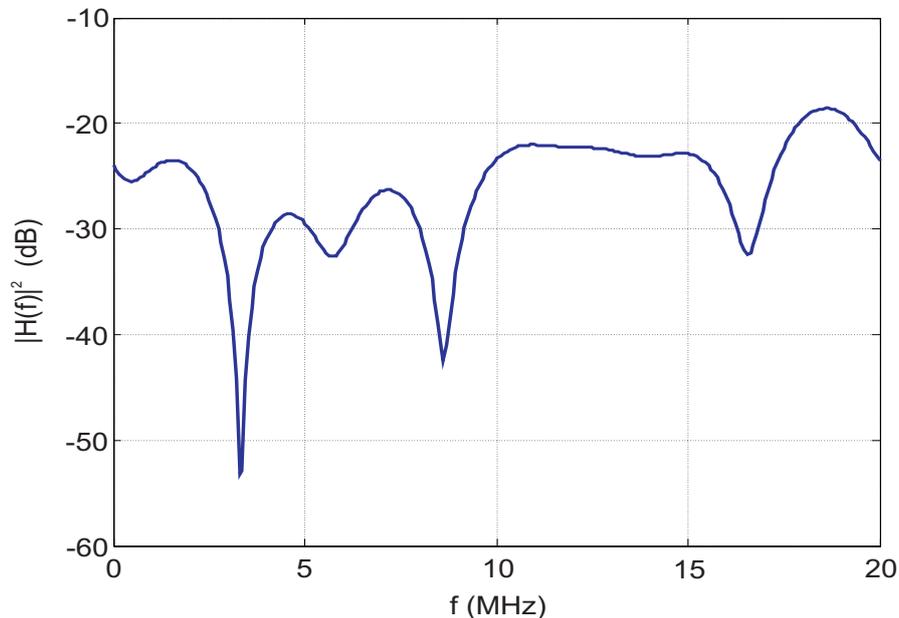
Fig. 4: Example channel transfer function of LTI multipath channel

2.2.2   Time-variant Multipath Channel Model

In contrast to the time-invariant channel behavior discussed in the preceding section, the radio channel varies rapidly if the scatterers and MTs inside the cell are assumed to be moving. Such a behavior can be analyzed by applying a linear time-variant system model [Bel63], [Pro00]. This approach leads to a description of the multipath channel by its *time-variant channel impulse response* $h(\tau,t)$. Now, the channel does not stay unchanged for the duration of the transmission. Instead, $h(\tau,t)$ describes the channel response at a time t with respect to a Dirac impulse sent at time $t-\tau$. This additional time-dependence of the channel leads to a higher complexity of the corresponding deterministic channel model. Thus, after introducing characteristic functions for a deterministic channel description, stochastic models for time-variant channels are developed.

10

The first characteristic function of a time-variant multipath channel in the baseband is its time-variant channel impulse response $h(\tau,t)$, which consists of N_P propagation paths, cf. (2.12). The variables $\alpha_p(t)$ and $\theta_p(t)$ are real-valued and represent attenuation and phase shift, respectively.

$$h(\tau,t) = \sum_{p=1}^{N_P} \alpha_p(t)\delta(\tau-\tau_p)\exp\left(j2\pi\theta_p(t)\right) \qquad (2.12)$$

In contrast to the time-invariant impulse response (2.9), attenuations $\alpha_p(t)$, and phase shifts $\theta_p(t)$ are time-dependent and lead to strong changes of the channel situation during an ongoing transmission.

The multipath channel influence on the transmit signal $x(t)$ can be described by a convolution with $h(\tau,t)$. This yields the receive signal $y(t)$.

$$y(t) = \int_{-\infty}^{\infty} x(t-\tau)h(\tau,t)d\tau$$
$$= x(t)*h(\tau,t) \qquad (2.13)$$

As in section 2.2.1, an equivalent channel description is given by Fourier transformation of $h(\tau,t)$. Transformation over τ leads to the *time-variant channel transfer function*

$$H(f,t) = \mathcal{F}_\tau\{h(\tau,t)\}$$
$$= \int_{-\infty}^{\infty} h(\tau,t)\exp\left(-j2\pi f\tau\right)d\tau. \qquad (2.14)$$

An example for a time-variant transfer function $H(f,t)$ is given in Fig. 5.
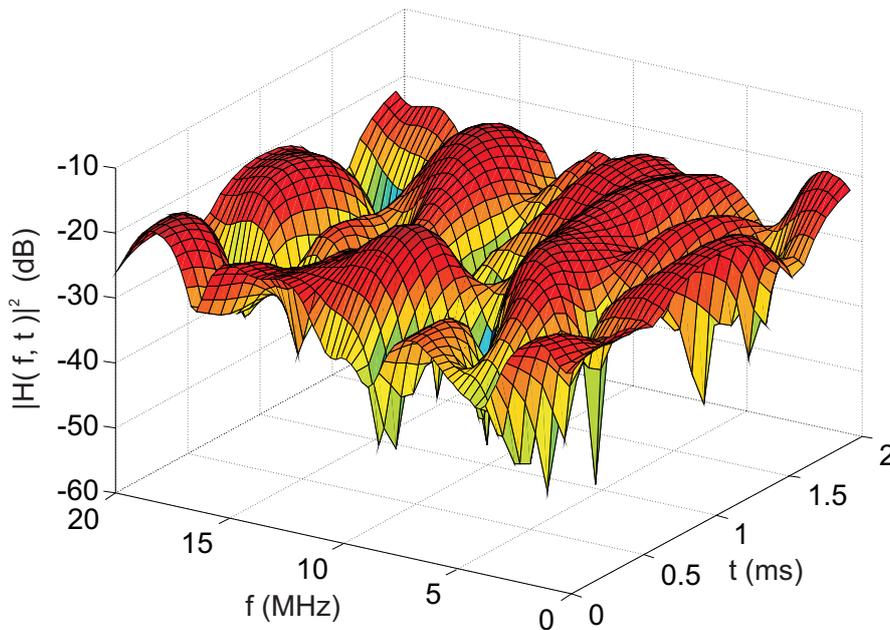
Fig. 5: Example channel transfer function of multipath channel with moving MT

11

The figure shows the strong frequency selectivity that is caused by the multipath propagation inside the channel. This effect is analogous to Fig. 4. The additional variation of the transfer function over time seen in Fig. 5 is due to the movement of the MT in relation to the BS.

Additional characteristic functions to describe the channel are the Fourier transforms over the time t of (2.12) and (2.14). These are called the *delay Doppler function* $V(\tau, f_D)$ and the *frequency Doppler function* $U(f, f_D)$, which are introduced as follows:

$$
\begin{aligned}
V(\tau, f_D) &= \mathcal{F}_t\{h(\tau, t)\} \\
&= \int_{-\infty}^{\infty} h(\tau, t) \exp(-j2\pi f_D t)\, dt
\end{aligned}
\tag{2.15}
$$

$$
\begin{aligned}
U(f, f_D) &= \mathcal{F}_\tau\{V(\tau, f_D)\} = \mathcal{F}_t\{H(f, t)\} \\
&= \int_{-\infty}^{\infty} H(f, t) \exp(-j2\pi f_D t)\, dt
\end{aligned}
\tag{2.16}
$$

Each of the introduced functions can be derived from each other by (consecutive) Fourier transform. This can be seen in Fig. 6. The equations (2.15) and (2.16) depend on the variable f_D which is referred to as the *Doppler frequency*. It describes a change in frequency due to movements of scatterers or relative motion between MT and BS.
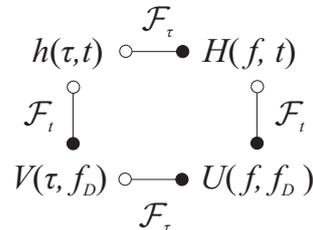
Fig. 6: Characteristic functions of time-variant multipath channel

The upcoming sections show that the introduced characteristic functions $h(\tau, t)$, $H(f, t)$, etc. can be used as a basis for describing the radio channel as a random process. Such an approach leads to a simplified channel model, which can be used efficiently to simulate transmission systems on a computer.

2.2.3   Stochastic Multipath Channel Model

The introduced means of channel analysis allow the description of a particular channel at a specific point in time. Of course, this description is only complete if all parameters like path delays τ_p, Doppler shifts f_D and attenuations α_p are known. As introduced earlier, the knowledge of these parameters depends on a detailed description of the complete system envi-

ronment, which has to include all properties and positions of scatterers and users inside the radio system. And additionally, the whole environment may vary rapidly over time.

The vast amount of data that would arise from such a deterministic channel description makes this approach impractical for computer simulations. Instead, stochastic channel models are employed, which treat each channel parameter as a random process with individual statistics. The statistics of each parameter can be adapted in order to approximate the properties of certain propagation environments and cell scenarios.

The desired stochastic channel description can be derived using the autocorrelation functions (ACF) of the characteristic channel functions introduced in the previous section.

Starting with the channel impulse response $h(\tau,t)$, its ACF for two channel delays τ_1 and τ_2 is given by

$$\varphi_{hh}(\tau_1,\tau_2,t_1,t_2) = E\left\{h(\tau_1,t_1) \cdot h^*(\tau_2,t_2)\right\}. \qquad (2.17)$$

The same approach holds for the functions $H(f,t)$, $V(\tau,f_D)$ and $U(f,f_D)$.

Since (2.17) still depends on individual points in time t_1 and t_2, further simplifications are required. Thus, a couple of assumptions are introduced in the following, which are generally fulfilled under realistic conditions. Without loss of generality, these assumptions are applied to the channel impulse response $h(\tau,t)$.

First the stochastic process, which describes $h(\tau,t)$ is assumed to be *wide sense stationary* (WSS). This means, that the expectation value of the process does not change over time and the ACF $\varphi_{hh}(.)$ does not depend on the individual points in time t_1 and t_2, but only on the time difference $\Delta t = t_2 - t_1$:

$$\varphi_{hh}(\tau_1,\tau_2,\Delta t) = E\left\{h(\tau_1,t) \cdot h^*(\tau_2,t+\Delta t)\right\} \qquad (2.18)$$

Furthermore it is assumed, that the scattering parameters (e.g. attenuation, phase shift) of each propagation path are uncorrelated (*uncorrelated scattering*, US). Under this assumption, the ACF φ_{hh} vanishes for $\tau_1 \neq \tau_2$:

$$\varphi_{hh}(\tau_1,\tau_2,\Delta t) = \begin{cases} \varphi_{hh}(\tau_1,\Delta t) & \text{for } \tau_1 = \tau_2 \\ 0 & \text{else} \end{cases} \qquad (2.19)$$

Multipath channel models fulfilling both of these assumptions are known as *wide sense stationary uncorrelated scattering* (WSSUS) channels [Bel63]. The ACF $\varphi_{hh}(\tau,\Delta t)$ of the channel impulse response now describes the channel statistics while only depending on two parameters: the delay τ and the time span Δt. Similar to section 2.2.1, a number of functions can be derived from (2.19), which are an alternative means to describe the channel statistics:

By applying the Fourier transform to the parameter τ, the *time-frequency-correlation function* $\phi_{HH}(\Delta f,\Delta t)$ can be derived

13

$$\phi_{HH}(\Delta f, \Delta t) = \mathcal{F}_\tau \{\varphi_{hh}(\tau, \Delta t)\} \tag{2.20}$$

whereas the transformation over Δt leads to the *scattering function* $\phi_S(\tau, f_D)$ with

$$\phi_S(\tau, f_D) = \mathcal{F}_{\Delta t} \{\varphi_{hh}(\tau, \Delta t)\}. \tag{2.21}$$

The scattering function describes the average output power of the radio channel [Pro00] depending on the delay τ and the Doppler frequency f_D. This function is especially important in deriving channel models for simulations, since it is directly proportional to the probability densities of τ and f_D [Pae99].

Analogous to the previous section, the introduced functions describing the channel statistics are connected to each other via the Fourier transform. An overview of these relations is given in Fig. 7. The figure also introduces another characteristic function: the *Doppler cross power density spectrum* $\phi_{UU}(\Delta f, f_D)$, which is noted here for the sake of completeness.
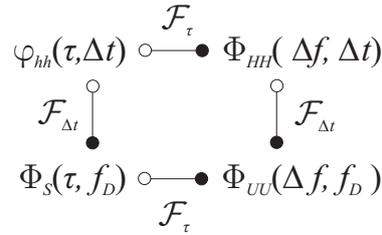
Fig. 7: Characteristic functions of channel statistics

In order to simplify the general dimensioning of a communication system, also single-value descriptions of a channel are introduced. The two most important values are the *coherence time* T_C and the *coherence bandwidth* B_C. The coherence time is defined as the smallest time span Δt for $\Delta f = 0$ where the time-frequency correlation function still has half of its maximum magnitude $|\phi_{HH}(0,0)|$, cf. (2.22).

$$|\phi_{HH}(0, T_C)| = \frac{1}{2}|\phi_{HH}(0,0)| \tag{2.22}$$

The definition (2.23) of the coherence bandwidth is quite similar.

$$|\phi_{HH}(B_C, 0)| = \frac{1}{2}|\phi_{HH}(0,0)| \tag{2.23}$$

It is a common approximation to assume the channel transfer function $H(f,t)$ to be constant inside an interval which is considerably smaller than the coherence time T_C or the coherence bandwidth B_C, respectively. Both parameters can be approximated by

$$T_C \approx \frac{1}{f_{D,\max}} \qquad B_C \approx \frac{1}{\tau_{\max}} \tag{2.24}$$

14

with $f_{D,\max}$ as the maximum occurring Doppler frequency and $\tau_{\max}$ as the maximum observed propagation delay in the considered radio channel.

2.2.4 Simulation of Multipath Channels

Multipath channels are often simulated by using random number generators on computers [Tra04]. In order to obtain a realistic simulation, the stochastic properties of these generators as e.g. the PDF must be matched to the statistics of the underlying channel model. In the following, a simulation model for a WSSUS channel is introduced. To this end, the stochastic channel descriptions from section 2.2.3 and in particular the scattering function $\phi_S(\tau, f_D)$ are used.

It can be shown [Pae99], that the scattering function $\phi_S(\tau, f_D)$ is proportional to the joint PDF of delay and Doppler shift.

$$\phi_S(\tau, f_D) \sim p(\tau, f_D) \tag{2.25}$$

The relation (2.25) connects the stochastic description of a physical channel environment $\phi_S(\tau, f_D)$ to a PDF, which in turn is a vital means to define random number generators for channel simulations.

Furthermore, based on measured channel properties it is often assumed that delay and Doppler are independent influences, which leads to the following description by two separate PDFs:

$$p(\tau, f_D) = p(\tau) \cdot p(f_D). \tag{2.26}$$

The PDF $p(\tau)$ describes the distribution of the channel delays and is assumed in this work to be a negative exponential distribution for mobile radio scenarios [Pae99]. For the Doppler PDF, an omnidirectional antenna and uniformly distributed incidence angles for the propagation paths are assumed, which leads to the well-known *Jakes distribution*. Both PDFs are introduced in the following:

In order to simplify channel simulation, the PDF of the delay $p(\tau)$ is limited to a maximum delay $\tau_{\max}$, at which the average output power of the channel has dropped to a negligible level. This leads to

$$p(\tau) = \begin{cases} a_0 \exp(-b_0 \tau) & \text{for } 0 \leq \tau < \tau_{\max} \\ 0 & \text{else} \end{cases}. \tag{2.27}$$

Thus, the parameter b_0 is chosen in such a way, that the PDF $p(\tau)$ at the maximum delay $\tau_{\max}$ in relation to its value at $\tau = 0$ has dropped to

$$\frac{p(\tau_{\max})}{p(0)} = \exp(-b_0 \tau) = 10^{-3} \tag{2.28}$$

15

which yields

$$b_0 = \frac{3\ln(10)}{\tau_{\max}}.$$ (2.29)

Normalization of the PDF is done by setting the value of the constant a_0 according to

$$a_0 = \frac{b_0}{1 - \exp(-b_0 \tau_{\max})}.$$ (2.30)

The PDF of the Doppler frequencies $p(f_D)$ is assumed to be Jakes-distributed and is given by (2.31) for a maximum Doppler frequency of $f_{D,\max}$.

$$p(f_D) = \begin{cases} \left[\pi f_{D,\max} \sqrt{1 - \left(\dfrac{f_D}{f_{D,\max}} \right)^2} \right]^{-1} & \text{for } |f_D| \leq f_{D,\max} \\ 0 & \text{else} \end{cases}$$ (2.31)

3 OFDM Transmission Technique

In the considered channel environment a wireless transmission technique must cope with frequency selectivity caused by multipath propagation. A technique perfectly adapted to these channel conditions is known as *Orthogonal Frequency Division Multiplex* (OFDM), which is introduced in the following. The underlying idea of OFDM is to split the overall channel bandwidth W into a large set of N orthogonal frequency bands referred to as *subcarriers*. Due to their orthogonality, each subcarrier can be used as an independent resource, which greatly simplifies compensation of the channel influence at the receiver side [Wei71]. If in addition a guard interval is used at the end of each OFDM symbol, intersymbol interferences due to multipath propagation can be completely avoided.

The following sections give a detailed introduction to the OFDM transmission technique.

3.1 Motivation

Nowadays, a wide variety of mobile digital transmission systems like GSM and DECT use single carrier (SC) techniques for low-rate data transmission [Gsm96][Dec08]. Future mobile systems must cope with increasing data rate demands and multipath environments [Roh99]. In such a scenario, the SC technique reveals a serious drawback known as intersymbol interference. The reason for this effect is discussed in this section and serves as a motivation for a transmission technique, which is robust against ISI: The OFDM transmission technique.

In SC systems, user data are transmitted as a time sequence of pulses $g(t)$ weighted by modulation symbols X_n, which carry the actual information. Such a modulated pulse is referred to as a transmit symbol. This leads to the transmit time signal

$$x(t) = \sum_{n=-\infty}^{\infty} X_n g(t - nT_{S,SC}).$$ (3.1)

In order to adapt the pulses $g(t)$ to meet the system's channel bandwidth W, the duration of a modulation symbol $T_{S,SC}$ in the considered SC system is chosen in accordance with

$$W \approx \frac{1}{T_{S,SC}}.$$ (3.2)

In broadband systems with high data rates, also the corresponding rate of modulation symbols is very high, leading to short symbol durations $T_{S,SC}$. If a time-invariant multipath channel with impulse response $h(t)$ (cf. (2.9)) is considered, the received signal $y(t)$ is given by

$$y(t) = \sum_{p=1}^{N_P} \alpha_p x(t - \tau_p).$$ (3.3)

If (3.3) holds and additionally a high-rate SC system with

$$T_{S,SC} < \tau_{\max}$$ (3.4)

is considered, it is obvious that ISI is observed at the receiver. The severance of ISI depends on the symbol duration $T_{S,SC}$ as well as on the maximum path delay $\tau_{\max}$ of the radio channel. The number of adjacent transmit symbols N_{ISI} disturbed by ISI is obtained from

$$N_{ISI} = \frac{\tau_{\max}}{T_{S,SC}}.$$ (3.5)

For high-rate transmission systems, N_{ISI} reaches a magnitude in the order of 10 to 100. Naturally this leads to a high computational effort at the receiver for canceling the influence of so many interfering symbols.

These severe ISI effects are the motivation for the OFDM transmission technique. Since the main cause of ISI is the short symbol duration $T_{S,SC}$ of SC systems, an obvious approach to reduce ISI is to increase $T_{S,SC}$. If an increase of $T_{S,SC}$ fulfills the condition

$$T_{S,SC} > \tau_{\max}$$ (3.6)

only a small part of the symbol is influenced by ISI, which can be cancelled at the receiver more easily. Of course, if SC systems are considered, such long symbol durations strongly decrease the available data rate. But at the same time, also the required channel bandwidth W of the system is reduced in the same proportion by which $T_{S,SC}$ is increased, cf. (3.2). Thus, in order to maintain the data rate, multiple of these narrowband SC transmissions can be established in parallel using multiple adjacent frequency bands. This concept, also known as *multicarrier* (MC) transmission, is the basis of the OFDM transmission technique.

For the time being the symbol duration of a MC system is referred to as $T_{S,MC}$ in order to avoid confusion with a SC system. The bandwidth occupied by each of the narrow frequency bands of a multicarrier system is given as

$$\Delta f = \frac{1}{T_{S,MC}}.$$ (3.7)

Assuming a number of N bands used for parallel transmission, the overall system bandwidth is defined by

$$W = N \cdot \Delta f = \frac{N}{T_{S,MC}}.$$ (3.8)

18

A comparison of (3.2) and (3.8) shows that the symbol duration $T_{S,MC}$ of a multicarrier system is by a factor of N larger compared to an equivalent SC system, leading to a much lower influence of ISI:

$$T_{S,MC} = N \cdot T_{S,SC}. \qquad (3.9)$$

Since the MC technique uses N frequency bands in parallel for data transmission, the overall data rate of a MC system using a bandwidth W is identical to that of a SC system with the same bandwidth. This means, an MC system such as OFDM is able to provide high data rates and robustness against ISI at the same time. Motivated by this introduction, the following section discusses the principles of the OFDM technique in extensive detail.

3.2   OFDM Principles

The OFDM transmission technique represents a special case of the MC technique introduced above. Like in an MC system, OFDM uses multiple parallel frequency bands referred to as subcarriers. In particular, a subcarrier signal consists of a complex-valued time signal $\exp(j2\pi k\Delta ft)$, which is modulated by data symbols. These subcarriers are equidistantly spaced on the frequency axis at intervals Δf. For the sake of brevity the duration of an OFDM transmit symbol is denoted by T_S from now on, which replaces the notion $T_{S,MC}$ used above. The duration of the transmit symbol T_S is chosen in relation to the subcarrier spacing Δf such that

$$\Delta f = \frac{1}{T_S} \qquad (3.10)$$

is fulfilled, cf. (3.7). In this way, all subcarriers are orthogonal to each other for the duration of a symbol T_S as shown in (3.11). *Thus, there is no interference between adjacent subcarriers and consequently each subcarrier k represents an independent resource for data transmission.*

$$\int_0^{T_S} \exp(2\pi\kappa\Delta ft) \cdot \exp(2\pi\iota\Delta ft)\, dt = \begin{cases} T_S & \text{for } \kappa = \iota \\ 0 & \text{else} \end{cases} \qquad (3.11)$$

This property simplifies parallel data transmission in OFDM: Inside the duration T_S of an OFDM symbol, every subcarrier k is modulated with a complex modulation symbol X_k. Due to their orthogonality, all N modulation symbols of an OFDM symbol can be extracted independently at the receiver.

The transmit time signal $x(t)$ of a single OFDM symbol is given by the superposition of all modulated subcarriers:

$$x(t) = \sum_{k=1}^{N} X_k \exp(2\pi\Delta ft(k-1)) \quad \text{for } 0 \le t \le T_S. \qquad (3.12)$$

19

W. l. o. g. the following analysis considers the processing of a single OFDM symbol, whereas in practical systems a whole sequence of symbols is processed.

Although the signal $x(t)$ used for the actual data transmission over the channel is time-continuous, the signal processing inside the transmitter is done in the time-discrete domain. This enables the use of contemporary digital signal processing. The time-discrete transmit signal is represented by the sequence x_l and is generated by sampling $x(t)$ at integer multiples of the time interval Δt given by

$$\Delta t = \frac{T_S}{N}. \tag{3.13}$$

The subcarrier spacing Δf and the sampling interval Δt are related through (3.10) and (3.13) such that

$$\Delta f \cdot \Delta t = \frac{1}{N} \tag{3.14}$$

holds. This relation, combined with the representation of $x(t)$ given in (3.12) yields the time-discrete transmit sequence x_l:

$$x_l = \sum_{k=1}^{N} X_k \exp\left(j2\pi \frac{(l-1)(k-1)}{N} \right) \text{ for } l = 1...N. \tag{3.15}$$

Thus, x_l is described by a sequence of complex numbers suitable for processing on a digital computer. Equation (3.15) also reveals that this processing can be achieved by executing an *Inverse Discrete Fourier Transform* (IDFT) on the N modulation symbols X_k, see also [Bin90]. Assuming N as a power of two, the IDFT can be realized efficiently by inverse *Fast Fourier Transform* (FFT).
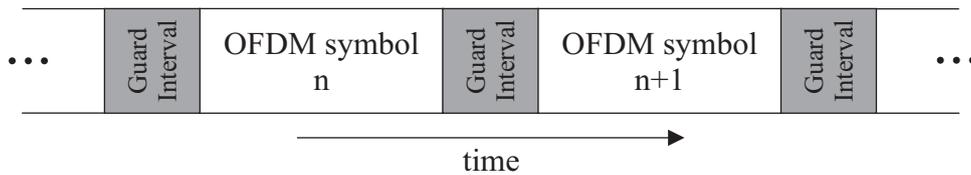
Fig. 8: Sequence of OFDM symbols with guard intervals

As shown in (3.9), the symbol duration T_S of an MC technique such as OFDM is N times larger than that of the corresponding SC technique. If the number of subcarriers N and thus the symbol duration T_S for an OFDM system is chosen such that

$$T_S \gg \tau_{max} \tag{3.16}$$

is fulfilled, signal distortions due to ISI are reduced to a minimum.

The effect of ISI can be completely avoided by prefixing a *guard interval* (GI) at the beginning of each OFDM symbol, cf. Fig. 8. To this end, the duration T_G of the guard interval must be chosen such that it covers the maximum delay τ_{max} of the channel:

$$T_G \geq \tau_{max} \tag{3.17}$$

At the receiver the GI is removed and solely the original OFDM symbol is processed. The overall time span for transmitting a single OFDM symbol is therefore

$$T = T_S + T_G. \tag{3.18}$$

Practical OFDM systems use a repetition of the ending part of the symbol as guard interval, thus producing a so-called *cyclic prefix*. This is necessary to maintain orthogonality between subcarriers in a multipath channel.
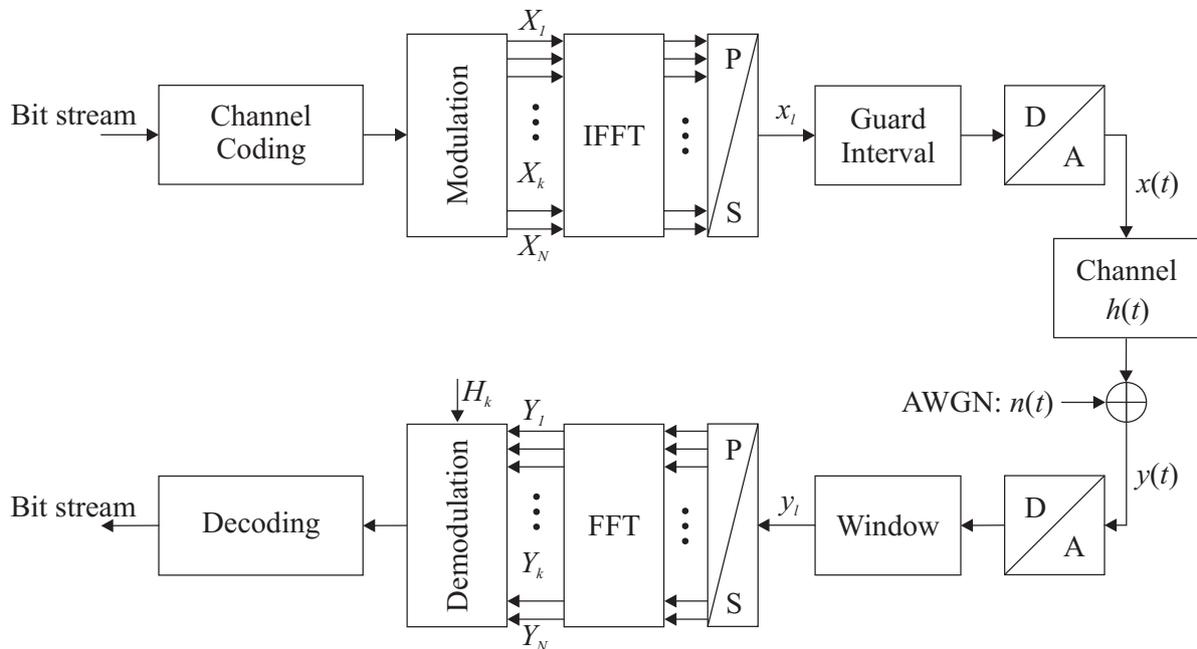
Fig. 9: OFDM transmission chain

The practical transmit signal generation inside an OFDM transmitter is explained using the transmission chain depicted in Fig. 9. First, the binary data to be transmitted are channel coded. This means, redundant bits are purposefully inserted into the bit stream in order to enable the receiver to detect and/or correct bit errors that occurred during transmission. Since channel coding is out of the scope of this work, the reader taking an interest in this topic is referred to [Pro00], [Bos99], [Vit67], and [Gal62].

Subsequently, the encoded bit stream is mapped onto complex modulation symbols X_k. A set of N modulation symbols forms an OFDM symbol, which is processed by an *N*-point IFFT yielding the time-discrete transmit sequence x_l. After appending the GI to x_l, the sequence is

digital-to-analog-converted to form the time-continuous transmit signal $x(t)$. Afterwards, this signal is transmitted over the channel.

After the discussion of the OFDM signal generation at the transmitter, the channel influence to an OFDM signal and its reception is discussed.

If the symbol duration T_S of the transmit signal $x(t)$ is chosen much smaller than the coherence time T_C of the multipath channel, i. e.

$$T_S \ll T_C \tag{3.19}$$

holds, then the channel can be described by an LTI system. The corresponding channel impulse response $h(t)$ is given by (2.9). Hence, the channel influence on an OFDM symbol $x(t)$ can be described by a convolution with the channel impulse response $h(t)$ and addition of white Gaussian noise $n(t)$. The noise $n(t)$ is predominantly produced by thermal agitation of electrons inside conductors and electronic devices. This yields the receive signal $y(t)$:

$$y(t) = x(t) * h(t) + n(t) \tag{3.20}$$

As introduced in (3.12), the transmit signal $x(t)$ is a superposition of complex subcarrier signals $\exp\left(j2\pi k\Delta ft\right)$ with complex amplitudes X_k. It can be shown that such signals are *Eigenfunctions* of the LTI channel $h(t)$ [Fli91]. This means, a signal $X_k \exp\left(j2\pi k\Delta ft\right)$ passing through the channel is changed solely in terms of its complex amplitude, but its frequency $k \cdot \Delta f$ stays unchanged. *Hence, despite multipath propagation, the orthogonality of the subcarriers is maintained.*

This fact, together with the complete avoidance of ISI effects by means of the guard interval allows a straightforward processing of the receive signal $y(t)$ at the receiver (cf. Fig. 9):

Upon reception, the received signal $y(t)$ is sampled again at intervals Δt. Subsequent removal of the guard interval yields the time-discrete sequence y_l. Executing an FFT operation on y_l obtains a set of complex amplitudes Y_k at the receiver side. These received amplitudes Y_k represent the modulation symbols X_k distorted by the channel $h(t)$. This procedure is the exact inverse of the signal processing at the transmitter.

Since the subcarrier signals are Eigenfunctions of the channel, the original modulation symbols X_k are solely affected by a complex channel transfer factor H_k with additional noise contribution N_k. Thus the received symbols Y_k at the FFT output are given by

$$Y_k = X_k \cdot H_k + N_k \tag{3.21}$$

with the factors H_k being samples taken from the channel transfer function $H(f)$ (cf. (2.11)) at integer multiples of the subcarrier spacing Δf : $f = k \cdot \Delta f$.

Since T_S is inversely proportional to Δf and is chosen in accordance with (3.16), the sampling of $H(f)$ occurs at very small frequency intervals Δf. Hence, the fading behavior of each subcarrier k can be assumed to be constant in the frequency domain, which is another confirmation of the fact that the channel influence on each symbol X_k can be described by a single factor H_k. In this respect the coherence bandwidth B_C is worth mentioning: Its definition given in (2.24) provides an indication for the maximum possible subcarrier spacing Δf inside an OFDM system:

$$\Delta f \ll B_C \approx \frac{1}{\tau_{\max}} \tag{3.22}$$

Considering again the symbols Y_k in (3.21), the equation shows that on a subcarrier k no interference from adjacent subcarriers is observed. Additionally, the guard interval prevents ISI between successive OFDM symbols. Thus, the equalization effort that must be spent at the receiver is much lower than in a comparable SC system as discussed in section 3.1. The relation (3.21) also reveals that the complete OFDM transmission chain can be modeled by a set of simple arithmetic operations. This greatly simplifies analysis and simulation of the system.

Based on these considerations, the distortions produced by the channel can be mitigated at the receiver by a single complex division yielding the estimate $\tilde{X}_k$ for the original symbol X_k:

$$\tilde{X}_k = \frac{Y_k}{H_k} = X_k + \frac{N_k}{H_k} \approx X_k \tag{3.23}$$

This approach assumes the channel transfer factors H_k to be known at the receiver (cf. Fig. 9). This knowledge is also called *channel state information* (CSI). It is generally obtained by using dedicated pilot signals for channel measurements. Since channel measurement is out of the scope of this work, the reader taking an interest is this topic is referred to [Toe07], [Col02], and [Hoe97].

Because CSI is essential for most of the transmission schemes considered in this work, perfect channel knowledge at both the transmitter and receiver side is assumed, unless stated explicitly otherwise.

3.3 OFDM System Design

The principles of the OFDM transmission technique discussed so far combined with the radio channel basics from the preceding section are all that is needed for the design of actual OFDM system parameters.

Since the radio channel is the environment a mobile transmission system has to cope with, it is important to consider the channel properties in order to achieve a successful transmission design.

The most decisive channel parameters for the design of an OFDM transmission system are the maximum path delay $\tau_{\max}$ and the maximum Doppler frequency $f_{D,\max}$. As explained in the sections 2.2.3 and 2.2.4, these two parameters are sufficient to give a general characterization of the channel behavior.

This becomes obvious by looking at the definitions of the coherence time T_C and the coherence bandwidth B_C cf. (2.24), since they are solely determined by the above-mentioned channel parameters. Thus, the boundary conditions for an OFDM transmission system, i.e. its design limitations are clearly outlined by the same parameters $\tau_{\max}$ and $f_{D,\max}$. This is shown in the following by giving design constraints for the most important OFDM system parameter: The symbol duration T_S. Since the system bandwidth W is generally fixed by third-party regulations, the single parameter T_S determines the subcarrier spacing Δf as well as the number of subcarriers N, cf. (3.8) and (3.10). Generally, the symbol duration depends on the channel parameters in such a way that it must be considerably larger than the maximum path delay $\tau_{\max}$ and much shorter than the inverse of the maximum Doppler frequency $f_{D,\max}$, which in turn approximates the channel's coherence time T_C:

$$\tau_{\max} \ll T_S \ll \frac{1}{f_{D,\max}} \tag{3.24}$$

The relation (3.24) is discussed as follows. First, the dependency on $\tau_{\max}$ is evaluated: A value of T_S, which is larger than $\tau_{\max}$ provides the narrowband characteristic of the individual subcarriers such that the subcarrier spacing Δf is much smaller than the coherence bandwidth $B_C \approx 1/\tau_{\max}$. Consequently, the channel influence on each subcarrier k can be described by a single transfer factor H_k, cf. (3.21). Additionally, a large T_S reduces the overhead introduced by the use of a guard interval T_G.

The preceding arguments suggest that it is beneficial to increase T_S indefinitely. But (3.24) shows, that the second characteristic channel parameter $f_{D,\max}$ - and thus the channel's time-variance - limits the duration of T_S: In order to allow simple signal equalization at the receiver, the considered channel $h(t)$ must possess LTI properties. Hence, the symbol duration T_S has to be much shorter than the coherence time of the channel T_C, which is approximately the inverse of $f_{D,\max}$. Based on this reasoning, the symbol duration is roughly outlined using (3.24). A more sophisticated range for T_S generally adopted by realistic system implementations is found in [Ald94] and gives the following limits:

$$4\tau_{\max} \leq T_S \leq \frac{0.03}{f_{D,\max}}. \tag{3.25}$$

4 Multiple Access in OFDM Systems

4.1 Motivation and Overview

The preceding chapters discussed the OFDM transmission technique and its robustness against various mobile radio channel effects. The focus lay on a point-to-point connection between a single transmitter and receiver, cf. Fig. 9. In such a scenario, a single receiver or user respectively is able to utilize the complete bandwidth W of the channel for an unlimited time period. Thus, the scenario represents a single-user system.

In systems with multiple users, the access to the channel must be divided into discrete resources, in order to allow each user access to the channel and to avoid interference between users. Since in this work solely multiuser systems are considered, the underlying medium access scheme is of vital importance for the system's efficiency. For this reason, various possible multiple access schemes are introduced and discussed in the following. In OFDM, there are several possibilities to achieve this so-called *Media Access Control* (MAC): The channel can be shared in time (*Time Division Multiple Access:* OFDM-TDMA), code (OFDM-CDMA) and frequency (OFDM-FDMA), cf. [Gal06], [Olo06], [Grue00], and [Roh97].

Since this work solely considers the case where BS and MTs are equipped with a single antenna each, the fourth possible MAC scheme known as *spatial division multiple access* (SDMA) is not discussed here. The reader taking an interest in this topic is referred to [Mac08], [Grue06], and [Tej06].

Due to the consideration of the OFDM transmission technique in this work, the smallest resource entity available to a user is a time-frequency block of subcarrier bandwidth Δf and symbol duration T, cf. (3.18). All MAC schemes introduced above assign integer multiples of this minimum resource to the users inside the system. The actual access paradigms utilized by these schemes are introduced in the following:

- OFDM-TDMA: In this access scheme, each user utilizes the complete bandwidth W of the system for a defined time period. Thus, the channel is shared between users in the time domain. A user's time period is also called a *time slot* and consists of an integer multiple of the OFDM transmit symbol duration T.

- OFDM-FDMA: Each user utilizes a subset of the N available subcarriers inside the system. The assigned subcarriers of a specific user do not have to be adjacent to each other. The subcarrier assignment stays valid for the complete duration of the transmission.

- OFDM-CDMA: This scheme separates multiple users by assigning an individual code to each user. The transmit signal of each user is multiplied with his respective code. The transmissions of all users take place simultaneously using the complete bandwidth

W for the whole duration of the transmission (cf. Fig. 10). Despite simultaneous transmission of all users on all time-frequency blocks, the separation of users at the receiver side is still possible, if the utilized codes are orthogonal to each other.

An overview over all three introduced MAC schemes is given in Fig. 10.
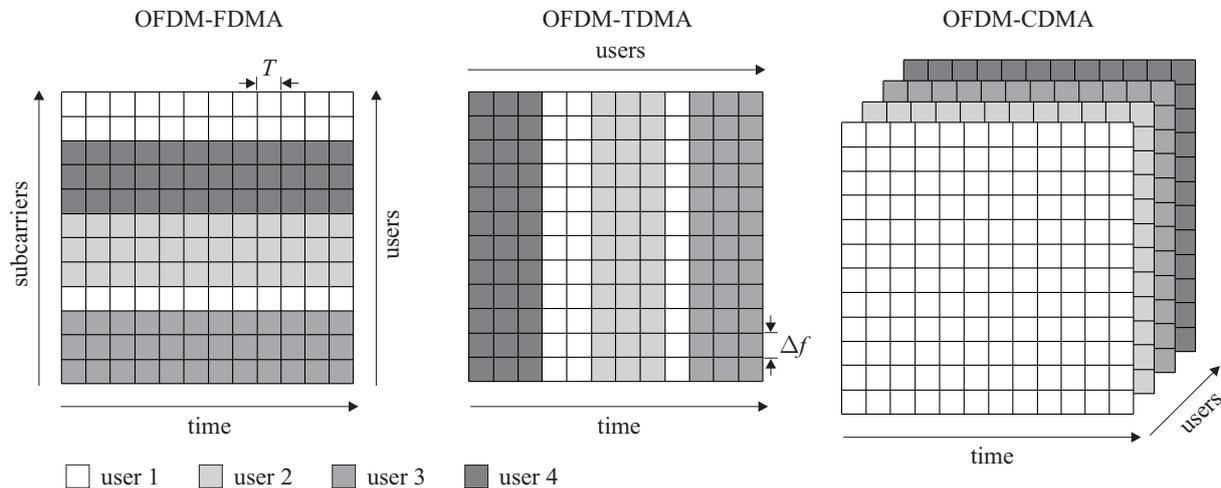
Fig. 10: Exemplary multiple access schemes for OFDM

Of all introduced MAC schemes, solely OFDM-FDMA (a. k. a. OFDMA) is considered throughout this thesis. The reasons for this decision are threefold:

First, OFDM-FDMA represents a straightforward MAC scheme for an OFDM system, since the IFFT/FFT processing inside the transmission chain already separates the bandwidth *W* into a set of orthogonal subcarriers, which are ready-to-use resources for a multiuser system. Additionally, the FFT processing enables a convenient implementation of the MAC scheme inside the digital processing domain.

