Scheduling Strategies for Frequency Selective Multi-User MIMO-OFDMA Systems

Diplomarbeit

zur Erlangung des akademischen Grades
Diplominformatiker

Humboldt-Universität zu Berlin
Mathematisch-Naturwissenschaftliche Fakultät II
Institut für Informatik

eingereicht von:	Johannes Marotzke
geboren am:	25.10.1984
in:	Berlin
Gutachter:	Prof. Dr. J-P. Redlich
	Prof. Dr.-Ing. E. Grass
Betreuer:	Dr. A. Zubow

eingereicht am: ...... verteidigt am: ......

Abstract

Modern wireless networks are faced with an increasing demand for higher data rates, better quality of service, and higher overall network capacities. In order to comply to these demands various limits, like the limited availability of the frequency spectrum and a complex space-time varying wireless environment must be overcome. Transmission schemes based on Multiple Input Multiple Output (MIMO) techniques and Orthogonal Frequency Division Multiple Access (OFDMA) are promising technologies that increase spectral efficiency, and counter the effects of the environment.

In case of MIMO, multiple antennas are used to improve communication performance. The most common MIMO modes either increase diversity to make signals more robust against fading and interference, or perform spatial multiplexing of multiple streams. Multi-User spatial multiplexing is especially promising as it can increases spectral efficiency in a linear manner. The channel access method associated with MIMO's spatial multiplexing mode is commonly called Space-Division-Multiple-Access (SDMA), where multiple Mobile Stations (MS) are spatially separated and allocated to the same time/frequency resources. OFDMA on the other hand, is a channel access mode where a wideband channel is divided into multiple orthogonal narrow-band sub-carriers, thus, separating multiple MSs in either frequency, and/or time. Combining SDMA and OFDMA, resources must be allocated in time, frequency, and space dimension to different MSs which results in a highly complex three-dimensional resource allocation problem. For any practical SDMA-OFDMA MAC scheduler the computational costs for solving the resource allocation problem can become a limiting factor and must be considered carefully.

The SDMA grouping problem of finding suitable MSs that are spatially orthogonal to each other, has been commonly acknowledged as being the main computational intensive problem that affects performance [YG06], [STKL01]. Suboptimal SDMA grouping algorithms, and different SDMA grouping metrics are often proposed, that sacrifice performance for reduction in complexity to different degrees. Furthermore, in combination with OFDMA the question arises if SDMA groups should be allocated per-subcarrier, or larger subsets thereof. Especially since OFDM systems are in general frequency-selective any SDMA group considered optimal on some sub-carriers might be far from optimal on others.

This work proposes a generic SDMA-OFDMA MAC scheduling solution that integrates current state-of-the-art suboptimal SDMA grouping strategies with varying complexity, a low complex Signal-to-Noise-plus-Interference-Ratio (SNIR) predictor for users within an SDMA group, and the option to perform *frequency selective scheduling* with varying degree of granularity. The proposed scheduler is analyzed in a typical urban macro-cell scenario by means of system-level packet-based simulations, with detailed MAC and physical layer abstractions based on the WiMAX 802.16e standard. In the course of this work, it will be shown that the proposed SNIR predictor allows a substantial reduction in SDMA grouping complexity with minimal performance loss for varying SDMA grouping algorithms, and that besides the complexity versus performance trade-off, there exists an additional trade-off between performance and MAC layer signaling overhead that can quickly invalidate the advantages of *frequency selective scheduling* in SDMA-OFDMA systems.

Contents

1 Introduction

Future wireless systems need to address the increasing growth of mobile data traffic and its resulting demands, which are heavily driven by video and web applications. The Cisco Visual Networking Index (VNI) Global Mobile Data Traffic Forecast [RRTW11] gives a good indication on these past and future demands. In 2010 alone, the global mobile traffic grew 2.6 fold in comparison to 2009, tripling the third year in a row with 237 petabytes per month it was over three times greater than the entire global internet traffic in 2000 (75 petabytes per month). This growth is driven by the increasing amount of available mobile devices, the success of smart-phones during the last few years, as well as the fact that wireless networks allow coverage of areas without wired infrastructure. The main question that arises is not if a predicted 26-fold increase till 2015 is possible, or even unavoidable, but how can the capacity of wireless systems be increased to cope with such demands?

A starting point is given by the theoretical limit of the channel capacity as defined by the Shannon-Hartley theorem. This equation is a tight upper bound on the channel *Capacity* in bits per second of a communication channel of a specified *Bandwidth* in the presence of noise.

$$Capacity = Bandwidth \cdot log_2(1 + SNR) \tag{1}$$

As can be seen, any improvement on the Signal-to-Noise-Ratio (SNR) results in logarithmic growth with only minor gains, whereas the bandwidth can increase the capacity linearly. Accordingly, increasing bandwidth is the most common approach in order to increase capacity. But, the bandwidth itself is not without limits. For example the coherence bandwidth, a statistical measurement over which the channel can be considered flat, gives an indication of such a limit [Gol05]. When the bandwidth is much larger than the coherence bandwidth the symbol rate increases and experiences frequency selective fading due to multi-path effects. One solution to this problem is to split the wideband channel into a set of orthogonal non-interfering narrow-band sub-carriers [TV05] called Orthogonal Frequency Division (OFDM).

Even though this is an effective approach, it does not lead to an unlimited capacity, especially since the computational cost increases with increasing bandwidth, as well as the fact that bandwidth itself is a limited resource with possibly expensive licenses [Dra10].

Another approach that shows a promising potential for increasing capacities beyond the Shannon-Hartley limit is based on using multiple antenna systems on receiver and transmitter side, and is commonly referred to as a Multiple-In-Multiple-Out (MIMO). Equation 2 shows the assumed capacity of a MIMO system[PFE08]. This equation is a rough approximation on the capacity of multiple antenna systems. It is based on the assumption that MIMO allows up to N spatial streams each equivalent to a Single-In-Single-Out (SISO) channel:

$$Capacity = N \cdot Bandwidth \cdot log_2(1 + SNR), \tag{2}$$

where $N = min(n_{RX}, m_{TX})$ represents the available degrees of freedom given by the minimum number of antennas of either side used for transmission or reception. This limitation of the "smallest device" is particularly severe for mobile devices like handhelds, or smart phones that might not be able to sustain multiple antennas. An extension to MIMO, which overcomes this limitation, is the so called Multi-User MIMO (MU-MIMO). MU-MIMO combines multiple streams for multiple Mobile Stations (MS), instead of utilizing multiple streams for one MS with multiple antennas.

In conclusion, and to answer the question in how to cope with the increasing future demands of wireless networks, the combination of a large bandwidth with Orthogonal Frequency Division Multiple Access (OFDMA) and Multiple Input Multiple Output (MIMO) techniques holds a great promise to do so.

1.1 Thesis

The objective of this work is to develop a feasible Multi-User MIMO-OFDMA capable downlink (DL) MAC scheduler based on current state-of-the-art solutions, and to evaluate its performance on MAC layer through simulation on a per-packet basis[1]. For this purpose the proposed scheduler is to be analyzed by means of system-level simulations with detailed MAC- and physical layer abstractions in compliance to an urban macro-cell scenario based on the WiMAX system evaluation methodologies.

During the course of this work, the complexity involved in scheduling the space, time and frequency dimensions available in a Multi-User MIMO-OFDMA system, led to the development of a low-complex SNIR prediction method[2] for multiple MSs that are to be scheduled simultaneously and separated in the spatial dimension. This method and its performance as part of this work will be evaluated in terms of throughput and complexity. Furthermore, this work includes the effects of a frequency selective environment and its possible advantages for MIMO-OFDMA systems on MAC layer.

1.2 Thesis Structure

This work, and its remainder are structured into five main sections whose contents are briefly described here. Section 2 gives a brief introduction into MIMO and OFDM/OFDMA systems, their advantages and how these techniques are supported by the WiMAX standard as well as a brief overview on related work focusing on complete SDMA-OFDMA scheduling solutions. Section 3 discusses some issues faced during the course of this work towards a practical MAC scheduler, and describes the chosen approach and its implementation in detail. The proposed scheduling algorithm is analyzed by means of simulations. Section 4 describes the underlying simulation framework, provides information about the level of detail considered in the modeling as well as the configuration parameters and settings used within this work. Section 5 focuses on the evaluation of the proposed low complexity SNIR prediction method and the complete MAC scheduling solution and its variations. The performance of the proposed SNIR prediction method is evaluated in detail in Sub-section 5.1. The system level performance of the complete scheduling algorithm and its variations are studied more thoroughly in Sub-section 5.2, and the possible gain achievable through frequency selective scheduling within an SDMA-OFDMA system is evaluated in

[1] Parts of this thesis and its results have been published in [ZMDMXC12, ZM12].

[2] Patent: "Method for operating a wireless network, a wireless network and a device", PCT/EP2012/050523, to appear, 2012.

Sub-section 5.3. Finally, Section 6 summarizes the main results and concludes this work.

2 Basics

The following sections gives a brief introduction into different state-of-the-art MIMO techniques, the advantages and disadvantages of multi-carrier system like OFDM, as well as a short overview of the WiMAX 802.16e PHY layer support of such techniques and current related work on SDMA-OFDMA scheduling proposals.

2.1 MIMO and spatial multiplexing

Multiple-Input-Multiple-Output (MIMO) refers to the use of multiple antennas at the transmitter and/or receiver. It provides several advantages compared to single antenna system like spatial diversity, multiplexing, or even interference reduction. Spatial diversity is simply achieved through more than one receive antenna. Since each antenna encounters different channel effects, the signals received are different variations of the same signal. Spatial multiplexing combines these diversity effects, on receiver and transmitter side in order to multiplex multiple signals into independent sub-streams. On receiver side each of these sub-streams has its own spatial signature due to multi-path effects, which can be used to demultiplex and decode each of these signals. In the same way these spatial signatures can be used to separate known sources for multiple streams, they can be applied to cancel known interference.

These techniques not only apply for the communication between two instances, but can also be adopted for multiple MSs. Instead of utilizing multiple streams for one MS with multiple antennas, multiple streams are utilized for multiple MSs.

2.1.1 Channel Model

The link of each receiver-transmitter antenna pair can be modeled as a Single-Input-Single-Output (SISO) channel. Whilst all links share the same path-loss, due to the closely spacing, they do not share the same fading statistics.

The general MIMO case is depicted in Figure 1. The transmitter side is equipped with an N element antenna array, and the receiver side with M-elements. Each SISO channel between an antenna pair (i, j) can be expressed with a channel coefficient $h_{i,j}$. The complete MIMO channel is defined as the $N \times M$ channel matrix $\mathbf{H}$, where $h_{i,j} = \mathbf{H}(i, j)$.

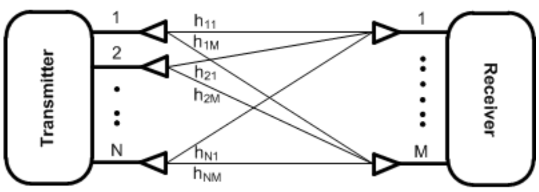

Figure 1: General MIMO layout.

The signal vector y at the receiver during time t becomes:

$$\bar{y}(t) = \bar{s}(t)\mathbf{H}(t) + \bar{n}(t), \tag{3}$$

where s is the transmitted signal. The dimensions of the vectors y and s depend on the number of possible spatial streams with at most $min(N, M)$ many.

2.1.2 Open-Loop and Closed-Loop techniques

There are many techniques that take advantage of the MIMO system in different ways. Most commonly they are categorized into two types depending on whether or not the channel state information (CSI) are available at the transmitter. These categories are known as open-loop and closed-loop MIMO.

Open-loop techniques adjust the signal on transmitter side independent of the channel conditions. One particular simple approach for two transmit antennas that is capable of exploiting the full transmit diversity, without channel knowledge, is known as Alamouti scheme [Ala98]. This scheme transmits two symbols simultaneously during two time slots over two antennas, and corrects the signal to enforce orthogonality. During the first time-slot both symbols are transmitted on separate antennas, and during the second time slot the antennas are alternated. This creates two copies separated in time and two copies separated in space for each symbol, resulting in a higher diversity at the receiver which improves the robustness of the signal. More general techniques of this space-time coding principal exist for more than two transmit antennas that act on blocks of data. Some of these techniques are referred to as space-time block codes [TJC99], or space-time trellis codes [TSMC98].

Closed-loop techniques adjust the signals on transmitter side according to the underlying channel. Due to their adaptive nature these techniques can simplify encoding compared to space-time coding [Tso06], improve error rate, and increase spectral efficiency. Essentially, in closed-loop systems any technique that can be applied on receiver side using channel state information to decode received signals (e.g. Maximum Ratio Combining, Equal Gain Combining, or Interference Reduction), can be used to pre-code signals on transmitter side with similar effects. Nevertheless, the greatest advantage of closed-loop techniques is the possibility of applying pre-coding on transmitter side with full channel knowledge, and additional decoding on receiver side. This enables spatial multiplexing of different data streams to different MSs, which increases spectral efficiency.