Second, in contrast to OFDM-CDMA, the subcarriers of an OFDM-FDMA system stay orthogonal even after passing through a WSSUS channel, which significantly reduces the equalization effort at the receiver side, cf. [Bur01].

But the most important reason with respect to the focus of this work is the possibility of OFDM-FDMA to exploit the frequency selectivity of the radio channel by channel-adaptive subcarrier allocation: In the considered downlink scenario of a multiuser OFDM-FDMA system, each user observes a different channel transfer function $H(f)$ since the positions of scatterers are individual for each channel. Based on instantaneous CSI from all users, the BS is able to select subcarriers which show a preferably high channel transfer factor for each user [Roh97]. In such a way, the system performance can be increased since subcarriers which undergo a deep fade in one user's channel, can be accessed by another user, who may observes a much higher channel transfer factor at the respective subcarrier.

After this motivation for the OFDM-FDMA multiple access scheme, its integration into the considered OFDM transmission system introduced in section 3 is described in detail.

4.2   OFDM-FDMA System Model

If multiple access schemes for the OFDM transmission technique are considered, an almost inevitable candidate is the FDMA scheme. The subcarriers of an OFDM system constitute a set of ideally orthogonal resources in the frequency domain, which can be utilized by an FDMA multiple access scheme in a straightforward way. The combination of OFDM transmission and FDMA multiple access is known in literature as OFDM-FDMA or OFDMA. In the following it will be referred to as OFDM-FDMA.

An important prerequisite for using OFDM-FDMA in a transmission system is that all users are synchronized to the base station. If this is the case, no interference between subcarriers of adjacent users is observed (so-called *Multiple Access Interference*, MAI) and thus the signals of all users can be ideally separated at the receiver side by a simple FFT operation. *The assumption of ideal synchronization in time- and frequency-domain is also made in this work*.

In contrast to the single-user OFDM transmission chain introduced in section 3, from now on an OFDM-FDMA system serving multiple users inside the cell area is considered. Therefore, the analytical model for an OFDM transmission given by (3.20) and (3.21) must be extended for multiple users:

In OFDM-FDMA systems the access of multiple users to the system is realized by assigning subsets of subcarriers from the overall set of N subcarriers to the individual users. Therefore each user can only utilize a small portion of the overall system bandwidth W. In order to enable unambiguous allocation of a subcarrier k to a user, all N_U users inside the system are identified by an index i ranging from 1 to N_U. A subset of subcarriers stays assigned to a user i at least for the duration T of an OFDM transmit symbol, but mostly a certain assignment covers the duration of multiple OFDM symbols.

The block diagram on the left-hand side of Fig. 11 shows the generation of an OFDM-FDMA transmit signal x_l inside the base station. In comparison with the single-user system depicted in Fig. 9, the OFDM-FDMA system contains a new processing block before the IFFT-processing: The allocation of subcarriers to each user. The procedure for the multiuser transmit signal generation is as follows:

The bit stream of each user i is channel encoded and the data of all N_U users are mapped onto N modulation symbols $X_{i,k}$ per OFDM symbol. The decision, which user is entitled to how many modulation symbols $X_{i,k}$ and thus subcarriers is done by the *subcarrier allocation* block (highlighted in Fig. 11). In the considered OFDM-FDMA subcarrier allocation scheme, only those $X_{i,k}$ of a user i represent non-zero values, which are actually allocated to a subcarrier k. Additionally, the allocation of a user i to a specific subcarrier k is done exclusively such that no other user is entitled to use the same subcarrier simultaneously. The further signal processing is identical to Fig. 9, which leads to a transmit signal x_l containing the signals of

all users inside the system. Thus, after calculation of the FFT at the receiver side of an arbitrary user i, the receive signal is given by (cp. (3.21))

$$Y_{i,k} = X_{i,k} \cdot H_{i,k} + N_{i,k}. \tag{4.1}$$

The separation of users at the receiver side is done by solely processing the modulation symbols $X_{i,k}$ on the subcarriers allocated to a specific user i. This fact shows the straightforward integration of frequency division multiple access into an OFDM transmission.
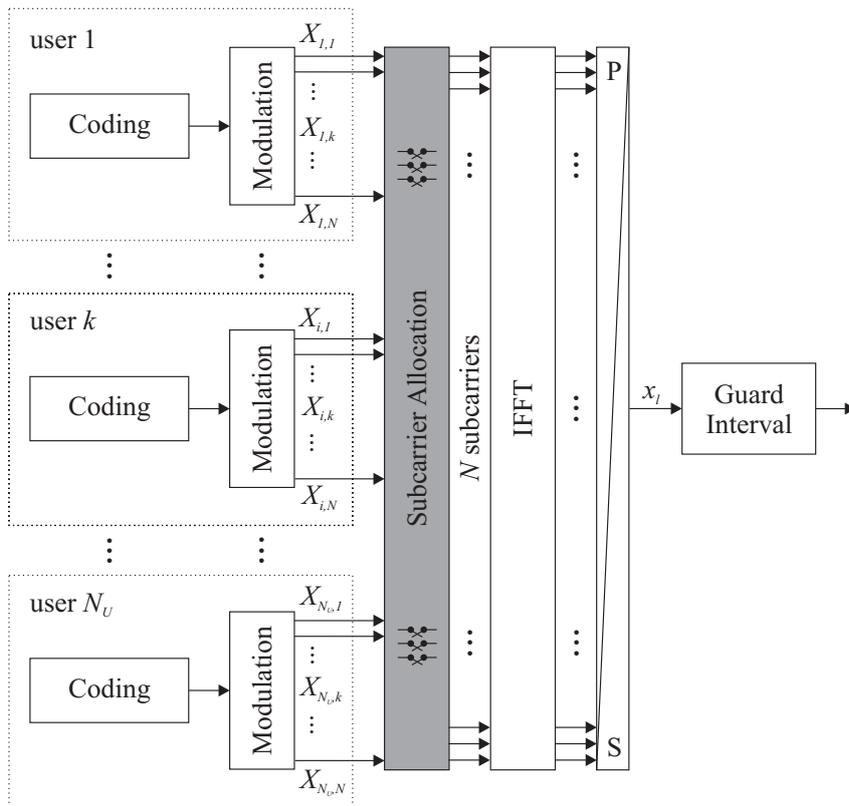
OFDM-FDMA Overview: **Example:**
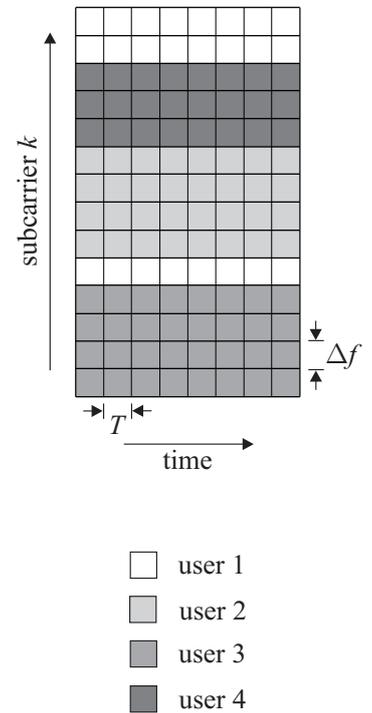
Fig. 11: Block diagram of multiuser OFDM-FDMA system with example subcarrier allocation

An exemplary subcarrier allocation for four users is shown on the right-hand side of Fig. 11. The example shows that a user i can use one or multiple subcarriers k exclusively. Since the radio channel of each user is assumed to be frequency selective, the transmission performance of each user strongly depends on the number and positions of his allocated subcarriers. *Therefore, the actual procedure of choosing suitable subcarriers for the users is of vital importance to the system performance and is considered in detail in this thesis.*

5 Channel and System Parameters

At this point, all analytical models for the description of a multiuser OFDM-FDMA transmission system are known to the reader. Before actual subcarrier allocation schemes for OFDM-FDMA and their respective performance can be discussed, certain channel and system parameters must be agreed upon. All parameters are chosen in order to describe a multiuser OFDM-FDMA transmission system.

In all simulations discussed in this work a unique set of channel parameters is used, thus keeping the results comparable to each other. As introduced earlier, the downlink transmission inside a cell from a BS to multiple MTs is considered. Schematic examples for the positioning of multiple MTs inside a cell are shown in Fig. 12. The used system and channel parameters are introduced in the following.

In the multiuser scenarios under consideration, it is assumed that each user not only wants to use voice services, but also broadband services like video streaming and file transfer. Since these services require high data rates, the underlying OFDM transmission system needs to cover a large channel bandwidth W. The parameter in Tab. 1 is chosen accordingly.

The parameters concerning multipath propagation effects are also collected in Tab. 1 and are based on [Ste07]. The PDF of channel delays $p(\tau)$ is defined by (2.27) to (2.30). It is assumed that the MTs are more or less stationary, so that Doppler influence can be neglected. Thus, for each transmitted symbol the influence of multipath propagation can be described by a time-invariant channel impulse response $h(t)$ or a channel transfer function $H(f)$ respectively. Since multiple users i are assumed, also the observed channel situation between the BS and user i is user-specific and denoted by $h_i(t)$ and $H_i(f)$.

Tab. 1: Small scale channel parameters

Max. path delay	$\tau_{\mathrm{max}} = 3.2\mu s$
Channel Bandwidth	$W = 20\mathrm{MHz}$
Number of propagation paths	30

The consideration of path loss and shadowing effects in the simulations depends on the underlying cell scenario. In this thesis, two different cell scenarios are used. Both assume a single-cell transmission system connecting a single BS to multiple MTs. The cell scenarios are depicted in Fig. 12. The left scenario is referred to as *ring scenario* since all users are located in

the same distance to the BS. Thus, it is assumed that each user observes identical shadowing and path loss factors $G_i^{SH} = 1$, $G_i^{PL} = 1$ as well as identical average *signal-to-noise-ratios* (SNR).
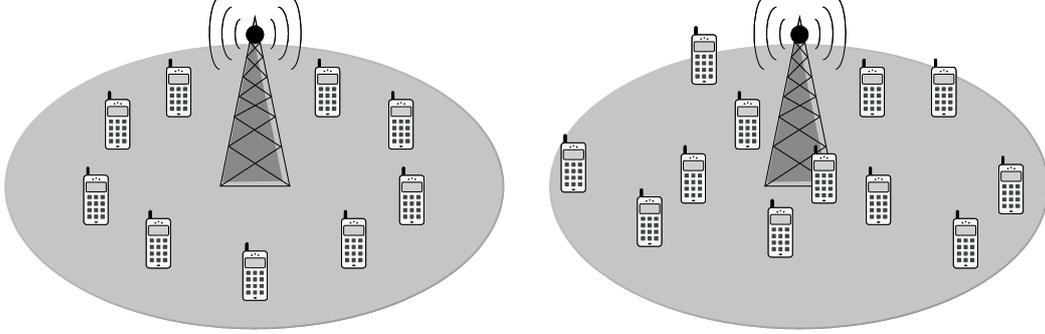
Fig. 12: Considered cell scenarios: ring (left) and uniform (right) distribution

Thus, the receive signal $Y_{i,k}$ of an individual user i on subcarrier k in the ring scenario is sufficiently described by small scale effects with added Gaussian noise $N_{i,k}$, cf. (5.1).

$$Y_{i,k} = H_{i,k} X_{i,k} + N_{i,k} \qquad (5.1)$$

Analogously, individual signal-to-noise-ratio values $SNR_{i,k}$ can be defined as given in (5.2) with σ_N^2 as the variance of the AWGN. In the scenario at hand, $E\{|H_{i,k}|^2\} = E\{|X_{i,k}|^2\} = 1$ holds. Thus the average signal-to-noise ratio of a user is determined by σ_N^{-2}.

$$SNR_{i,k} = \frac{|H_{i,k}|^2}{\sigma_N^2} \qquad (5.2)$$

Summarizing, the only disturbances in the ring scenario are multipath fading and noise. The corresponding channel parameters are given in Tab. 1.

The second scenario under consideration is referred to as *uniform scenario*. From the BS's point of view, the situation in the uniform scenario is totally different, since strong variations in receive power between users are observed. These stem from the large dynamic in the distance between BS and MTs, which causes user-specific large scale influences G_i^{SH} and G_i^{PL}. In the uniform scenario, G_i^{SH} and G_i^{PL} are no longer constant values, but random variables dependent on the distance d_i between MT and BS. The applied analytical description of G_i^{SH} and G_i^{PL} is given by the equations (2.2) to (2.4) and (2.7) to (2.8), respectively. The necessary large scale parameters are listed in Tab. 2. Consequently, the received signal $Y_{i,k}$ in the uniform case now incorporates path loss and shadowing and is described by (5.3).

$$Y_{i,k} = G_i^{SH} G_i^{PL} H_{i,k} X_{i,k} + N_{i,k} \qquad (5.3)$$

30

The corresponding $SNR_{i,k}$ values are given in (5.4). The average received power of a user is defined solely by path loss and shadowing and thus $E\left\{\left|H_{i,k}\right|^2\right\} = E\left\{\left|X_{i,k}\right|^2\right\} = \sigma_N^2 = 1$ is assumed.

$$SNR_{i,k} = G_i^{SH} G_i^{PL} \left|H_{i,k}\right|^2 \tag{5.4}$$

In order to guarantee a sufficient power margin especially for users near the cell border, a minimum SNR is defined at the maximum distance d_{max} from the BS. This minimum SNR is set to 12dB by neglecting all influences aside from path loss. This value gives a sufficient margin even if strong shadowing or deep multipath fading are observed on a specific subcarrier. The values for the references G_0 and d_0 are set accordingly, cf. Tab. 2. These parameters lead to a dynamic of $64\text{dB} - 12\text{dB} = 52\text{dB}$ of the path losses observable inside the cell area.

In summary, the data transmission in the uniform scenario is disturbed by all introduced channel effects (cf. section 2) with corresponding parameters given in Tab. 1 and Tab. 2.

Tab. 2: Large and medium scale channel parameters

Shadowing mean value	$\mu_{dB} = 0\text{dB}$
Shadowing variance	$\sigma_{dB} = 4\text{dB}$
Path loss exponent	$\alpha = 2.6$
Reference distance	$d_0 = 100\text{m}$
Path loss reference factor	$G_{0(dB)} = 12\text{dB}$
Minimum distance	$d_{min} = 1\text{m}$
Maximum distance	$d_{max} = 100\text{m}$

As a general rule, the design of any OFDM transmission system must follow the observed channel conditions. The most important OFDM parameters are the symbol duration T_S and the duration of the guard interval T_G. In the scenarios introduced above, T_S and the number of subcarriers N are depending on each other since the overall system bandwidth is fixed to $W = 20\text{MHz}$. Additional constraints are that T_S must be chosen in accordance with (3.16) and N should be a power of two. Suitable choices for T_S and N are the ones listed in Tab. 3. The given guard interval complies with (3.17). Since the channel parameters introduced in

Tab. 1 and Tab. 2 lead to strong path loss over the whole bandwidth and additionally to fading on individual subcarriers, a robust convolutional coding scheme is applied. *The OFDM system parameters given in Tab. 3 are the basis for all simulations in this thesis. Each time additional parameters are necessary for the interpretation of results, they will be introduced in the appropriate section.*

Tab. 3: Selected system parameters

Number of subcarriers	$N = 256$
Channel Bandwidth	$W = 20\text{MHz}$
Subcarrier spacing	$\Delta f = 78.125\text{kHz}$
Symbol duration (without GI)	$T_S = 12.8\mu\text{s}$
Guard interval	$T_G = 3.2\mu\text{s}$
Overall symbol duration	$T = 16\mu\text{s}$
Convolutional Coder:	
Constraint length	$CL = 7$
Code rate	$CR = 1/2$
Interleaver	Random

Furthermore, an overview considering the two introduced cell scenarios and corresponding simulations is given in Fig. 13. The depicted diagram shows the outline for all system considerations and simulations that are discussed in this work from now on. It gives an insight how the topics covered in the following are connected. The italic captions inside the boxes refer to the respective chapters in this work. Although not all terms given in Fig. 13 are introduced at this point, the diagram provides a helpful reference as the reader proceeds through the chapters.
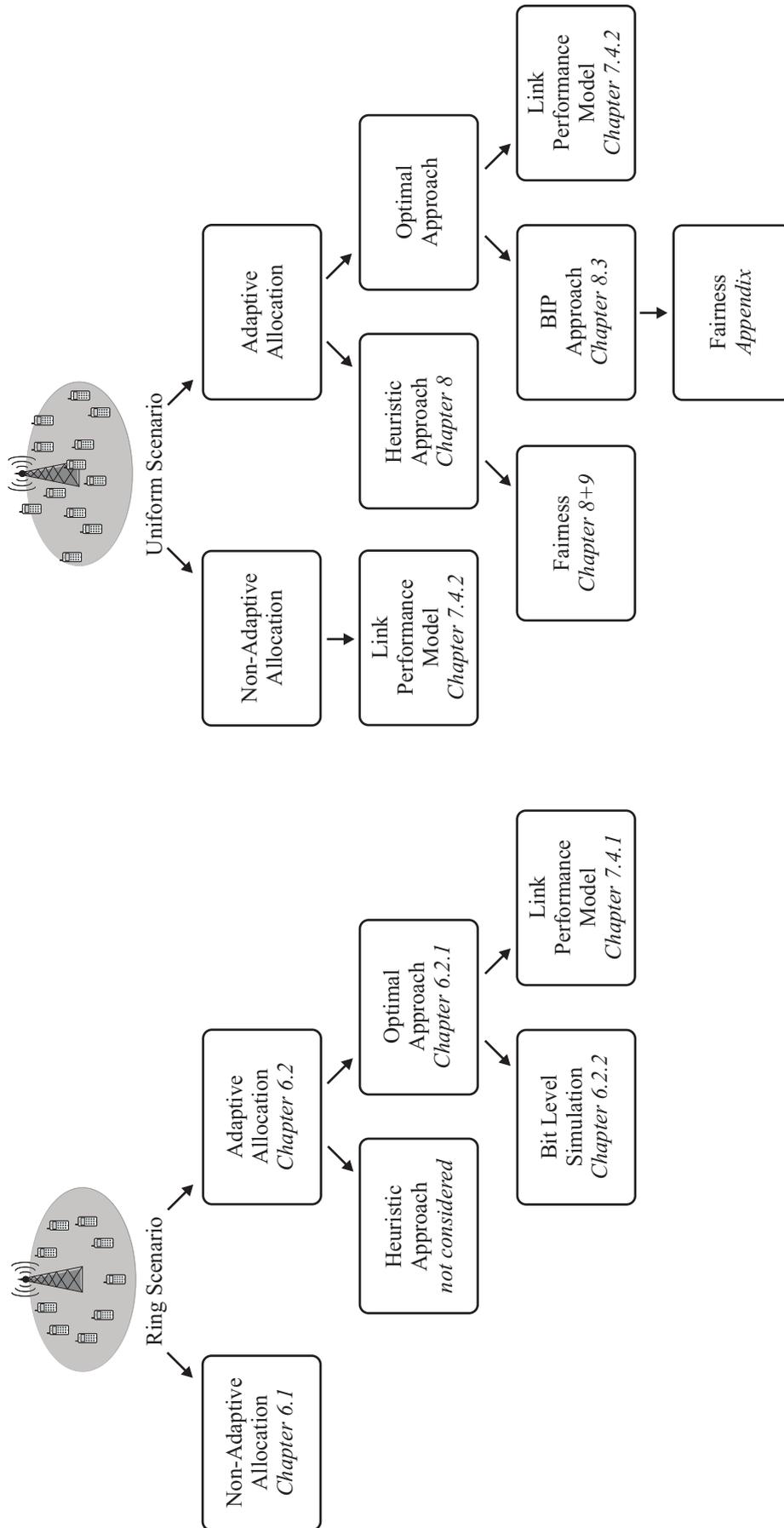
Fig. 13: Overview of considered system scenarios and corresponding simulations

6 OFDM-FDMA Multiple Access Schemes

After the introduction of the utilized OFDM system parameters, the focus now returns to the considered multiple access scheme OFDM-FDMA and its adaption to the underlying cell scenarios.

The principle of OFDM-FDMA is to allocate subsets of the available subcarriers to the users inside the cell. In the considered multiuser system, the actual subcarrier assignment has a strong influence on the system performance. Thus, various approaches for subcarrier allocation in multiuser OFDM-FDMA systems are discussed in the following.

Future mobile communication systems will have to satisfy challenging requirements with respect to data rate and flexibility. Systems based on OFDM-FDMA technology are found very promising to fulfill both requirements at the same time [Roh04][Roh01]. This is especially true for multiuser scenarios, where each user observes a radio channel with an individual fading situation.

The radio channel of mobile communication systems often shows frequency selectivity due to multipath propagation. In OFDM systems the bandwidth Δf of a subcarrier is in general significantly smaller than the coherence bandwidth B_C of the mobile radio channel. Hence, each subcarrier of the system shows the characteristics of a flat-fading narrowband channel. The SNR of such a subcarrier can be unsuitably low if the channel transfer function shows a deep fade at the position of the considered subcarrier. On the other hand, constructive interference of propagation paths can also lead to an especially favorable SNR value for a certain subcarrier. Thus, the variance of the SNR values is quite strong. This effect of frequency selective environments can be utilized in OFDM-FDMA systems to achieve large diversity gains. The inherent flexibility of OFDM-FDMA in selecting and assigning subsets of subcarriers gives rise to two fundamental possibilities for exploiting the channels frequency selectivity:

- A diversity gain can be achieved in combination with channel coding by assigning those subcarriers to a user, which are uncorrelated in terms of channel influence. This is known as frequency diversity. Since no knowledge about the channel state is necessary, this approach is termed *non-adaptive*.

- The subcarriers assigned to a user can be selected adaptively based on the instantaneous channel situation. In this case, subcarriers with preferably high SNR are selected. This is known as *multiuser diversity* (MUD). Since channel state information is needed, this approach is termed *adaptive*.

In this chapter, adaptive as well as non-adaptive approaches for the selection and assignment of subcarriers in OFDM systems are introduced and evaluated. In section 6.1, basic non-adaptive assignment schemes are introduced followed by adaptive schemes in section 6.2.

This section discusses an optimum algorithm used for adaptive subcarrier assignment. Both assignment schemes are compared in terms of bit error performance.

6.1 Non-Adaptive Subcarrier Allocation

The medium access of multiple users in OFDM-FDMA is based on the allocation of subcarriers to each user. Due to the frequency selectivity of the radio channel, the positions of the assigned subcarriers inside the system's bandwidth have a decisive impact on the link performance of each user. In the non-adaptive case, basically three approaches exist to allocate the subcarriers of an individual user along the frequency axis:

- Blockwise subcarrier allocation. This approach groups adjacent subcarriers into blocks

- Equidistant subcarrier allocation. Here, the subcarriers are distributed equidistantly over the entire system bandwidth. This ensures a maximum mutual distance between subcarriers.

- Random subcarrier allocation. In this approach the subcarriers are chosen randomly from the set of available subcarriers.

Fig. 14 gives an overview of all three approaches. Besides these allocation approaches there exist hybrid approaches, which can be constructed by combining the three allocation techniques introduced above.
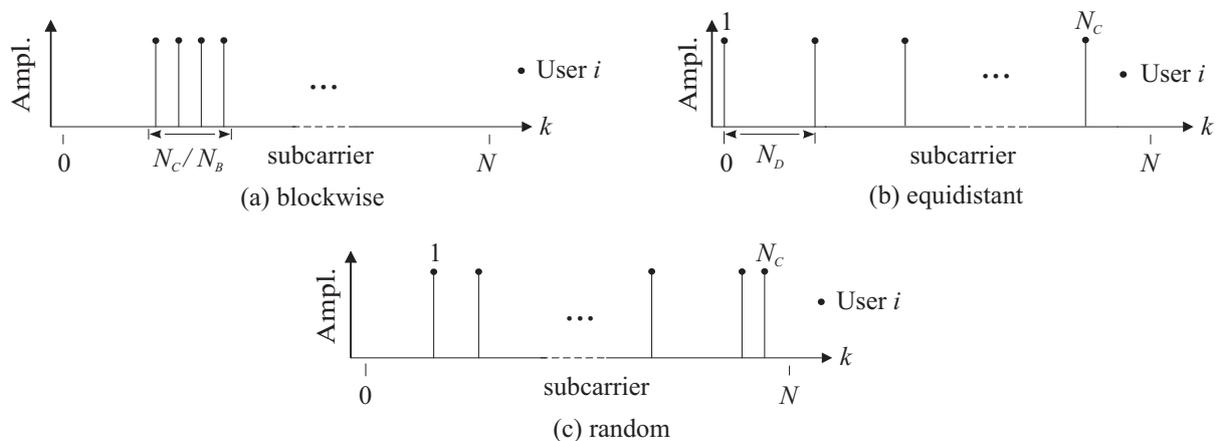
Fig. 14: Considered subcarrier allocation approaches: a) blockwise, b) equidistant, c) random

The first two techniques, namely blockwise and equidistant allocation, are well-known techniques and considered in [Gal06], [She03], and [Grue00]. In the blockwise approach each user gets assigned one or multiple blocks of adjacent subcarriers. The number of blocks per user is indicated by N_B and an identical block size for each user is assumed. If for simplicity reasons a single block per user is assumed, a block size of $N_C = N / N_U$ subcarriers follows. In case of the equidistant approach, each user gets N_C evenly spaced subcarriers with a distance of $N_D = N / N_C$ between them.

Between both allocation approaches a duality exists considering the statistics of channel transfer factors $H_{i,k}$ inside the subcarrier blocks and between the subcarrier blocks, respectively:

In the blockwise case and considering the WSSUS channel model introduced in a previous section, the channel transfer factors inside a block are highly correlated, since the subcarriers are directly adjacent to each other. Thus, a user's link performance mainly depends on the subcarrier statistics inside a block. By contrast, the subcarriers in the equidistant approach generally show uncorrelated channel transfer factors, due to their large mutual distance.

Considering practical system implementations, the blockwise approach has the advantage that the necessary overhead to signal the resource allocation result to each user is decreasing for an increasing block size [Grue00]. Random and equidistant allocation do not share this advantage.

The equidistant approach has a unique property that is especially useful in the uplink transmission direction [Ste06][Fra05]: If an equidistant subcarrier allocation is combined with a discrete Fourier spreading of the transmit symbols, the signal processing at the transmitter side becomes extremely simple. Additionally, the resulting transmit signal shows all characteristics of an SC signal and thus, well-known pulse-shaping techniques can be applied. This leads to a tx signal with an almost constant envelope, but which can still be processed using the familiar FFT at the receiver side.

Thus, this particular equidistant approach is beneficial in the uplink direction, since it poses low requirements on the processing capability and amplifier characteristics of the MT [Ste06]. Although the uplink transmission is out of the focus of this work, this particular advantage of the equidistant approach is worth mentioning.

The third case of random subcarrier allocation poses no constraints on the actual process of subcarrier selection. The subcarriers allocated to each user are merely selected randomly. The statistical properties of a user's N_C subcarriers are similar to those in the equidistant case. Random subcarrier allocation approaches are generally not applied in actually deployed systems and are merely used as a reference for other allocation approaches [Gal06].

The processing of a user's N_C subcarriers at the transmitter and receiver side in all three considered allocation approaches is done jointly. This means, the data of user i to be transmitted is interleaved and encoded altogether and the resulting data symbols $X_{i,k}$ are loaded onto the user's subcarriers. *Since the focus of this chapter lies on subcarrier allocation, it is assumed that on all subcarriers of a user an identical modulation and coding scheme (PHY mode) is applied.* For an introduction to interleaving, coding, and modulation [Pro00], [May00], [Bos99], and [Bla84] are recommended for further reading.

6.1.1 Adjusting Allocation Parameters

Although non-adaptive subcarrier allocation schemes do not need instantaneous channel state information (CSI) at the transmitter side, some general considerations concerning the subcar-

rier allocation have to be taken into account when setting the basic system parameters. The first parameters to mention are the number of subcarriers N_C and the number of blocks N_B per user. They determine the block size in the blockwise approach. In the equidistant approach, an important parameter is the spacing N_D between the subcarriers of a user. Further parameters are the total number of available subcarriers N and the number of users N_U inside the system. In systems with non-adaptive allocation and given radio channel statistics, these parameters define the statistics of the $H_{i,k}$ allocated to each user and thus the system performance.

In blockwise allocation approaches it is assumed that the bandwidth of a single block is smaller than the coherence bandwidth B_C (cf. (2.24)) of the radio channel. Thus, if an identical block size for all users is assumed, it follows:

$$\frac{N_C}{N_B} \le \frac{B_C}{\Delta f} \tag{6.1}$$

If (6.1) is met, all subcarriers inside the block show identical channel transfer factors. The overall behavior of the block can be described by a flat-fading narrowband channel. If the block size matches B_C, adjacent blocks of neighboring users will be nearly uncorrelated.

By contrast, the subcarriers in an equidistant approach are distributed over the whole bandwidth. If the inequality

$$N_D = \frac{N}{N_C} \ge \frac{B_C}{\Delta f} \tag{6.2}$$

is met, the distance between individual subcarriers of a user is larger than the coherence bandwidth and therefore the channel transfer factors of these subcarriers are highly uncorrelated. This is the general goal in equidistant allocation, since in this way the frequency diversity of the channel is fully exploited. In this case, the channel transfer factors $H_{i,k}$ a user i observes will follow a Rayleigh distribution [Pae99]. The subcarrier sets of the users in the system will show nearly identical channel properties. This is especially true for neighboring users, since all their subcarriers will be directly adjacent to each other.

6.1.2 Subcarrier Allocation Task

Since in the considered non-adaptive approaches no instantaneous CSI is present in the system, the only goal of a non-adaptive subcarrier allocation is the exploitation of the frequency diversity.

In case of a blockwise allocation, a user observes flat-fading inside a block. To achieve a diversity gain, each user must be provided with multiple blocks, which have to be sufficiently uncorrelated, i. e. separated in frequency. As explained earlier, the data of a user is interleaved and encoded in one stream and distributed over the user's assigned blocks. If a large number

of blocks is considered, bit errors tend to be uncorrelated and can be corrected by channel coding.

To increase the system's diversity, two possibilities exist: Allocate multiple blocks per user, or perform frequency hopping. In the case of frequency hopping, the subcarrier block of a user will be moved along the frequency axis following a predefined sequence and in regular time intervals. Diversity is achieved by distributing the data stream over multiple blocks in time. Since this approach leads to considerable delays in the data transmission, it will not be pursued further.

The second possibility to improve diversity is to assign multiple blocks per user. Because a fixed number of N_C subcarriers belongs to each user, the block size decreases as the number of blocks increases. Since each block shows flat-fading behavior, the blocks of a user are ideally placed in an equidistant manner along the frequency axis. This will lead to an uncorrelation between blocks. In an extreme case, the block size is reduced to a single subcarrier, which corresponds to the equidistant approach.

6.1.3   Simulation Results

This section evaluates simulation results for various non-adaptive subcarrier allocation approaches. An OFDM-FDMA transmission system with $N = 256$ and $N_U = 16$ is considered. Also, the parameters from Tab. 3 apply. The modulation scheme applied on all subcarriers is QPSK. The number of subcarriers per user is constantly $N_C = 16$ in all simulations. The data transmission of each user is modeled by a complete OFDM-transmission chain as described in section 4. This means, the data of each user is encoded, interleaved and transmitted from the BS over the radio channel, cf. Fig. 11. The subcarrier allocation block highlighted in Fig. 11 is realized by either blockwise or equidistant allocation in the following simulations.

The inverse operations including Viterbi decoding take place at the receiver side. The channel follows a WSSUS model with parameters given in Tab. 1.

The channel is assumed to be stationary for the duration of 20 OFDM symbols. After that, a new channel realization is generated. Subsequent channel realizations are assumed to be uncorrelated. In the frequency domain, the correlation between adjacent subcarriers is quite high. The WSSUS parameters in Tab. 1 lead to a coherence bandwidth of approximately $B_C \approx 312.5\text{kHz}$ (cf. (2.24)), which corresponds to the bandwidth occupied by four subcarriers. This correlation in the frequency domain can also be seen in Fig. 4, which shows a realization of the considered WSSUS channel. The considered scenario is that of the ring cell (cf. Fig. 12 (left)).

The first allocation approach to be considered is the blockwise approach. The blocks of each user are placed equidistantly along the frequency axis. This ensures sufficient decorrelation between blocks in all considered cases. The block sizes are identical between users. Additionally, all considered block sizes fulfill (6.1), which leads to a maximum block size of four for the considered WSSUS model. A block size of one results in an equidistant allocation ap-

proach. The subcarrier spacing in this case is $N_D = 16$ in the scenario at hand. For reasons of reference, also a random approach is considered, where the individual N_C subcarriers are allocated randomly along the frequency axis.

The results are depicted in Fig. 15. The plotted *bit error rate* (BER) represents the average BER of all users inside the system. Since the ring scenario is considered, the long-term performance is identical for each user and thus the BER performance of an individual user corresponds to the average performance of all users. The SNR given on the x-axis is the average SNR a user observes over the system bandwidth W. Since the WSSUS channel solely models small-scale effects, this average SNR is the same for all users in the ring scenario.

As indicated in the previous chapter, it becomes obvious from Fig. 15 that an increasing number of blocks N_B for each user also increases the frequency diversity and thus improves the performance. This is especially true for the considered WSSUS channel with its high correlation in the frequency domain, cf. Fig. 4.

The system where $N_B = 16$ offers the best performance, since the large number of individually placed subcarriers offers maximum diversity. The diversity gain is not lowered by correlation of the subcarriers because they are placed with maximum mutual distance on the frequency axis. This scenario, where N_B equals 16 corresponds to the equidistant allocation approach, since also (6.2) is fulfilled.
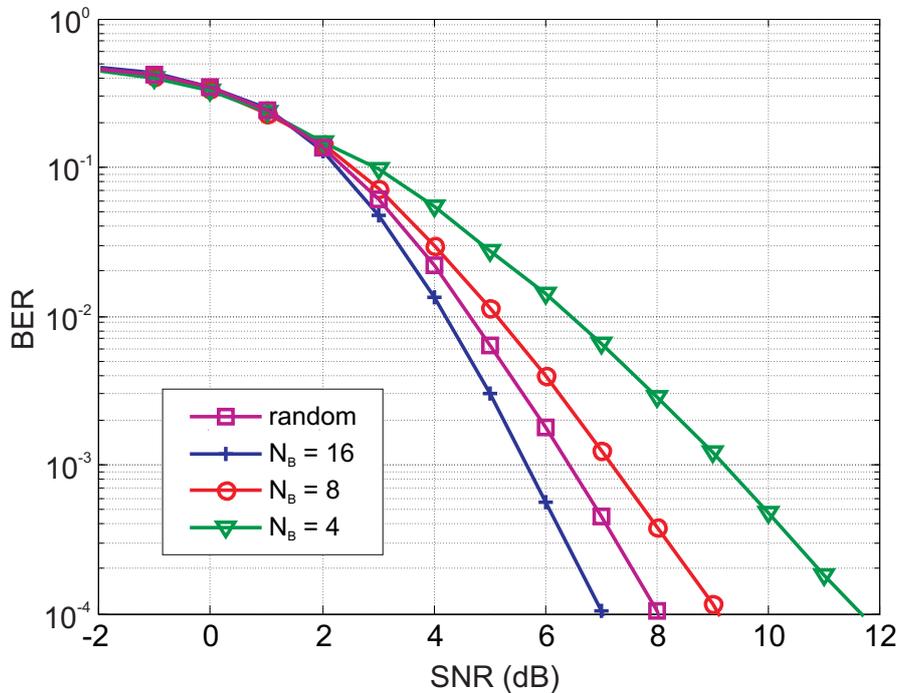
Fig. 15: Average BER for non-adaptive allocation; 16 users, QPSK, CR=1/2

As N_B decreases, the system performance declines accordingly due to declining diversity. This means, analogous to the decreasing N_B, the number of subcarriers inside a block gets larger. If $N_B = 4$ is considered, there are also four subcarriers inside one block. From the channel model it follows that all subcarriers inside a block show identical channel transfer factors. If now the channel exhibits a deep fade at the position of the block, four subcarriers are affected simultaneously. Thus, the performance of the system with $N_B = 4$ is degraded considerably. As a consequence, the curve for $N_B = 4$ (triangular markers) shows the worst performance of all systems.

An exceptional position is assumed by the random allocation (square markers): Here the subcarriers are also allocated individually like in the equidistant approach. Unlike the equidistant approach, the random allocation does not guarantee a maximum mutual distance between subcarriers. Thus, the allocated subcarriers are not sufficiently decorrelated in the considered scenario, which results in a slightly lower performance.

These results show that a large number of resources per user helps to exploit the frequency diversity of the channel. But they also show that once the highest level of diversity is reached, no further performance improvement is possible with non-adaptive approaches.

Due to this fact, the next section will show a much more promising approach for subcarrier allocation in OFDM-FDMA systems, which exploits the systems inherent multiuser diversity by employing CSI: Adaptive subcarrier allocation.

6.2 Adaptive Subcarrier Allocation

The preceding chapter showed that exploiting the frequency diversity of the radio channel is only a limited means for improving system performance. This is due to the fact that frequency diversity solely relies on channel statistics and does not use instantaneous channel state information. If instead the instantaneous SNR values $SNR_{i,k}$ for every user are known at the BS, the performance can be increased even further: Now each user can be assigned to its best subcarriers. In the following, such a channel adaptive assignment is called *adaptive subcarrier allocation* [Ste07][Gro03].

Fig. 16 shows a simple example for an adaptive subcarrier allocation: Based on the CSI at the BS, each user gets assigned to its best-performing subcarriers. The number of assigned subcarriers N_C is identical for each user. Since N_C is not always as naturally balanced as in this example, the allocation algorithm itself must provide for fairness in the allocation process [Roh05].