In general the increased antenna diversity can be used in two ways, either to make signals more robust against fading and interference, or for spatial multiplexing. Techniques like space-time coding increase diversity and therefore robustness, whereas, spatial multiplexing increases spectral efficiency. In terms of capacity, improving robustness outperforms spatial multiplexing at low Signal-to-Noise-plus-Interference-Ratio (SNIR), and at high SNIR, at which point a system becomes bandwidth limited, spatial multiplexing outperforms diversity techniques [WGSB07].

2.1.3 Channel State Information

The greatest problem faced by closed-loop MIMO communication is acquiring the channel state information (CSI) at the transmitter side. Perfect channel knowledge would need a continuous feedback link between the receiver and the transmitter and might not be a feasible solution for a practical system as it would consume additional capacity. As an alternative there are various methods available that limit the feedback through using only channel statistics or quantized versions of the channel state. Even though a few bits of limited feedback can provide substantial improvements in performance [LHSH04], they are far more prone to channel estimation error and channel evolution. Another common approach is based on the reciprocity of the channel. Here, a transmitter derives the channel state information from signals received along the opposite link [TV05]. Reciprocity does not mean the channels between transmitter and receiver are identical, it means the channel coefficients are the same. Factors like interference, noise, or differences in hardware may vary strongly per link and must be accounted for during calibration of the coefficients [GPK09]. Furthermore, reciprocity does not exist in the frequency-domain. This means that *time division duplexing* (TDD) systems must be calibrated carefully in the frequency domain for reciprocity, and *frequency division duplexing* (FDD) systems must use direct feedback of the receiver.

2.1.4 Multi-User MIMO techniques

In case of MU-MIMO scheme multiple antennas are used to send signals to different or multiple MSs at the same time using the same frequency. For this purpose knowledge of the Channel State Information (CSI) of each MS is required, that can be utilized in different ways.

On one hand, there is Dirty Paper Coding (DPC), which is a technique that adapts the signal according to the interference amongst MSs, using it constructively and completely reducing its effects. Instances of DPC techniques include Costa precoding [Cos83], Tomlinson-Harashima precoding [WC98] and the vector perturbation technique [PHS05]. However, DPC relies on non-linear precoding techniques, which is impractical in commercial wireless systems due to its high complexity.

On the other hand, there are beamforming techniques which allow the separation of multiple MSs through assigning beamforming directions and treating other MSs as noise. Common techniques either completely reduce the effects of the interferer, through nulling (e.g. Zero-Forcing [Mac08]) which is done even at high costs to the own signal, or through maximizing the SIR (e.g. Minimum Variance Distortion-less Response, Maximum SIR, Min Mean Square Error) [Gro05].

Beamforming techniques, as compared to DPC, are suboptimal strategies [YG06] that try to reduce the mutual interference between MSs, whereas, DPC uses the interference constructively. Nevertheless, due to the high computational cost of DPC, they are valuable alternatives for any real time system.

2.1.5 SDMA Grouping

Spatial-Division-Multiple-Access (SDMA) is a channel access method that is more frequently associated with beamforming techniques that rely on the spatial separation of different MSs. The spatial correlation determines the effectiveness of placing nulls between multiple MSs, as well as being able to optimize their signal. It has been shown that in case MSs are orthogonal to each other beamforming can reach the same capacity as DPC [SCA+06]. In scenarios, where more MSs exist than number of antennas a sophisticated SDMA grouping algorithm that decides which MSs to serve simultaneously with minimal interference, would overcome the limitations of beamforming techniques and achieve the same capacity as DPC.

This so called SDMA grouping problem is known to be NP-complete [STKL01], with computational cost that increase exponentially as the number of MSs increases. In order to organize the SDMA groups efficiently with low complexity, suboptimal grouping algorithms are required. Multiple suboptimal SDMA grouping algorithms have been proposed in literature [Mac08, STKL01].

One of the SDMA grouping algorithms commonly used, which has been shown to perform well with low complexity is a greedy algorithm called the Best Fit Algorithm (BFA) [STKL01]. It constructs an SDMA group through starting with the MS with the weakest signal, i.e. lowest SNR/SNIR, and sequentially extends the group by admitting the MS providing the highest increase for a given grouping metric. Once the group size reaches a target size, or no more MSs exist to increase the grouping metric, the SDMA group is fully constructed. Another closely related algorithm with even lower complexity is the First Fit Algorithm (FFA). The only difference, compared to the BFA, is that instead of admitting the MS providing the highest increase for the grouping metric it adds the first MS to the SDMA group that holds the slightest gain [STKL01].

In literature a broad number of SDMA grouping algorithms exist. A comprehensive overview is given in [Mac08]. Another grouping algorithm related to this work, with a higher complexity as the FFA/BFA is called the Cluster-Based-Algorithm [ZMDMXC12].

2.1.6 SDMA Metrics

As previously mentioned, grouping algorithms require a *grouping metric* in order to compare candidate SDMA groups with each other. In general, a grouping metric makes use of the *Channel State Information* (CSI) in order to map the characteristics of the spatial channels of the MSs to a scalar value [Mac08]. The most commonly used grouping metrics are the *group capacity* and the *group minimum SNIR*. The former considers the Shannon-Hartley capacity of an SDMA group, and the latter returns the lowest SNIR of an MS in a given SDMA group. Both grouping metrics rely on the actual beamforming weights and/or the power allocation which involves complex vector/matrix operations which comes at the expense of an increased complexity. In order to decrease complexity, grouping metrics that are based only on the spatial correlation and on the channel gains of the MSs, involving much simpler vector/matrix operations, have been proposed [FN96, STKL01, Mac08].

2.2 OFDM/OFDMA

Orthogonal Frequency Division Multiplexing (OFDM) is a frequency division multiplexing scheme (FDM), which divides a wideband channel into multiple orthogonal sub-carrier frequencies. Unlike conventional frequency division multiplexing schemes the orthogonality between the carriers allows a close spacing between multiple sub-carriers, without causing interference, leading to a high spectral efficiency. This is possible due to the fact that each sub-carrier frequency is orthogonal to its adjacent sub-carriers, meaning, the peak of one sub-carrier coincides with the null of its adjacent neighbor.

One of the greatest advantages of OFDM systems over single-carrier transmission schemes in wideband channels is the fact that it turns a wideband single-carrier signal into multiple narrow-band signals. Each sub-carrier is modulated separately by multiple lower rate data streams that, as a sum, hold a total data rate similar to a conventional single carrier scheme without the need of complex equalization filters. Therefore, OFDM can cope with severe channel conditions like frequency selective fading, which can be considered flat for a single narrow-band sub-carrier, through using sufficient spacing.

2.2.1 Frequency Diversity

Figure 2 shows the channel gain over a frequency band of 10 MHz split into multiple narrow-band sub-carriers. In a typical wideband system the total bandwidth can be much larger than the coherence bandwidth of the channel resulting in frequency-selective channel gains per sub-carrier. This diversity in frequency creates additional advantages for multi-user schemes like OFDMA, which is a multi-user version of OFDM where subsets of sub-carriers can be assigned to different MSs. An OFDMA system can benefit from frequency-selectivity in two distinct ways [Lee07].

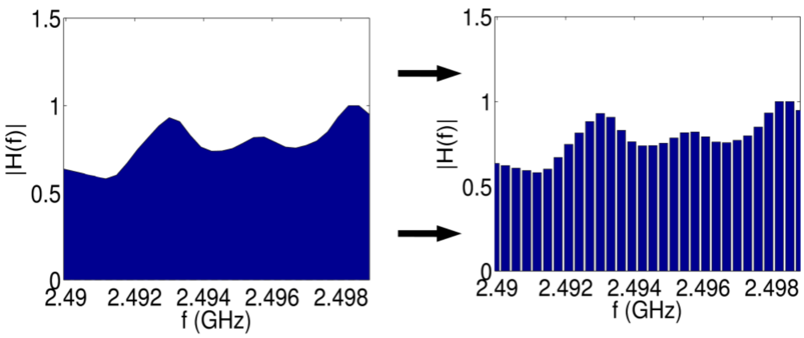

Figure 2: Wideband single-carrier frequency spectrum split into multiple narrow-band sub-carriers.

The first approach allocates each MS a subset of the total available sub-carriers spread across the entire frequency band. The performance gain achieved this way, is obtained by using multiple sub-carriers whose path gains are independently faded, rather than using adjacent sub-carries with similar faded path gains. This approach is commonly referred to as *frequency diversity gain* (FDG).

The second approach, called *frequency-selective scheduling* (FSS), is achieved by allocating each MS adjacent sub-carriers located within a sub-band of limited bandwidth having the most favorable channel conditions within the entire frequency band.

FSS in comparison to FDG, heavily depends on accurate channel state information. Without channel knowledge FDG reduces the probability of allocating sub-carriers that lie within the same deep-fade, where FSS would fail. But, with channel knowledge, FSS selects peaks, avoiding deep-fades all together. This means there is a gray line between the performance of FDG and FSS. For example in scenarios with high mobility where the channel state is quickly changing FDG has a greater advantage. However, as mobility is reduced and the channel turns more stable there is a turning point where the FSS gain grows beyond FDG.

2.3 SDMA-OFDMA support in WiMAX

This work is based on the real world use-case of the WiMAX IEEE 802.16-2009 standard [80209]. This section provides an overview of possible PHY layer features, the TDD/OFDMA frame structure, and the included channel feedback support of WiMAX for MIMO techniques.

2.3.1 PHY Layer Features

WiMAX 802.16e-2005 defines three different PHY layers: single-carrier transmissions, OFDM, and OFDMA. The first two are pure time division multiplexing access (TDMA) schemes, OFDMA on the other hand, uses both time and frequency dimensions for resource allocation. The frequency ranges include 2–11 GHz and 23.5–43.5 GHz, and supports different bandwidths between 1.25 MHz and 20 MHz. For robustness and reliability WiMAX supports BPSK, QPSK modulation as well as higher order schemes like 16-QAM and 64-QAM. The uplink (UL) and downlink (DL) can be duplexed either in time (TDD) or in frequency (FDD). In FDD full-duplex is supported for Mobile Stations (MS) capable of transmitting and receiving simultaneously, and half-duplex for MSs without such feature. In TDD on the other hand, the uplink and downlink share the same frequencies and are separated in time into different sub-frames that are not necessary divided into equal parts.

2.3.2 OFDM/OFDMA

An OFDM/OFDMA system uses multiple subcarriers for transmission. In WiMAX there are three OFDM subcarrier types called data, pilot, and null subcarriers each with its own purpose. The data subcarriers are used for actual data transmission, whereas a pilot subcarrier can be used for various estimation and synchronization purposes, and the null subcarriers are plainly used as guard bands. All available subcarriers are grouped into subsets called subchannels. These subchannels are the smallest frequency resource unit. An OFDM frame can be considered as a matrix where rows and columns are subchannels and OFDM symbols, respectively.

In comparison to OFDM, OFDMA is the multiple access scheme where different subchannels can be allocated to different MSs. The mapping of physical subcarriers to logical subchannels may be performed through randomly assigning subcarriers distributed across the entire frequency spectrum, or through assigning adjacent subcarriers to each subchannel. The former method called *partial usage of subcarriers* (PUSC), provides *frequency diversity*. The latter, scheme is called *adaptive modulation and coding* (AMC) and allows the exploitation of multi-user diversity through *frequency selective scheduling*. Both schemes have advantages and disadvantages as mentioned in Section 2.2.1.

2.3.3 Channel Feedback

Closed-loop MIMO techniques rely heavily on the availability of the channel's state information. In WiMAX there are different ways to acquire CSI, including the channel coefficients. Either by relying on channel reciprocity by using the uplink channel estimation as the downlink channel state, or through using direct feedback where the MS transmits the estimated channel state to the base station. The first method is only applicable for TDD, whereas FDD must use direct feedback due to the channel separation in frequency. Since a MS can utilize a different number of receive antennas than transmit antennas, direct feedback might also be necessary for TDD. Optionally, the standard includes a signaling mechanism where a MS can be requested to transmit a channel sounding waveform on the uplink for channel estimation in TDD. Also the uplink sounding waveform can be transmitted in combination with the channel coefficients estimated by the MS, thus allowing its use in TDD as well as in FDD.

2.4 Related Work

There exists a substantial amount of literature on SDMA grouping algorithms, grouping metrics and beamforming techniques. A broad overview can be found in [Mac08], where various solutions have been classified and compared for use within an SDMA-OFDMA system. However, only a few proposals evaluate the performance of SDMA in real OFDMA systems like 802.16e. This section discusses recent proposals on SDMA-OFDMA scheduling solutions for WiMAX.