In a single-user system, the application of adaptive subcarrier allocation is straightforward: According to the user's rate demands, a sufficient number of subcarriers is reserved and the BS allocates those subcarriers to the user, which show the highest SNR values. Thus, like the non-adaptive subcarrier allocation, the adaptive approach also exploits the channel-inherent frequency diversity.
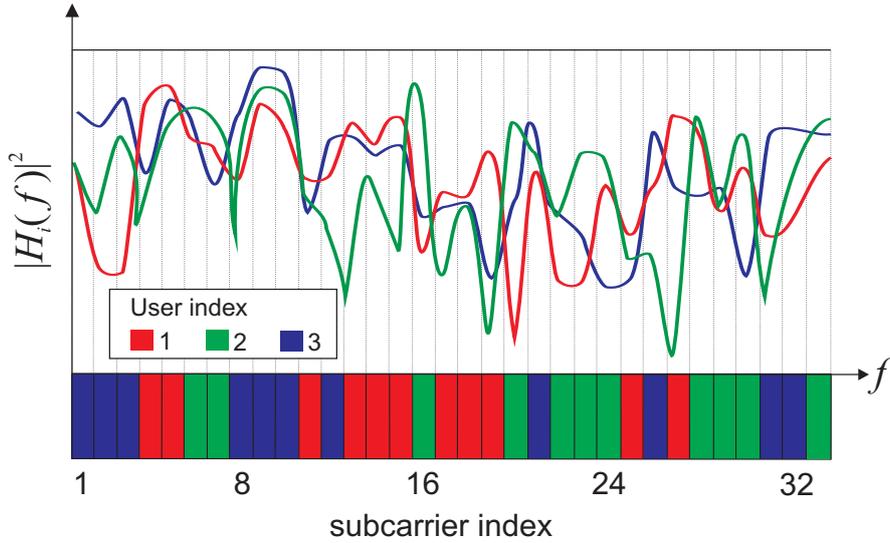
Fig. 16: Adaptive allocation example for 33 subcarriers and three users

In addition to frequency diversity, another type of diversity is observed in adaptive systems if multiple users are considered. This is the so-called *multiuser diversity*. Due to the statistical independence between radio channels of different users, a specific subcarrier k can possess a completely different channel transfer factor $H_{i,k}$ for each user i inside the system. This effect can be observed in Fig. 16. For adaptive approaches this effect is beneficial since a subcarrier with poor SNR for one user most likely shows a much better SNR for a different user. But MUD also raises issues not considered in the non-adaptive approaches: If the number of users increases, also the probability for two users observing the same subcarrier as their best resource is increasing. If this happens, a fair balance between users with respect to the number and quality of subcarriers must be found. This fairness strongly depends on the employed subcarrier allocation algorithm. Fairness is intrinsically tied to adaptive subcarrier allocation and is discussed in detail in further chapters. In general, the fairness aspect limits the performance achievable by utilization of MUD. Thus, in adaptive allocation approaches the system performance is not only determined by the channel statistics, but also by the number of users inside the system [Olo06].

Adaptive allocation schemes require detailed knowledge about the instantaneous channel situation of each user. The more detailed this knowledge is, the higher the performance gain through adaption will be. But on the other hand the effort that is necessary to measure and update the CSI increases. This is especially true in systems subject to time-variant channels, where the channel state must be updated in intervals shorter than the channel's coherence time. Since an adaptive system using outdated or even false CSI will eventually show the performance of a non-adaptive system. The requirement for detailed CSI is also one of the biggest drawbacks in comparison with non-adaptive schemes. Still, as shown in [Gro05], the performance gain through adaptive allocation generally outweighs the overhead produced by CSI updates.

In order to evaluate the maximum performance gain achievable by adaptive subcarrier allocation, ideal channel knowledge at the BS will be assumed in the following.

The non-adaptive schemes introduced in section 6.1 have in common that they don't need a complex algorithm for the allocation of subcarriers. Once the type of approach (e.g. block-wise) is decided, the actual mapping of subcarriers to users can be done according to a predefined scheme for every OFDM symbol and thus extremely low processing power is needed.

In case of adaptive subcarrier allocation, much higher effort must be spent for the calculation of an allocation result, basically due to the extensive processing of CSI. The computational complexity of an adaptive subcarrier allocation strongly depends on the employed algorithm, which is why these algorithms can be divided into two different classes:

- Optimal algorithms. These algorithms give an optimal solution to the subcarrier allocation task e.g. in terms of maximized system throughput. The optimization is based on the given CSI. Unfortunately an optimum solution mostly leads to excessive computational complexity.

- Heuristic algorithms. These algorithms also adapt the allocation according to CSI. The solution found is not necessarily optimal but requires much less computational effort.

This division in algorithm classes already indicates that computational complexity is both an issue and a possible drawback in adaptive allocation tasks.

Optimal subcarrier allocation algorithms generally are formalized using an objective function that is to be maximized or minimized respectively. To this objective, certain constraints can be added. The choice of the actual objective function depends on the performance goal of the system. For example, maximizing the sum of the system throughput is an objective that is often applied. The constraints are used to formalize restrictions inherent to the considered communication system. A possible constraint is e.g. the allowed number of subcarriers per user N_C.

If the objective and constraints can be expressed by a system of linear equalities and inequalities the optimization task is a so-called *linear programming problem* (LP problem) [Nem89]. A detailed introduction to formulation and solving of linear programming problems is given in later sections. At this point, it should be emphasized that most algorithms for the solution of LP problems are based on systematic searches inside the set of all feasible solutions for the problem at hand. Considering the problem of exclusively allocating N subcarriers to N_U users, which has a maximum of $(N_U)^N$ possible solutions, a search for the optimal solution can take excessive time even for relatively small values N and N_U. Although for some classes of optimization problems efficient solving algorithms exist, the application of optimal algorithms to adaptive subcarrier allocation is mostly used to establish upper performance bounds for given systems. In real-time systems, optimal allocation algorithms are generally not applied due to their computational complexity.

Nevertheless, the following chapters do consider optimal algorithms to show the performance limits of the introduced multiuser systems and to give a benchmark for the also considered

heuristic algorithms. The drawbacks of optimal algorithms in terms of complexity provide the motivation for the discussion of so-called suboptimal algorithms or heuristic algorithms.

If an optimal algorithm is substituted by a heuristic algorithm, the objective and constraints stay valid, but the search for a solution changes significantly. In general, heuristics try to find good solutions to given problems in a short time-span. A common approach to do this is to limit the search inside the set of possible problem solutions by eliminating a large part of this set. This can be done by using a-priori knowledge about the considered system. In the case of subcarrier allocation algorithms, the heuristic algorithm can include the fact that it is very likely for a user to allocate its best subcarrier. Then, the algorithm can start a search for a local optimum based on this assumption. Such local optima are mostly found with a relatively low effort in computation time and often provide good solutions to the problem at hand.

A drawback of heuristic algorithms is that there is no guarantee for a certain quality of the found solution. Also the margin to the optimal solution is often unknown. Thus, the quality of an applied heuristic must be evaluated empirically by comparing solutions to multiple problem instances to the corresponding optimum solutions.

In the following, an example for an optimum allocation algorithm is introduced and its performance is assessed by simulation results. The immediate focus for the time being lies on the comparison between general adaptive versus non-adaptive subcarrier allocation approaches. In order to give an upper performance bound for adaptive allocation, an optimum allocation approach is considered. A heuristic allocation approach is introduced in a later chapter.

6.2.1   Optimal Subcarrier Allocation

The task of optimal subcarrier allocation for a multiuser OFDM-FDMA system is formalized by an LP problem. The objective chosen for the considered ring scenario is to maximize the sum of allocated $SNR_{i,k}$ values, see (6.3). This corresponds to the allocation of subcarriers with preferably high signal-to-noise-ratio. The $SNR_{i,k}$ values follow the definition given in (5.2). The objective of maximizing the sum of $SNR_{i,k}$ values is chosen, since the system performance is evaluated in terms of bit error rate. Naturally, maximized $SNR_{i,k}$ values for each user will minimize the corresponding bit error rates and thus improve the system performance. A direct minimization of the BER by using it as objective is not possible, since the BER values are not known at the time the resource allocation is executed.

The allocation of a subcarrier k to user i is indicated by setting the corresponding decision variable $x_{i,k} = 1$. As a constraint, each user must get the same data rate. In the considered ring scenario, this can be achieved by providing each user with the same number of subcarriers N_C. The fact that each subcarrier in OFDM-FDMA is allocated exclusively to a user is formalized by (6.4). The complete formulation of the LP problem is given by (6.3) to (6.5). For the solution of this special case of an LP problem, efficient algorithms are known [Kuh55].

But still, the computational effort necessary for an optimum allocation is considerably higher than in non-adaptive approaches.

Another means to reduce this effort besides using heuristics is to bundle adjacent subcarriers to blocks. If the bandwidth covered by each block is smaller than the coherence bandwidth B_C, the performance of all subcarriers inside the block can be described by a single channel transfer factor $H_{i,k}$ or $SNR_{i,k}$ value respectively. This decreases the number of available resources and thus dramatically reduces the set of possible solutions of the LP problem. If the block size is increased beyond B_C, no longer flat-fading can be assumed inside the block. Additionally, the number of blocks covering the system bandwidth decreases, which limits the adaption capabilities of the allocation algorithm. In order to assess this effect, various block sizes are considered in the following.

The second effect to be assessed is multiuser diversity, which was introduced in section 6.2. This assessment is carried out by assuming various numbers of users N_U inside the cell. The adaptive subcarrier allocation is done subcarrierwise in this case to guarantee maximum diversity gain. The number of available subcarriers N is apportioned equally between the users. Results are given in the following section.

The complete LP-formulation of the applied optimum allocation algorithm is given below:

Objective:

$$\arg\max_{x_{i,k}} \sum_{i=1}^{N_U} \sum_{k=1}^{N} SNR_{i,k} x_{i,k} \quad \text{with } x_{i,k} \in \{0,1\} \tag{6.3}$$

Constraints:

Exclusive allocation of all subcarriers:

$$\sum_{i=1}^{N_U} x_{i,k} = 1 \quad \text{for all subcarriers } k \tag{6.4}$$

N_C subcarriers per user:

$$\sum_{k=1}^{N} x_{i,k} = N_C \quad \text{for each user } i = 1...N_U \tag{6.5}$$

6.2.2 Simulation Results

This section gives results for the LP-based optimal allocation approach introduced by (6.3) to (6.5). The considered system model as well as the utilized OFDM transmission chain is the same as in section 6.1.3 in order enable comparisons between adaptive and non-adaptive approaches. Thus, the simulation results are based on bit level simulations, which model the complete transmission chain. The system provides a total of $N = 256$ subcarriers. Again, QPSK modulation and coding with $CR = 1/2$ on all subcarriers is assumed.

First, the effect of different block sizes on the performance of an adaptive subcarrier allocation scheme is evaluated: A scenario with $N_U = 16$ users and $N_C = 16$ is considered. The corresponding results for various block sizes and thus various numbers of blocks per user N_B are shown in Fig. 17. As a reference, the curve for random subcarrier allocation from Fig. 15 is also drawn (square markers). The BER and SNR definitions from section 6.1.3 also apply.

Fig. 17 shows that the performances for the optimum adaptive allocation with $N_B \geq 4$ are almost identical since in these cases the size of the individual blocks is smaller than B_C. Here, the definition of B_C from (2.24) is applied. Thus, the individual subcarriers inside a block show almost identical SNR and can be treated as a single resource by the allocation algorithm. The results show, that this approach leads only to a negligible loss in multiuser diversity. The performance loss starts to be significant if $N_B = 2$ is assumed. In this case, the block size exceeds B_C and also the small number of blocks limits the flexibility of the allocation. This leads to a loss of about 1.3 dB compared to $N_B = 16$.

On the other hand, a decrease of N_B by a factor of two already leads to an exponential reduction of the set of LP solutions (cf. section 6.2) and thus to a considerable simplification of the allocation task.

In summary, the assembling of adjacent subcarriers to blocks is useful both in terms of performance and computational complexity, as long as the block size is kept below B_C. In this case, the small performance loss is outweighed by the significant reduction in complexity.

Apart from the effect of the block size, the more striking result of Fig. 17 becomes apparent in the direct comparison of the adaptive allocation results (e.g. cross-marked curve) with the non-adaptive (random) allocation result (square markers). The random allocation considered here is identical to the one carried out for the results shown in Fig. 15 (square markers). This means, the subcarriers allocated to a user i are merely selected at random from the N available subcarriers.

If the adaptive allocation is carried out subcarrierwise ($N_B = 16$), a significant performance gain of 7.0 dB is achieved compared to the random allocation approach. This impressively shows the beneficial effect of MUD, when each user observes an independent channel state with strong frequency selectivity. From the set of N resources available to each user, the N_C resources selected by the optimum allocation approach predominantly show $SNR_{i,k}$ values high above the average SNR level. This leads to the observed huge performance gain because each user mostly uses its best subcarriers for transmission.
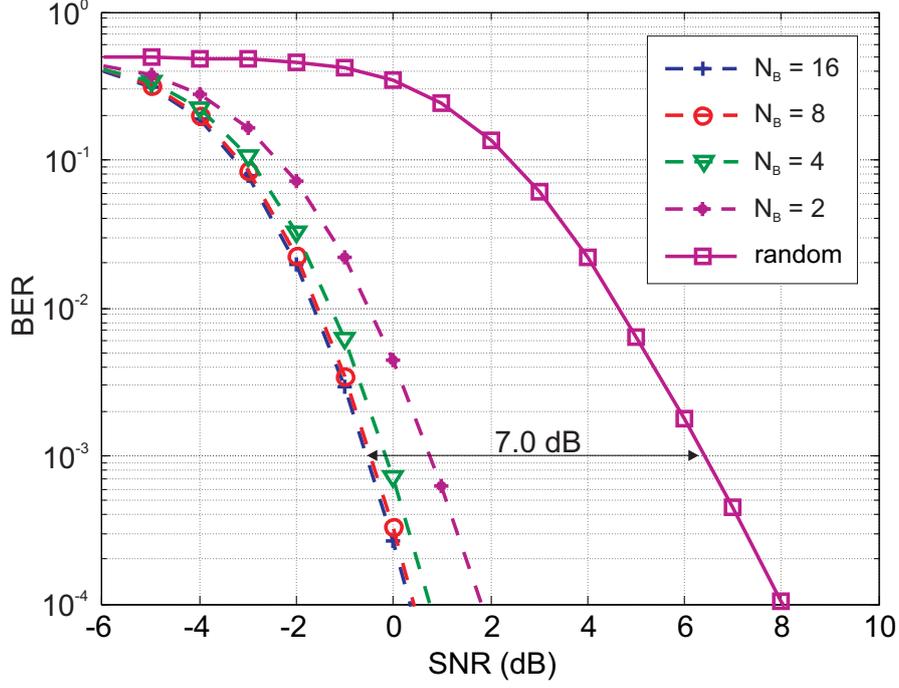
Fig. 17: Average BER for optimal allocation, 16 users, QPSK, CR=1/2

The diversity effect of multiple users becomes even more obvious, if the system performance of the adaptive allocation is evaluated for various user numbers N_U. This is depicted in Fig. 18. There, BER curves are shown for N_U = 4, 8, and 16. In each case, the $N = 256$ available subcarriers are divided equally between the users according to $N_C = N / N_U$. This means e.g. for $N_U = 4$ that each user obtains 64 subcarriers. For higher user numbers, the number of subcarriers per user decreases accordingly.

Each plot in Fig. 18 (a to c) shows a curve representing the optimal allocation scheme introduced by (6.3) to (6.5) and also a random allocation curve for reference.

A survey of Fig. 18 shows that the performance gap between an adaptive and a random subcarrier allocation increases for an increasing user number N_U. Consequently, the curves depicted e.g. in Fig. 18 a) for $N_U = 4$ show the lowest observed performance difference of 4.2dB. The plot shown in Fig. 18 c) is equivalent to the adaptive/random comparison in Fig. 17 and shows a gain of 7.0dB, which is the highest observed gain of all considered plots.

Two diversity effects are responsible for these performance gaps: Frequency diversity and multiuser diversity. The latter effect is dominating in this case.

From the two allocation approaches utilized in Fig. 18, solely the adaptive approach exploits MUD by mostly selecting that user i, which observes the highest $SNR_{i,k}$ on subcarrier k. If e.g. $N_U = 4$, the probability to find an exceptionally high $SNR_{i,k}$ on a particular subcarrier is naturally lower than for $N_U = 16$ where much more channel realizations are available to select from. Thus, a much lower average SNR is needed to achieve a BER of 10^{-3} if 16 users are

46

present inside the cell, cf. Fig. 18. The MUD manifests itself by the progressing left-shift of the adaptive allocation curve (dashed) in Fig. 18 a) to c). The performance gain achievable due to MUD is only limited by the channel statistics and the constraint that all users obtain an equal number of subcarriers N_C.
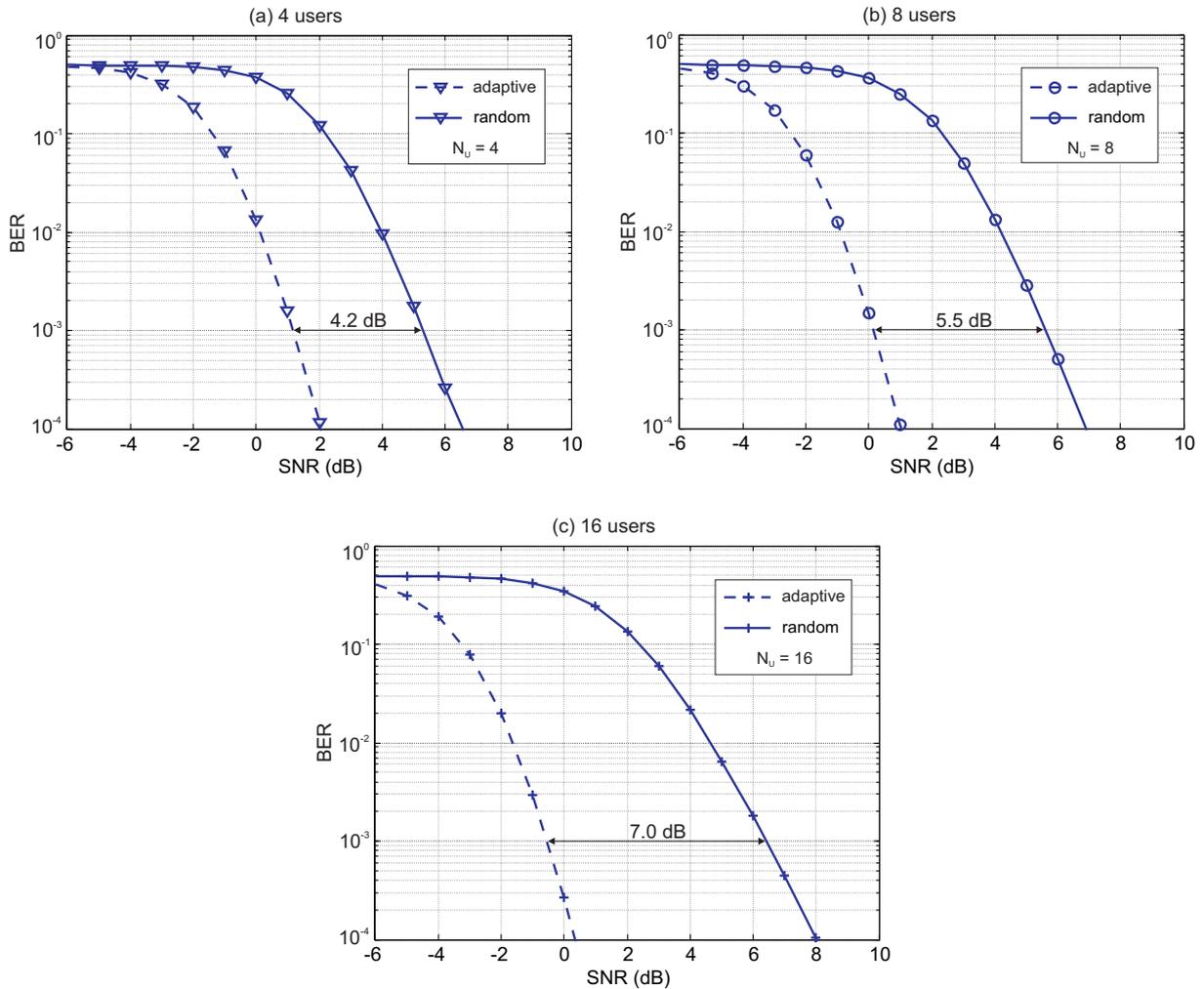
Fig. 18: Average BER for optimal and random allocation, QPSK, CR=1/2, various user numbers

Hence, a high number of users inside the cell causes allocation conflicts if two users observe a particular subcarrier as especially favorable. These conflicts are resolved by limiting N_C to a fixed value but this limit also impedes the achievable performance gain of the adaptive allocation approach. Thus, increasing the number of users N_U beyond $N_U = 16$ does not lead to a higher system performance.

The second effect contributing to the gap between random and adaptive allocation is the above-mentioned frequency diversity. Its impact is less dominant than that of the MUD but it is still noteworthy:

The effect of frequency diversity solely affects the random allocation and does that in a similar manner as pointed out in section 6.1. A detailed scrutiny of Fig. 18 a) to c) shows that the BER curve of the random approach is more and more shifted to the right as the number of users N_U increases from 4 to 16. This is due to the decreasing number of subcarriers each user has available for his data transmission. If $N_U = 16$ is assumed, only $N_C = 16$ subcarriers are available for each user. Thus, every subcarrier showing a low SNR value in this case has a much stronger performance impact than in the 4-user-case, where 64 different subcarriers are available per user.

The described frequency diversity effect is only limited by the channel statistics, which determines the correlation between individual subcarriers. Since this correlation is quite significant in the considered channel model, the frequency diversity stagnates for user numbers $N_U < 4$. In such cases, the number of subcarriers N_C per user is so high that the accompanying correlation between subcarriers substantially reduces the diversity.

In summary, two effects are responsible for the impressive performance gain of adaptive allocation compared to the random approach, namely MUD and frequency diversity. In the considered scenario, the MUD effect is dominating and reacts conversely to a changing N_U as does the frequency diversity. Thus, in scenarios utilizing adaptive allocation approaches, high user numbers are of benefit.

Due to the obvious advantage of adaptive subcarrier allocation with respect to system performance, this kind of approach will be pursued in further detail. Up to now, solely the influence of multipath propagation and thus frequency selectivity was considered in the multiuser scenario. The resulting effect of all users observing identical average SNR is not a sufficient model for scenarios, where the users are distributed over the cell area. In such a scenario, also channel effects like path loss and shadowing must be considered. Therefore, in the following chapters the additional influence of these channel effects on adaptive allocation approaches is evaluated.

7 Adaptive Resource Allocation in different Single-Cell Scenarios

Link adaption is an important element of system design procedures. The previous chapter showed a remarkable performance gain for adaptive resource allocation in comparison with non-adaptive allocation. This gain was achieved assuming the simple ring scenario, which solely considers multipath fading and leaves out path loss and shadowing channel effects.

In contrast, this chapter also considers the complete channel influence including path loss and shadowing. This corresponds to the uniform scenario with respect to the user distribution inside the cell.

Thus, the focus of this chapter lies on a general performance comparison between an adaptive and a non-adaptive (random) subcarrier allocation approach for a multiuser OFDM-FDMA system considering the *uniform scenario*.

In order to simplify the system simulation of the uniform cell scenario, a link performance model is introduced in the following. This model is validated against the simulation results from section 6.2.2.

The results presented in this chapter show a considerable increase in system performance also for the uniform scenario if adaptive subcarrier allocation is applied.

7.1 System Model

In the following, both system models introduced in chapter 5 are considered. This means, various resource allocation schemes are evaluated for the ring scenario as well as for the uniform scenario (cf. Fig. 12). In this case, the ring scenario is considered for mere validation purposes.

In both scenarios, perfect CSI is assumed at the BS. The BS allocates subcarriers for the downlink to the MTs inside the cell. To this end, random and optimum resource allocation are considered. To keep the results comparable between scenarios, all users in both scenarios are provided with the same number of subcarriers $N_C = N / N_U$. Since only the effects of adaptive subcarrier allocation are to be evaluated, identical modulation and coding schemes (PHY mode) are applied on each subcarrier. Each subcarrier k is allocated exclusively to a user i.

7.2   Subcarrier Allocation Task

Two subcarrier allocation schemes are considered in the following: a non-adaptive and a channel-adaptive scheme.

In order to keep simulation results comparable, all simulations utilize identical PHY modes for each user. Also the PHY modes are not adapted to the instantaneous channel situation.

The non-adaptive subcarrier allocation scheme is straightforward and randomly allocates to each user a number of N_C subcarriers from the available N subcarriers. Thus, multiuser diversity has no influence on the system performance. This allocation scheme is referred to in the following as *random approach* and is identical to the scheme introduced in section 6.1.

The adaptive scheme also assigns N_C subcarriers to each user, but it takes into account each user's individual channel situation. This channel situation is described for each subcarrier *k* of user *i* by a signal-to-noise ratio $SNR_{i,k}$. The adaptive allocation is implemented as an optimization task, which maximizes the sum of the $SNR_{i,k}$ of all assigned subcarriers. This optimization task is formalized as given by (6.3) to (6.5). In order to maximize the gain by MUD, the block size is set to one, i. e. the subcarriers are allocated individually. This adaptive allocation scheme is referred to in the following as *adaptive approach*. The introduced allocaton approaches (adaptive & random) are equivalent to the approaches used in section 6.2.2.

7.3   Link Performance Model

According to the introduced system model with an identical number of subcarriers per user and identical PHY modes, each user is provided with the same data rate. Thus, the performance of the random and adaptive approach is compared in terms of bit error rate for each link.

In section 6 all simulations were carried out using *bit level simulation*, where all components of the OFDM chain (e.g. coding, modulation, FFT, etc.) were processed inside the simulation system.

In this chapter, a higher layer of abstraction is utilized for system simulations in order to keep the simulation complexity for the uniform scenario reasonable. To this end, a *link performance model* is introduced, which maps the set $\{SNR_{i,k}\}$ of individually allocated $SNR_{i,k}$ of each user *i* to a user-specific bit error rate $BER_{E,i}$. Thus, the signal processing inside the transmission chain as applied in 6.1.3 and 6.2.2 is replaced by a simple mapping procedure: $\{SNR_{i,k}\} \mapsto BER_{E,i}$.

The performance of the system is evaluated in terms of BER in order to keep the results of the link performance model comparable to the results obtained in chapter 6. In the following, $BER_{E,i}$ represents the bit error rate *estimated* by the link performance model, while a value $BER_{M,i}$ is a *measured* bit error rate derived from bit level simulations.

It is shown in the following that such a simplified link performance model is sufficient to accurately describe the essential system behavior and performance. This is done by comparing the *measured* BER results from chapter 6 to the BER results *estimated* by the introduced link performance model.

The considered link performance model performs the actual mapping $\{SNR_{i,k}\} \mapsto BER_{E,i}$ in two steps:

First, all $SNR_{i,k}$ values allocated to user i are mapped to a single so-called *effective SNR* value: $SNR_{eff,i}$. This $SNR_{eff,i}$ value accounts for the instantaneous channel quality and considers both the average *SNR* of user i and the frequency selectivity of the channel. The mapping $\{SNR_{i,k}\} \mapsto SNR_{eff,i}$ is done according to

$$ SNR_{eff,i} = -\beta \ln \left[\frac{1}{N} \sum_{k=1}^{N} \exp \left(-\frac{SNR_{i,k} x_{i,k}}{\beta} \right) \right]. \tag{7.1} $$

The sum in (7.1) contains only N_C non-zero elements determined by the decision variables $x_{i,k}$ (cf. (6.3)), which represent the subcarriers allocated to user i. In principle, if all values $SNR_{i,k}$ of a user are identical, i. e. flat fading is observed, the $SNR_{eff,i}$ value will match the $SNR_{i,k}$ values. However, if one subcarrier undergoes a deep fade, the effective SNR will also decline. As (7.1) suggests, the relation between $SNR_{i,k}$ and $SNR_{eff,i}$ is highly non-linear, since the relation between the SNR on the subcarriers and the bit error performance is also non-linear. Because the robustness of the transmission system to frequency selective fading strongly depends on the applied modulation and coding scheme, the parameter β is used to adapt the mapping function (7.1) in this respect. Further details about the mapping function and link performance models can be found in [Cam06], [Bru05], and [Eri03].

The second step is the mapping of $SNR_{eff,i} \mapsto BER_{E,i}$. This is done by using the AWGN performance curve ($SNR \mapsto BER$) of the considered system. This curve is depicted in Fig. 19 together with a simple mapping example: If in the considered system an $SNR_{eff,i}$ of 2 dB is observed for a certain user, the transmission of this user reaches a bit error rate of $BER = 10^{-2}$. The AWGN curve in Fig. 19 is produced by bit level simulation of the OFDM transmission system outlined in Tab. 3, assuming QPSK modulation coded with a convolutional code of $CR = 1/2$ and a pure AWGN channel.

The AWGN performance curve is a suitable choice for this mapping since in the AWGN channel the parameter β has no impact on $SNR_{eff,i}$ and additionally $SNR_{eff,i}$ equals the instantaneous values $SNR_{i,k}$. Thus, the system's BER performance in any channel situation can be derived from the system's AWGN performance using (7.1).
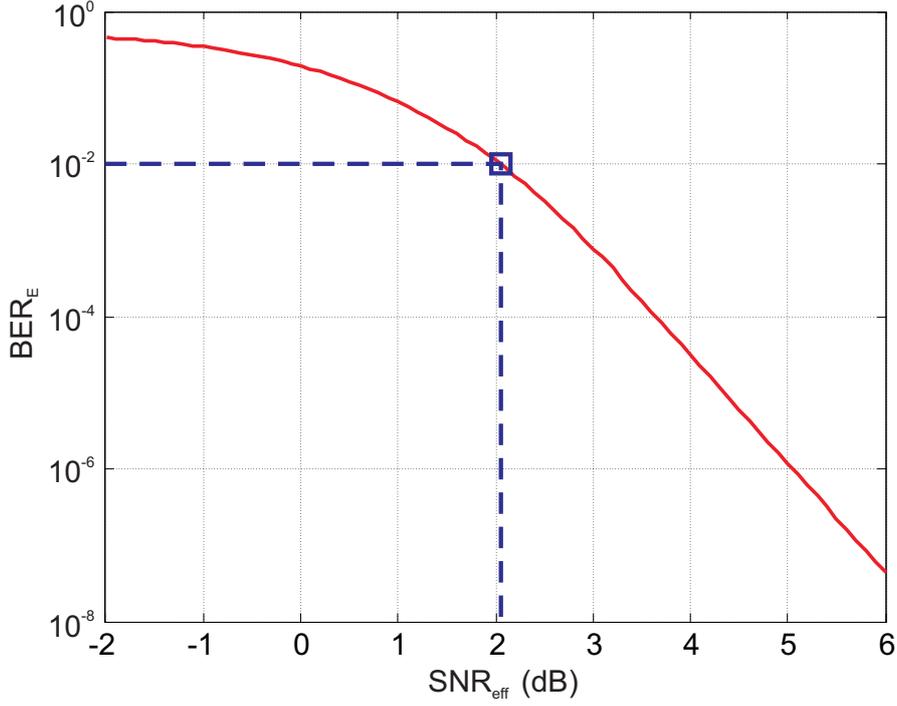
Fig. 19: AWGN performance curve used for mapping of effective SNR to estimated BER

To match the chosen link performance model to the considered transmission system, a suitable value for β must be computed. As indicated above, β solely depends on the applied modulation and coding scheme but is independent of the applied channel model. Calculation of β is achieved by minimizing the difference between the measured and estimated bit error rate in a least-squares sense. The minimization is formalized as follows:

$$\beta_{opt} = \arg\min_{\beta} \sum_{c=1}^{C} \left| \Delta e_c \left(\beta \right) \right|^2 \tag{7.2}$$

$$\Delta e_c \left(\beta \right) = \log_{10} \left[BER_{E,c} \left(\beta \right) \right] - \log_{10} \left[BER_{M,c} \right] \tag{7.3}$$

Here, the value of C determines the number of considered channel situations used for the computation of β. The BER difference between estimation and measurement is evaluated in the logarithmic domain (cf. (7.3)) since this yields a better match between estimation and simulation for the BER region of interest.

The $SNR_{i,k}$ values serving as input to (7.1) are derived from the stochastic channel model given in chapter 5. The link performance model is valid both for the uniform and the ring scenario. Depending on the actually considered scenario, the used definition of $SNR_{i,k}$ is either (5.2) or (5.4). Actually, the definition in (5.2) is just a special case of (5.4) with $G_i^{SH} = 1$ and identical path loss G_i^{PL} for all users, whose value depends on the radius of the ring. The complete set of channel parameters can be found in Tab. 1 and Tab. 2.

In the considered channel model the overall bandwidth is divided into $N = 256$ subcarriers. All users employ QPSK modulation and a convolutional coder with code rate $CR = 1/2$ (cf. Tab. 3).

To fit the parameter β of (7.1) to the PHY-mode considered above, bit level simulations considering the complete transmission chain are processed for a single user scenario (w.l.o.g.: user 1) and setting $G_1^{SH} = G_1^{PL} = 1$. This corresponds to the ring scenario from Fig. 12 (left) with a single user. Since $E\left\{\left|H_{i,k}\right|^2\right\} = 1$ and $E\left\{\left|X_{i,k}\right|^2\right\} = 1$ holds, an average SNR value $SNR = \left(\sigma_N^2\right)^{-1}$ can be defined for the system. This leads to the $BER_{M,1}$ curve (circular markers) in Fig. 20. The curve representing $BER_{E,1}$ (square markers) is produced using the described mapping procedure $\left\{SNR_{i,k}\right\} \mapsto BER_{E,i}$ for a single user. The value for β is optimized using (7.2) and (7.3), which gives a value of $\beta = 2.11$ for the applied PHY mode.

The resulting BER curves for the estimated $BER_{E,1}$ and the $BER_{M,1}$ measured by bit level simulation are shown in Fig. 20. The $BER_{E,1}$ curve of the link performance model (square markers) fits very close to the measured curve $BER_{M,1}$ (circular markers) especially for relevant values $BER \leq 10^{-3}$.
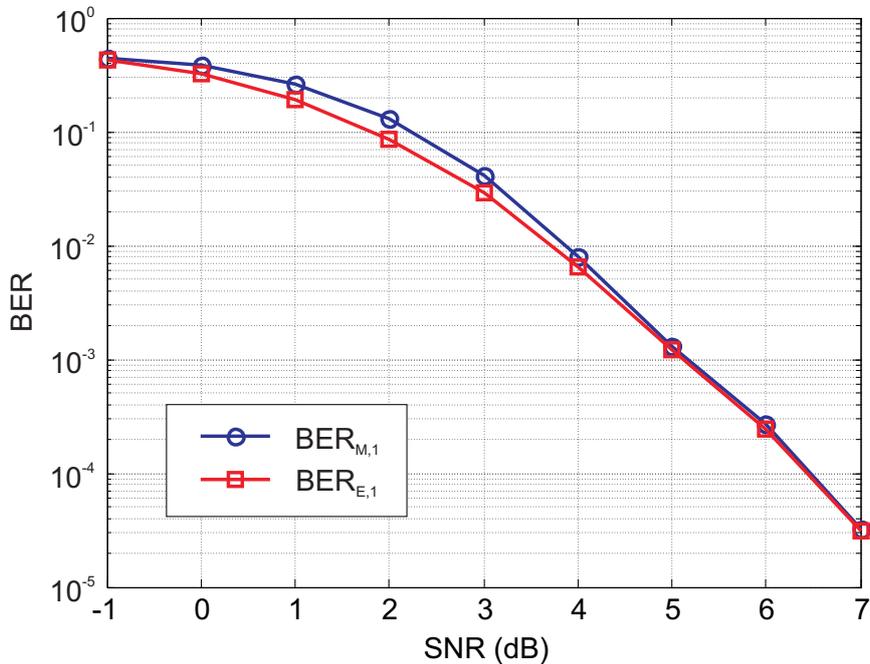
Fig. 20: Difference between measured BER_M and estimated BER_E

The value found for β in the single user scenario solely depends on the chosen PHY-mode (QPSK, $CR = 1/2$) and is also valid for the multiuser scenarios discussed in the following

sections. In the multiuser scenarios, the statistics of the $SNR_{i,k}$ for the individual users will be different from the single user case, but as shown in [Bru05] and validated by own simulations, this has no impact on the chosen value for β.

7.4 Performance of Adaptive Subcarrier Allocation

Using the introduced link performance model, the performance of both random and adaptive subcarrier allocation in any multiuser scenario can be evaluated in terms of bit error rate per user $BER_{E,i}$. In all following simulations in the sections 7.4.1 and 7.4.2 a fixed number of users $N_U = 16$ is considered. The other channel parameters are identical to those introduced in section 7.3.

The simulation results are evaluated in such a way that the BER performance is plotted over an average SNR. In all cell scenarios, an average SNR for each user i can be defined as SNR_i, based on the normalizations $\mathrm{E}\left\{\left|H_{i,k}\right|^2\right\} = \mathrm{E}\left\{\left|X_{i,k}\right|^2\right\} = 1$ and $E\left\{G_i^{SH}\right\} = 1$ (cf. (5.2) and (5.4)).

Equation (7.4) gives the definition for SNR_i, which represents the average quality of a user's instantaneous channel without considering its frequency selectivity.

$$SNR_i = \frac{G_i^{PL}}{\sigma_N^2} \tag{7.4}$$

Due to its proportionality to the path loss factor G_i^{PL}, the SNR_i strongly depends on the actual position of each user i inside the cell.

7.4.1 Simulation Results for Validation of the Link Performance Model

This section is intended as a validation for the accuracy of the recently introduced link performance model. Thus, performance results of a random and an adaptive subcarrier allocation approach are shown, which are based on the same ring cell scenario and system parameters as the results presented in section 6.2.2. Only this time, the results are gathered using the link performance model given by (7.1) instead of using bit level simulations.

Thus, the results presented in this section allow a direct comparison with the results from section 6.2.2. It is shown in the following that the introduced link performance model is an accurate means to describe the impact of the considered subcarrier allocation schemes on the transmission system.

The considered ring scenario is described in section 7.1 and depicted in Fig. 12 (left), meaning that all MT have the same distance to the BS.

Using the link performance model from section 7.3 and the fact that all users i observe an identical SNR_i, the performance evaluation is straightforward. For a SNR region of interest, a

sufficiently high number of channel situations is calculated for each user. The $SNR_{i,k}$ values found after application of both subcarrier allocation approaches can be mapped to $BER_{E,i}$ values for each user (cf. (7.1) and Fig. 19). Since all users experience the same average channel performance, these $BER_{E,i}$ results can be averaged for identical SNR_i.

The adaptive subcarrier allocation approach follows the LP-optimization described by (6.3) to (6.5). The random approach is identical to the one introduced in section 6.1. The allocation is done subcarrierwise, i. e. the block size is set to one.
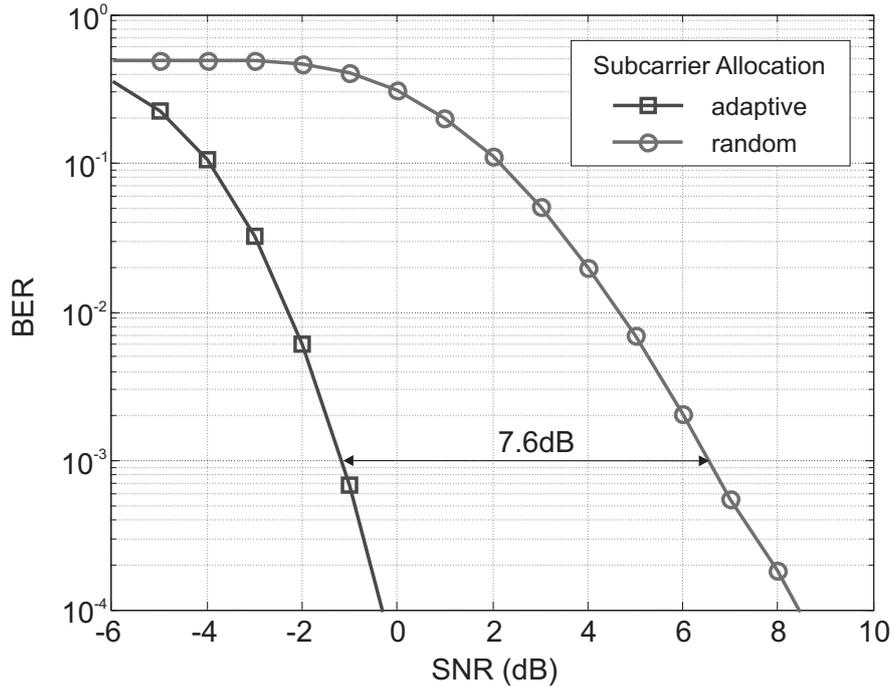
Fig. 21: Average BER for ring cell scenario, 16 users, QPSK, CR=1/2

The results of these simulations are shown by the performance curves depicted in Fig. 21. The ordinate axis shows the estimated BER_E averaged for all 16 users. The abscissa shows the averaged SNR values according to (7.4).

A subcarrier allocation, which uses the adaptive approach (square markers) yields a performance gain of 7.6dB compared to the random subcarrier allocation (circular markers). This means, the SNR_i a user i needs to achieve a certain bit error rate can be 7.6 dB lower in average if a channel adaptive subcarrier allocation scheme is applied. Thus a processing gain of 7.6 dB is achieved. These results gathered by the link performance model are consistent with the results achieved by bit level simulations (cf. Fig. 18 c).

A detailed comparison of Fig. 18 c) and Fig. 21 shows that the link performance model indicates a slightly higher processing gain (0.6dB) for the adaptive allocation scheme compared to the bit level simulations. Further scrutiny reveals that the curves for the random scheme in Fig. 18 c) and Fig. 21 almost exactly match. Only the curve for the adaptive scheme in Fig. 21

55

indicates a better performance than in the bit level simulations. This is due to the changed statistics of the $SNR_{i,k}$ values, which is caused by the channel adaptive allocation of subcarriers. In this case, the mapping $\{SNR_{i,k}\} \mapsto BER_{E,i}$ is a little less accurate.

Nevertheless, the introduced link performance model is accurate enough to describe the influence of various subcarrier allocation schemes on the system performance.

Thus, in the following section the performance impact of channel adaptive subcarrier allocation is evaluated for the uniform scenario using the link performance model.

7.4.2   Simulation Results for the Uniform Scenario

In the uniform scenario (cf. Fig. 12, right) the users are distributed uniformly over the cell area. The corresponding system and channel parameters given in section 5 are applied in the simulation. The only difference being that the maximum distance from an arbitrary MT to the BS is set to $d_{\max} = 400m$. In such a way, the range of observed SNR_i values allows easier comparison with previous results.

Due to the uniform user distribution inside the cell, each user observes a different path loss factor G_i^{PL} and thus an individual SNR_i. The performance evaluation for the optimum adaptive- and the random allocation approach in the uniform scenario is therefore slightly different from the ring scenario.
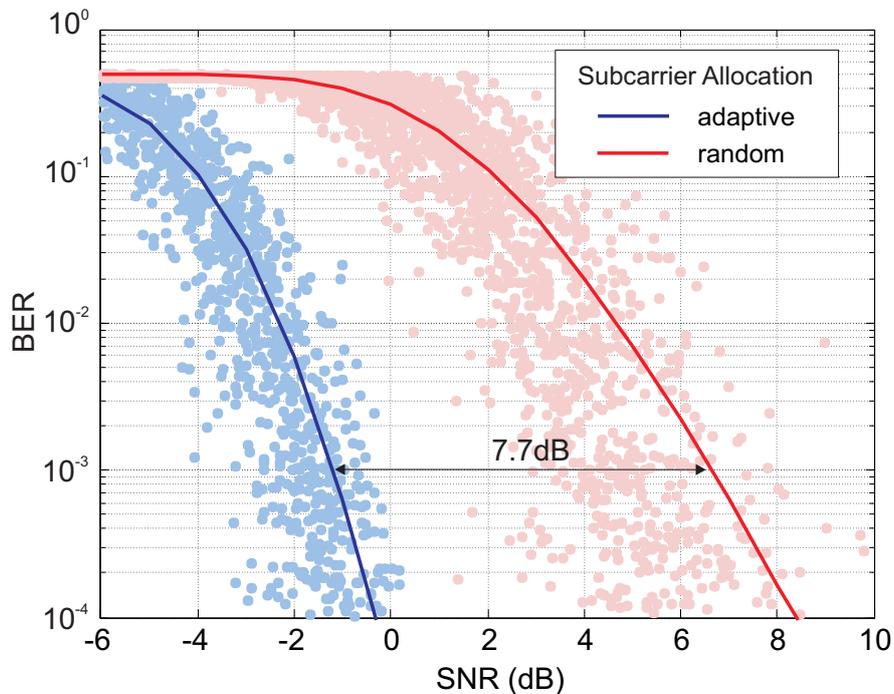
Fig. 22: Average BER for uniform cell scenario, 16 users, QPSK, CR=1/2

For the 16 users inside the cell, a large number of cell instances together with associated channel situations are simulated. For each user i in each cell instance, an individual $BER_{E,i}$-value can be computed for both allocation approaches (adaptive and random) according to section 7.3. The applied subcarrier allocation approaches are the same as in section 7.4.1.

By collecting these $SNR_i \mapsto BER_{E,i}$ data points for all users and all cell instances in a plot (cf. Fig. 22), it is possible to evaluate the average performance for both adaptive and random allocation approach, namely by averaging all $BER_{E,i}$ values that belong do different users, but possess identical SNR_i values. These averaged bit error rate values yield the solid curves in Fig. 22. Again, the ordinate axis shows the estimated BER_E. The abscissa shows the SNR_i values averaged over all users (cf. (7.4)).

Each dot in Fig. 22 represents the $BER_{E,i}$ for an individual user in a distinct cell instance. The light blue dots are the results of the adaptive subcarrier allocation, the light red dots represent a random allocation. The solid curves show the averaged bit error rates for the respective dot set. Of course, the figure shows only a limited section of the large SNR range inside the uniform scenario. This particular SNR range was chosen to enable a direct comparison between the figures Fig. 21 and Fig. 22.

The performance gain in the uniform scenario due to adaptive subcarrier allocation is considerable and almost matches the value reached in the ring cell scenario (cf. Fig. 21 and Fig. 22).

The close resemblance between the results from the ring and the uniform cell stems from the fact that the results of the simulations are evaluated on a user-by-user basis. In the ring cell scenario, the SNR_i corresponds to a certain ring radius and thus to a certain distance of a user to the BS. Hence, a user in the uniform scenario, which observes an identical distance to the BS, will experience an identical SNR_i and therefore in average a similar BER performance.

Another interesting finding of Fig. 22 is the different scattering of data points for the random and the adaptive allocation scheme. In the adaptive case, the scattering is considerably lower. This is due to the fact that the adaptive allocation preferably selects subcarriers with high $SNR_{i,k}$ values and thus reduces the variance of the system performance.

7.5   Conclusion

The presented results show that generally a large performance gain can be expected from an adaptive subcarrier allocation approach if the radio channel provides MUD. This gain is independent from the user distribution inside the cell, if the performance is evaluated on a user-by-user basis. Although the comparison of both cell scenarios in terms of BER is quite unusual, it shows the universal applicability of adaptive schemes in MUD scenarios.

So far in all simulations, an identical number of subcarriers for each user and identical PHY modes were assumed. These assumptions allowed straightforward comparisons between vari-

ous cell scenarios and allocation approaches. But at the same time, these assumptions limit the performance gains achievable using the introduced adaptive allocation approach. This is especially true for the uniform scenario: Due to the strongly varying path loss effect on each user, this scenario offers even higher MUD than the ring scenario and thus a high potential for adaptive resource allocation schemes. If e.g. an individual number of subcarriers was allowed for each user, the number of allocated subcarriers could be increased for users who suffer from strong path loss in order to keep their link performance.

To exploit this potential, a higher flexibility in the resource allocation of each user is necessary. Thus, the following chapters will introduce subcarrier/resource allocation schemes, which are more suitable for distributed user scenarios.

The abandonment of the restricting assumptions with respect to number and utilization of resources per user not only enables a higher flexibility of resource allocation but also increases the complexity of the resource allocation task (cf. [Kim01]) and raises the issue of fairness. So far, this issue was avoided by assuming $N_C = \text{const}$ for all users.

In this context, fairness deals with the fair distribution of resources between users. The decision, which system parameters are included into this resource distribution process strongly influences the performance of the overall system and the individual users as well. Thus, the aspect of fairness is covered in detail in the subsequent chapters.

8 Adaptive Resource Allocation under Fairness Considerations

The previous results showed large performance gains for an adaptive resource allocation and for all considered cell scenarios.

This chapter introduces a new aspect to the discussion of adaptive resource allocation schemes: The aspect of fairness [Roh05]. This aspect must be considered because in multiuser systems, the channel quality of a certain user does not always correspond to the user's transmission demands. This creates a classical tradeoff situation for the system provider: On the one hand, the provider wants to utilize a cell economically by assigning most resources to users with good channel quality, on the other hand, also users with low SNR (e.g. users at the cell border) must be served in order to provide good coverage.

In the considered context, fairness concerns resource allocation in such a way, that each user gets a fair share of the available system resources.

The previous results on adaptive resource allocation were based on BER as a performance measure. This led to a good comparability of various resource allocation schemes. But as seen in section 7.4.2, the expressiveness of this performance measure is reaching its limits if systems with a strong variance in user performance are concerned.

Thus, from now on, the objective of the adaptive resource allocation is changed from maximizing the sum SNR to maximizing the system throughput. Also, in order to evaluate the upper performance bounds of adaptive resource allocation the constraint of allocating equal numbers of subcarriers N_C is dropped.

Since simple throughput maximization is a highly unfair approach, different approaches to provide a certain level of fairness inside the system are considered in the following.

The discussion of adaptive resource allocation is also extended by the aspect of complexity. Therefore, two different approaches to solve the task of adaptive resource allocation are introduced and compared in terms of complexity and performance. The first approach uses an optimal algorithm for resource allocation, the second approach uses heuristics.

The optimal algorithm is the so-called *Binary Integer Programming* (BIP), which is a well known technique to solve optimization tasks with constraints, where the decision variables are binary integers. This technique leads to optimal solutions at the cost of exponential computation complexity.

The second approach will be called *Heuristic Evaluation* (HE) and is a heuristic approach which leads to non-but-close-to optimal solutions. The HE approach is a new allocation scheme and is covered in elaborate detail in this work. One advantage of the heuristic ap-

proach is its considerable smaller computational complexity. In the following chapters, the considered subcarrier allocation task and two solution approaches (BIP and HE) are discussed and the results will be compared in detail.

8.1 Subcarrier Allocation Task

Since it was the uniform cell scenario that brought up the aspect of fairness, this scenario is covered exclusively from now on. The discussion of fairness for the ring scenario is trivial and thus this scenario is abandoned.

The considered cell scenario is depicted in Fig. 23, which corresponds to the right-hand side of Fig. 12. The system parameters from chapter 5 apply. The radio channel is assumed to suffer from multipath fading, which leads to a strong frequency selectivity. Additionally, the varying distances between BS and MT cause path loss and shadowing effects.
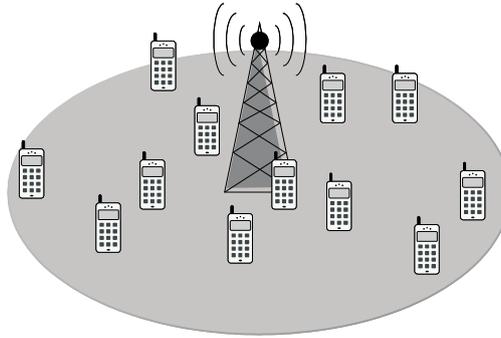
Fig. 23: System Model

The user-specific subcarriers are allocated to each user by the BS in an exclusive and individual way. In contrast to chapter 7, the constraint of allocating an equal number of N_C subcarriers to each user is dropped. This increases the flexibility of the allocation process and allows the process to take into account the individual channel situations of the users more readily.

Also the objective of the resource allocation is changed from maximizing the sum SNR to maximizing the system throughput. This is done because in contrast to the BER performance measure, the throughput can be maximized directly and gives a much better insight into the system performance if the uniform scenario is considered. This is especially true if various numbers of users inside the cell are evaluated. The corresponding system parameters are introduced as follows:

Each subcarrier k allows the transmission of a certain data rate $R_{i,k}$, which is individual for each user i, since each user will observe a different signal-to-noise ratio $SNR_{i,k}$ on the subcarrier k. The $SNR_{i,k}$ is defined according to (5.4). The data rate $R_{i,k}$ is calculated as the information-theoretical channel capacity on the considered subcarrier, see (8.1). The sum of

data rates $R_{i,k}$ that are allocated to user i yields the overall data rate R_i for this user. The allocation of a subcarrier k to a user i is denoted by the binary decision variable $x_{i,k}$. Since OFDM-FDMA is considered, a subcarrier is always allocated to one user exclusively, see (8.3). Additionally, (8.2) introduces the overall system throughput R as the sum of the user data rates R_i for all users.

$$R_{i,k} = \log_2\left(1 + SNR_{i,k}\right) \tag{8.1}$$

If the subcarrier allocation would solely be based on the objective (8.2), the solution to the optimization task would be trivial: The users, which are closest to the BS and thus observe high SNR would get all subcarriers, while users at the cell border would be left without service. Thus, a minimum data rate R_B is introduced that guarantees basic connectivity for all users, cf. (8.4).

The objective of the considered subcarrier allocation problem is to maximize the overall system throughput R while satisfying the rate constraints R_B of each user. The rate constraint R_B is introduced as a means to guarantee a minimum level of fairness inside the system.

A possible application for this specific subcarrier allocation problem would be a wireless service provider who tries to maximize its cell throughput since billing is based on the throughput of users. In order to provide service inside the whole cell coverage and thus reach a high number of customers, a minimum data rate R_B is guaranteed for each user, which allows basic services like voice telephony. This means, each user i should get at least a data rate R_B to be satisfied. Of course, the individual user's rate R_i can be higher than R_B in order to maximize the system throughput. The only degree of freedom in this optimization task is the allocation of suitable subcarriers to each user. This optimization task is formalized as follows:

Objective function:
Maximize system throughput

$$R = \sum_{i=1}^{N_U} R_i = \sum_{i=1}^{N_U}\sum_{k=1}^{N} R_{i,k} x_{i,k} = \max_{x_{i,k}} \quad \text{with } x_{i,k} \in \{0,1\} \tag{8.2}$$

Constraints:
Exclusive allocation of all subcarriers:

$$\sum_{i=1}^{N_U} x_{i,k} \leq 1 \text{ for all subcarriers } k \tag{8.3}$$

Minimum data rate for each user:

$$R_i = \sum_{k=1}^{N} R_{i,k} x_{i,k} \geq R_B \quad \text{for each user } i = 1...N_U \tag{8.4}$$

$$\text{with } R_{i,k} = \log_2\left(1 + SNR_{i,k}\right)$$

Equation (8.2) describes the objective function. The constraint (8.3) states, that each subcarrier is assigned to one user exclusively and can not be shared between users. Constraint (8.4) defines the rate constraint R_B, which must be fulfilled or exceeded by each user. This constraint guarantees a minimum level of fairness for the subcarrier allocation task at hand. In the following, the formulation of the above optimization task for a specific channel realization and therefore with given values $R_{i,k}$ will be called a problem instance.

Two different approaches to solve the considered optimization task will be introduced in the following paragraph. The first approach uses binary integer programming (BIP); the other employs heuristics (HE).

8.1.1 BIP Approach

The Binary Integer Programming approach is used to solve optimization tasks where the decision variables $x_{i,k}$ are binary integers [Nem89]. In the presented context, the decision variable $x_{i,k}$ declares if a subcarrier k is allocated to a user i or not. The BIP approach solves the optimization task formulated by (8.2) to (8.4) and gives an optimal solution if one exists. In the considered context, an optimum solution is a subcarrier allocation for the $x_{i,k}$ which maximizes the overall cell throughput R (8.2) and at the same time fulfills both constraints (8.3) and (8.4). The search for suitable values $x_{i,k}$ leads to a binary search tree, in which each node represents a possible subcarrier allocation for the cell. With the adept application of the constraints and partial solutions, some branches of such a search tree can be neglected, but still most problem instances consume a lot of computation time for the search through the tree [Mao08]. A thorough introduction to Binary Integer Programming can be found in [Nem89].

Due to the user distribution inside the cell (cf. Fig. 23), users who are close to the BS will observe a high average SNR, while users at the cell border observe a low average SNR. Together with the objective to maximize the cell throughput R, this will always lead to a subcarrier allocation, where users at the cell border will get a minimum number of subcarriers (with preferably high SNR) in order to achieve the basic data rate R_B, cf. (8.4). Users which are close to the BS will get as many subcarriers as possible in order to maximize the overall cell throughput R. Thus, the maximization of the cell throughput R and the fulfillment of the R_B constraint are two conflicting tasks, since subcarriers assigned to users with low SNR in order to reach R_B make only a limited contribution to the throughput R. This aspect is discussed in section 8.2.

The advantage of the BIP approach is that it finds a provable optimum solution to an optimization task, if such a solution exists. Its drawback is the high computational effort that has to be spent in finding this solution [Mao08].

The following paragraph will introduce a computational efficient alternative to the BIP approach: The heuristic HE approach.

8.1.2 HE Approach

The reason for the high computational complexity of the BIP approach is its search for a global optimum, which means that for each change in the allocation the influence to the final solution must be calculated.

In [Mao08], [Zho07], [Hoo04], and [Gro03] it is shown, that solving an optimization problem by using heuristics can yield close-to-optimum results if the specific structure of the problem is well understood. At the same time, heuristics generally require much less computational effort than optimum solutions. Based on these facts, a heuristic allocation approach is introduced in the following, that relaxes the computational burden of the resource allocation task by breaking up this overall task into smaller parts.

The Heuristic Evaluation approach (HE) executes the resource allocation task by means of an iterative procedure, cf. [Ste08]. In each iteration step only a partial solution to the overall allocation task is considered. It is shown in the following, that these partial solutions can be calculated very efficiently. This computational efficiency is a big advantage in comparison to the BIP approach.

Like the BIP approach, the HE approach tries to find a solution to the optimization task defined by (8.2) to (8.4), but this solution – in contrast to the BIP solution – is not guaranteed to be optimum. However, the solutions found by HE are generally very near to optimality and are found with a lot less computational effort. This will be shown in later sections.

The objective function (8.2) and constraint (8.3) also apply to the HE approach but the basic rate constraint (8.4) is applied in a slightly different way. Still, each user requires the data rate R_B, but the constraint R_B serves as a stopping criterion for the iteration process. The HE approach is explained in detail in the following.

As seen in the BIP section the algorithm mainly must solve two conflicting tasks: Maximize the overall cell capacity R and give a basic rate R_B to each user. In the BIP approach, these tasks are performed jointly. In the HE approach, they will be regarded as two separate tasks.

Thus the HE approach is split into two successive phases:

Phase 1:

This first phase is used to give every user the basic data rate R_B. This is done in an iterative way. In each iteration step, each user gets a single subcarrier. Define the set X as the set of all available $x_{i,k}$ in the system. Then in each iteration a subset A of X with cardinality N_U is set to one, such that (8.3) is fulfilled. Thus, each user gets exactly one subcarrier per iteration and also

$$\sum_A R_{i,k} x_{i,k} = \max \tag{8.5}$$

63

holds. Note well, that the objective given by (8.5) is just a special case of the LP objective (8.2). Thus, in each iteration of the HE, an LP problem is solved, but considering a much smaller set of feasible solutions.

The subcarriers and corresponding $R_{i,k}$ selected in each iteration step are marked as allocated and are not considered in further iteration steps. This shows that in HE the objective (8.2) is applied locally instead of globally.

After each iteration step, the algorithm checks, which users i already have reached R_B. These users leave the iteration. All other users continue with the next iteration step. The iteration stops if all users have achieved R_B or if all subcarriers have been assigned. If there are still subcarriers to be assigned after phase 1, phase 2 starts.

An important remark: If the number of users N_U equals the number of subcarriers N, then the HE approach finishes after one iteration. If also all found R_i fulfill $R_i \geq R_B$ (e.g. if R_B is small), then an optimum solution is found by the HE and this solution exactly matches the solution found by the BIP approach. However, such a case is quite unlikely.

Phase 2:

Since all users are already supplied with R_B at this point, the constraint (8.4) can be neglected from now on. Thus, the remaining subcarriers are solely used to further increase the objective (8.2). This leads to a very simple allocation for the remaining subcarriers: The user i, who observes the highest $R_{i,k}$ on a subcarrier k gets assigned to this subcarrier.

This description of the two phases of the HE approach reveals its limited viewpoint especially in phase 1. The maximization in each iteration step considers at most N_U subcarriers and neither future nor past allocation decisions are taken into account. Thus in the HE approach it is likely, that each user gets its best available subcarrier in each iteration step without considering the average performance of the individual users. Hence, the HE approach generally does not find optimum solutions to the optimization task, but, as the next paragraphs will show, the solutions found are close to those of the BIP approach. As a final remark it is stated, that the HE approach is purely deterministic. This means for a given set of $R_{i,k}$, a unique solution is found by the HE approach.

8.2 Comparison of BIP and HE Approach

In the following, the general process of subcarrier allocation is evaluated for both the HE and the BIP approach. Differences in the solutions found for the subcarrier allocation task are discussed. Also, a brief comparison in terms of computational complexity is done.

8.2.1 Subcarrier Allocation Process

In this paragraph the behavior of both considered subcarrier allocation approaches is compared by considering a given channel realization for a limited number of users. An exemplary channel realization for three users can be seen in Fig. 24, where the allocation is done inside a system with 32 subcarriers. The lines in the upper part of the plot show the frequency selective channel in terms of rate per subcarrier. Thus, each section of each plotted line represents a value $R_{i,k}$. Below, a pair of rows with colored bars shows the subcarrier allocation done by the HE and BIP approach, respectively. Each colored bar represents a subcarrier allocated to the corresponding user. The upper row illustrates the subcarrier allocation done by the HE approach.
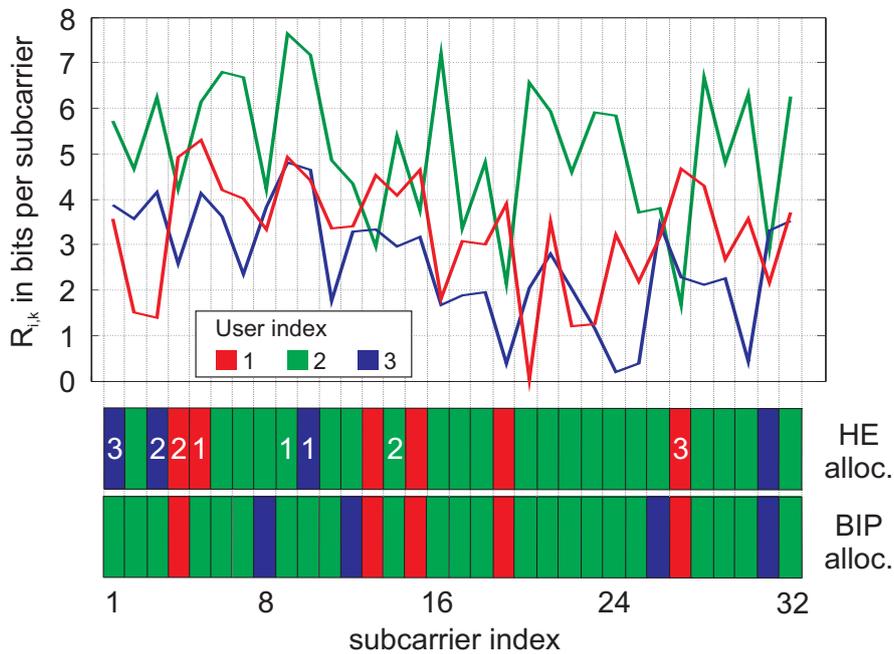
Fig. 24: Examples of subcarrier allocations by HE and BIP approach

The lower row shows the allocation done by the BIP approach, which is discussed in the following. The allocation is done based on the objective function and constraints given in (8.2) to (8.4). The data rate R_B is set to 12 net bits per OFDM symbol (0.75 Mbps) in this example.

It is obvious that user 2 (green), due to his good channel performance gets the most subcarriers and thus maximizes the cell throughput R. User 3 (blue) just gets enough subcarriers to fulfill the rate constraint R_B. This is the general behavior of the algorithm towards users with poor channel conditions. The majority of subcarriers is assigned to such users, who observe the maximum performance on a particular subcarrier. This behavior matches the requirements defined by (8.2) to (8.4).

The BIP allocation result in Fig. 24 reveals that the maximization of the objective and the fulfillment of the R_B constraint are two conflicting tasks. This conflict is solved by the BIP

approach by assigning only those subcarriers to the low-performance user 3, which also would provide low performance to the other users.

While the solution to the assignment problem seems quite obvious for a limited system such as shown in the example, solving the problem gets more demanding as the number of users inside the cell increases. Then, a higher share of subcarriers must be assigned to users with low performance to fulfill their constraints. These subcarriers are no longer available to high performance users, which otherwise could push the overall cell throughput.

After this introduction to the optimum assignment solution by the BIP approach, now the HE approach will be introduced and compared to the BIP approach.

If the described HE approach is applied to the cell- and channel-situation shown in Fig. 24, the subcarrier allocation shown in the upper colored row is obtained. As introduced in 8.1.2, the HE approach executes phase one until all users have reached R_B. For example, user 3 (blue) reaches $R_B = 12$ net bits per OFDM symbol after exactly three iterations. The assigned subcarriers are marked by their iteration number in Fig. 24. The sequence of assignment is given by the subcarrier indices 10, 3 and 1. In the depicted example, all users are provided with R_B after three iterations. Note that user 2 needs only two iterations to reach R_B. All other subcarriers are assigned inside phase two.

At the first glance both the assignment of the BIP approach and the HE approach look quite alike. But the differences become obvious especially for user 3, which is the user with the worst average performance inside the cell. Only a few subcarriers of this user's channel provide a suitable $R_{i,k}$ but at the same time the other users have a much higher performance on the same subcarriers.

In case of the BIP approach the overall throughput is maximized by allocating subcarriers to user 3, which are also observed as poor by the other users. This can be seen in Fig. 24 e.g. at subcarriers 8 and 26.

In contrast, the HE approach in its phase one allocates only three subcarriers in each iteration, which is a limited viewpoint. Thus, no future or past allocation decisions are considered for the allocation at hand. Therefore, in the HE approach it is likely, that each user gets its best available subcarrier in each allocation without considering the average performance of the users. This can lead to allocation decisions different from the BIP approach. In the example at hand, user 3 indeed gets its almost best subcarriers in phase one. This is indicated by the colored bars marked with '1', '2', and '3', which is the order of allocation of the respective subcarriers.

In the first iteration, the subcarriers 5, 9 and 10 are allocated since the sum rate $R_{3,10} + R_{2,9} + R_{1,5}$ of the assigned subcarriers is the maximum possible sum rate for the allocation of three subcarriers at this point in the iteration process. Thus this particular subcarrier assignment fulfills (8.5). The same holds for the following iterations. After three iterations, every user got its rate R_B. Afterwards, the remaining subcarriers are allocated in phase two.

These examples give a good insight of the working principle of the two allocation approaches. They show that both subcarrier allocation approaches seek to maximize the cell throughput R. The actual solution however may be different. It is also obvious, that the chosen optimization task defined by (8.2) to (8.4) does not provide a high degree of fairness to the users, since the rates R_i assigned to the individual users can cover a wide range starting from R_B with no upper bound.

The given examples show the process of the subcarrier allocation but are not suitable for a comparison between the average performances of BIP and HE approach. Therefore, section 8.3 gives results of system level simulations for both approaches.

8.2.2   Computational Complexity

After the introduction of the working principle of the BIP and HE approach, a brief comparison of their respective computational complexity is given.

As already indicated in section 8.1.1, realizable algorithms for solving a binary integer program mostly use a binary search tree containing all possible solutions for the problem instance at hand. Thus, the complexity of a BIP solver mainly depends on the number of possible combinations of $x_{i,k}$ values. Since a vector representation $\mathbf{x}$ of the decision variables $x_{i,k}$ has $N_U N$ entries, also the computational complexity of calculating a BIP solution depends on $N_U N$. It can be shown (cf. A.2) that an upper bound for the complexity of a BIP solution is given by

$$O\left(2^{N_U N}\right) \tag{8.6}$$

using the $O(.)$-notation introduced in [Pap82].

In contrast to this, the complexity of the HE approach is much lower and is basically determined by its phase 1. In this phase, a local optimum is calculated, whose complexity is upper bounded by

$$O\left(N_U^2 N^2\right) \tag{8.7}$$

Since the complexity of phase 2 given by $O\left(N_U N\right)$ is much lower, the overall complexity of the HE approach is upper bounded by (8.7). A thorough derivation of these results can be found in A.3.

8.3   Simulation Results

In all simulations a cell scenario as shown in Fig. 23 is assumed, where the users are uniformly distributed inside the cell. The overall number of subcarriers is $N = 32$. The performance of both approaches (BIP and HE) is simulated for various user numbers and basic rates

R_B. For each user number, 10.000 cell scenarios were simulated. The results are discussed in the following.

In the simulations, the uniform cell scenario introduced in section 5 is applied. A single cell with radius $d_{max} = 100m$ is considered. The channel model includes both small scale and large scale effects. Given these parameters, the overall signal-to-noise ratio $SNR_{i,k}$ on each subcarrier is analytically described by (5.4). The complete set of channel parameters can be found in section 5. The system in this reference considers a channel divided into 256 subcarriers. To allow for a reasonable computation time especially for the BIP approach, all simulations in this document are based on a subset of 32 subcarriers from this channel. The subcarriers are situated equidistantly over the whole channel bandwidth to retain the channel's frequency selectivity.

8.3.1 Average System Throughput based on BIP and HE

In this paragraph, the cell throughput R of the BIP- and the HE-approach is compared. For the subcarrier allocation, the objective and constraints are defined by (8.2) to (8.4). For the simulation, the parameters introduced above are applied. The results are depicted in Fig. 25.
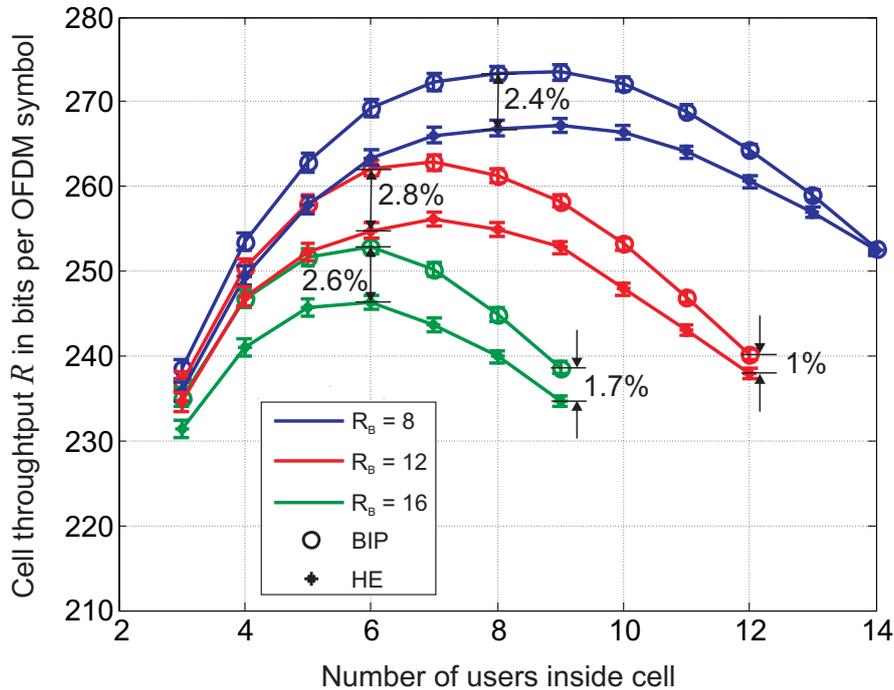
Fig. 25: Comparison between throughputs of BIP and HE approach

The basic rate R_B for each user is defined as "net bits per OFDM symbol" and varies between 8, 12 and 16 bits per OFDM symbol. The achieved rates R_i of all users are summed up and form the cell throughput R given on the y-axis.

The curves for $R_B = 16$ stop at 9 users, because for these user numbers a feasible solution for the allocation task is found in less than 95% of all cases. For higher user numbers at this rate, the cell starts to get overloaded. Similar conditions can be observed for $R_B = 12$ and $R_B = 8$.

Between curves of the same allocation approach but for different R_B, a big difference in cell throughput R can be observed. This is due to the increasing amount of subcarriers that is needed for a user to fulfill an increased R_B. This means an increasing amount of resources must be tied to users with probably bad channel conditions, which makes only a limited contribution to R.

If the curves for the BIP and HE approach are compared for the same basic rate R_B, it can be seen that the largest difference between both approaches amounts to less than 3%. This is for medium numbers of users inside the cell. For high and low numbers, the difference between both approaches is much less. The small difference for low user numbers is caused by the low multiuser diversity (MUD) inside the system in this case. If e.g. three users are inside the cell, in most cases one user observes a very good channel compared to the other users. Thus this user gets almost all subcarriers, as seen in Fig. 24. Hence the throughput R for low user numbers is limited by the average channel quality defined by the channel model.

For an increasing amount of users, R is rising for both approaches, since the MUD crops up, cf. section 6.2.2. At a certain point for each R_B, the MUD and thus also the throughput reach their maximum. At this point the largest difference between BIP and HE approach is observed. This difference is due to the different allocation strategies in both approaches, as explained in section 8.2.1.

For further increasing user numbers the throughput decreases for both approaches. Since the overall number of subcarriers inside the system is fixed, it gets more and more demanding to fulfill the rate constraint R_B for all users. The overall throughput is no longer bounded by the performance of the best users inside the cell, but by the task to satisfy the R_B constraint for all users. For a further increasing number of users, the result of the optimization gets to the point where almost all users get rates just above the basic rate R_B.

An interesting observation is that in these heavy load scenarios, the performance of the HE approach is almost the same as that of the BIP approach. The difference in cell throughput R is e.g. only 1% for $R_B = 12$ and 12 users inside the system. For $R_B = 8$, the throughput R even converges. As already introduced in section 8.1.2, the HE approach is able to find an optimal solution if N_U equals N. Since in each iteration, the HE approach allocates N_U subcarriers while maximizing (8.5), the solution to (8.5) equals the solution to (8.2) if $N_U = N$ and $R_i \geq R_B$. In case of a heavily loaded cell, N_U gets near to N and thus the HE solution gets nearer to the optimal BIP solution.

Summarizing the above results, it is obvious that the throughput R found by the HE approach amounts to over 97% of the throughput achieved by BIP approach while at the same time saving several orders of magnitude in computation time. Especially in scenarios where the

number of users is either very low or very high, the HE approach gives almost the same throughput R as the BIP approach.

8.3.2 Fairness Level based on BIP and HE

Another important system aspect is the fairness level inside the cell. The subcarrier allocation task as formulated in (8.2) to (8.4) does consider fairness only in so far as each user requires a data rate R_B. This leads to very unbalanced rates R_i, especially if only a few users are present inside the cell. Only if the cell gets more loaded with users, the user rates R_i converge to R_B due to shortage of free resources.

This effect is depicted in Fig. 26 by showing various *Cumulative Distribution Functions* (CDF) with N_U as parameter. *The CDF of the achieved user rates R_i is introduced in the following as a measure for the level of fairness inside the system.* For each rate value R_i the CDF gives the probability that a user *i* has reached a rate less or equal to R_i.
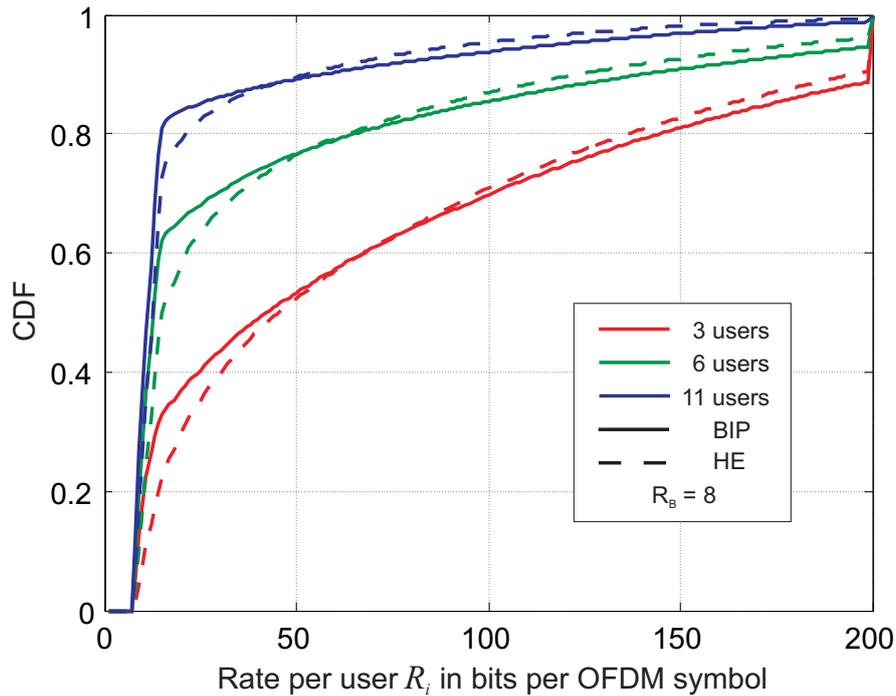
Fig. 26: CDF of user rates for BIP and HE. Varying N_U. Minimum rate R_B = 8.

Thus, the CDF is a monotonously increasing function over R_i. Its slope is a measure for the system's fairness. For example a steep slope reaching from probability zero to one shows a high level of fairness, since all users obtain similar rate values R_i.

If for the time being the solid curves representing the BIP approach in Fig. 26 are considered, the steep slope for 11 users (blue curve) inside the system is striking. It shows that over 80% of all users obtain a rate R_i in the range of 8 to 15 bits per OFDM symbol. Only a low per-

centage of users gets higher rates, but these lie in a much higher range of up to 200 bits per OFDM symbol. This is analogous to the findings in section 8.2.1, that most users get a rate close to R_B while the few users with very high SNR get a disproportionate fraction of the available resources. All in all, the fairness level for 11 users inside the cell is quite high, since the cell is considerably loaded in this scenario, which forces the allocation algorithm to use the majority of resources to guarantee R_B for each user. Only a small fraction of resources can be used for maximizing the cell throughput.

This situation changes, if the number of users N_U is decreased (solid green and red curve). The decreasing slopes of the CDF curves indicate a lower level of fairness, which gets especially obvious for $N_U = 3$: Here, mostly a single user gets a rate close to R_B, while for the other two users a wide range of rates R_i is possible. Especially the high probability for rates higher than 200 bits per OFDM symbol is notable. The curve for $N_U = 6$ has a similar shape.

Thus, the level of fairness of the allocation approach based on (8.2) to (8.4) strongly depends on the number of users N_U inside the cell. The same findings hold for the HE approach, which are indicated by the dashed curves in Fig. 26.

Hence also in terms of fairness, there is only a slight difference between the HE and the BIP approach (cp. solid and dashed curves). A close comparison reveals only minor differences in the probability of lower rates R_i for the HE approach, which is due to its limited optimization viewpoint as indicated in 8.2.1. Since the level of fairness in the HE approach comes so close to that of the BIP approach, the HE approach will be the basis for the remaining fairness discussion:

The figures Fig. 27 and Fig. 28 show the CDF of the HE approach for two basic rates $R_B = 8$ (Fig. 27) and $R_B = 12$ (Fig. 28). The shape of the CDF curves with N_U as parameter is similar in both figures. This affirms the fact that the level of fairness for the HE approach strongly depends on the number of users inside the system. Also, this effect is independent of the basic rate R_B, which can be shown by comparing Fig. 27 and Fig. 28.