Nascimento et al. [NR10] proposed a joint utility packet scheduler and SDMA-based resource allocation architecture for 802.16e. Similar to this work, [NR10] uses the Exponential Effective SIR Mapping (EESM [EES05]) method to capture frequency selective channels and takes advantage of the scalar product as low complex metric to build SDMA groups consisting of uncorrelated MSs. The proposed scheduling solution therein assigns MSs to beams, adding subsequent MSs by choosing the one with the lowest spatial correlation in a First-Fit kind of manner while testing its compatibility through the actual EESM SNIR whilst steering beams towards the estimated direction of arrival. QoS was achieved through a prioritized assignment similar to the proportional utility, using a sorted priority list to select the next MS to be assigned. In order to reduce complexity, the authors reduced the SDMA capabilities to one third of the OFDMA frame leaving the rest for non-SDMA transmissions. Performance was evaluated under the assumptions of a full queue traffic model. The results showed substantial gains in comparison to non-SDMA.

Similar to the previous work, Yao et al. [HCCW08] evaluated the MAC performance of an SDMA-OFDMA system where the OFDMA frame is separated into SDMA and non-SDMA capable zones. The scheduling solution provided, prioritizes MSs based on their channel conditions and packet delays. Specifically, MSs are grouped according to their channel conditions, whereas, within a group, MSs are prioritized according to packet deadlines. Depending on the intra-beam interference amongst MSs, they are assigned to a beam with favorable conditions or moved to the regular non-SDMA zone. The improvement achievable through using SDMA was evaluated for FTP and VoIP services. It showed that for FTP services with large packets and delay tolerance using SDMA improved performance significantly.

Both of these approaches assume an idealized Line of Sight (LoS) channel that allows beamstearing based on the estimation of the angle of direction of an MS. But they neglected the advantage of MIMO systems being able to take profit of LoS components as well as possible scattering and multi-path effects.

3 Proposed Scheduling Algorithm

This section is split into two sections. The first section describes and discusses the resources allocation problem in frequency, time, and space in gernerall terms. Whereas Section 3.2 focuses on the proposed algorithm and its design in detail.

3.1 Resource Allocation Problem

The resource allocation of multiple dimensions provided by a MIMO-OFDMA based system results in a highly complex problem. Alone the resource allocation of an SDMA/TDMA system comes with an exponentially increasing scheduling complexity [STKL01]. In combination with OFDMA, or rather multiple subcarriers, the complexity further increases. Nevertheless, it is possible to divide the problem into subproblems which can be solved independently with lower complexity, using suboptimal solutions. A common approach to simplify the problem is based on the assumption of adequate synchronization of the frequency and time dimensions through which neighboring resources can be guaranteed to be interference free of one another. The frequency and time resources can, therefore, be allocated independently, whereas the spatial dimension can not guarantee the same in-dependency for which its allocation is solved separately. Section 3.1.1 describes the allocation of the spatial dimension and some aspects that need to be considered by a practical solution. Section 3.1.2 focuses on the joint frequency-time allocation, and briefly formulates the task of assigning resources to the best SDMA groups in terms of a given utility function.

3.1.1 Spatial Dimension

This section briefly describes the allocation of the spatial dimension and its role in combination with an OFDMA system. Most notably this section elaborates two aspects: i) The SDMA grouping problem, its components in theory and in practice, ii) The practical need of performing SDMA grouping across multiple subcarriers beyond the coherence bandwidth, and how to achieve it without neglecting the frequency selective properties of the channel.

The SDMA grouping problem, as described in Section 2.1.5, refers to the task of selecting MSs with spatial channels that are close to orthogonal, or at least highly uncorrelated, so that a spatial multiplexing gain can be achieved. The spatial correlation amongst these MSs that where selected to transmit in the same time-frequency resource unit are crucial to the overall capacity. The problem of finding the best SDMA group is known to be NP-complete [STKL01], and is similar to the well known knapsack problem [Cal04]. Since an exhaustive search over all possible SDMA group combinations would be infeasible, suboptimal SDMA algorithms are necessary in order to reduce the overall computational costs at a tolerable performance loss. The theoretical structure of an SDMA grouping algorithm consists of two parts: the grouping algorithm and a grouping metric [Mac08]. Both parts allow a reduction in complexity. In a practical system however, there is another issue that must be accommodated for. Scheduled MSs must fulfill a minimum requirement, like a minimum SNIR threshold for the lowest Modulation and Coding Scheme (MCS), to ensure a successful transmission. This minimum requirement can only be assured through estimating the SNIR using the actual precoding. In case a sub-optimal low complex grouping metric is used it must be guaranteed, in a final step, that there are no MSs within the SDMA groups that fail the minimum requirement.

In conclusion to these observations, the allocation of the spatial dimension in a practical systems consists of three parts as depicted in Figure 3. Namely, the SDMA grouping algorithm, the Grouping Metric, and a final step that ensures the minimum requirement for all SDMA groups. The task of the final step can reach from simply removing MSs of an SDMA group to collecting all drop-outs of all SDMA groups and returning them to the SDMA grouping algorithm.

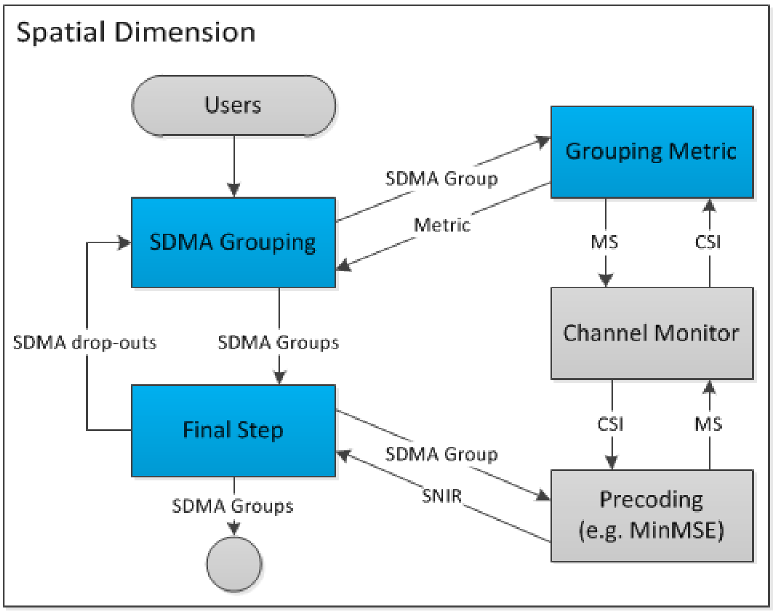

Figure 3: The required components of a practical scheme for the allocation of the spatial dimension (emphasized in blue). In addition to the theoretical components of an SDMA grouping algorithm, a practical approach must ensure in a final step that every MS within an SDMA group fulfills a minimum requirement.