In summary, the preceding sections showed the high similarity between the HE and BIP approach in terms of system throughput and fairness. In the current implementation both approaches provide only a low level of fairness, which additionally depends on the actual number of users N_U inside the cell. Thus in chapter 9, possibilities to increase the fairness inside the system and their impact on the cell throughput will be discussed.
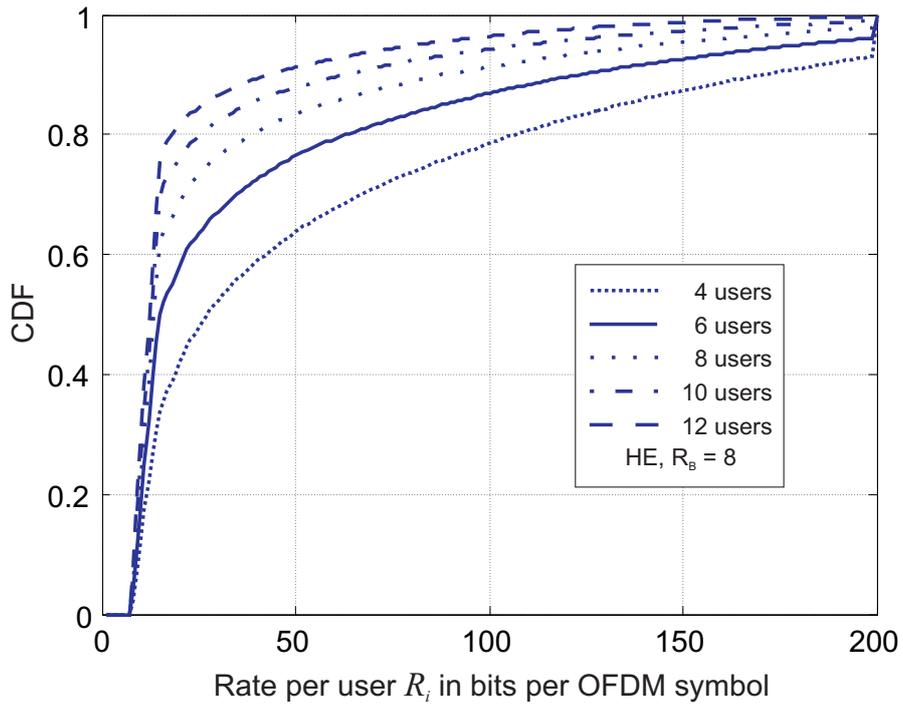
Fig. 27: CDF of user rates for HE. Varying N_U. Minimum rate $R_B = 8$.
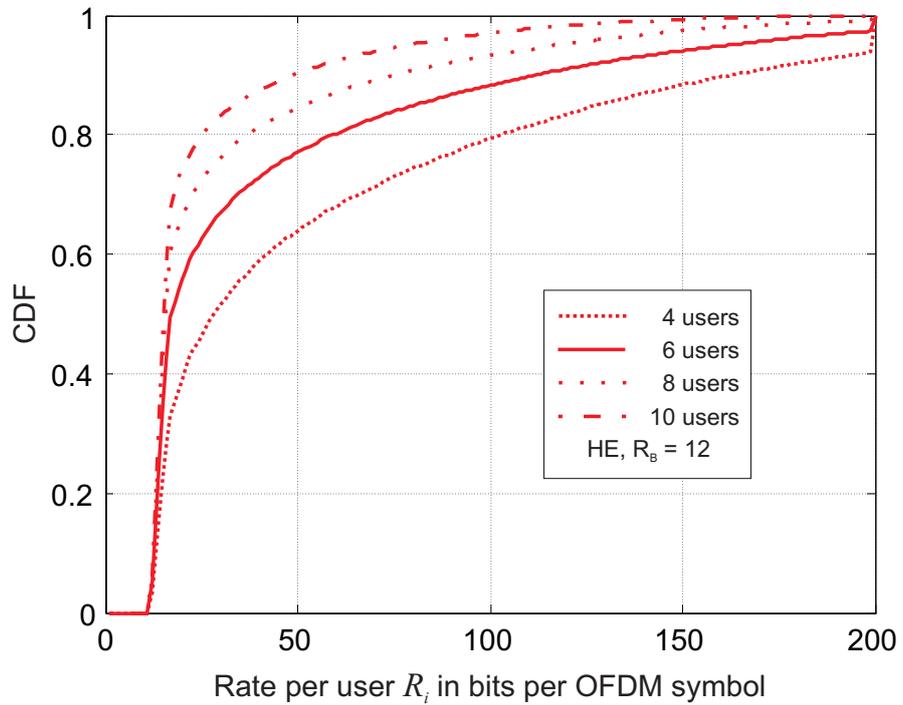
Fig. 28: CDF of user rates for HE. Varying N_U. Minimum rate $R_B = 12$.

8.3.3 Simulation Results for Discrete PHY Modes

The simulation results of the chapters 6 and 7 were based on the link performance of an OFDM transmission chain in terms of BER, even though various levels of abstraction where used. By contrast, the results presented in the sections above are based on an information theoretic model for the data rate R_i of each user (cf. (8.1)).

In order to show that the results obtained so far using (8.1) are well suited to evaluate the system performance, this chapter shows results obtained by applying PHY modes instead of information theoretical rates. The comparison of the results in this section and section 8.3.1 shows that there is of course a difference in throughput quantity, but not of quality. An advantage of the information theoretical model of section 8.3.1 is that it is independent of any modulation and coding technique that may be applied.

The results of the previous section were based on a subcarrier allocation where the data rate of each selected subcarrier was given by information theoretical rates

$$R_{i,k} = \log_2\left(1 + SNR_{i,k}\right).$$

In actually deployed communication systems, the selected subcarriers will be loaded with specific PHY modes corresponding to the subcarrier's SNR. Hence in the following, the continuous rates $R_{i,k}$ are substituted by discrete rates $\hat{R}_{i,k}$ due to the granularity of PHY modes. In this section, such PHY modes are applied to the considered system.
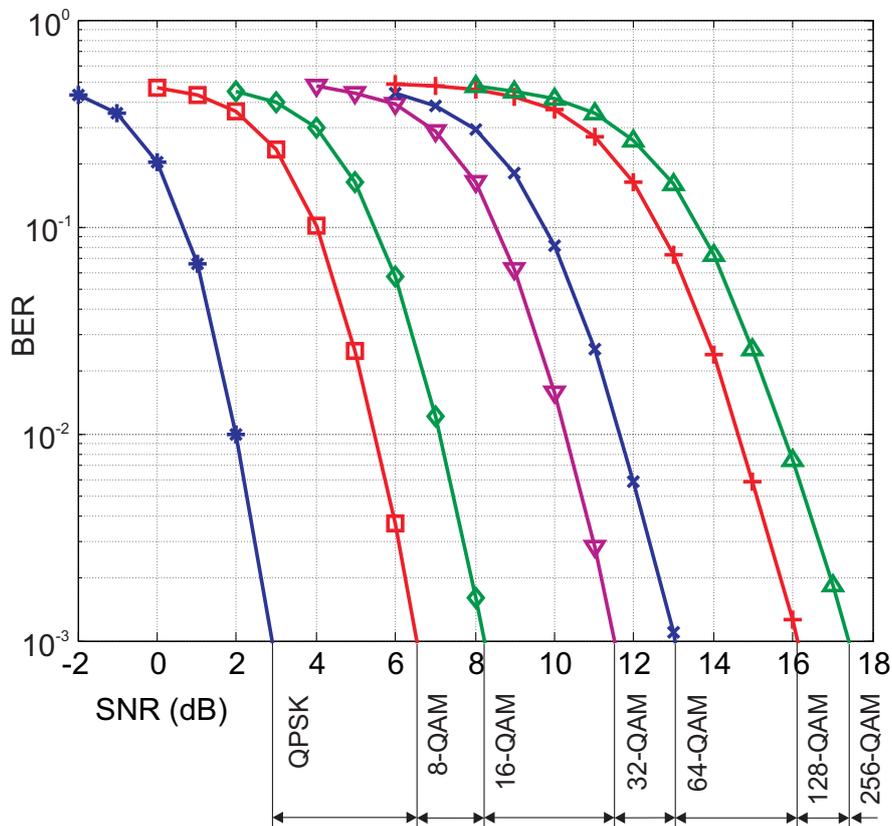
Fig. 29: BER curves for PHY-Mode selection

All PHY modes assume a convolutional coder with code rate $CR = \frac{1}{2}$. In order to adapt the utilized PHY modes on each subcarrier to the current channel condition, multiple modes are considered. To each PHY mode a specific SNR threshold is defined, which determines when the utilization of the respective PHY mode is appropriate. The SNR thresholds SNR_{min} corresponding to each PHY mode are set for a minimum bit error rate of $BER \leq 10^{-3}$. If the $SNR_{i,k}$ on a subcarrier falls below an SNR_{min} threshold, the next lower PHY mode will be applied to this subcarrier. Fig. 29 shows the BER curves for all considered PHY modes and the SNR_{min} thresholds are selected accordingly. The applicable PHY modes and data rates $\hat{R}_{i,k}$ are listed in Tab. 4. The optimization task (cf. (8.2), (8.3) and (8.4)) stays unchanged for both the BIP and HE approach. The only difference is that now the continuous information theoretical rates $R_{i,k}$ are substituted by their discrete counterpart $\hat{R}_{i,k}$.

For this new PHY scenario, system level simulations were run based on the same system- and channel model as in the above section. The results are depicted in Fig. 30.

Tab. 4: SNR-Thresholds for the PHY mode selection

QAM	4	8	16	32	64	128	256
$\hat{R}_{i,k}$ in net bits per subc.	1	3/2	2	5/2	3	7/2	4
SNR_{min} (dB)	3.1	6.6	8.3	11.8	13.1	16.1	17.5

If compared with the results from Fig. 25 it stands out that all throughputs R are approximately cut in half. This is due to the application of realistic PHY modes and channel coding, which causes a performance gap compared to the transmission system based on information-theoretical capacities. Also in the PHY scenario the differences between the BIP and HE approach are much smaller. The difference in terms of throughput R between the two allocation approaches amounts to considerably less than 1%. The reason for this small throughput difference lies in the rate limitation by the maximum possible PHY mode. The advantage of the BIP approach to identify and assign especially suitable subcarriers is now upper bounded by this effect.

The general shape of the curves in Fig. 25 (rising for low and falling for high user numbers) is similar to the curves in Fig. 30 since the same effects considering the fulfillment of constraints and the maximization of throughput R apply.

A property specific to the PHY scenario is its system inherent upper bound to the throughput R. Given the code rate $CR = \frac{1}{2}$, it can be derived from Tab. 4 that the maximum possible R-value for a system with $N = 32$ subcarriers is 128 net bits per OFDM symbol. This value is

almost reached by the BIP approach for $R_B = 8$ and a number of 6 users inside the cell. This shows the system inherent limitations for all subcarrier allocation approaches.
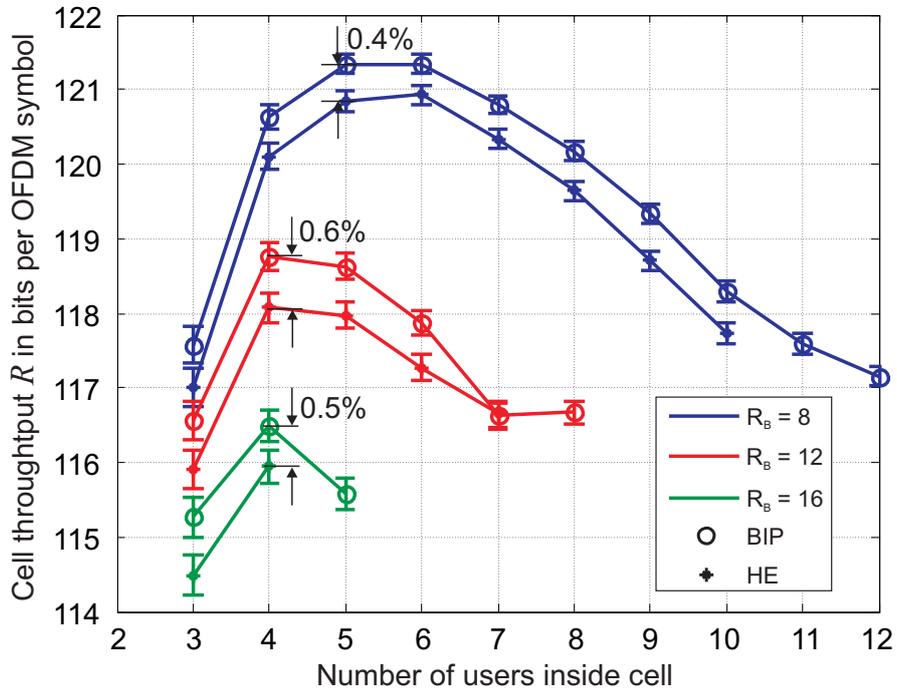
Fig. 30: Comparison between throughputs of BIP and HE approach with discrete PHY modes
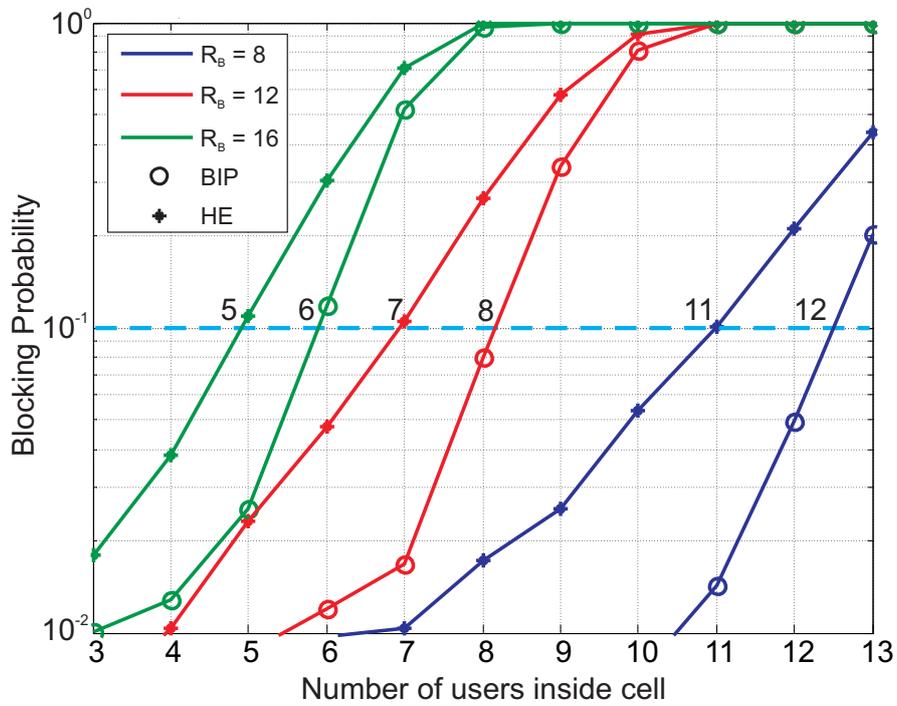
Fig. 31: Blocking probabilities for BIP and HE and for various numbers of users

The discrete PHY modes not only limit the maximum achievable R but also the maximum number of users inside the cell. Thus the results shown in Fig. 30 exclusively consider user numbers, where the subcarrier allocation leads to a feasible solution in more than 95% of all cases.

For higher user numbers, the probability of an infeasible solution grows rapidly as can be seen in Fig. 31. This figure shows the blocking probability for the BIP and HE approach for an increasing number of users. The blocking probability is defined as the probability that no feasible solution for the subcarrier assignment can be found. A reasonable assumption would be to keep this probability below 10^{-1}. When this threshold is applied, the maximum number of users that can be served inside the cell can be read from Fig. 31 (see dashed line). In summary, the BIP approach is able to serve one additional user compared with the heuristic HE approach.

The comparison of the results from this section and section 8.3.1 shows that the results obtained by using (8.1) instead of technical PHY modes are quite similar, if the results are compared with respect to resource allocation approaches. This is especially obvious, if the change of cell throughput over the number of users N_U is considered (cf. Fig. 25 and Fig. 30). Differences are basically made out in terms of absolute performance numbers. These are generally due to the employed code rate CR and the maximum degree of modulation. *Since the information theoretical model based on* (8.1) *allows the evaluation of resource allocation approaches independently of used modulation and coding techniques, this approach will be pursued further in the following.*

The results presented in the chapters 8.3.1 and 8.3.3 have demonstrated the differences and similarities between the HE and the BIP approach. It was shown that the HE approach achieves almost the same cell throughput R as the BIP approach. This result was independent of the application of information theoretic rates $R_{i,k}$ or of discrete PHY modes $\hat{R}_{i,k}$. It was also shown that a big advantage of the HE is its computational efficiency. *For these reasons, the basis for the remaining discussion will be solely the HE approach.*

Independent of the applied allocation approach, the results to the subcarrier allocation task are mainly determined by its formulation, which was specified in (8.2), (8.3) and (8.4). The presented results show, that the current formulation sacrifices system fairness for the sake of system throughput R (cf. Fig. 25 and Fig. 26). Hence the following sections will introduce various approaches to adjust the tradeoff between fairness level and throughput by introducing additional constraints to the subcarrier allocation task.

9 Fairness Enhancements considering the HE Approach

In broadband communication systems, the balance between system throughput and fairness poses a classical tradeoff situation to the system operator. The preceding sections show that a subcarrier allocation based on the maximization of throughput according to (8.2) and (8.4) generally allows only a very limited level of fairness. From a user's perspective, fairness is a very desirable property for a communication system, since every user wants a fair share of the system resources. Also for a system operator it can be beneficial to improve the fairness inside the system and thus to increase customer satisfaction or to balance system loads.

The general goal of this fairness discussion is to achieve a better balance between the individual rates R_i of the users. This balance should ideally be independent of the current cell situation. The previous results show significant shortcomings of the considered allocation approaches in this respect. The key to a balance of rates R_i and thus to increasing fairness lies in the modification of the applied resource allocation approaches. Such modifications are discussed in the following.

Also, the following sections introduce means to adjust the fairness inside the cell such that a balance between fairness and throughput can be achieved. This adjustment can be done e.g. by introducing additional constraints to the subcarrier allocation task. Another possibility is to modify the process of subcarrier allocation itself.

The following sections illustrate these possibilities of modifying the allocation task. Additionally, the level of fairness gained by these modifications will be quantified. Finally, simulation results for the modified HE approach are presented. Due to presented reasons, all further pursued approaches are based on heuristics.

9.1 Adding Constraints to the Subcarrier Allocation Task

An optimization task to find an optimal subcarrier allocation was introduced in (8.2) to (8.4). The objective was to maximize the cell throughput R. The constraints given in this optimization task solely refer to exclusive assignment of subcarriers and the rate constraint R_B. With these constraints applied, the solution found for a problem instance in most cases resembles the example depicted in Fig. 24: The user with the best average channel performance gets most of the resources, while other users are supplied with the minimum rate R_B.

This leads to very unbalanced rate distributions inside the cell, since only a few users utilize the better part of the cell's capacity. To increase the fairness of the system and to provide each user with a fair share of resources, an additional constraint to the optimization task will be introduced in the following.

This constraint defines the maximum number of subcarriers $K_{\max}$ each user can claim. It is valid for every user. The complete formulation for the optimization task with the additional constraint is given as follows:

Objective function:

Maximize system throughput

$$R = \sum_{i=1}^{N_U} R_i = \sum_{i=1}^{N_U}\sum_{k=1}^{N} R_{i,k} x_{i,k} = \max_{x_{i,k}} \quad \text{with } x_{i,k} \in \{0,1\} \tag{9.1}$$

Constraints:

Exclusive allocation of all subcarriers:

$$\sum_{i=1}^{N_U} x_{i,k} \leq 1 \text{ for all subcarriers } k \tag{9.2}$$

Minimum data rate for each user:

$$R_i = \sum_{k=1}^{N} R_{i,k} x_{i,k} \geq R_B \quad \text{for each user } i = 1...N_U \tag{9.3}$$

$$\text{with } R_{i,k} = \log_2\left(1 + SNR_{i,k}\right)$$

Maximum number of subcarriers per user:

$$\sum_{k=1}^{N} x_{i,k} \leq K_{\max} \quad \text{for each user } i = 1...N_U \tag{9.4}$$

This optimization task can be carried out using the already introduced approaches BIP or HE. To prevent confusion with the HE and BIP approaches defined by (8.2) and (8.4), the allocation approaches modified in the way specified by (9.4) are called $K_{\max}$ constrained HE and BIP approach respectively.

The following discussion of the allocation process and simulation results is limited to the $K_{\max}$ constrained HE approach. This is due to the very similar performance compared with BIP while using much less computational effort. Results for the $K_{\max}$ constrained BIP can be found in Appendix B.

Equation (9.4) requires only minor changes to the HE approach, but with a remarkable effect to the allocation results. This will be shown in the next section.

9.2 Subcarrier Allocation Process

In the following, the impact of the additional constraint introduced by (9.4) on the subcarrier allocation task will be illustrated by an example. The resulting allocation found by HE and $K_{\max}$ constrained HE are discussed. The necessary modifications of the underlying algorithm for HE as introduced in 8.1.2 are straightforward and therefore not discussed here.

The same channel example as in section 8.2.1 is used below. The resulting allocation of sub-carriers for the modified HE approach is shown in Fig. 32. As in Fig. 24, three users are located inside the system. The optimization described by (9.1) - (9.4) is applied. In this example, K_{max} is set to $K_{max} = 10$. It is obvious, that the allocation of subcarriers to each user leads to a completely different result, compared to the allocation without the constraint K_{max}.

In Fig. 32 for the K_{max} constrained HE, the number of allocated subcarriers is identical for each user. The constraint K_{max} is exactly met. Thus, each user gets a fair share of the available resources. This is a strong contrast to the allocation result of the unconstrained HE (cf. Fig. 32, lower row), where user 2 gets almost all resources due to its superior channel quality.

In the considered scenario, the constraint $K_{max} = 10$ allows only the allocation of a total of 30 from the overall 32 subcarriers. Thus, the K_{max} constraint plays a dominant role in the allocation. If the number of users increased, all N resources would be claimed and it would not be possible for each user to claim exactly K_{max} subcarriers. The more users are joining the system, the more the K_{max} constraint is losing its impact and the R_B constraint is starting to dominate.
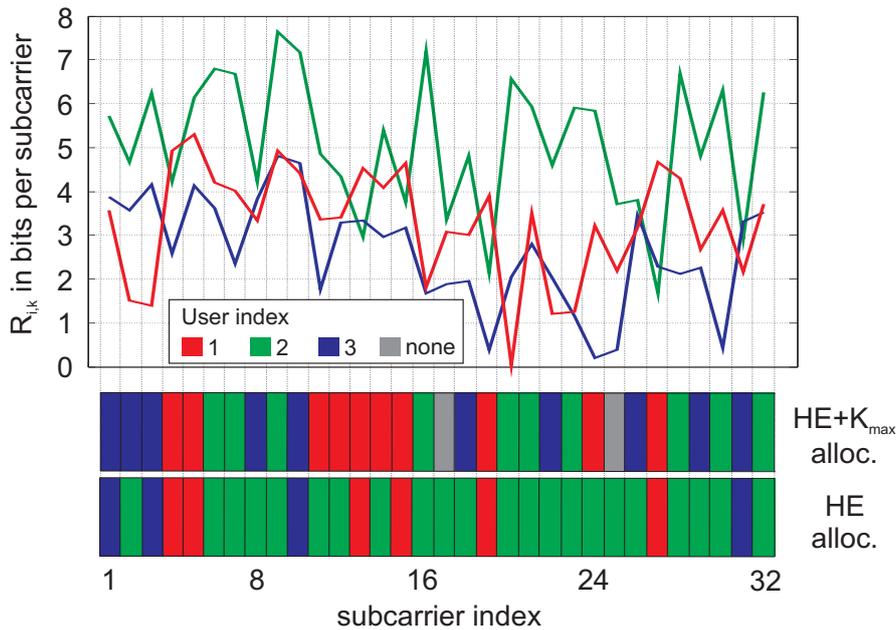
Fig. 32: Exemplary subcarrier allocation by HE constraint to $K_{max} = 10$ and unconstrained HE

As this example shows, the additional K_{max} constraint allows a much higher level of fairness than the optimization problem described by (8.2) to (8.4). Quantitative results for the degree of fairness and the system throughput are given in the chapter on simulation results.

9.3 Simulation Results

In Fig. 32 it was hinted that limiting the maximum number of subcarriers for each user by a value K_{max} can improve the fairness inside the system. This is now confirmed by further simulation results. The system and channel model applied in the following simulations was already introduced in section 8.3.

In the section below, the influence of the constraint K_{max} on fairness and system throughput is evaluated for the HE approach and its modification. Comparisons between the performance of HE and BIP can be found in Appendix B.

9.3.1 Fairness Level based on K_{max} constrained HE

In the following, the various degrees of fairness provided by applying the constraint K_{max} to the HE approach are evaluated. As a fairness measure, again the CDF of the achieved user rates R_i is used. Fig. 33 shows CDF graphs for a cell with 6 users. The overall number of subcarriers is $N = 32$ and the rate constraint $R_B = 8$ is assumed. Each graph is calculated for a different value of K_{max}. All allocations for this graph are solely calculated by the HE approach and its modifications. The graphs for the BIP approach are not shown, since they give almost exactly the same results (cf. B.2). Also no curves for different values of R_B are shown, since their shapes are generally identical with those in Fig. 33.

The graph for $K_{max} = 32$ represents the HE allocation with no effective K_{max} constraint (cf. Fig. 27, solid curve), since each user is theoretically allowed to claim all resources inside the system. This corresponds to the subcarrier allocation introduced in chapter 8.1. This graph has a distinguishing run. It can be observed that most users get a minimum rate and that there is a broad distribution of high rates. If no constraints are set for the maximum number of subcarriers per user, most users get the minimum rate R_B, while a few users with convenient channel conditions can accumulate very high rates. These high rates show also quite a high variance, which is indicated by the flat slope of the curve for $K_{max} = 32$ in Fig. 33 for rate values between 50 and 200 bit per symbol. Therefore, this subcarrier allocation is highly unfair. The small increase of slope near to the end of the graph stems from users, which reached a rate of over 200 bits per symbol. These were counted as users with 200 bits per symbol in order to keep the graph clearly arranged.

The other extreme is the graph for $K_{max} = 5$. Since there are six users sharing $N = 32$ inside the system, applying the constraint $K_{max} = 5$ leads to a maximum of 30 allocated subcarriers. In order to maximize the system throughput, all users will get 5 subcarriers in any allocation. This makes the system highly fair, which is visualized by the steep slope of the corresponding curve.
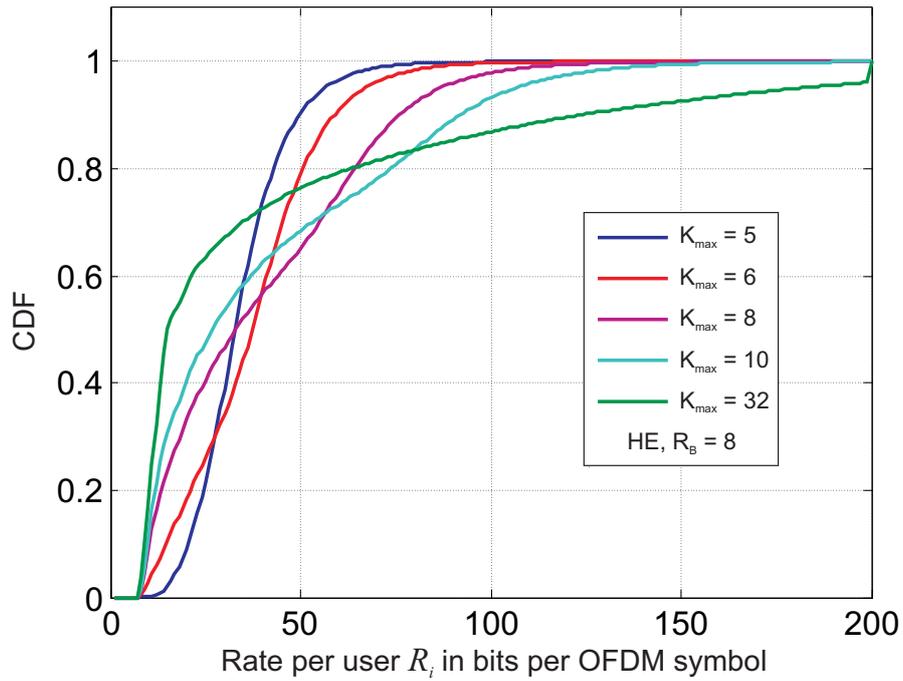
Fig. 33: CDF of user rates (HE approach) for $R_B = 8$ and various values of K_{max}, 6 users inside the cell
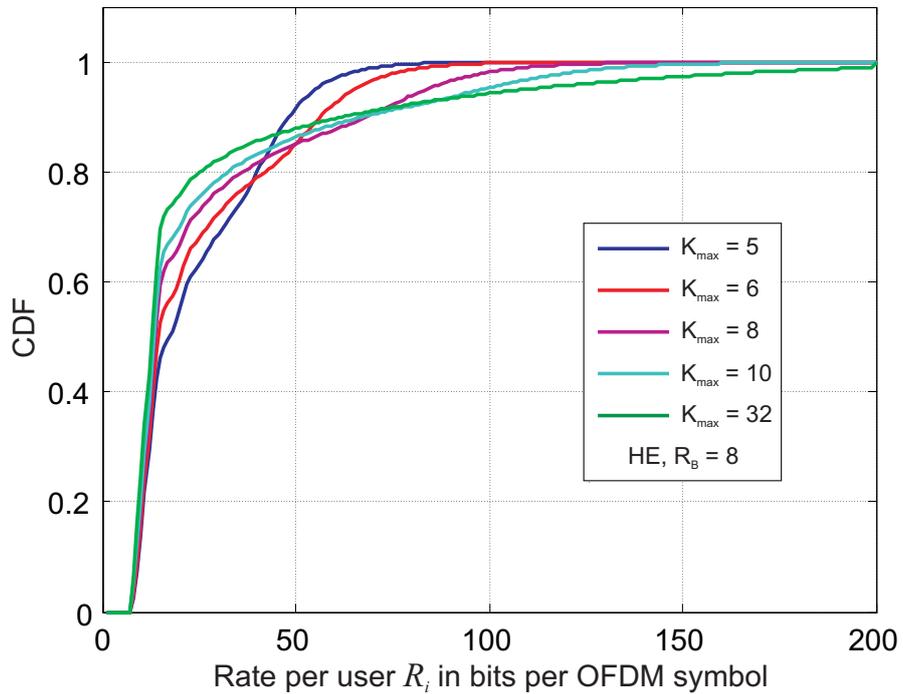
Fig. 34: CDF of user rates (HE approach) for $R_B = 8$ and various values of K_{max}, 10 users inside the cell

If K_{max} increases such that not all users can reach this maximum number of subcarriers, there will be some users, which have to give up subcarriers in favor of users with better channel conditions. With increasing K_{max} this leads to a quantum of users with a minimum rate R_B

while there is an increasing amount of users with very high rates. Thus, the rates R_i are less balanced. The growing amount of minimum rate users for growing values of K_{max} is also visible in Fig. 33. It reflects the fact that only users with good channel conditions reach K_{max} and thus high rates, while for users with low channel quality the constraint R_B becomes dominant. Thus, for an increasing value K_{max}, the level of fairness inside the system decreases.

From the above results can be concluded that the constraint K_{max} is a suitable parameter for tuning the fairness inside the system.

The results in Fig. 33 indicated that the level of fairness depends on the number of users who reach K_{max}. In reverse, this means that the fairness level does not solely depend on K_{max}, but also on the number of users N_U inside the cell. This becomes apparent in a comparison between Fig. 33 and Fig. 34. In Fig. 34, CDF curves based on the same parameters as in Fig. 33 are shown but with the number of users increased to $N_U = 10$. In this highly loaded system, the number of allocated subcarriers of very few users is limited by K_{max}, so that the CDFs converge to the unconstrained HE approach (solid green curve). Still, the parameter K_{max} is able to influence the fairness level of the system, but with a much lower impact than in the less loaded system ($N_U = 6$). In summary, the fairness in the K_{max} constrained HE is much higher than in the unconstrained case (cf. 8.3.1) and can additionally be adjusted by the parameter K_{max}. But still, the fairness level of the K_{max} constrained HE depends on N_U.

9.3.2 System Throughput based on K_{max} constrained HE

After the discussion of the rate distribution for varying values of K_{max}, the dependence of the system throughput on K_{max} will be evaluated for the HE approach. The corresponding curves of throughput R are shown in Fig. 35. All users require a minimum rate $R_B = 8$. The constraint K_{max} varies. The solid line in Fig. 35 represents the HE approach for $K_{max} = 32$ and thus, K_{max} has no effect in this curve. Hence the uppermost throughput curve in Fig. 25 is identical with the solid curve in Fig. 35.

The dashed curves in Fig. 35 represent lower values of K_{max}. It is obvious that the application of a value $K_{max} < N$ leads to strong cuts in system throughput R. Especially for low user numbers only a fraction of the available subcarriers is allocated. As long as the condition $K_{max} \cdot N_U \leq N$ is met, each user gets K_{max} subcarriers and thus the system throughput is linearly increasing with the user number inside the cell. If this condition can not be met due to increasing user numbers, the slope of the curve starts to decline. The shape of the dashed curves now resembles the solid curve, since now the R_B constraint gains dominance. But still the constraint K_{max} provides a strong bound to the cell capacity. Only if the user number in-

creases so much that R_B is the only active constraint in the optimization, the dashed and solid curves will converge. This is shown in Fig. 35 for user numbers greater than eleven.
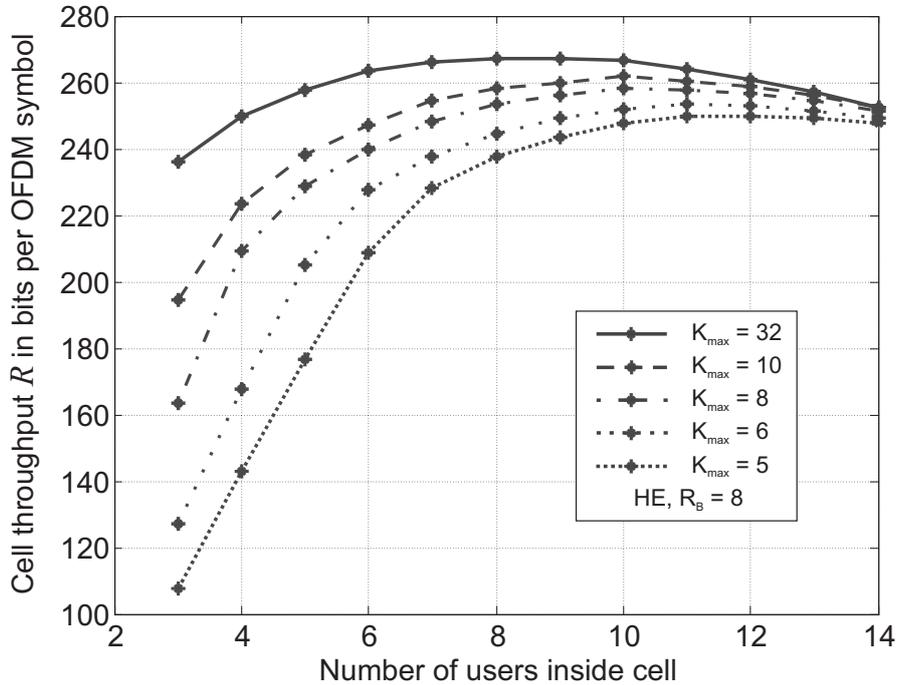
Fig. 35: HE approach with additional constraint K_{max}

If e.g. the curves for $K_{max} = 5$ and $K_{max} = 6$ are compared, an increase in throughput is observed for growing K_{max}. An increasing number of subcarriers per user gives more freedom in the subcarrier assignment and thus a higher throughput: The more K_{max} approaches N, the more the dashed curves will approach the solid curve.

The results presented in Fig. 33 and Fig. 35 reveal that the constraint K_{max} gives a remarkable gain of fairness at the expense of system throughput. The decision if throughput or fairness prevails is a tradeoff situation, which can be adjusted by a suitable choice of K_{max}.

The preceding results show the throughput performance of the K_{max} constrained HE for a fixed R_B. In the following, also R_B is a varying parameter. Fig. 36 shows throughput curves for various values R_B: 8, 12 and 16 bits per OFDM symbol for each user, respectively. For each R_B, K_{max} is set in such a way that even users with a very poor channel can claim enough subcarriers to be supplied with R_B. This is necessary, since the K_{max} constraint is already applied in phase 1.
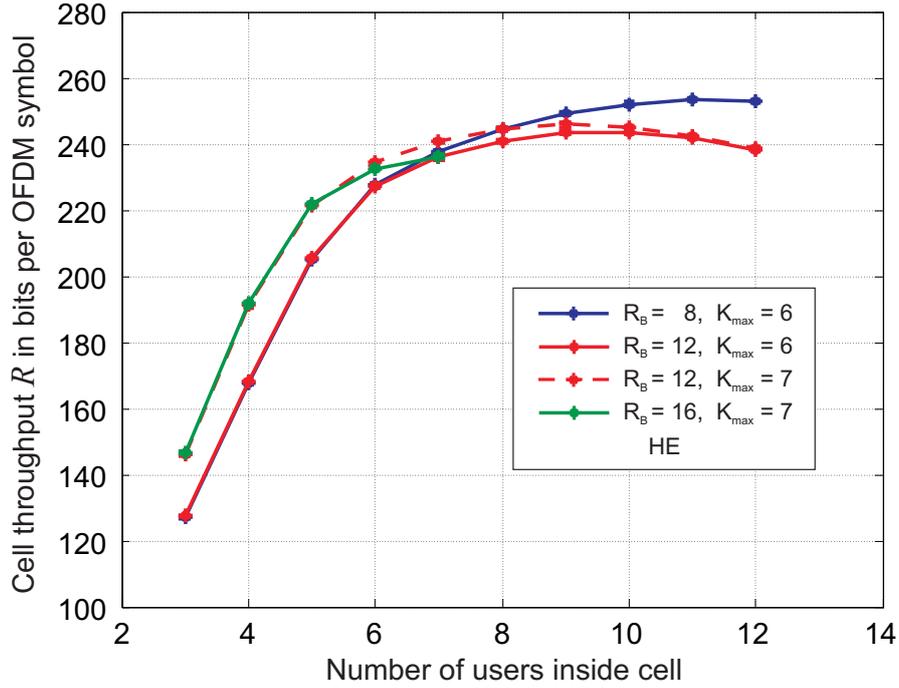
Fig. 36: HE approach for various values of R_B and K_{max}

The general shape of the curves is very similar for any R_B. As long as the condition

$$K_{max} \cdot N_U \leq N \qquad (9.5)$$

is met, the curves run straight. The distance between the curves in this area is determined by K_{max}: Curves with identical K_{max} value coincide, since each user gets exactly K_{max} resources and thus R_B does not influence the allocation procedure.

For higher user numbers the slope of the curves declines. If curves with identical K_{max} but different R_B are considered (e.g. solid red and solid blue), the curves start to split up. For low R_B, the number of subcarriers still to assign after phase one is higher than for high R_B. This gives the allocation process a higher flexibility in phase two, which leads to a higher throughput.

If curves with identical R_B but different K_{max} are considered (solid red and dashed red), two observations are made: First, there is a constant throughput offset for low user numbers, since the higher K_{max} allows a larger number of resources to be allocated. Second, for high user numbers, the curves tend to converge, since not all users get assigned to K_{max} resources. These two effects were already observed in Fig. 35 for $R_B = 8$ and are now confirmed to be generally applicable.

In summary, the K_{max} constrained HE approach offers less system throughput R than the unconstrained approach from section 8.3.1. The throughput depends strongly on the value

$K_{\max}$ and shows an even stronger dependency on the actual number of users in the cell. Together with the findings with respect to fairness, the $K_{\max}$ constrained HE approach is a good example for the tradeoff between fairness and throughput.