Combining SDMA with OFDMA means combining a narrow-band technique with a wideband channel consisting of multiple narrow-band subcarriers. As was mentioned in Section 2.2, an OFDMA system can cope with severe channel conditions like frequency-selective fading, which can even hold a *frequency-selective scheduling* gain. To the best of the authors knowledge the question on how a frequency selective environment effects performance of SDMA in OFDMA based systems has not been discussed in literature. One could assume that this is because there is no need for it since most work focuses on performing SDMA grouping on the smallest block of frequencies for which full CSI can be obtained, and doing so means scheduling the best MSs for each frequency block. In other words, combining SDMA with OFDMA implies performing *frequency-selective scheduling*. From a theoretical point of view the optimal SDMA-OFDMA resource allocation strategy in a frequency-selective channel is to perform SDMA grouping for the smallest possible frequency range. However, from a practical point of view this is not reasonable. Every packed data burst needs to be signaled by an entry in the MAC header, i.e DL-MAP in case of WiMAX, resulting in substantial signaling overhead. To the best of the author's knowledge there is only one practical proposal that addresses this issue. In [Cal04] it was suggested to reduce signaling overhead through allocating the same SDMA groups for an integer number of adjacent subcarriers. The more subcarriers are being allocated to the same SDMA group, the lower the needed signaling overhead.

For practical purposes we will refer to a set of adjacent sub-carriers that are used for SDMA grouping as sub-band. Figure 4 depicts how the OFDMA DL frame can be divided into SB many sub-bands. Depending on the value SB ≥ 1, SDMA grouping is performed either on one sub-band[3] across the complete available frequency spectrum, or for each available sub-band separately. Possible frequency-selective properties of the underlying sub-carriers within a sub-band can be collapsed into a single scalar value by using techniques like the Exponential Effective SIR Mapping (EESM [EES05]). This method maps a set of sub-carrier SNIR measurements into a single effective SNIR measurement, which can be used by the SDMA grouping algorithm.

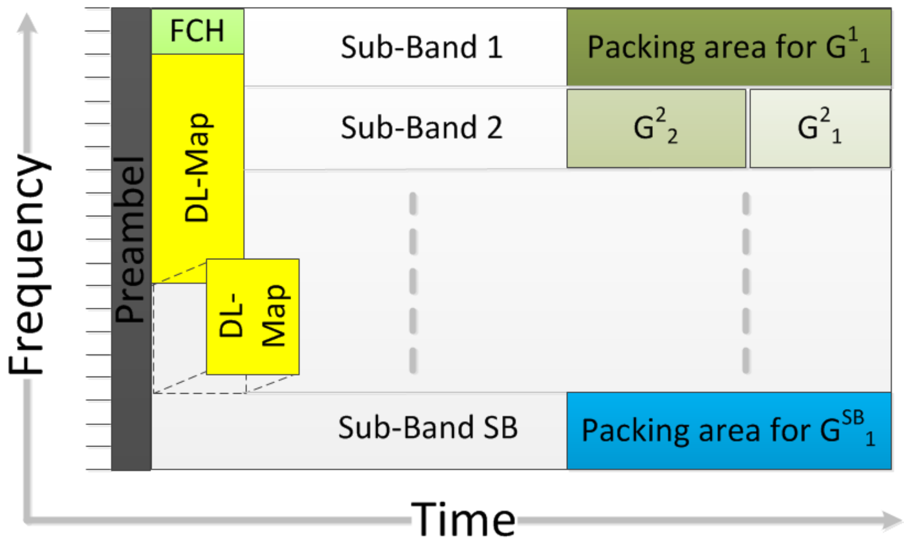

Figure 4: Proposed partitioning of the OFDMA downlink frame into sub-bands. G_j^i refers to the j-th SDMA group on the i-th sub-band.

In short, scheduling the resources of a spatial layer in a practical system means: i) ensuring a minimum requirement for all scheduled MSs, and ii) being able to schedule the same SDMA group across multiple sub-carriers in an effort

[3]A sub-band refers to a set of adjacent subcarriers

to minimize MAC layer overhead.

3.1.2 Frequency-Time Allocation

The main objective of a DL SDMA-OFDMA MAC scheduler is to assign resources according to a given QoS utility. This section briefly describes how the frequency and time resources should be scheduled.

The previous section summarized the allocation of the spatial dimension and introduced the term sub-band. A sub-band refers to a given frequency range associated to a set of SDMA groups generated by the SDMA grouping algorithm. For each sub-band a number of compatible SDMA groups are given, where each MS u within an SDMA group is associated to a packet list P_u, which is individually tagged with a utility value u to account for QoS (e.g. priority of a packet). Given a set of SDMA groups per sub-band, the scheduler must select a subset of the given SDMA groups to be allocated that maximize the sum utility of the associated packets carried by the frame, as well as allocating the needed resources for the DL map whilst minimizing the amount of unused resources. Furthermore, it is important to select the best SDMA groups across frequencies to ensure an FSS gain.

The main reason why the QoS utility is considered during this part of the scheduling process instead during the SDMA grouping, is due to the fact that this work focuses on a packet-switched network. Which means, different packets of the same MS can be tagged with different utility values. Therefore, the actual priority is based on packets and not on the MS. It would be difficult to distinguish during the SDMA grouping which packets should represent the MSs priority, especially as a small packet that barely consumes resources can hold a very high utility, but takes no advantage of additional spatial resources.

3.2 Design

This section describes the greedy DL SDMA-OFDMA MAC scheduler that has been implemented. It employs and combines various state-of-the art techniques of future OFDMA and SDMA systems into a practical greedy scheduling solution. It is a generic solution that is capable of integrating different SDMA grouping algorithms with varying degrees of complexity. The grouping metric is based on the SNIR values achievable by members of an SDMA group. Multiple SNIR values per carrier are collapsed into a single SNIR equivalent value using the ESSM mapping technique. Additionally the total number of adjacent subcarriers can be grouped into subsets of variable sizes, thus, allowing different degrees of granularity upon which the SDMA grouping algorithms can operate. This way SDMA groups can either span the whole frequency spectrum or subsets thereof. Since computing an accurate prediction of the achievable SNIR includes complex calculation of the precoding vectors, an alternative to do so, with low complexity is proposed. Finally, an OFDMA frame construction algorithm is introduced that allows the selection of a subset of the possible SDMA groups in order to maximize the utility within a frame, therefore, achieving a QoS defined by the system operator. The selection is performed across frequencies to achieve a frequency selective scheduling gain (FSS). The time dimension is allocated iteratively by the frame construction algorithm to ensure a gradual allocation of MAP signaling and resources as to minimize possibly unallocated resources.