Although an improvement in the fairness level compared to section 8.3.1 is achieved, the fairness level still depends on the number of users N_U. This issue is addressed in the next chapter by evaluating specific rules for the choice of values for the $K_{\max}$ constraint.

9.3.3 Suitable Choice of K_{max} Value

The preceding chapters 9.3.1 and 9.3.2 showed that the $K_{\max}$ constrained HE approach allows the adjustment of the system's fairness level and throughput by means of the parameter $K_{\max}$. In the simulation results discussed so far, always a fixed value $K_{\max}$ was assumed, which was chosen independently on the number of users N_U inside the cell. This resulted in a strong coherence between N_U and the throughput and fairness level respectively, cf. Fig. 33, Fig. 34, and Fig. 35.

The reason for this unwanted coherence is found in (9.5), which shows that if small values for $K_{\max}$ and N_U are assumed, not all N available subcarriers are utilized for data transmission. This leads to an unwanted waste of resources and to a low throughput.

On the other hand, if $K_{\max} \cdot N_U > N$ holds, the $K_{\max}$ constraint has only a limited influence on the allocation procedure, which causes the fairness level to decline.

The solution to this issue is to set $K_{\max}$ in such a way, that always all subcarriers are utilized and at the same time a reasonable balance between throughput and fairness is maintained. Thus, $K_{\max}$ is chosen in accordance with

$$K_{\max} = \left\lceil \frac{N}{N_U} \right\rceil \tag{9.6}$$

where $\lceil . \rceil$ stands for the rounding up operation. In this way, $K_{\max}$ is kept integer and all subcarriers are used for transmission.

The introduced allocation approach is referred to as HE approach with variable $K_{\max}$.

Fig. 37 shows the CDFs for this new HE approach where each curve represents a different number of users N_U inside the cell area. For comparison, Fig. 38 shows CDF curves for the HE approach assuming a fixed value $K_{\max} = 5$ for any N_U. The remaining simulation parameters are identical to those assumed in section 9.3.1.

A first comparison of both figures shows that the HE with variable $K_{\max}$ mostly achieves a higher level of fairness than the approach with a fixed $K_{\max}$. The only exception are the solid

curves ($N_U = 6$) in Fig. 37 and Fig. 38: Here, the HE approach with fixed $K_{max} = 5$ achieves a higher fairness level (Fig. 38, solid curve), since in this case the other HE approach (Fig. 37, solid curve) utilizes $K_{max} = 6$ according to (9.6).
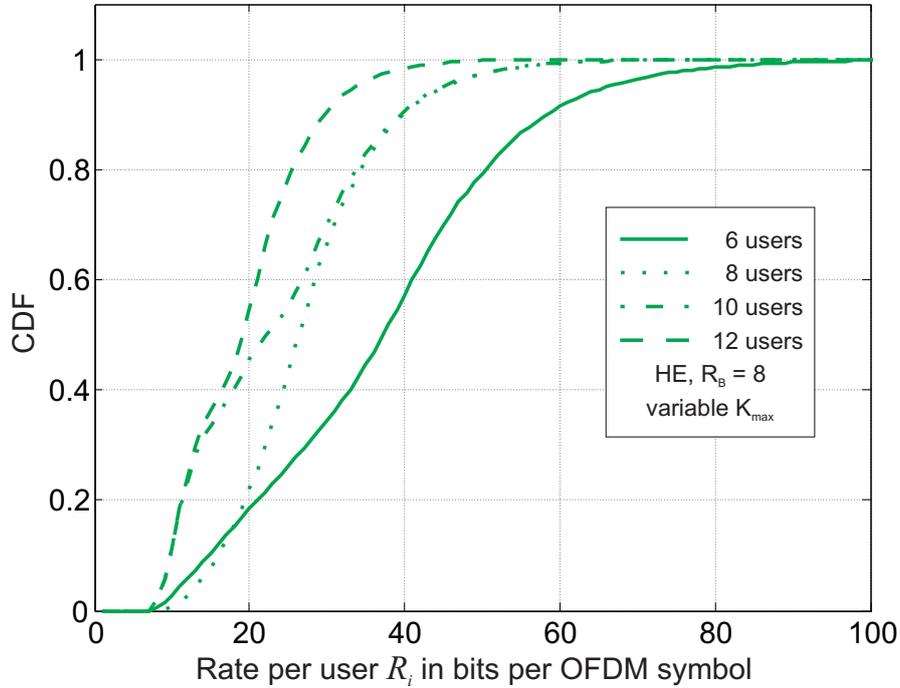
Fig. 37: CDF of HE approach with K_{max} adapted according to N_U ; varying user number
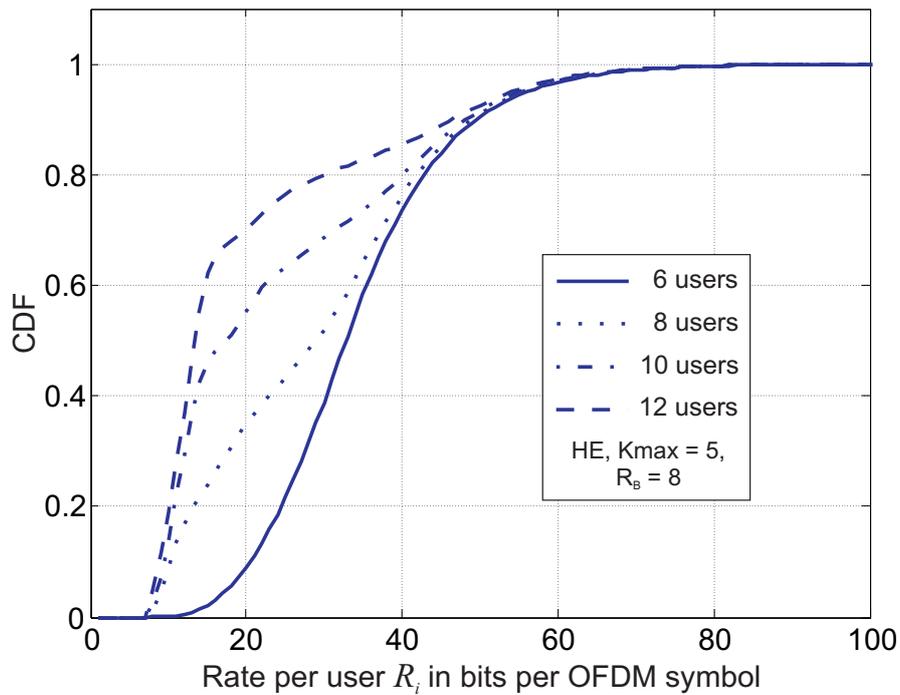
Fig. 38: CDF of HE approach with $K_{max} = 5$; varying user number

A more elaborate scrutiny of Fig. 37 reveals that the fairness level is much less dependent on N_U than in Fig. 38 although a variance between the slopes of the different curves is still visible. The steepest slope of all curves is observed for $N_U = 8$ (dotted curve in Fig. 37). This is because the considered system parameters N and N_U yield an integer result of (9.6) even without rounding up. Thus, the actual allocation process will assign exactly K_{max} subcarriers to each user, which implicates a very high level of fairness as discussed in preceding sections. In all other cases, the rounding operation in (9.6) yields $K_{max} \cdot N_U > N$ and thus allows a certain margin in the actual number of subcarriers each user gets. As a consequence, the fairness level is decreased a little in those cases as can be seen in Fig. 37 by comparing the curve for $N_U = 8$ to all remaining curves in the figure.

After discussion of the fairness level, the throughput achieved by the HE approach using variable K_{max} is evaluated. The familiar throughput curves are found in Fig. 39, which depicts the throughput curve for the HE approach discussed in this chapter (square markers), as well as a curve for fixed $K_{max} = 5$ (circular markers) and $K_{max} = 32$ (cross markers). In the latter case, the K_{max} constraint has no impact and thus the curve represents the original HE approach from section 8.1.2.
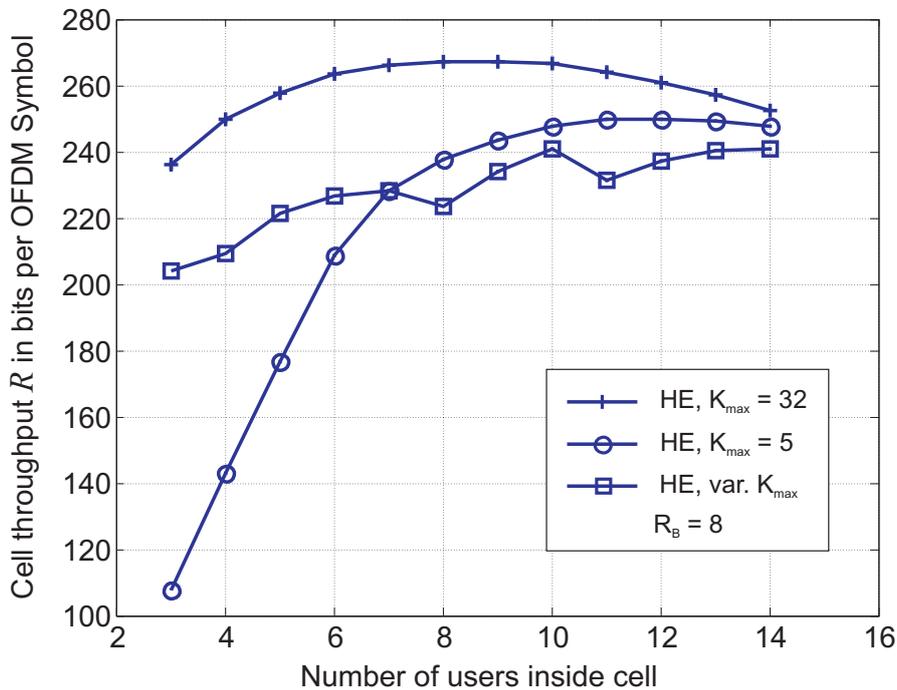
Fig. 39: Throughput of various HE approaches: fixed K_{max} = 5, 32 and variable K_{max}

The square-marked throughput curve shows again that the recently introduced HE approach is a lot less dependent on N_U if compared to the HE with constant K_{max} (circular-marked curve). The huge loss in throughput at low user numbers due to unused subcarriers is success-

fully avoided by setting K_{max} in accordance to (9.6). Also, since the adaption of K_{max} to N_U maintains a high fairness level for all user numbers, the convergence of the corresponding throughput curve to the unconstrained HE is less prominent.

The small notches in the variable K_{max} curve at $N_U = 4$, 8 and 11 represent situations where (9.6) exactly or almost exactly yields integer solutions without rounding. Since this leads to highest fairness levels, the corresponding throughput is decreased due to the well-known tradeoff. This unsteady shape of the throughput curve indicates a small drawback of the introduced HE approach with variable K_{max}: The fact that K_{max} must be integer limits the range of possible adjustments to fairness and throughput.

In summary, this chapter showed that the K_{max} constrained HE approach can be modified to show a stable fairness level over a wide range of user numbers N_U. This is achieved by setting K_{max} in such a way that for all cell loads, the complete number N of subcarriers is utilized and at the same time K_{max} is kept low enough to allow each user a fair amount of resources.

This goal was reached by introducing the rule (9.6) for choosing K_{max}. It was shown that this rule yields superior performance in terms of throughput and fairness compared to using a fixed value for K_{max}.

But still, there are limitations to the adaptivity of K_{max}. Its main drawback is that K_{max} must be an integer value. This property is assured by the rounding operation in (9.6), which leads to small irregularities in throughput and fairness if the term N/N_U is far from integer.

Thus, the following chapters introduce and evaluate further possibilities to adapt and/or improve the fairness level inside the system.

9.4   Increasing Fairness Level by Modification of HE Approach

In the previous section, the effect of adding constraints to the subcarrier allocation task on the fairness of the system is discussed. The results show a remarkable fairness gain. But the level of fairness can be improved even further if not only the constraints of the allocation are changed, but instead complete parts of the subcarrier allocation process are revised. The HE approach allows especially simple possibilities for modification since it is divided into two phases. Thus, each phase can be modified separately.

So far, the overall goal of the allocation was to maximize the cell throughput R. Now we will change this goal for phase 2 of the HE approach.

To prevent confusion with the HE approach applied in the preceding chapters, the modified approach will be referred to as *HE2*.

The HE2 approach considered in the following has two main characteristics:

- Phase 1 is kept unchanged (cf. section 8.1.2)
- Phase 2 has a completely changed goal: It is an iterative procedure where in each iteration the user with so far lowest R_i will get assigned to its best subcarrier

In this way, the rates R_i are balanced between the users. Since phase 1 was already introduced in section 8.1.2, only the modifications of phase 2 are explained in detail.

Phase 2:

After completion of phase one, all users are supplied with a rate R_B or higher. Those subcarriers, which are not yet assigned, are used to balance the individual rates R_i of the users. This is done in an iterative approach, where in each iteration step a single subcarrier k is allocated. In the current step, the user with instantaneous lowest R_i gets the subcarrier on which he observes the highest $R_{i,k}$. The iteration ends when all N subcarriers are assigned.

In contrast to the HE approach with additional constraint K_{max} (cf. section 9.2 and 9.3), in HE2 there is no free parameter like K_{max} to influence the allocation.

After this introduction, simulation results for HE2 are presented. The performance of HE2 in terms of throughput and fairness is compared to the previously introduced HE approach.

9.5   Simulation Results

The presented results are based on the channel model introduced in section 8.3, which is also used in the previous sections.

The section below evaluates the performance of HE2 in terms of throughput and fairness. Additionally, comparisons with previously considered HE approaches are drawn.

9.5.1   Fairness Level based on HE2 Approach

The HE2 algorithm is intentionally designed to provide a high level of fairness between users. This level is quantified by evaluating the CDF of the HE2 approach. In Fig. 40, the CDF of the HE2 approach is shown for six users inside the cell (purple curve) and $R_B = 8$ bits per OFDM symbol. Also drawn are curves for various HE approaches, one with unconstrained K_{max} (red) and two with constrained K_{max} (blue and green). Curves for different values of R_B are not shown, since they generally possess the same shape as in Fig. 40.

If the curve of HE2 is compared with the curves of the other HE approaches it becomes clear that the HE2 approach gives a much higher fairness, since the users inside the cell receive almost identical rates R_i. The R_i values are even less distributed than in the HE approach with $K_{max} = 5$, where in average each user gets the same amount of subcarriers. For illustra-

tive purposes, also the curve for $K_{max} = 32$ is shown, which shows the widest distribution of rates R_i. This is due to the unconstrained number of subcarriers per user in this approach.

The comparison of the four curves also shows the strong impact of the actual implementation of phase 2. Thus, the partition of the subcarrier allocation process in two phases is a powerful means to adjust the system performance. The fact that low-performance users are preferred in the allocation process in HE2 gives the highest fairness level of all considered allocation approaches so far.

Since a general tradeoff between fairness and throughput exists, the following section evaluates the system throughput R achievable by HE2.
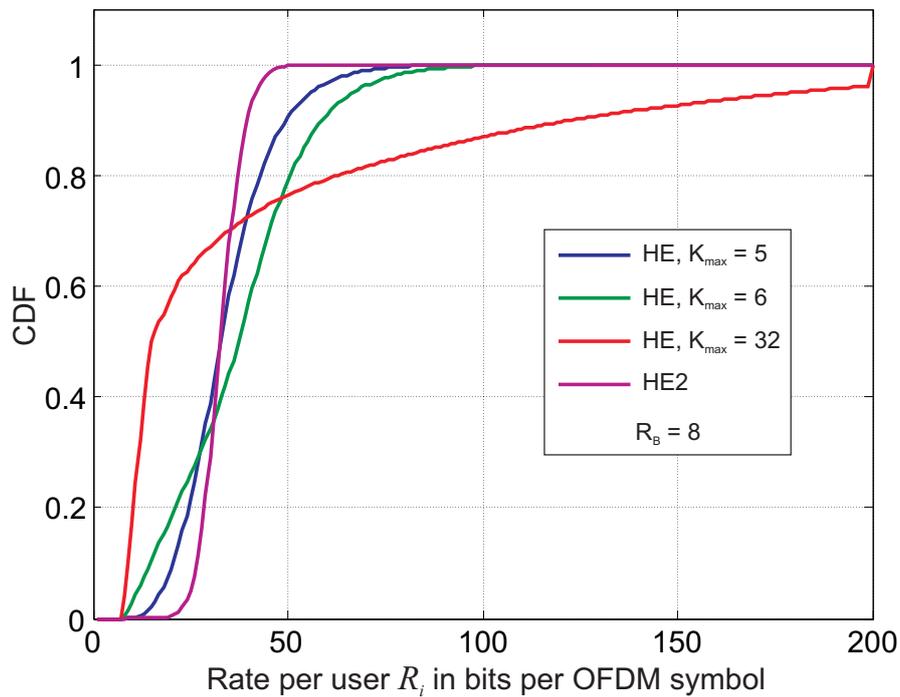
Fig. 40: CDF of user rates for $R_B = 8$; 6 users inside cell, HE and HE2 approach

9.5.2   Throughput of HE2 Approach

Considering the cell throughput R, the expected performance of the HE2 approach is lower than in the other HE approaches, which explicitly maximize R. A comparison between HE and HE2 considering the cell throughput is shown in Fig. 41: The upper curve shows the cell throughput for an $R_B = 8$ calculated by the HE approach as introduced in Fig. 25. The lower curves show the cell throughput calculated by HE2 for various values of R_B.

The most obvious difference between the curves is, that the HE2 approach loses at most 25% of cell throughput R compared to the HE approach that tries to maximize R. This is due to the different assignment strategy in phase 2: In the HE2 approach, users with low R_i are favored in the subcarrier assignment process. Since these users have a quite low average channel per-

formance, the overall throughput R only grows slowly. Hence, the performance loss in terms of throughput is quite substantial for the whole range of users inside the cell.

Another important observation is that the curves of the fair HE2 approach almost coincide with each other. The reason for this behavior is also found in phase 2: The sequence of the iterative subcarrier allocation in phase 2 depends on the instantaneous rates R_i of the users. Since in each iteration step the user with the lowest R_i gets an additional subcarrier and thus increases his R_i, in the next iteration a different user will be the one with lowest R_i. In this way, the set of subcarriers will be partitioned quite evenly between the users.
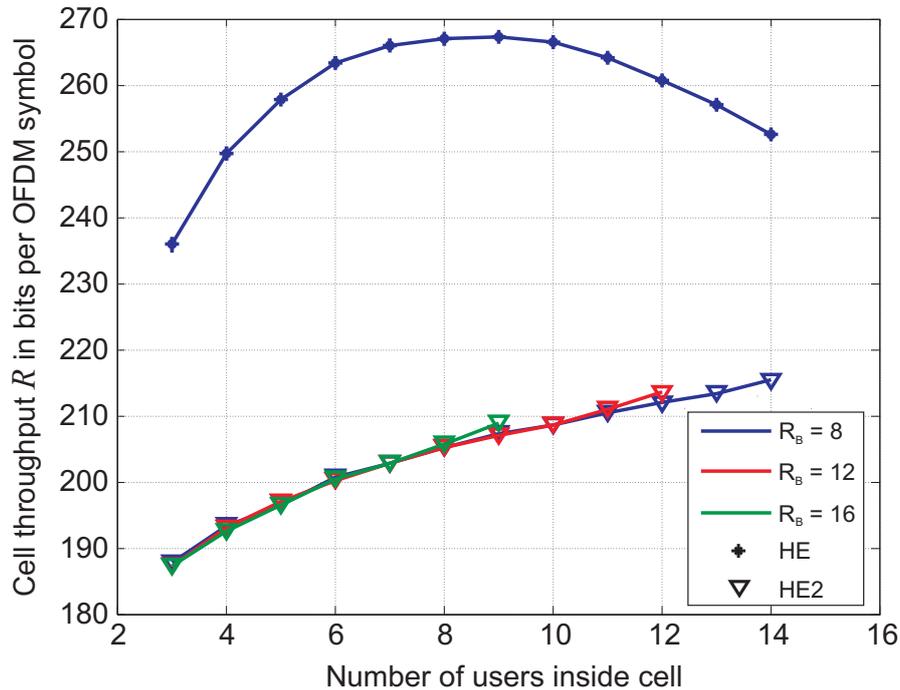
Fig. 41: Comparison between HE approaches: unconstrained HE and fair HE2 approach

Since the allocation is based predominantly on the instantaneous R_i and thus almost each user is taking a turn in the iteration, the cell throughput gained in phase 2 reflects the average channel properties inside the cell. Therefore, the amount of throughput gained in phase 2 solely depends on the number of subcarriers still available after phase 1 and hence on N_U as well as on R_B.

The relation between throughputs gained in phase 2 and N_U is highlighted particularly by Fig. 42. This figure shows the contributions of phase 1 and phase 2 to the overall throughput R separately. It is obvious that the throughput added by phase 2 (triangular markers) decreases with increasing user number N_U. The opposite effect is observed for phase 1, since this part of the algorithm tries to allocate a rate $\geq R_B$ to each user. This leads to an increase of system

throughput R proportional to N_U. Both effects more or less cancel out each other, thus leading to the slowly increasing throughput curves of Fig. 41.

Observing a HE2 curve of a specific R_B value (cf. Fig. 41), a small increase in the slope for a high user number N_U is visible. This situation occurs, because for a high N_U almost all available subcarriers are already allocated in phase 1 and thus often phase 2 is not even executed. Since phase 1 has the much larger contribution to R in high N_U scenarios (cf. Fig. 42), the slope of the throughput curve increases slightly at this point.
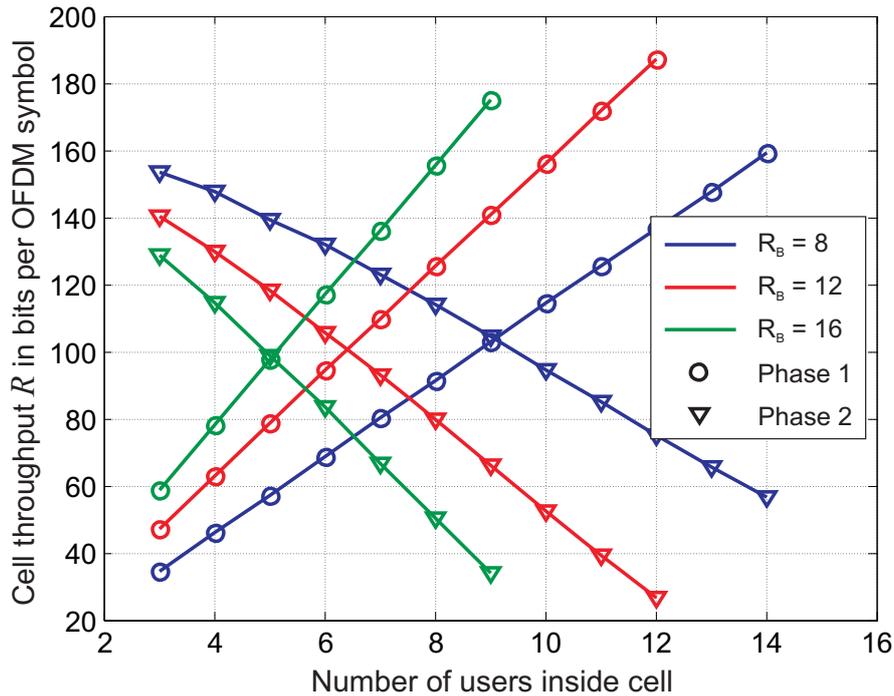
Fig. 42: Cell throughput after execution of phase 1 and phase2 of HE2 approach

In essence, the HE2 approach trades in throughput for fairness as already expected. But in order to rank this approach more precisely, detailed comparisons with the K_{max} constrained HE approaches from sections 9.2 and 9.3 are carried out in the following.

9.5.3   Comparison of HE2 and K_{max} constrained HE

The preceding section discussed the general properties of HE2 and showed its performance with respect to fairness and throughput. Since in sections 9.2 and 9.3.3, two variants of the K_{max} constrained HE approach were introduced, which were also designed to increase the level of system fairness, the chapter at hand gives a more elaborate comparison between HE2 and the K_{max} constrained HE approaches. The differences between these approaches in terms of fairness are considered in particular.

Fig. 33 shows that the fairness of the K_{max} constrained HE depends on the value of K_{max}. Hence, this parameter can be used to adjust the system's fairness level. On the other hand, Fig. 38 reveals a dependency between the fairness level and N_U. This fact led to the introduction of the HE approach with variable K_{max} in section 9.3.3. There, the fairness level is kept almost independent of N_U by adjusting K_{max} according to the instantaneous N_U. The effect of this approach on the fairness level is visible in Fig. 37. Although a considerable improvement is obvious compared to Fig. 38, still a slight dependency between fairness and N_U is observed.

In contrast to this, Fig. 43 shows the CDF curves of HE2 for various values of N_U. In this approach, the distribution of rates R_i is almost independent of N_U. Also, the slope of each curve in Fig. 43 is much steeper than that of any curve in Fig. 37, which indicates a much higher level of fairness.
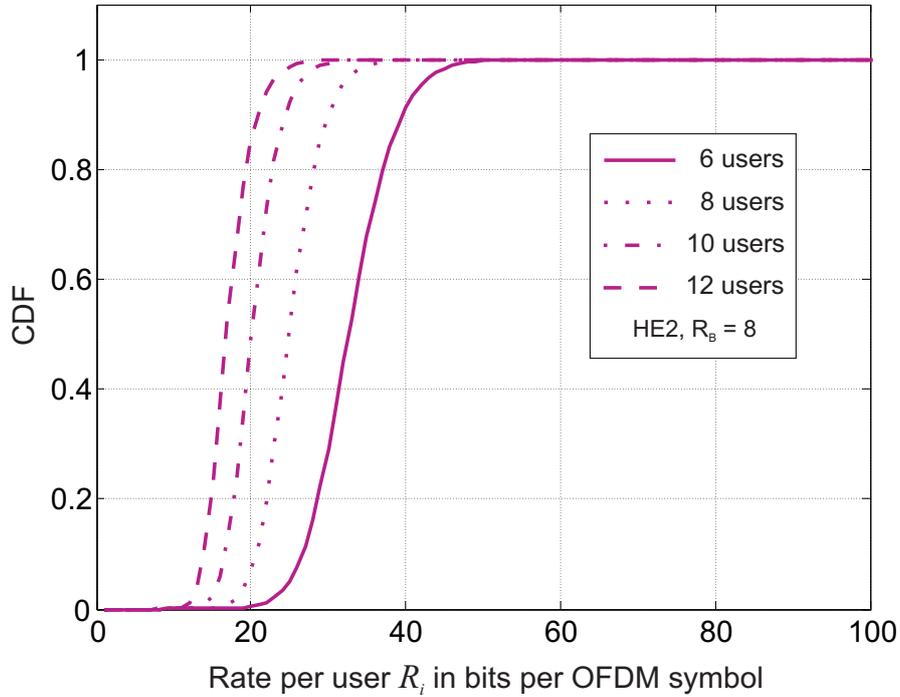
Fig. 43: CDF of HE2 approach; varying user number

As the user number increases, the average Rate R_i for each user gets smaller, since the resources are divided up between more users. This leads to a shift of the CDF curves to the left. Nevertheless, the slope of the CDF stays steep, showing a low variance on R_i and thus a fair system.

The difference in fairness level between the K_{max} constrained HEs and the HE2 approach is due to the respective allocation strategy: On the one hand, the constrained HE tries to balance

the number of allocated subcarriers per user using K_{max}, which leads to an indirect balancing of the user rates R_i. On the other hand, the HE2 approach directly balances the rates R_i by using a modification of phase two. This leads to an advantage in terms of fairness.

After the consideration of fairness, as a second performance measure for the subcarrier allocation approaches the system throughputs are compared in the following paragraph.

The comparison is done on the basis of Fig. 44, where all four approaches, namely the pure HE, which explicitly maximizes throughput (cf. Fig. 25), the K_{max} constrained HE, the HE with variable K_{max}, and HE2 are represented. It is not surprising that the HE approach for maximizing R achieves the highest throughput over the complete range of users.

In the low N_U range, the K_{max} constrained HE approach with $K_{max} = 5$ shows the lowest throughput of all three approaches. The reason being that not all of the N available subcarriers are actually allocated. E.g. for 3 users inside the system, only 15 subcarriers are actually used.
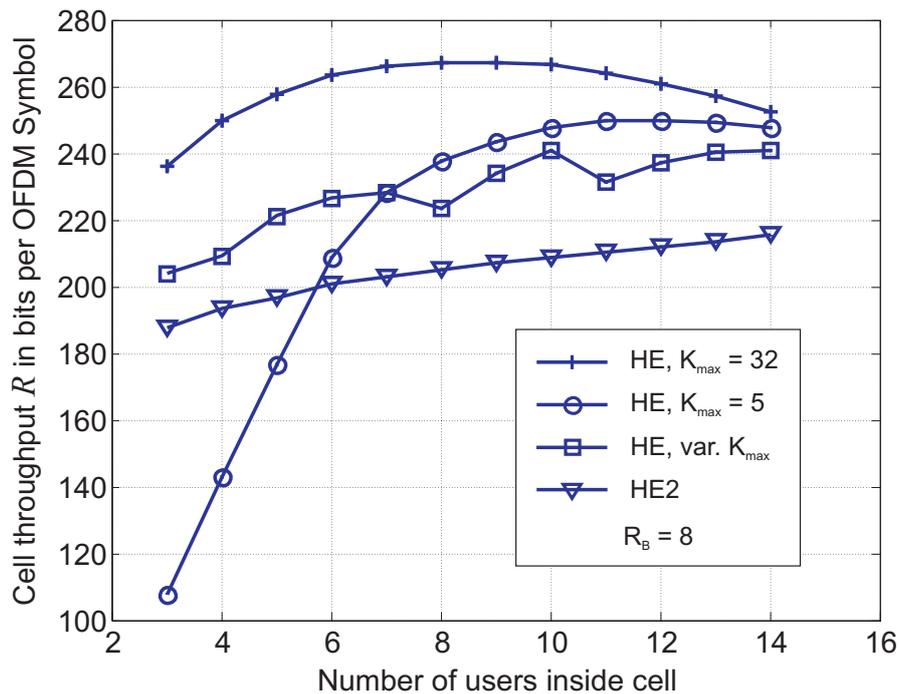
Fig. 44: Throughput of various HE approaches: unconstrained, K_{max}-constrained and HE2

If larger user numbers are considered, the gap between the curves for $K_{max} = 5$ and $K_{max} = 32$ almost closes. The reason for this phenomenon lies in the decreasing impact of K_{max} for increasing N as explained in section 9.3.1. This high increase in throughput also indicates the

simultaneous loss in fairness for increasing N and the convergence to the K_{max} unconstrained HE approach.

The square-marked curve represents the HE approach with adaptive K_{max}. This approach achieves higher throughput than HE2, although its increase is slightly unsteady. This fact was already discussed in section 9.3.3.

Considering the HE2 approach, the progression of its curve is more or less constant over the user number. Its general shape being already explained in section 9.5.2, it is noteworthy that the HE2 approach shows the highest fairness level at the cost of the lowest throughput so far.

An especially interesting aspect of Fig. 44 is found, if all four throughput curves are evaluated at $N_U = 6$. This point corresponds to the fairness situation depicted in Fig. 40 for all four allocation approaches. The approach with variable K_{max} is in this case represented by the curve for $K_{max} = 6$ in Fig. 40. The comparison of both figures particularly illuminates the tradeoff between fairness and throughput:

While the throughput of K_{max} constrained HE and HE2 is almost identical, this also holds for their fairness level. On the other hand, the unconstrained HE shows a much higher throughput than the other approaches for $N_U = 6$, but its fairness level is much lower.

Compared with all other allocation approaches discussed above, the HE approach with adaptive K_{max} shows intermediate performance in terms of throughput and fairness level.

Summing up the results for this section, the K_{max} constrained HE approach has the possibility to adjust fairness and throughput by changing K_{max}, but shows at the same time a dependence on N_U if K_{max} is kept constant. This behavior is mitigated by adapting K_{max} according to N_U, although still a slight dependency on N_U remains since adaption of K_{max} is only feasible in integer steps.

On the other hand, the HE2 approach shows almost constant fairness and throughput performance for a wide range of users, but does not allow adjusting one of these performance measures. But indeed the HE2 approach proves that modification of phase 2 of the HE approach is a suitable means to influence the fairness of a system.

Thus, the following chapter also considers a modification solely of phase 2. This time, the drawbacks of HE2 and the HE approach with variable K_{max} are avoided by allowing a smooth fairness adjustment, which is also independent of N_U.

9.6   Increasing Fairness Level by Use of Utility Function

The approaches to increase the fairness level discussed so far show either a strong dependency on the number of users or allow only a limited or even no adaptation of the fairness level.

Now a resource allocation approach is introduced, which maintains a desired fairness level independently of the number of users N_U inside the system. Also, the level of fairness can be adapted by modifying a single system parameter.

As in the preceding chapters, the considered resource allocation process is based on the HE approach. The modifications applied to the HE are limited to phase 2 as in section 9.4.

The difference to the modifications done in section 9.4 is that the allocation of subcarriers to users is based on a so-called *utility function* $U(R)$. This function is directly dependent on the system throughput R.

The following sections introduce the concept of utility and show how it can be integrated beneficially into the HE approach where it allows the flexible adjustment of the fairness level.

9.6.1   Introduction to the Concept of Utility

In general, a *utility function* $U(R_i)$ is a mapping of a user's throughput R_i into a *utility* or rather *quality* value that represents the level of satisfaction of the considered user *i*.

The concept of a utility function is widely-used in literature as a means to describe an application's or user's behavior in adaptive networking environments. Originally introduced for wireline transmissions [She95], utility functions are also common in wireless mobile systems as described by [Boc06], [Jia05], [Gao01], [Bia98], and [Lee95].

Utility functions are mostly used to provide a system-wide level of fairness. This is generally done by maximizing the aggregate utility $\sum_i U(R_i)$ for all users *i* inside the system. It is shown in [She95] that such an approach increases the system's fairness level.

Thus, in the context of this work, the utility function (UF) is used as a means to provide fairness between users. In order to do that, a UF must fulfill certain criteria, which are introduced in the following, cf. [Gao01]. These criteria enable the network operator to specify a UF according to desired throughput and fairness constraints. The necessary criteria are given as follows:

Consider a real function $U(R_i)$ subject to a user's data rate R_i. In order to represent a UF, $U(R_i)$ is assumed to be *continuous*, *differentiable*, *increasing*, and *strictly concave* for values $R_i \geq 0$. The fact that $U(R_i)$ is concave implicitly provides a certain level of fairness between users, assuming the user satisfaction solely depends on R_i.

In [She95], also other function types for $U(R_i)$ are introduced, which are not strictly concave. These functions are used to model the behavior of applications, which are subject to certain delay constraints. Since in the work at hand the performance of a user *i* is exclusively measured in terms of rate R_i, all considered utility functions are strictly concave, cf. Fig. 45.
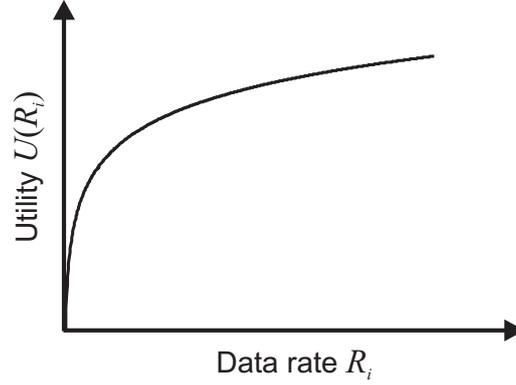
Fig. 45: Typical shape of concave utility function

It is shown in [Jia05] by means of user surveys that UF fulfilling the introduced criteria are well suited to describe the user satisfaction for popular wireless services like web browsing. This also becomes clear intuitively by the following example:

If $U(R_i)$ is concave, the increase of utility (i.e. satisfaction) decreases over R_i. Hence, a user with a low rate R_i observes a higher boost in satisfaction if he gets an additional resource, as compared to another user who is already provided with a high rate and thus won't observe a significant gain of utility. This behavior of the concave UF directly corresponds to the subjective performance perception of a user who is browsing the internet. Hence, it is beneficial to increase the rate of a low-performing user instead of boosting the rate of an already satisfied user. Thus, the gain of utility a specific user achieves can be used as an indicator for the fairness of a particular subcarrier assignment. The example is understood easily by looking at the example UF given in Fig. 45. This intuitively shows the concept of fairness behind the utility function.

After this conceptual description of utility functions, their analytical background is introduced: The goal of the utility-based resource allocation considered in the following is to find a subcarrier allocation scheme that maximizes the overall utility of the system, which is given as

$$\sum_{i=1}^{N_U} U(R_i) = \max_{R_i} . \tag{9.7}$$

This means the sum of utilities corresponding to each user *i* must be maximized.

The choice of the actually used UF must comply with the design criteria mentioned above. Additionally, the utility function should contain a parameter ρ, which allows to adjust the fairness level inside the system because this was the motivation to apply a UF in the first place.

Realizations of UF found in literature are mostly of exponential or logarithmic form [Boc06], [Jia05], since these forms readily fulfill the necessary design criteria. *In the remainder of this work solely exponential UF are considered due to a strong advantage that will be explained soon.*

The UF considered from now on is of the following form:

$$U(R_i) = a + b(R_i + c)^\rho$$

$$\text{for } a,b,c,\rho \in \mathbb{R}; \; -\infty < \rho < 1; \; R_i \geq 0$$

(9.8)

The parameters *a*, *b*, and *c* are real constants used to set the origin and asymptotic behavior of the UF. However, the most important parameter of (9.8) is the real constant ρ. This parameter is used to adjust the fairness level inside the system as will be shown later.

It is trivial to show that (9.8) fulfills the criteria of continuity, differentiability, and is also increasing as well as strictly concave over R_i if all parameters are suitably chosen.

The property of concavity is given, if for the second derivative of $U(R_i)$

$$\frac{d^2 U}{dR_i^2} < 0$$

(9.9)

holds. In general, the second derivative of (9.8) is given by

$$\frac{d^2 U}{dR_i^2} = b(\rho^2 - \rho)(R_i + c)^{\rho - 2}.$$

(9.10)

It is easy to show, that (9.10) fulfills (9.9) when *b*, *c*, and ρ are chosen accordingly.

Equation (9.10) also reveals that the second derivative of $U(R_i)$ still depends on the parameter ρ. Thus, also the slope of the satisfaction-gain a user achieves for an increase in R_i depends on ρ.

This leads to the finding that the parameter ρ actually provides a means to adjust the fairness between users.

An advantage of the exponential UF over other known utility approaches (e.g. logarithmic) is that the change of its slope ranges from extremely concave ($\rho \rightarrow -\infty$) to perfectly linear ($\rho \rightarrow 1$). This allows a wide adjustment of the fairness level, as will be shown in later sections.

9.6.2   Suitable Choice of Utility Parameters

For the choice of the parameters *a*, *b*, *c*, and ρ two separate cases must be considered: The first case is the most common and is distinguished by assuming $\rho < 0$. In this case, the utility curve shows a shape similar to Fig. 45.

In order to fulfill the criteria valid for a UF (especially increase and concavity) the remaining parameters are chosen as follows:

Case 1: $\rho < 0$

$$a > 0, \ b = -a, \ c = 1 - R_B \qquad (9.11)$$

It is straightforward to show that (9.8) fulfills all criteria of a UF if (9.11) holds. In the considered case, a fulfills the task of an upper bound to $U(R_i)$, while b and c are chosen in such a way that $U(R_i) \to 0$ if R_i approaches R_B. This is done since the utility-based subcarrier allocation takes place in phase 2 of the HE approach, where all users can be assumed to have been allocated a rate $R_i \geq R_B$.

Fig. 46 shows a collection of UF for a common range of ρ, which conforms to case 1.
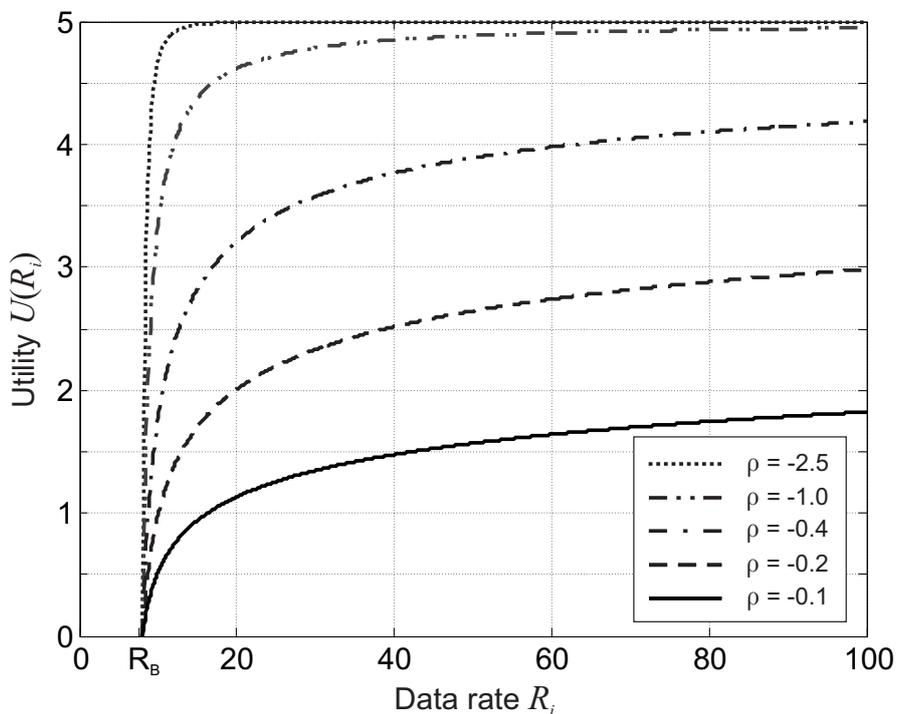
Fig. 46: Typical curves for exponential utility with negative ρ, R_B = 8. Parameters a = 5, b = -5, c = -7.

The second case of (9.8) to be considered is to assume a range of ρ given by $0 < \rho < 1$. As $\rho \to 1$, the UF assumes an almost linear shape. The corresponding impact on the fairness level of the system is shown in the next section. For the time being, the remaining parameters of (9.8) are chosen in order to fulfill all UF criteria:

Case 2: $0 < \rho < 1$

$$a < 0, \ b = -a, \ c = 1 - R_B \qquad (9.12)$$

The slight modifications on a and b compared to (9.11) are done to ensure the increase of the UF over R_i. In practical systems, the signs of a and b can simply be interchanged in order to

switch between case 1 and case 2. Typical UF curves for case 2 with positive values ρ are depicted in Fig. 47.
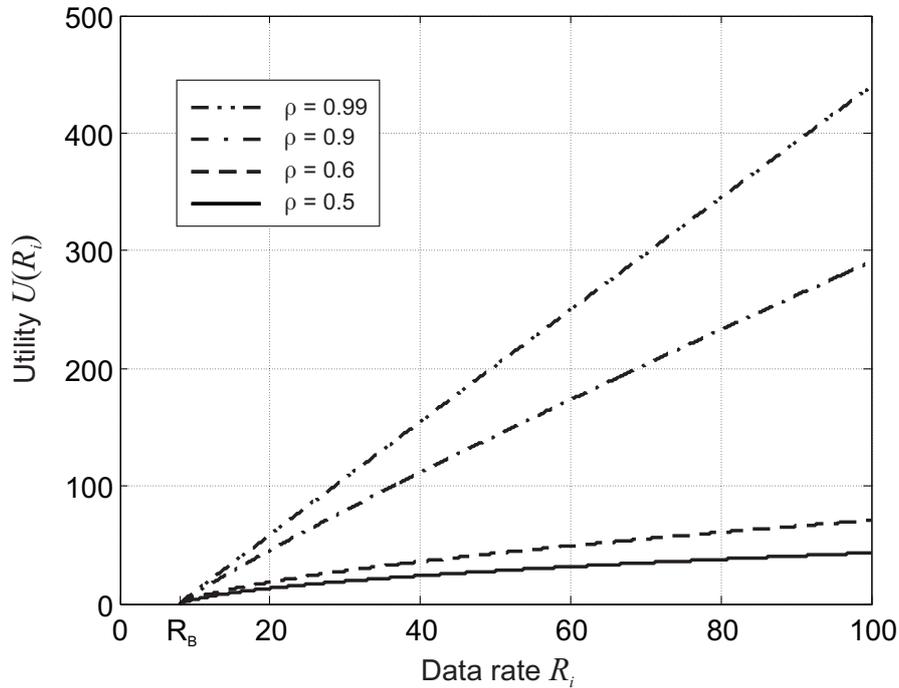
Fig. 47: Typical curves for exponential utility with positive ρ, R_B = 8. Parameters a = -5, b = 5, c = -7.

The example curves depicted in Fig. 46 and Fig. 47 impressively show the wide range of influence of the parameter ρ on the utility function. Thus, based on the findings of the previous section, a strong impact of the parameter ρ on the system's fairness level is to be expected and is actually proved in the upcoming sections.

After this introduction to the exponential utility function its integration into the second allocation phase of the HE approach is discussed in the following.

9.6.3   Subcarrier Allocation Process considering Utility

In section 9.4 it was shown how the system's fairness level is increased strongly by way of modifying phase 2 of the HE approach. This line of resource allocation is pursued further, only this time the main drawback of the previously introduced HE2 approach is avoided:

Besides the excellent fairness level provided by the HE2 approach, its fairness level remains fixed and is not adjustable.

Thus, the modification discussed in this section includes a utility function to be applied in the resource allocation process. It was shown in section 9.6.1 and 9.6.2 that UF are able to describe a user's satisfaction and are at the same time easily adjustable to the actual quality expectations of the user.

The integration of UF into phase 2 of the HE approach is done as follows. Phase 1 again remains unchanged. The resulting modification of the HE approach is referred to as *HE3* from now on. The HE3 approach considered in the following has two main characteristics:

- Phase 1 is kept unchanged

- The allocation of subcarriers in phase 2 is done subcarrier by subcarrier. A subcarrier is allocated to a particular user if this allocation decision maximizes the overall utility of the system.

Phase 2:

After completion of phase 1, all users are supplied with a rate R_B or higher. The not yet assigned subcarriers are used to balance the individual rates R_i of the users. Also in HE3, the allocation process is iterative, hence in each iteration step a single subcarrier k is considered and assigned to a specific user i. This is done in such a way, that in each iteration the utility $U(R_i)$ reached so far by each user i is calculated using the exponential utility function introduced in (9.8) and applying the parameters given in (9.11) or (9.12) respectively.

Regarding the instantaneously considered subcarrier k of a particular iteration, the potential utility $U(R_i + R_{i,k})$ is calculated for each user i. The value $U(R_i + R_{i,k})$ represents the utility of a user i that would be achieved if the subcarrier k was allocated to him.

W.l.o.g., for the time being the potential utility value of user 1 with $U(R_1 + R_{1,k})$ is considered. This value is summed up together with the values $U(R_i)$ of the remaining users $i \in [2, N_U]$. Such a sum is calculated for each user, yielding N_U new utility values representing the new potential system utility. The largest of these values is chosen and the considered subcarrier k is allocated to the user i responsible for this highest utility gain.

Thus, phase 2 is a heuristic approach pursuing the goal of maximizing the system's utility given in (9.7). The introduced algorithm does not find the global maximum of (9.7) since the overall system's utility depends on the sequence in which the subcarriers are allocated.

However, as the simulation results will show, the overall performance of the HE3 approach is by no means lower than that of the other HE approaches. Actually, it will be shown that the HE3 approach can be adjusted from extreme fairness levels back to an exact equivalent of the original HE.

9.7   Simulation Results

The preceding sections gave an introduction to the concept of utility and its integration into a resource allocation process as a means to influence the system's fairness level. The theoretical findings are now fortified by simulation results. The system and channel model applied in the following is identical to that of section 8.3 and all consecutive sections.

The utilized utility function corresponds to the exponential approach discussed in sections 9.6.1 and 9.6.2. In the paragraphs below, the influence of the utility parameter ρ on the fairness level and system throughput is evaluated for the recently introduced HE3 approach. Also, performance figures of the HE3 are compared to the performance of the earlier discussed HE approaches.

9.7.1 Fairness Level based on HE3 Approach

The utility-based HE3 approach allows an exceptional wide range of fairness levels. Similar to the approach given in section 9.3.1, the level of fairness is dependent on an adjustable system parameter. Only this time, no interdependencies with the number of users N_U are observed, as will be demonstrated later. The fairness level inside the system is evaluated in terms of the CDF of the achieved user rates R_i.

Fig. 48 shows the strong impact of the utility parameter ρ on the system's fairness level. The lowest fairness levels are observed for positive values of ρ approaching $\rho \rightarrow 1$. In this case, the utility curves show an almost linear shape (cf. Fig. 47). Thus, the gain in utility for the allocation of a specific subcarrier is almost independent of the previously allocated rates $R_{i,k}$ and solely depends on the absolute rate gain connected to the considered subcarrier. Especially if $\rho = 1$, the maximization of utility (9.7) degrades to

$$\sum_{i=1}^{N_U} R_i = \max_{R_i}, \tag{9.13}$$

which corresponds to a straightforward throughput maximization as in the first HE approach.

This is confirmed by the CDF curve for $\rho = 0.9$ (cf. Fig. 48), which shows almost the same shape as the CDF curve for the unconstrained HE approach (cf. Fig. 33). If $\rho = 1$ was set, both curves would match exactly.

Reducing the value of ρ more and more increases the fairness level until for $\rho = -2.5$, the CDF curve in Fig. 48 shows an almost rectangular shape. The corresponding utility curve $U(R_i)$ is given by the dotted curve depicted in Fig. 46: For very low rates R_i, a steep increase in utility is visible, whereas for all rates above a certain level the utility function remains practically constant. Hence, users with low rates R_i are preferred by the subcarrier allocation process and users who reached a certain rate level are not likely to get any more subcarriers, unless they are the users with the actual lowest accumulated rate. This fact corresponds to a highly fair allocation, which also shows in the CDF curve for $\rho = -2.5$ in Fig. 48. Values $\rho < -2.5$ are not considered in this work since the corresponding utility functions reach an almost constant utility value for quite small values R_i and thus lose their impact on the resource allocation process.
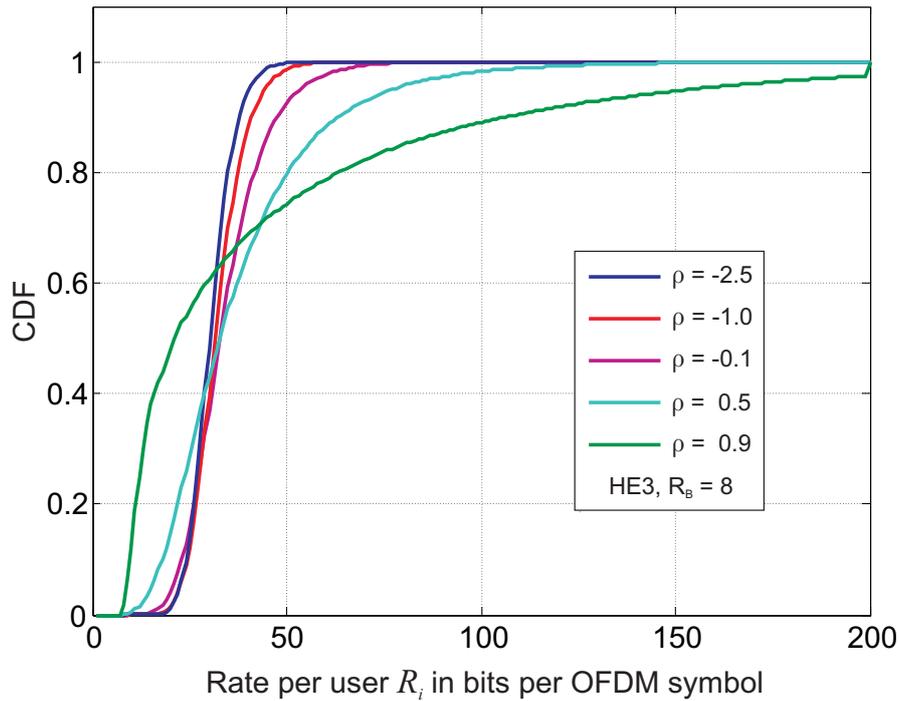
Fig. 48: CDF of user rates (HE3 approach) for R_B = 8 and various values ρ; 6 users inside the cell

Between these two extreme cases from $\rho = 0.9$ to $\rho = -2.5$, a gradual changeover is visible in Fig. 48. Thus, the parameter ρ of the exponential utility function is an excellent means to adjust the system's fairness level.
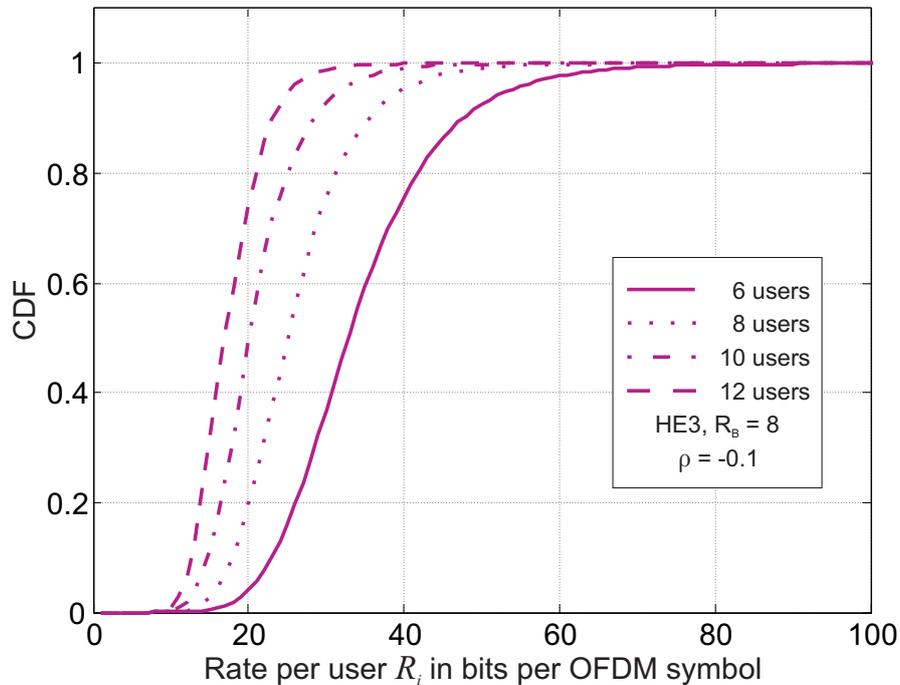
Fig. 49: CDF of HE3 approach with ρ = -0.1; varying user number

After these revealing results, the focus is moved to the stability of the fairness level for a changing number of users N_U. To this end, Fig. 49 shows a set of CDF curves for a constant parameter $\rho = -0.1$ and varying N_U ranging from 6 to 12 users.

The slope of the curves remains almost constant for all N_U, indicating a stable fairness level. Thus, the HE3 approach shows a similar stability in terms of fairness as the HE2 approach discussed previously (cf. Fig. 43).

Compared to the stability of the HE with variable K_{max} (cf. Fig. 37), the independence of the fairness level from N_U is much stronger in the HE3 case.

In summary, the utility-based HE3 approach offers a widely adjustable fairness level depending on a single parameter ρ, which ranges from strong rate balance $(\rho \to -\infty)$ to throughput maximization $(\rho \to 1)$. At the same time, the HE3 approach avoids the drawback of the fairness level being dependent on N_U.

These arguments make the HE3 approach the best resource allocation approach discussed so far in terms of fairness adaptivity. Detailed comparisons with other resource allocation approaches considering absolute fairness levels are carried out in the upcoming section 9.7.3.

The following section discusses the performance of the HE3 allocation approach focusing on the system's throughput and its dependence on the parameter ρ.

9.7.2 Throughput of HE3 Approach

After the discussion of fairness aspects, the dependence of the system throughput on the parameter ρ will be evaluated for the HE3 approach. The corresponding throughput curves are shown in Fig. 50. The minimum rate per user is set to $R_B = 8$. Six curves show the performance of the HE3 approach for varying values of ρ. As a reference, also the throughput curve of the unconstrained HE approach (i.e. $K_{max} = N = 32$) is depicted. Thus, the uppermost curve in Fig. 50 is identical to the uppermost HE curve in Fig. 25.

A superficial examination of Fig. 50 shows that the throughput curves are approaching the HE curve (cross markers) for increasing ρ. If $\rho = 0.9$ holds, the HE3 and the HE are a close match. If ρ was set to $\rho = 1$, both curves would even be identical. For the sake of clarity, this case is not depicted.

As hinted in the previous chapter, this striking resemblance is caused by the almost linear shape ($\rho = 0.9$) of the utility function $U(R_i)$. In case of $\rho = 1$, the utility function actually shows a perfect linear shape. Thus, the objective of Phase 2 of the HE3 approach becomes equal to (9.13), which means that each subcarrier k is allocated to that user who observes the highest $R_{i,k}$ on the instantaneously considered subcarrier. This is the exact equivalent of the

second phase of the original HE approach (cf. 8.1.2) and hence, the throughput curves for $\rho \to 1$ converge to the unconstrained HE curve (cross markers).
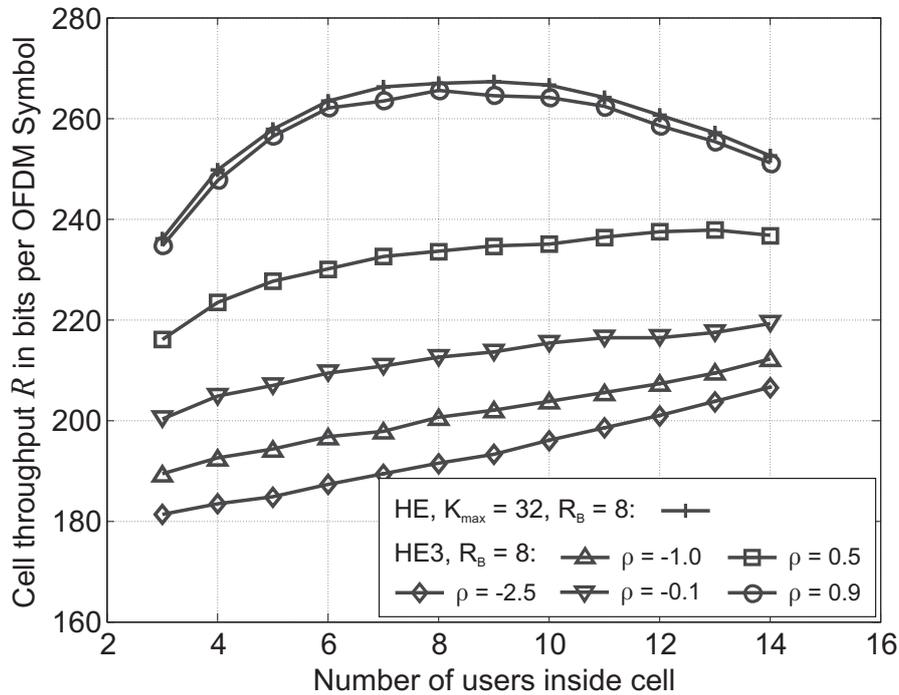
Fig. 50: Throughput of various HE approaches: unconstrained HE and HE3

Quite the opposite is true for $\rho < 0$: In this case, the throughput curves show a moderate but steady increase over N_U. Also, the overall throughput level decreases as ρ advances farther below zero. The flat shape of these considered curves (all angular marked) is caused by the same mechanisms as explained in section 9.5.2, where the throughput performance of HE2 is discussed: In these high-fairness approaches, users with low average performance are preferred by the subcarrier allocation. Thus, the overall system throughput is low.

Additionally, the subcarrier allocation result depends more strongly on the average channel properties inside the cell and not so much on the diversity between users. Therefore, the throughput curves for negative values of ρ are monotonously increasing over the user number N_U (e.g. triangular-marked curve) and are less bent than in case of pure throughput maximization (cf. cross-marked curve).

All in all, combining the findings of this section and of section 9.7.1, the HE3 approach allows an excellent adjustment of the balance between throughput and fairness by using a single parameter ρ.

The following chapter will contrast these findings with the resource allocation approaches discussed in earlier sections.

9.7.3 Comparison of Discussed HE Approaches

The preceding sections gave extensive performance figures for the HE3 approach in terms of fairness and throughput. In the following, these results are compared with those of the HE approaches discussed in the sections 9.3 to 9.6. This includes the K_{max} constrained HE, the HE2, the HE3, and the unconstrained HE approach for reference purposes.

First, the above-mentioned resource allocation approaches are compared in terms of their fairness level. To this end, Fig. 51 shows the CDF curves of various HE modifications. The original HE from section 8.1.2 is not considered due to its limited fairness level.
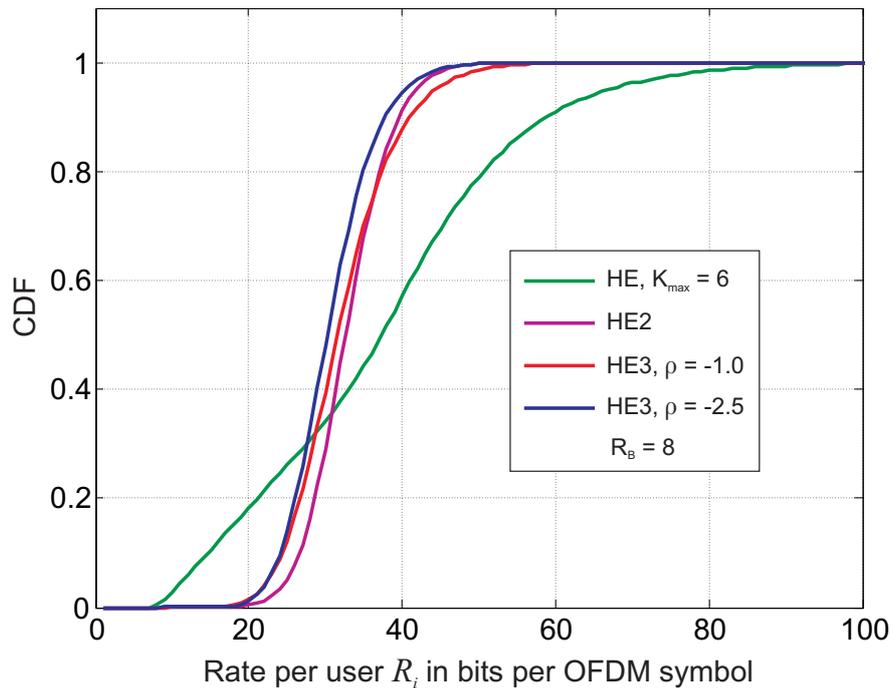
Fig. 51: CDF of user rates for various HE approaches, 6 users inside the cell

The lowest performance in terms of fairness is shown by the K_{max} constrained HE (green curve). Since a scenario with a fixed number of users $N_U = 6$ is considered, the resulting curve is equivalent to the HE with variable K_{max} constraint. The reasons for its limited fairness contribution were already discussed elaborately in section 9.3.3. Thus, the green curve merely serves as a reference in the considered figure.

The most interesting part of Fig. 51 is the comparison between the CDF curves of the HE2 and HE3 approach:

It shows, that the fairness level of HE2 (purple curve) and HE3 with a utility parameter of $\rho = -1.0$ (red curve) are almost equivalent. The HE2 offers only a slightly better fairness level, visible by the somewhat steeper slope of the curve. Further analysis shows that utility

parameters close to $\rho = -1.0$ yield the closest matches to the HE2 CDF curve in a least-square sense.

An adjustment of the HE3 fairness level, i.e. increasing the curve's slope, in order to match that of the HE2 is only possible by further decreasing the fairness parameter down to $\rho = -2.5$ (blue curve). Fig. 51 shows that this leads to an almost identical slope for HE2 and HE3, but at the cost of a slightly lower average rate R_i for the users. This is visible by the left-shift of the blue curve with respect to the purple HE2 curve.

The reason for this loss in throughput is found in the shape of the considered UF for $\rho = -2.5$ (cf. dotted curve in Fig. 46): If the aggregate throughput inside the system has reached a certain level, the utility curve becomes almost horizontal, such that only marginal utility differences are observed between various allocation alternatives. From a certain point, the allocation decisions become almost random. Thus, a slight loss in throughput must be accepted compared to the HE2 approach, which follows an unambiguous decision criterion independently of the aggregate throughput level. If the parameter ρ of HE3 was further reduced to $\rho < -10.0$, the UF would assume an almost rectangular shape leading to a completely random subcarrier allocation process. Thus, such a scenario would keep the overall fairness level, but would yield an even stronger loss in throughput than the scenario applying $\rho = -2.5$.

In summary, a decrease of ρ far below $\rho = -2.5$ does not produce any gain in terms of fairness but leads to considerable throughput losses.

Another aspect in terms of fairness is the stability of the fairness level for varying user numbers. Here, the stability of HE2 and HE3 is compared. Comparisons for other HE approaches are found in preceding chapters.

Fig. 52 shows CDF curves for HE2 and HE3 considering varying user numbers N_U ranging from $N_U = 6$ to $N_U = 12$. The considered utility parameter for HE3 is $\rho = -1.0$, since its fairness level and throughput performance is similar to that of the HE2 approach.

A first survey of Fig. 52 reveals that the fairness level of HE3 (blue curves) stays almost constant over N_U. Only a slight increase of slope is notable for an increasing number of users due to a stronger impact of MUD.

Comparing the curves for HE2 (red curve) and HE3 (blue curve), a small difference in slope between both approaches as already observed in Fig. 51 is visible. This difference is sustained for all values of N_U although it decreases proportionally to the increase in N_U.

In summary, the HE3 approach is almost equal to the HE2 approach in terms of fairness level and fairness stability, assuming that the utility parameter ρ is chosen appropriately. Although the HE2 approach shows a slightly higher fairness level, the utility-based HE3 approach has still the advantage of adjustable fairness.
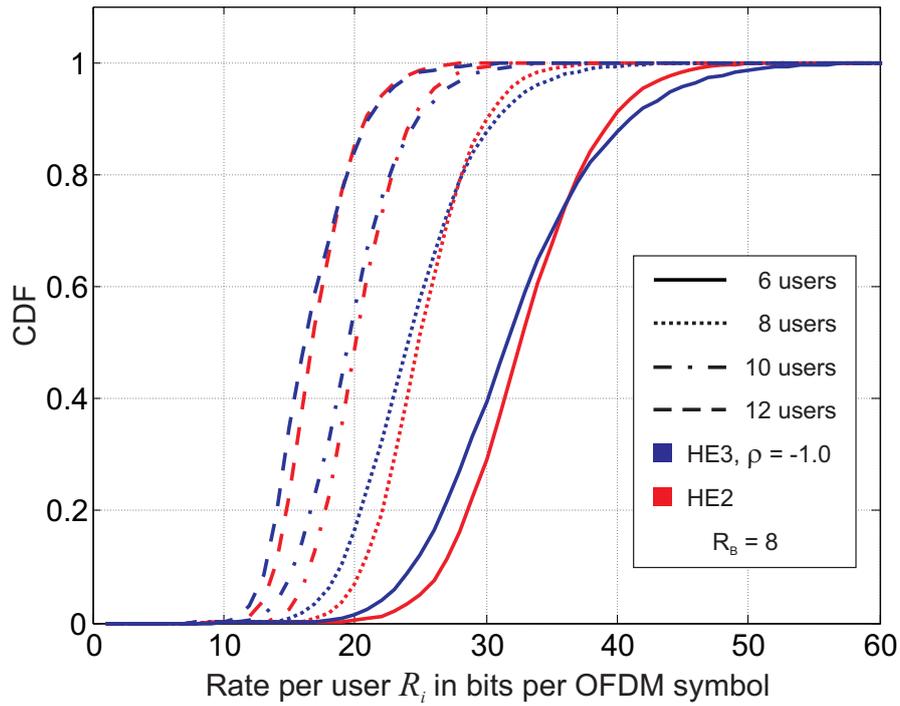
Fig. 52: CDF of user rates for HE2 and HE3 approach, varying user number
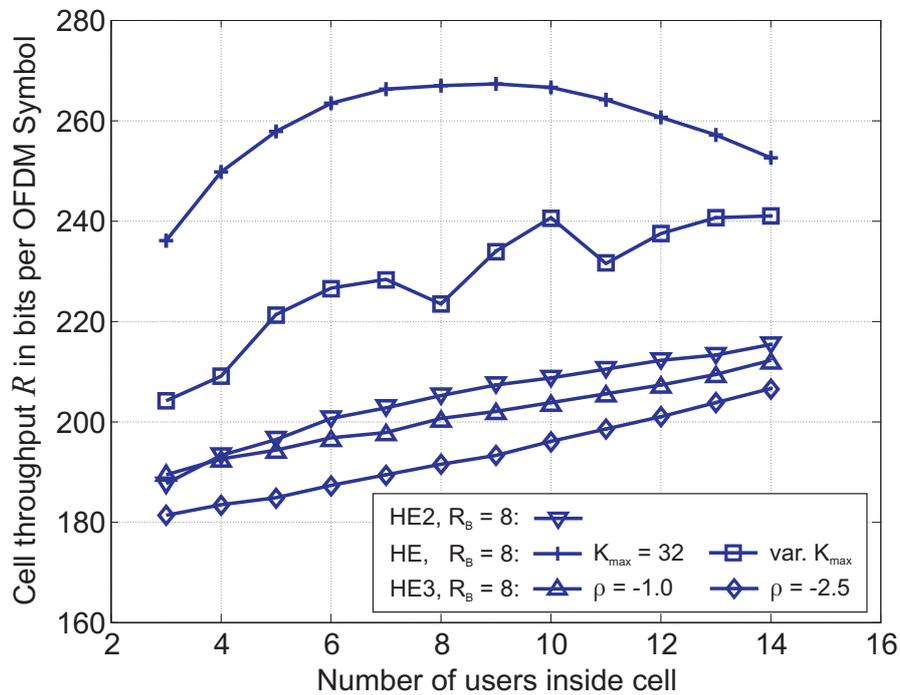
Fig. 53: Throughput of various HE approaches: HE, HE2, and HE3

Keeping these results in mind, a final comparison of all HE resource allocation approaches discussed so far is done in terms of system throughput. Fig. 53 gives a thorough overview of

the discussed approaches and allows a qualitative ranking of their system throughput. As always, the unconstrained HE serves as a benchmark to the other approaches.

The comparison between the HE2 approach (downward-pointing triangle) and its utility based counterparts (triangle and diamond) are of particular interest. As already indicated in Fig. 52, the throughput of an HE2 is very similar to its HE3 counterpart if $\rho = -1.0$ is chosen. In this case, the difference in throughput R between both approaches amounts to four percent at most. For low user numbers, their throughput even matches.

If the utility parameter ρ is further decreased to $\rho = -2.5$, only a slight gain in terms of fairness is achieved (cf. Fig. 51) but at the cost of considerable throughput losses over the whole range of N_U. The loss of the HE3 approach in terms of throughput compared to HE2 is partly caused by the almost rectangular shape of the UF for $\rho < -2.5$. This explains the notable throughput offset between the triangular marked curve ($\rho = -1.0$) and the diamond-marked curve ($\rho = -2.5$).

But this does not explain sufficiently the throughput-difference between the HE2 curve and the HE3 curve for $\rho = -1.0$. In the latter case, the UF still shows a sufficient slope even for high throughput values (cf. Fig. 46). Thus, another influence must be responsible for the obvious throughput loss. In the considered case, this loss is also partly caused by the different search depths of both algorithms: In the HE2 case, for the considered user, the best available subcarrier is chosen leading to a search over all available subcarriers. In contrast, the HE3 approach merely iterates from subcarrier to subcarrier and chooses a suitable user for each subcarrier based on the current utility value. This leads to allocation decisions with slightly less aggregate throughput. The reason for using the latter procedure in HE3 is its direct deduction from the original HE approach, which allows a straightforward comparison (cf. Fig. 50).

In principle, an adaption of the HE3 iteration process towards the HE2 process is feasible, but would only yield minor performance gains. Thus, this line of inquiry is not pursued since the focus of the current evaluations lies on the flexibility of the HE3 approach.

The behavior of the HE3 approach is concluded by the following finding: At a certain level of fairness, no significant gains are achievable in this respect. Further decrease of the utility parameter ρ in order to improve the fairness level yields only fractional fairness gains but on the other hand leads to heavy throughput losses.

Summarizing the discussion of the heuristic resource allocation approaches introduced so far, it is obvious that the HE3 approach offers the highest flexibility in balancing the throughput and fairness level inside a transmission system. Its slight performance loss in high-fairness scenarios compared to the HE2 is more than compensated by its wide range of fairness adaptivity.

The additionally considered HE approach with variable K_{max} constraint is only advantageous with respect to its simplicity of implementation but lacks adaptivity as well as stability in scenarios with varying user numbers.

The original HE approach (section 8.1.2) merely serves as a benchmark for maximum achievable throughput and can be reproduced by the HE3 approach. In general and also not surprisingly, all considered resource allocation schemes can not break the fundamental tradeoff between the system's throughput and fairness level.

10 Summary

This thesis considered the task of subcarrier allocation to multiple users in a single-cell OFDM-FDMA system. Particular attention was directed to the two aspects of allocation efficiency and fair sharing of channel resources. In this context, adaptive and non-adaptive allocation schemes were compared, which revealed strong performance advantages for adaptive schemes in frequency selective channel environments.

A particular performance gain was achieved, if the subcarrier allocation task was formalized as an optimization problem, which was then solved by well-known and efficient algorithms. This approach works excellent in frequency selective channel environments solely impaired by multipath propagation.

If additionally path loss and shadowing effects are present, the subcarrier allocation in multi-user systems proved to be a particular challenge: These effects cause highly individual link qualities between each user and the base station. Thus, also the number of resources each user claims in order to achieve his data rate demands is highly individual.

On the one hand, this individuality dramatically increases the computational complexity of finding an optimum subcarrier allocation. This led to the introduction of a novel heuristic allocation approach (HE), which was compared to well-known optimum algorithms such as binary integer programming (BIP).

It was shown that the heuristic HE approach finds near-to-optimum subcarrier allocations in multiuser systems so that the overall cell throughput achieved by the HE approach amounts to approximately 98% of that of the BIP approach. This minor performance gap is more than compensated by the low computational complexity of the HE approach, which saves orders of magnitude in computation time if compared to the BIP scheme.

On the other hand, the high dynamic in link quality between users raised the issue of fairness: This issue addresses the classical tradeoff between supplying each of a preferable high number of users with a fair share of the system capacity on the one hand and maximizing the system throughput by privileging users with especially good link qualities on the other hand.

Although this tradeoff situation cannot be evaded, various subcarrier allocation schemes were introduced, which find a desirable balance between the above-mentioned extremes of absolute fairness and maximum throughput.

In this work, various approaches to influence the fairness level inside a multiuser OFDM-FDMA transmission system were investigated and compared. All of these approaches were

based on the previously introduced HE subcarrier allocation scheme due to its superior qualities in terms of computational complexity.

A total of four different HE schemes were considered, all with individual advantages and occasional shortcomings.

The first considered subcarrier allocation approach was the original HE with its objective to maximize the overall throughput of the system and the concurrent guarantee of a basic data rate R_B for each user. The latter constraint offers a minimum level of fairness between users and also allows some sort of adaptivity by changing the basic rate R_B. On the downside, the setting of a uniform rate R_B limits the number of users that can be supported by the system.

Thus, in a first modification of the HE the minimum rates were kept at a relatively low scale while an additional constraint was introduced to improve the fairness level. This constraint implies an upper limit K_{max} to the number of allocated subcarriers per user. After integration of this constraint into the HE, the fairness level rises considerably and at the same time this leaves the allocation algorithm with enough flexibility to find a suitable allocation solution for a large number of users. An additional advantage of this so-called K_{max} constrained HE compared to the original HE is that the parameter K_{max} can be used to adjust the fairness level or the system throughput respectively. Further evaluations showed that this adjustment must be done taking into account the actual number of users inside the cell. This fact and the integer character of the K_{max} constraint posed certain limits to the fairness adaptivity of the K_{max} constrained HE approach.

In order to evaluate the limits of the HE approach in terms of system fairness a further modification – the so-called HE2 approach – was developed, which abandoned the paradigm of throughput maximization and aimed at improving the rates of users with low link qualities. This was done by exploiting the two-phase characteristic of the HE approach. The HE2 algorithm works as follows:

In the first phase every user is supplied up to the basic rate R_B. In the second phase, all remaining resources are allocated in such a way, that the user with the instantaneously lowest rate is provided with another subcarrier. This procedure ensures an almost perfect rate balance between users providing a benchmark in terms of fairness for all previous allocation approaches. Besides this property, the fairness level of HE2 is fixed since it is inherently set by the specified allocation procedure.

The previously introduced subcarrier allocation approaches offered either limited or no possibilities to adjust the throughput-fairness-balance inside the system. Thus, a further modification of the HE approach, which employs the concept of utility was developed. This so-called utility-based HE (HE3) uses a suitable indicator- or utility function to measure the subjective link quality perceived by each user. The novel combination of the HE approach with a made-to-measure utility function allows to cover the full tradeoff range between maximum through-

put and absolute fairness. This is achieved by application of an exponential utility function and using its exponent ρ as a fairness parameter.

Adjustment of this parameter allows a continuous changeover between the maximum throughput performance of the original HE approach and the outstanding fairness level of the HE2. Although the utility-based HE3 does not completely reach the fairness level of HE2, its unique flexibility more than outweighs this minor drawback.

Summarizing the preceding discussion, it was shown that the introduced heuristic subcarrier allocation approaches only lead to minor losses in system performance if compared to optimum allocation solutions. The presented approaches with their huge advantages in terms of computational efficiency show a high potential for implementations of future wideband communication systems.

The discussed fairness issues on the one hand showed the straightforward adaptability of the HE approach towards higher fairness levels. Especially, the introduced HE2 approach proved that by its outstanding rate balance. The also considered K_{max} constrained HE offered the highest simplicity of implementation but is by far outperformed by the utility–based HE if flexible adjustment of the throughput-rate-balance is concerned. In this respect, the utility-based HE approach showed the best performance of all discussed subcarrier allocation approaches.

For this reasons, the utility-based HE is considered to be a promising candidate for implementation in a future mobile communication system.