3.2.1 Overview

An overview is given in Figure 5. There are four modules highlighted that play a substantial part in allocating the time, frequency and space resources. The first three are related to the spatial dimension, whereas the last module allocates the time and frequency resources in order to maximize the utility carried within a frame. The first module represents the SDMA grouping algorithm that utilize the EESM mapped SNIR representation of the channel, provided by the second module (SNIR Prediction), to compare different SDMA group combinations. The SNIR Prediction module estimates the SNIR per subcarrier either through applying the correct complex precoding weights, or through a low complexity heuristic. These modules are described in Section 3.2.2 and 3.2.3, respectively. In case the low complex metric is applied for SDMA grouping, it might occur that SDMA groups are malformed, leaving members of a group in outage due to high intra-group interference. The Adaptive Modulation and Coding module, that sets the MCS of PHY bursts to ensure error free transmissions, has been extended to compensate for the inefficiency of the low complex SNIR prediction and is explained in Section 3.2.4. The last Section 3.2.5 concludes the proposed scheduling solution, as it describes the final OFDMA frame construction, the frequency selective scheduling of the available SDMA groups, and how QoS is achieved.

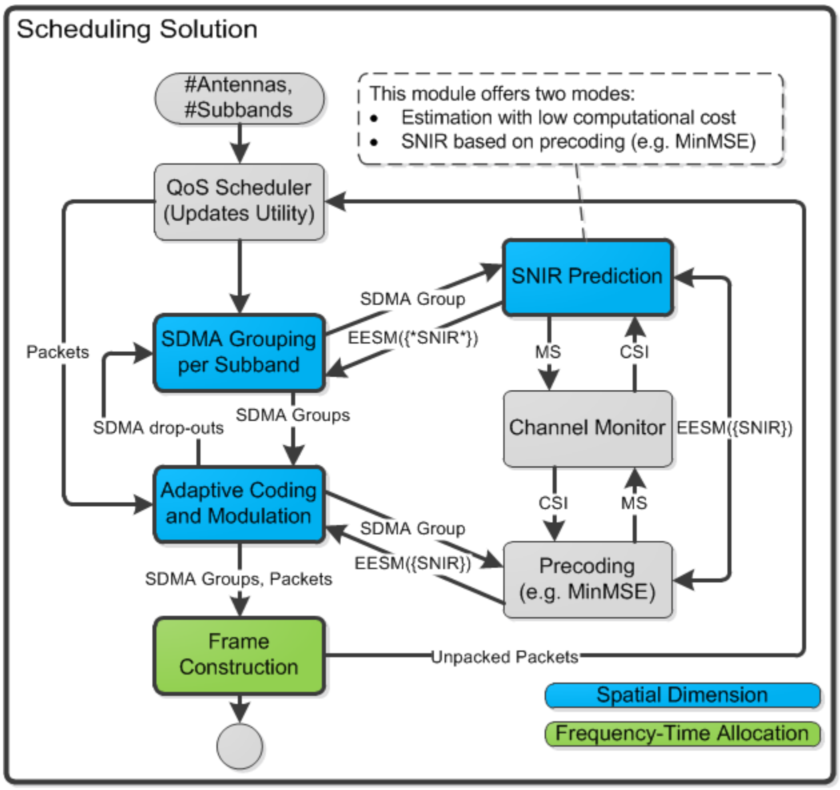

Figure 5: Overview of the proposed scheduling solution.

3.2.2 SDMA Grouping

This section gives a short description of the implemented state-of-the-art SDMA grouping algorithms, the SDMA grouping metric, and how these are applied to multiple sub-carriers. Furthermore, a simple observation is made on how SDMA grouping over the whole frequency spectrum, in comparison to doing so over subsets of sub-carriers, effects the complexity.

The SDMA group formation problem has a high computational complexity. Thus, the main characteristic that distinguishes SDMA grouping algorithms from one another are the worst case number of comparisons carried out to reach a sufficient performance. The SDMA grouping module includes three such algorithms. In ascending order of complexity they are the First Fit (FFA), Best Fit (BFA) and the Cluster-based Grouping Algorithm (CBA) [ZMDMXC12].

The FFA has the lowest complexity. It takes a sorted list of MSs according to their SNR and starts with the first one to start a group. Then the next MS within the list is admitted test-wise to the group, at which point the new group is checked if it full-fills the minimum SNIR requirement for each MS, if so the new MS is permanently admitted to the SDMA group. In case a MS is not compatible the next MS in line is added test-wise. Ones the maximum group size is reached, or no MS could be admitted without breaking the minimum SNIR requirement the SDMA group is fully constructed. For each subsequent time slot, this procedure is repeated until the initial pool of MSs is empty and a list of compatible SDMA groups has been generated.

The BFA extends the principle of the FFA, instead of adding the first MS which does not break the minimum SNIR requirement of the group it adds the best compatible MS of all available based on a SDMA group performance metric.

The final, and most complex grouping algorithm included in the grouping module is the CBA. In comparison to the FFA/BFA it takes a completely different approach. Instead of optimizing a single SDMA group it tries to optimize a complete set of SDMA groups. Initially k MSs are randomly distributed among $\lceil \frac{k}{N} \rceil$ groups, assuming that an N-element antenna array can serve at most N MSs within a SDMA group. After this step the SDMA group set is iteratively improved through exchanging MSs between groups until every MS is above the minimum SNIR requirement. The exchange is performed in two steps, only if it improves the sum of a given metric. In the first step every MS pair i, j, that lie in different SDMA groups, are swapped and tested if this exchange improves the sum of both group metrics. If it does and no MS is below the minimum SNIR requirement they are reassigned. In the second step a single MS is allowed to jump into another SDMA group in case this increases the metric. In addition, if at the end of one iteration some MSs are below the minimum SNIR requirement they are separated into their own SDMA group, which increases the number of SDMA groups at most by one, upon each iteration. A more detailed description of this algorithm is given in [ZMDMXC12].

These three grouping algorithm have two things in common. They all build SDMA groups containing MSs that lie above the minimum SNIR requirement, to ensure that every MS can be served at least in the lowest MCS, and they all use the same SDMA grouping metric, namely the group capacity:

$$C_{cap}(G) = \sum_{u \in G} log_2(1 + \mathrm{SNIR}_{lin}(u)), \tag{4}$$

where G is the given SDMA group and $\mathrm{SNIR}_{lin}(u)$ the function which returns the achieved SNIR of MS u in linear units.