11 Abbreviations and Symbols

11.1 Abbreviations

ACF	Auto Correlation Function
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BIP	Binary Integer Programming
BS	Base Station
CDF	Cumulative Distribution Function
CDMA	Code Division Multiple Access
CSI	Channel State Information
FDMA	Frequency Division Multiple Access
FFT	Fast Fourier Transform
GI	Guard Interval
HE	Heuristic Evaluation
HE2	Heuristic Evaluation, 2nd Version
HE3	Heuristic Evaluation, 3rd Version
IDFT	Inverse Discrete Fourier Transform
IP	Integer Programming
ISI	Intersymbol Interference
LOS	Line-Of-Sight
LP	Linear Programming
LTI	Linear Time-Invariant
LTV	Linear Time-Variant
MAC	Media Access Control
MAI	Multiple Access Interference
MC	Multicarrier
MT	Mobile Terminal
MUD	Multiuser Diversity

NP	Nondeterministic Polynomial (time)
OFDM	Orthogonal Frequency Division Multiplexing
OFDM-CDMA	OFDM transmission combined with CDMA multiple access
OFDM-FDMA	OFDM transmission combined with FDMA multiple access
OFDM-TDMA	OFDM transmission combined with TDMA multiple access
PDF	Probability Density Function
PHY	Physical
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase-Shift Keying
SC	Single-Carrier
SDMA	Spatial Division Multiple Access
SNR	Signal-to-Noise Ratio
TDMA	Time Division Multiple Access
UF	Utility Function
US	Uncorrelated Scattering
WSS	Wide-Sense Stationary
WSSUS	Wide-Sense Stationary, Uncorrelated Scattering

11.2 Symbols

α	Path loss exponent
α_p	Complex attenuation factor of pth propagation path
β	Mapping parameter of link performance model
β_{opt}	Optimal value of mapping parameter
Δe_c	Error between estimated and measured BER for channel c
Δf	Subcarrier spacing
Δt	Sampling interval
θ_p	Phase shift of pth propagation path
μ_{dB}	Expectation value in dB scale
ρ	Utility exponent
σ_{dB}^2	Variance in dB scale

σ_N^2	Noise variance
τ	Channel delay
$\tau_{\max}$	Maximum channel delay
$\varphi_{hh}(.)$	ACF of function $h(.)$
$\phi_{HH}(\Delta f, \Delta t)$	Time-frequency correlation function
$\phi_S(\tau, f_D)$	Scattering function
$\phi_{UU}(\Delta f, f_D)$	Doppler cross power density spectrum
ω	Upper bound for the value of an entry in $\mathbf{x}$
$\mathbf{A}$	Constraints matrix for IP
A	Subset of $x_{i,k}$ that represents allocated subcarriers
$\mathbf{b}$	Constraints vector for IP
$BER_{E,c}$	Estimated bit error rate of channel c
$BER_{E,i}$	Estimated bit error rate of user i
$BER_{M,i}$	Measured bit error rate of user i
$BER_{M,c}$	Measured bit error rate of channel c
B_C	Coherence bandwidth
c	Index of channel realization
$\mathbf{c}$	Cost vector for IP
C	Number of channel realizations
CL	Constraint length
CR	Code rate
d	Distance between MT and BS
d_0	Reference distance
$d_{\max}$	Maximum distance between MT and BS
$d_{\min}$	Minimum distance between MT and BS
$E\{.\}$	Expectation value
f	Continuous frequency
f_D	Doppler frequency

$\mathcal{F}\{.\}$	Fourier transformation
$g(t)$	Transmit pulse
G_0	Reference path loss factor
G_i^{PL}	Channel path loss factor of user i
G_i^{SH}	Channel shadowing factor of user i
$h(\tau,t)$	Channel impulse response
$H(f,t)$	Channel transfer function
$H_{i,k}$	Channel transfer factor for user i on subcarrier k
i	User index
k	Subcarrier index
K_{max}	Maximum number of subcarriers per user
l	Discrete time index
L	Length of input to represent IP instance
m	Row index of $\mathbf{A}$
M	Number of iterations in HE
n	Column index of $\mathbf{A}$
$n(t)$	noise time signal
N	Number of subcarriers inside system bandwidth
$N_{i,k}$	Noise contribution for user i on subcarrier k
N_B	Number of subcarrier blocks per user
N_C	Number of subcarriers per user
N_D	Distance between subcarriers of a user in equidistant allocation
N_{ISI}	Number of adjacent symbols disturbed by ISI
N_P	Number of propagation pathes
N_U	Number of users
$O(.)$	Function of computational complexity
p	Index of propagation path
$p(.)$	Probability density function

P_{RX}	Receive power
P_{TX}	Transmit power
R	System throughput
R_B	Lower bound for data rate
R_i	Data rate of user i
$R_{i,k}$	Data rate of user i on subcarrier k
$\hat{R}_{i,k}$	Discrete data rate value of user i on subcarrier k
$SNR_{eff,i}$	Effective signal-to-noise ratio of user i
SNR_i	Average signal-to-noise ratio of user i
$SNR_{i,k}$	Signal-to-noise ratio for user i on subcarrier k
$SNR_{\min}$	Minimum signal-to-noise ratio for certain PHY mode
t	Continuous time
T	Overall OFDM symbol duration (incl. guard interval)
T_C	Coherence time
T_G	Duration of guard interval
T_S	Duration of OFDM block (without guard interval)
$T_{S,MC}$	Duration of transmit symbol of multicarrier transmission
$T_{S,SC}$	Duration of transmit symbol of single-carrier transmission
$U(f, f_D)$	Frequency Doppler function
$U(R)$	Utility function
$V(\tau, f_D)$	Delay Doppler function
W	System bandwidth
$x(t)$	Transmit time signal
$x_{i,k}$	Allocation decision variable for user i on subcarrier k
x_l	Discrete-time transmit sequence
$\mathbf{x}$	Vector of $x_{i,k}$
X	Set of all $x_{i,k}$

$X_{i,k}$	Transmit symbol for user i on subcarrier k
X_n	Transmit symbol at time instance n
$y(t)$	Receive time signal
$Y_{i,k}$	Receive symbol for user i on subcarrier k

12 References

[Ald94]   M. Aldinger: *Multicarrier COFDM scheme in high bitrate radio local area networks*, Proc. 5th IEEE Pers., Indoor and Mobile Radio Comm. (PIMRC '94), vol. 3, The Hague, Netherlands, September 1994, pp. 969-973.

[Bel63]   P.A. Bello: *Characterization of Randomly Time-Variant Linear Channels*, IEEE Transactions on Communication Systems, Vol. 11, No. 4, pp. 360-393, Dezember 1963.

[Bia98]   G. Bianchi, A.T. Campbell, and R.R.F. Liao: *On Utility-Fair Adaptive Services in Wireless Networks*, Proc. 6th Intl. Workshop on Quality of Service (IEEE/IFIP IWQOS '98), Napa Valley, CA, May 1998, pp. 256-267.

[Bin90]   J.A.C. Bingham: *Multicarrier Modulation for Data Transmission: An Idea whose Time Has Come*, IEEE Communications Magazine, Vol. 28, No. 5, pp. 5-14, May 1990.

[Bla84]   R. E. Blahut: *Theory and practice of error control codes*, Addison-Wesley, Reading, MA,1984.

[Boc06]   H. Boche, M. Schubert: *Resource Allocation in Multiantenna Systems – Achieving Max-Min Fairness by Optimizing a Sum of Inverse SIR*, IEEE Transactions on Signal Processing, Vol. 54, No. 6, pp. 1990-1997, June 2006.

[Bos99]   M. Bossert: *Channel coding for telecommunications*, 1st edition, Wiley, Chichester, 1999.

[Bru05]   K. Brüninghaus, et al.: *Link Performance Models for System Level Simulations of Broadband Radio Access Systems,* Proc. 16th IEEE Pers., Indoor and Mobile Radio Comm. (PIMRC '05), vol. 4, Berlin, Germany, September 2005, pp. 2306-2311.

[Bur01]   A. Bury: *Efficient Multi-Carrier Spread Spectrum Transmission*, PhD thesis, Technical University of Dresden, Dresden, Germany, 2001.

[Cam06]   A. Camargo, A. Czylwik: *PER prediction for bit-loaded BICM-OFDM with hard-decision Viterbi decoding*, Proc. 11th Intl. OFDM Workshop (InOWo '06), Hamburg, Germany, August 2006, pp. 26-30.

[Col02]   S. Coleri, M. Ergen, A. Puri , and A. Bahai: *Channel Estimation Techniques Based on Pilot Arrangement in OFDM Systems*, IEEE Transactions on Broadcasting, Vol. 48, No. 3, pp. 223-229, September 2002.

[Dec08]   ETSI TS 300 175-2: *Digital European Cordless Telecommunications (DECT), Common Interface (CI) Part 2: Physical layer*, Ver. 2.2.1, 2008.

[Eri03]   Ericsson: *System-level evaluation of OFDM - further considerations*, 3GPP TSG-RAN-1 35, Document R1-03-1303, Lisbon, Portugal, November 2003.

[Fli91]   N. Fliege: *Systemtheorie*, 1st edition, Teubner, Stuttgart, 1991.

[Fra05]   T. Frank, A. Klein, E. Costa, and E. Schulz: *Interleaved Orthogonal Frequency Division Multiple Access with Variable Data Rates*, Proc. 10th Intl. OFDM Workshop, InOWo '05, Hamburg, Germany, Aug. 2005, pp. 179-183.

[Gal62]   R. Gallager: *Low-density parity-check codes*, IRE Transactions on Information Theory, Vol. 8, No. 1, pp. 21-28, January 1962.

[Gal06]   D. Galda: *Ein Beitrag zum Vielfachzugriff, zur Synchronisation und Kanalschätzung in einem OFDM-basierten Mobilfunksystem*, PhD thesis, Hamburg University of Technology, Hamburg, Germany, 2006.

[Gao01]   X. Gao, T. Nandagopal, and V. Bharghavan: *Achieving Application Level Fairness Through Utility-Based Wireless Fair Scheduling*, Proc. IEEE Global Telecommunications Conference (GLOBECOM '01), San Antonio, TX, November 2001, pp. 3257-3261.

[Gar79]   M.R. Garey, D.S. Johnson: *Computers and Intractability: A Guide to the Theory of NP-Completeness*, 1st edition, W.H. Freeman & Co, New York, 1979.

[Got90]   T.D. Gottschalk: *Concurrent implementation of Munkres Algorithm*, Proc. 5th Distributed Memory Computing Conference, Charleston, SC, April 1990, pp. 52-57.

[Gro03]   J. Gross, H. Karl, F. Fitzek, and A. Wolisz: *Comparison of heuristic and optimal subcarrier assignment algorithms*, Proc. Intl. Conference on Wireless Networks (ICWN '03), Las Vegas, NV, June 2003, pp. 249-255.

[Gro05]   J. Gross, H.F. Geerdes, H. Karl: *Performance Analysis of Dynamic OFDM-FDMA Systems with Inband Signaling*, TKN Technical Report TKN-05-001, Technical University of Berlin, Berlin, Germany, 2005.

[Grue00]   R. Grünheid: *Vielfachzugriffsverfahren für die Multiträger-Übertragungstechnik*, PhD thesis, Hamburg University of Technology, Hamburg, Germany, 2000.

[Grue06]   R. Grünheid: *Adaptive Resource Allocation Schemes in MIMO-OFDM based Cellular Communication Systems,* Habilitation thesis, Hamburg University of Technology, Hamburg, Germany, 2006.

[Gsm96]   ETSI TS 300 573 (GSM 05.01): *Digital Cellular Telecommunication System (Phase 2), Physical layer on the radio path, General Description*, Ver. 4.6.0, 1996.

[Hoe97]   P. Hoeher, S. Kaiser, and P. Robertson: *Two-Dimensional Pilot-Symbol-Aided Channel Estimation by Wiener Filtering*, Proc. International Conf. on Acoustics, Speech, and Signal Processing (ICASSP '97), Vol. 3, Munich, Germany, April 1997, pp. 1845-1848.

[Hoo04]   L.M.C. Hoo, B. Halder, J. Tellado, and J.M. Cioffi: *Multiuser Transmit Optimization for Multicarrier Broadcast Channels: Asymptotic FDMA-Capacity region and Algorithms*, IEEE Transactions on Communications, Vol. 52, No. 6, pp. 922-930, June 2004.

[Jia05]   Z. Jiang, Y. Ge, and Y. Li: *Max-Utility Wireless Resource Management for Best-Effort Traffic*, IEEE Transactions on Wireless Communications, Vol. 4, No. 1, pp. 100-111, January 2005.

[Kim01]   I. Kim, H.L. Lee, B. Kim, and Y.H. Lee: *On the Use of Linear Programming for Dynamic Subchannel and Bit Allocation in Multiuser OFDM*, Proc. IEEE Global Telecommunications Conference (GLOBECOM '01), San Antonio, TX, November 2001, pp. 3648-3652.

[Kuh55]   H.W. Kuhn: *The Hungarian Method for Assignment Problems*, Nav. Res. Log. Quart. 2, Washington, 1955, pp. 83-98.

[Lee95]   K. Lee: *Adaptive Network Support for Mobile Multimedia*, Proc. ACM Mobilcom 1995, Berkeley, CA, November 1995, pp. 62-74.

[Lin91]   J.P. Linnartz: *Effects of Fading and Interference in Narrowband Land-Mobile Networks*, PhD thesis, University of Technology Delft, Delft, Netherlands, 1991.

[Mac08]   T. F. Maciel: *Suboptimal Resource Allocation for Multi-User MIMO-OFDMA Systems*, PhD thesis, Darmstadt University of Technology, Darmstadt, Germany, 2008.

[Mao08]   Z. Mao, X. Wang: *Efficient Optimal and Suboptimal Radio Resource Allocation in OFDMA System*, IEEE Transactions on Wireless Communications, Vol. 7, No. 2, pp. 440-445, February 2008.

[May00]   T. May: *Differentielle Modulation und Kanalcodierung in breitbandigen OFDM Funkübertragungssystemen*, PhD thesis, Hamburg University of Technology, Hamburg, Germany, 2000.

[Nem89]   G.L. Nemhauser, et al.: *Optimization*, 1st edition, Elsevier Sience B.V., Amsterdam, 1989.

[Olo06]   S. Olonbayar: *Multiple Access Techniques and Channel Estimation in OFDM Communication Systems*, PhD thesis, Hamburg University of Technology, Hamburg, Germany, 2006.

[Pae99]   M. Pätzold: *Mobilfunkkanäle*, Vieweg, Braunschweig, 1999.

[Pap82]   Ch.H. Papadimitriou: *Combinatorial Optimization*, 1st edition, Prentice Hall, Englewood Cliffs, 1982.

[Pro00]   J.G. Proakis: *Digital Communications*, 4th edition, McGraw Hill, New York, 2000.

[Roh97]   H. Rohling, R. Grünheid: *Performance comparison of different multiple access schemes for the downlink of an OFDM communication system*, Proc. 47th Vehicular Technology Conference (VTC '97), Vol. 3, Phoenix, AZ, May 1997, pp. 1365-1369.

[Roh99]   H. Rohling, T. May, K. Brüninghaus, and R. Grünheid: *Broadband OFDM Radio Transmission for Multimedia Applications*, Proc. of the IEEE, Vol. 87, No. 10, pp. 1778-1789, October 1999.

[Roh01]   H. Rohling, R. Grünheid, and D. Galda: *OFDM – An Attractive Candidate for the Air-Interface of the 4th Generation Mobile Communication System*, 39th Annual Allerton Conference on Communication, Control, and Computing, Urbana-Champaign, IL, October 2001.

[Roh04]   H. Rohling, D. Galda, and E. Schulz: *An OFDM Based Cellular Single Frequency Communication Network*, Proc. Wireless World Research Forum (WWRF '04), Beijing, China, 2004.

[Roh05]   H. Rohling, R. Grünheid: *Cross Layer Considerations for an Adaptive OFDM-based Wireless Communication System*, Wireless Personal Communications, Vol. 32, No. 1, pp. 43-57, January 2005.

[She95]   S. Shenker: *Fundamental Design Issues for the Future Internet,* IEEE Journal on Selected Areas in Communications, Vol. 13, No. 7, pp. 1176-1188, September 1995.

[She03]   M. Shen, G. Li, and H. Liu: *Design Tradeoffs in OFDMA Traffic Channels*, Proc. International Conf. on Acoustics, Speech, and Signal Processing (ICASSP '04), Vol. 4, Montreal, Canada, May 2004, pp. 757-760.

[Ste06]   M. Stemick, H. Rohling: *OFDM-FDMA Scheme for the Uplink of a Mobile Communication System*, Wireless Personal Communications, Vol. 40, No. 2, pp. 157-170, January 2007.

[Ste07]   M. Stemick, S. Olonbayar, H. Rohling: *PHY-Mode Selection and Multiuser Diversity in OFDM Transmission Systems*, Frequenz, Vol. 61, no. 1-2, pp. 7-10, Jan./Feb. 2007.

[Ste08]   M. Stemick, H. Rohling: *Resource Allocation and Fairness in OFDM-FDMA based Systems*, Proc. 13th Intl. OFDM Workshop (InOWo '08), Hamburg, Germany, August 2008, pp. 90-94.

[Tej06]   P. Tejara, W. Utschick, G. Bauch, and J.A. Nossek: *Subchannel Allocation in Multiuser Multiple-Input-Multiple-Output Systems*, IEEE Transactions on Information Theory, Vol. 52, No. 10, pp. 4721-4733, October 2006.

[Toe07]   N. Tönder: *Kanalschätzung, Demodulation und Kanalcodierung in einem FPGA-basierten OFDM-Funkübertragungssystem,* PhD thesis, Hamburg University of Technology, Hamburg, Germany, 2007.

[Tra04]   W. Tranter, et al.: *Principles of communication system simulation with wireless applications*, Prentice Hall, Upper Saddle River, NJ, 2004.

[Vit67]   A. Viterbi: *Error bounds for convolutional codes and an asymptotically optimum decoding algorithm*, IEEE Transactions on Information Theory, Vol. 13, No. 2, pp. 260-269, July 1967.

[Wea13]   R.A. Weagant: *Some Recent Radio Sets of the Marconi Wireless Telegraph Company of America*, Proc. Institute of Radio Engineers (IRE), Vol. 1, No. 4, pp. 43-61, December 1913.

[Wei71]   S.B. Weinstein, P.M. Ebert: *Data Transmission by Frequency-Division Multiplexing Using the Discrete Fourier Transform*, IEEE Transactions on Communication Technology, Vol. Com-19, No. 5, pp. 628-634, October 1971.

[Zho07]   C. Zhou, G. Wunder, and Th. Michel: *Utility maximization for OFDMA Systems over discrete sets*, Proc. IEEE Intl. Conference on Communications (ICC '07), Glasgow, UK, June 2007, pp. 5726-5731.

Appendix A Computational Complexity

In this chapter, some basic concepts for assessing the computational complexity of algorithms are introduced. One of these is the concept of *NP-completeness*, which plays an important role in this respect. It is also shown, that the considered BIP approach for the optimal allocation of subcarriers to users is an NP-complete problem and thus needs exponential execution time. On the other hand, the complexity of the HE approach is assessed. The comparison of both approaches reveals the computational efficiency of the HE approach.

A.1. NP-Completeness

The term NP-completeness is introduced in [Gar79] as a complexity class for *decision problems*. Those problems have only two possible solutions: yes or no. Several optimization problems also belong to the class of NP-complete problems, since an optimization problem can be derived from a decision problem.

The acronym NP refers to *Nondeterministic Polynomial* (Time), which means a problem belongs to the class NP, if it is solvable by a nondeterministic algorithm in polynomial time. Solvability in polynomial time is generally a desirable feature of a problem, since this means its complexity can be expressed by a term similar to $p(n) = a_k n^k + \ldots n^2 + a_1 n + a_0$. In this term, k is integer, a_k are real coefficients and n gives some measure of the problem size (e.g. number of subcarriers to assign). But it is an unfavorable situation if this solution has to be achieved by a non-deterministic algorithm, since algorithms executed on computers have deterministic behavior. A nondeterministic algorithm mostly contains some "guessing"-stages and is therefore hardly predictable.

Inside the class NP, there are the NP-complete problems. It is *assumed* that these problems can only be *solved inefficiently*, which means a deterministic algorithm needs exponential (e.g. 2^n) time to find a solution. *All known* deterministic algorithms aiming at the solution of these NP-complete problems need exponential time. Research to find algorithms with polynomial time is ongoing since the 1970's without much success. Thus, if a NP-complete problem must be treated, exponential runtime algorithms should be assumed.

Another aspect of a NP-complete problem is: If a problem instance *I* of a NP-complete problem is given and also a solution *S* for *I*, only polynomial time is needed to check if *S* is really a solution of *I*. Of course, this does not mean at all that polynomial time was needed to find *S*.

Another well-known term in the context of NP is *NP-hardness*. The class of NP-hard problems contains all problems, which are at least as hard to solve as NP-complete problems, but must not necessarily be part of the NP class. The Binary Integer Programming problem considered in this thesis is NP-hard, since in this case it is solved as an optimization task. On the other hand, the BIP decision problem as introduced in [Gar79] is NP-complete.

A.2. Complexity of Integer Programming

In [Gar79] Integer Programming (IP) is introduced as an NP-complete problem. In the context at hand an integer programming problem is given by

$$
\begin{aligned}
&\text{minimize } \mathbf{c}'\mathbf{x} \\
&\text{subject to } \mathbf{Ax} \leq \mathbf{b} \\
&\text{with } \mathbf{x} \in \mathbb{Z}^n.
\end{aligned}
\tag{A.1}
$$

If $\mathbf{x} \in \{0,1\}$ is assumed then equation (A.1) is referring to a binary integer program. In equation (A.1) the vector $\mathbf{x}$ contains the decision variables, $\mathbf{A}$ and $\mathbf{b}$ the constraints and $\mathbf{c}$ is the cost vector. It does not matter if in the constraints an equality or inequality is used. The formulation in [Gar79] assumes that $\mathbf{c}, \mathbf{A}$ and $\mathbf{b}$ consist solely of integer numbers. The BIP approach introduced in chapter 8.1.1 and defined by (8.2), (8.3) and (8.4) can also be formalized using (A.1). It can be shown by *restriction* (cf. [Pap82]) that the BIP approach is a generalization of (A.1) and is NP-complete, if considered as a decision problem. The BIP optimization problem is at least as hard to solve as the decision problem and is NP-hard.

In [Nem89] more specific bounds on the computational complexity of integer programming are given. The measure of complexity chosen in [Nem89] is the length of input L necessary to present a problem instance to a computer. Let length L be defined as

$$
L = m \cdot n \cdot \log(\theta)
$$

with m and n as the dimensions of the constraint matrix $\mathbf{A}$. The scalar θ is the entry with maximum magnitude from either $\mathbf{A}, \mathbf{c}$ or $\mathbf{b}$. By $\log(\theta)$ an upper bound to the number of bits needed to encode one input entry is given. Having defined L, it can be stated that the amount of memory space and computation time needed to solve an IP is a function of L. It is also stated that an IP has a maximum of $(\omega+1)^n$ candidates for an optimal solution. We define $\omega = ((m+n)n\theta)^n$ as an upper bound for the value of an entry in $\mathbf{x}$. It is obvious, that ω can be expressed as a function of L. In the special case of a BIP, $\omega=1$ holds. Thus, a BIP has a maximum of 2^n solutions and thus a brute-force-search for the optimum value would have exponential runtime. The formulation of the BIP approach as in (8.2) leads to $n = N_U N$ and thus to an upper complexity bound of $O(2^{N_U N})$.

Nevertheless, there are approaches to solve BIP (e.g. branch-and-bound), which have reasonable runtime for a lot of BIP problem instances. Thus, the exponential runtime is an upper bound on the runtime for a BIP.

A.3. Complexity of Heuristic Subcarrier Allocation

The previous chapter gave an upper bound of $O(2^{N_U N})$ for the complexity of the BIP approach. This is an extremely large computational effort for calculating a subcarrier assignment. The effort is considerably reduced by the HE approach, which uses heuristics to find a suitable subcarrier assignment. In the following, an upper bound for the computational complexity of this approach will be derived. This is done by formalizing each iteration of HE as an instance of the so-called *assignment problem*. A thorough introduction to this well known problem of combinatorial optimization can be found in [Pap82]. In chapter 8.1.2, the HE approach was introduced as a two-phased algorithm. Since the main computational effort is spent in phase 1, this phase is discussed first and in more detail.

Phase 1 is divided into multiple iterations. In each iteration, each user gets assigned exactly one subcarrier. The set A of assigned subcarriers also fulfills (8.3). Due to these constraints, the subcarrier allocation inside an iteration can be described by the so-called *assignment problem*, cf. [Kuh55]. Although the assignment problem can also be brought to the form of (A.1), its constraint matrix is *totally unimodular* (cf. [Pap82]) and **b** solely contains ones. This special case of a BIP problem is not NP-hard and is solvable in polynomial time. Efficient algorithms to solve the assignment problem are introduced in [Kuh55], [Pap82] and [Got90]. For a set of N_U users and N subcarriers and assuming w. l .o. g. that $N_U \leq N$, the solution to the assignment problem has a total complexity of $O(N_U^2 N)$, cf. [Pap82] and [Got90]. This is an upper bound for the complexity of a single iteration in phase 1. After each iteration it is checked, if all users have reached R_B . If this is not the case, the next iteration begins, considering the not yet assigned subcarriers and not yet satisfied users. Thus, the overall complexity of phase 1 is the sum of the complexities of all iterations. In the following, the number of users in each iteration is assumed to be constantly N_U (e.g. for large R_B). Without loss of generality N is set to be an integer multiple of N_U, and so the number of iterations follows to be

$$M = \frac{N}{N_U}. \tag{A.2}$$

After each iteration, the number of available subcarriers is decreased by N_U and the complexity of the assignment problem in the next iteration is decreased accordingly. Inserting (A.2) into the expression $O(N_U^2 N)$, the complexity of the m th iteration is given as follows:

$$O\big((M+1-m)N_U^3\big) \quad \text{for iteration } m=1...M \tag{A.3}$$

The total complexity of phase 1 follows from summing up the terms (A.3) for all M iterations:

$$\sum_{m=1}^{M} \left(M+1-m\right) N_U^3$$

$$= \frac{M}{2} \left(M+1\right) N_U^3$$

$$= \frac{1}{2} \left(N_U^2 N^2 + N_U^2 N\right) \qquad \text{(A.4)}$$

$$\triangleq O\left(N_U^2 N^2\right).$$

Equation (A.4) serves as an upper bound for the complexity of phase 1.

The derivation of the complexity of phase 2 is also straightforward: Assuming the worst-case scenario that phase 1 finishes successfully after a single iteration, still $N - N_U$ subcarriers are left for assignment to N_U users in phase 2. Since the overall throughput R must be maximized, user i is assigned to subcarrier k if he observes the maximum $R_{i,k}$ on k with respect to the other users. This assignment is accomplished by a simple search through all $R_{i,k}$, which takes at most $\left(N - N_U\right)N_U$ operations. Thus phase 2 has a complexity of $O(NN_U)$, which is almost negligible compared to phase 1. Combining the computational effort for phase 1 and 2, the overall complexity of the HE approach is given by $O(N_U^2 N^2)$.

The above derivations show that an instance of the subcarrier allocation problem takes exponential time if solved by the BIP approach and polynomial time for the HE approach. Therefore the HE approach has a tremendous advantage if computational complexity is considered.

Appendix B   *K~max~* constrained BIP approach

This appendix discusses fairness and throughput performance of the K_{max} constrained BIP approach in contrast to the HE approach evaluated in section 9.3. Also, it is shown how closely the results of the heuristic HE approach match the results of the BIP approach if the K_{max} constraint is applied. This is another argument in favor of the computational efficient HE approach.

B.1. Subcarrier Allocation Process

In the following, the impact of the additional constraint introduced by (9.4) on the BIP approach is illustrated by example. The same channel example as in section 8.2.1 is used below. The resulting allocation of subcarriers for the K_{max} constrained BIP and HE approach is shown in Fig. 54. As in Fig. 24, three users are located inside the system. The optimization

described by (9.1) - (9.4) is applied. In this example, K_{max} is set to $K_{max} = 10$. It is obvious, that the allocation of subcarriers to each user leads to a completely different result, compared to the allocation without the constraint K_{max}. In Fig. 54, the number of allocated subcarriers is identical for each user. The constraint K_{max} is exactly met.
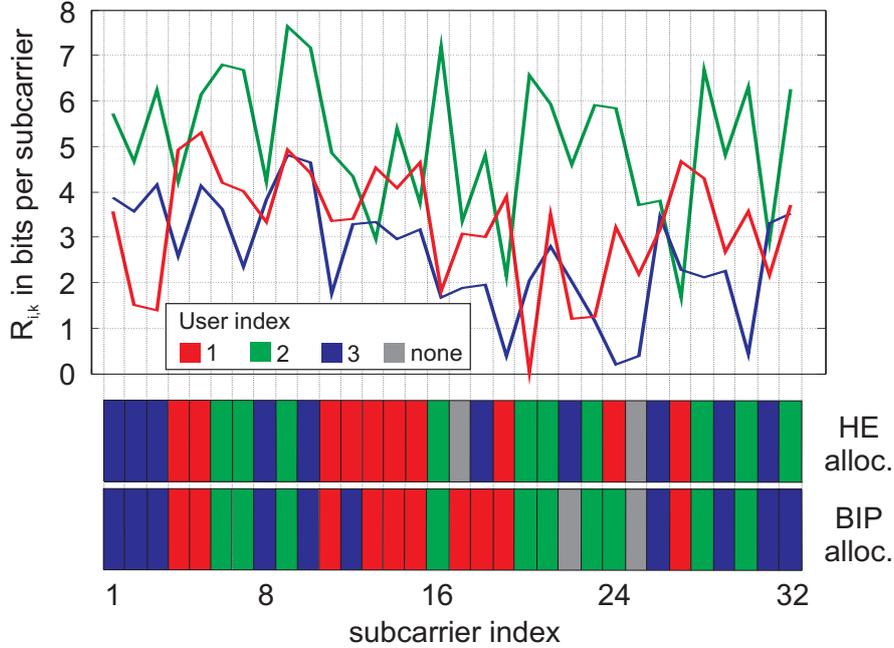
Fig. 54: Exemplary subcarrier allocation by HE and BIP constraint to K_{max} = 10

In the considered scenario, the constraint $K_{max} = 10$ allows only the allocation of a total of 30 from the overall 32 subcarriers. Thus, the K_{max} constraint plays a dominant role in the allocation. If the number of users would increase, all N resources would be claimed and it would not be possible for each user to claim exactly K_{max} subcarriers. The more users are joining the system, the more the K_{max} constraint is losing its impact and the R_B constraint is starting to dominate.

B.2. Simulation Results

The level of fairness achievable with the K_{max} constrained BIP approach is evaluated in Fig. 55 using CDF curves. Exactly the same simulation scenario and parameters as for Fig. 33 (K_{max} constrained HE approach) apply. The comparison of both figures shows that K_{max} constrained BIP and HE provide almost the same level of fairness.

The same holds if the achievable system throughput is considered. This is proven by a comparison of the figures Fig. 35 and Fig. 56: The throughput curves for $K_{max} = 5$ and $K_{max} = 6$ show also an almost identical shape for the HE and the BIP approach.

A direct comparison of the system throughput for both considered approaches and various constraint values is given in Fig. 57. This figure shows that the performance gap between HE and BIP is even smaller in the K_{max} constrained case than in the unconstrained scenario considered in section 8.3.1.
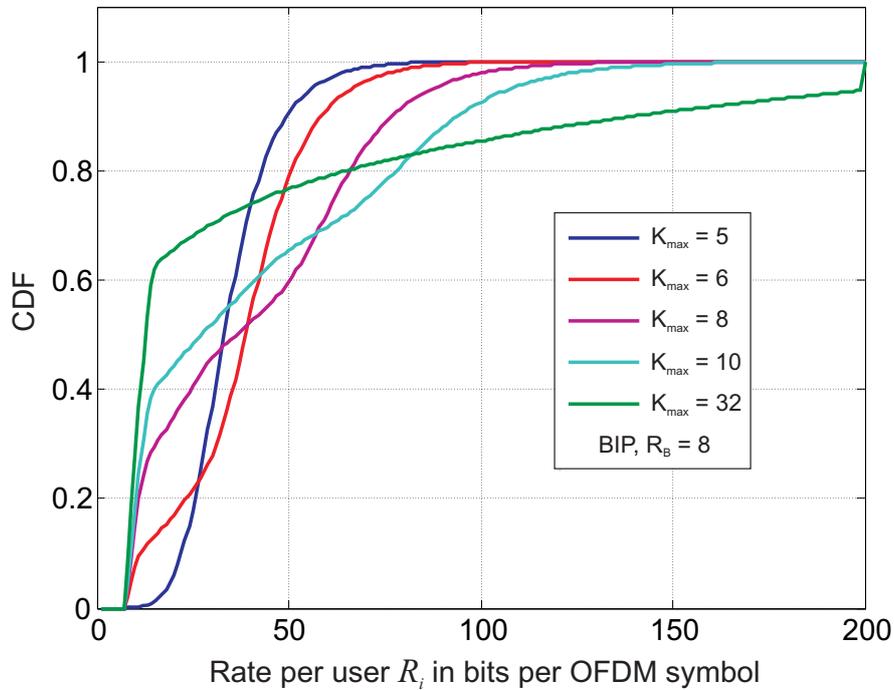
Fig. 55: CDF of user rates (BIP approach) for $R_B = 8$ and various values of K_{max}, 6 users inside the cell
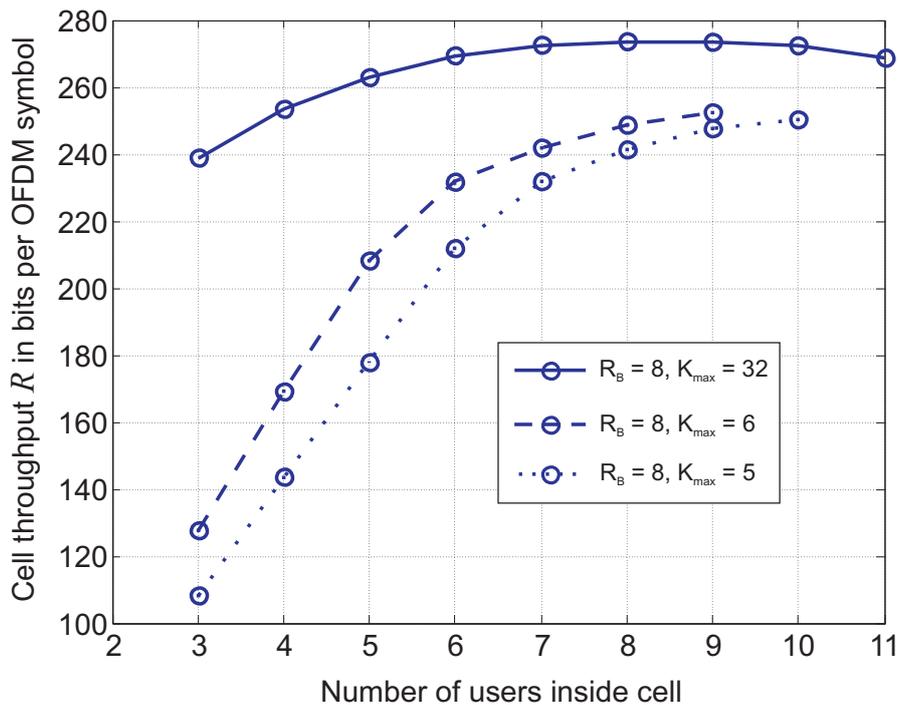
Fig. 56: BIP approach with additional K_{max}

The minimum rate R_B is set to 8, 12 and 16 bits per OFDM symbol for each user, respectively. For each R_B, K_{max} is set in such a way that even users with a very poor channel can claim enough subcarriers to be supplied with R_B. The curves with circular markers represent the BIP approach, star-shaped markers represent the HE approach.
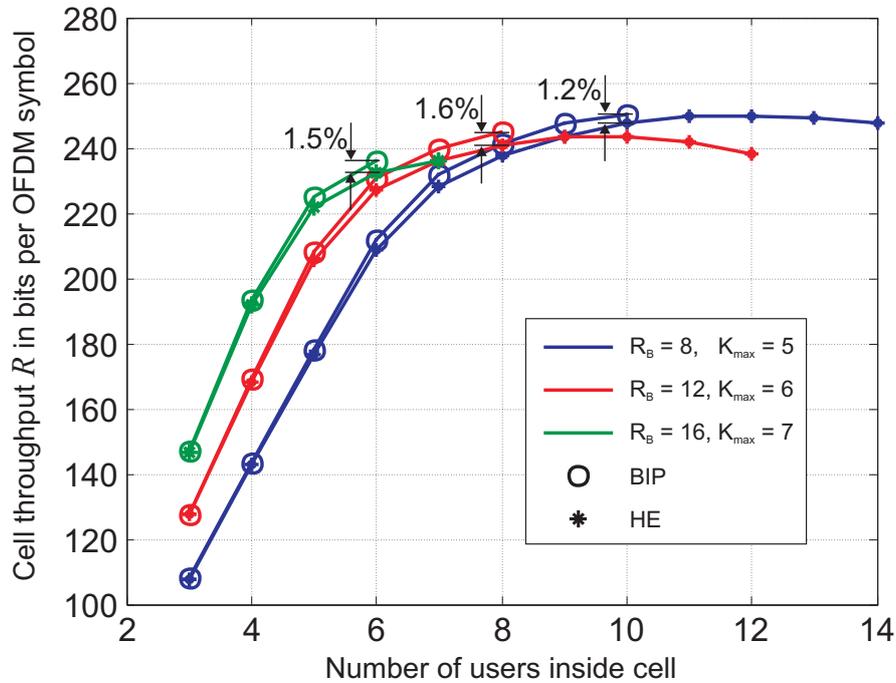
Fig. 57: Comparison between BIP and HE approach with additional constraint K_{max}

For low user numbers, the curves for HE and BIP are almost congruent. This means, the heuristic HE allocation gives almost the same performance as the optimized BIP allocation. Thus, the HE approach obviously gains on the BIP approach in this scenario if compared to its performance in section 8.3.1.

The reason why the HE approach is able to close the gap lies in its allocation procedure: The weak point of the HE approach is the iterative subcarrier allocation in phase 1, as already pointed out in section 8.1.2. But in the scenario at hand (low user number), only a part of the available subcarriers is allocated and all users get the same amount of subcarriers. This means, the constraint K_{max} is active for all users. Therefore, the chances are high that in the optimum allocation all users get their "best" subcarriers during the allocation. The only reason, why a user should not get its best subcarrier is that another user has a much higher performance on the same subcarrier. And this is almost exactly the strategy, which is applied in phase 1 of the HE approach. Therefore, the performance of BIP and HE is almost identical for a low number of users.

For higher numbers of users, the performance of the HE approach declines with respect to the BIP approach, since now, the constraint K_{max} is not necessarily active, which complicates the search for a global optimum.

But since the numbers of assigned subcarriers per user are much more balanced in the scenario with the additional K_{max} constraint, the performance gap between BIP and HE approach is smaller than in the scenario where solely the R_B constraint is applied, cf. section 8.3.1. This is due to the fact that the iterative allocation of one subcarrier per user in phase 1 of the HE approach favors such balanced subcarrier numbers.

If the relative differences between the cell throughputs of BIP and HE approach in Fig. 57 are compared, the difference between both approaches amounts to 0.6% for low user numbers and to a maximum of 1.6% for higher user numbers.

Thus, in a scenario where K_{max} and R_B constraints are applied simultaneously, there is almost no difference between the performance of the HE and the BIP approach, especially for low cell loads.

Another argument in favor of the HE approach in this context is again its low computational complexity. Especially for high user numbers the additional constraint K_{max} increases the time to find an optimal solution for the BIP approach. In contrast to this, the runtime of the HE approach is not at all increased, since the number of iterations in phase one is upper bounded by N.

In addition, the HE approach offers further possibilities for modifications, since it is split in two phases. Therefore it is possible to change the algorithm inside one or both phases to adjust the allocation results with respect to throughput or fairness.

Curriculum Vitae

Personal Data

Name	Martin Stemick
Date of Birth	22nd of February, 1979
Place of Birth	Uelzen, Germany

Education

June 1998	Abitur at Herzog-Ernst-Gymnasium, Uelzen
October 1999- August 2004	Study period and diploma in Electrical and Communications Engineering at Hamburg University of Technology, Hamburg, Germany
February 2004- August 2004	Diploma thesis at Volkswagen AG, Wolfsburg, Germany

Employment History

November 2004- August 2009	Scientific Staff at Hamburg University of Technology, Hamburg, Germany
Since September 2009	Employee of Rohde & Schwarz PMR GmbH, Bad Münder, Germany

Military Service

September 1998- June 1999	Service as motor mechanic in Lüneburg, Germany