The SDMA grouping algorithms need a scalar value as comparison metric. For this purpose a set of per-subcarrier SNIR measurements are collapsed into an effective SNIR measurement using the EESM method [EES05]. The SDMA grouping algorithms operate either on all available subcarriers, or an arbitrary subset thereof. Figure 6 depicts the mapping of SC many subcarriers onto SB equally sized subsets. On each subset $s \in 0 \cdots SB$, also referred to as sub-band, SDMA groups are generated from a given set of available MSs U_s. Due to the frequency selectivity of the channel not every MS u can be served on every sub-band. As a consequence a subset U_s does not necessarily contain all MSs, but only those that have an SNR value which is above the minimum SNR threshold. These subsets are generated prior to the SDMA grouping.

Complexity

As mentioned, the SDMA grouping module operates on SB many sub-bands of adjacent sub-carriers separately, and as a result increases the computational cost linearly with SB. Intuitively one might think that this fact is a limiting factor for the maximum range of SB. But, this is not the case, creating SDMA groups spanning all sub-carriers, and grouping MSs per sub-carrier is at most of equal computational cost.

Let us assume an arbitrary grouper performs at most E many evaluations, let the cost for estimating the performance of an SDMA group per subcarrier be a constant C, and may the total number of subcarriers be given by SC. Thus, the total cost of SDMA grouping over all subcarriers would be $E \times (C \times SC)$. Now, consider splitting the total numbers of subcarriers into SB many equally sized subsets. Due to the frequency selective channel, the grouping must be done per subset separately. So there are $SB \times E$ evaluations of SDMA groups, but only SC/SB many subcarriers that play a role per evaluation. In total the cost would be $SB \times E \times (C \times SC \times \frac{1}{SB})$, as can be seen these are the same cost as grouping over all subcarriers: $E \times (C \times SC)$.

Therefore, SB can be chosen as large as SC, meaning there is no complexity penalty for performing SDMA grouping per subcarrier. In fact, increasing SB, on one hand increases the number of SDMA group evaluations, and on the other hand reduces the cost per evaluation with the same magnitude.

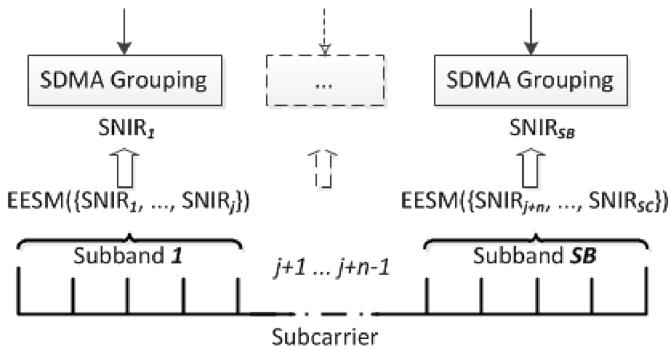

Figure 6: All available subcarriers SC are divided into SB many equally large subsets of subcarriers. The SNIR values within these subsets are collapsed into a single SNIR representation using the EESM method. Upon these values the SDMA grouping algorithm generate SDMA groups for every sub-band separately.

3.2.3 SNIR Prediction

As noted in Section 3.2.2 the SDMA groups are evaluated using the group capacity (eq. 4) based on the estimated SNIR of each member of the group. The SNIR includes the intra-group interference and can be accurately estimated using the applied precoding weights as in Equation 9. As explained earlier, the SDMA grouping collapses multiple subcarrier SNIRs into a scalar value using the EESM mapping technique. However, these calculations require complex vector/matrix operations to be performed per subcarrier with high computational cost. As an alternative this work proposes a SNIR Prediction algorithm that obtains an SNIR equivalent representation of each subcarrier without performing expensive pre-coding computations.

Figure 7 illustrates the basic difference between calculating the SNIR using a precoding matrix and the proposed SNIR Prediction algorithm. When

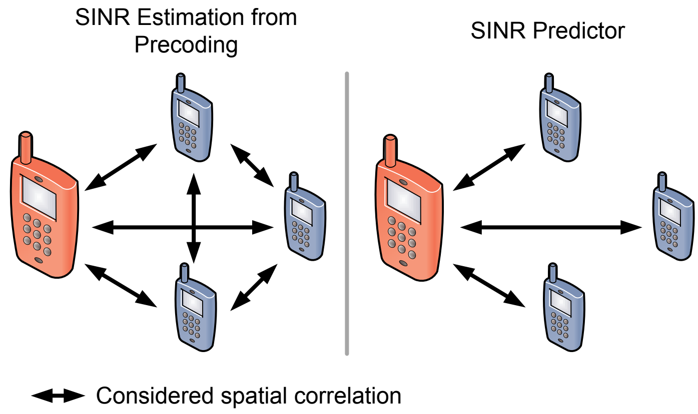

Figure 7: Illustration of two method for calculating the SNIR of a MS: (i) from the actual precoding matrix derived by any arbitrary MIMO technique, (ii) using the SNIR prediction algorithm.

calculating the SNIR of an MS (red) within an SDMA group the proposed SNIR Prediction method considers only the spatial correlation between the MS of interest and the other MSs within that group. Hence, unlike the case of using a precoding matrix, in the SNIR Predictor case the spatial correlations between the other MSs within that group are not considered. Therefore, the SNIR Predictor has a lower complexity but introduces an estimation error in the per-subcarrier SNIR.

In particular, the SNIR for any SDMA group G and MS $i \in G$ on frequency resource block b is estimated in the following way:

$$\tilde{\gamma}_b(u, G) = SNR_{\mathrm{MRC}}(u) \cdot \begin{cases} 1 - \chi_u^b, \text{ if } \chi_u^b \leq 1 \\ \beta, \text{ else} \end{cases} \tag{5}$$

where $\chi_u^b = \sum_{u' \in G, u' \neq u} \tau_{u,u'}^b$, i.e. the sum of the squared spatial correlations between MS u and the other MSs in group G on frequency resource block b.

The idea of the SNIR Predictor is to consider an initial SNIR for a MS u in an SDMA group G and frequency resource block b that is obtained assuming a non-SDMA transmission (MRC beamforming, refer to Equation 11). Then, this initial SNIR estimation is corrected by the expected intra-SDMA group interference caused by the other MSs in the group. In particular, the applied correction factor is either χ_u^b if $\chi_u^b \leq 1$, or a constant factor β, which in this case set to 13.5 dB (0.0443 in linear), if $\chi > 1$. The reason for the constant factor is that numerical simulations showed that for $\chi > 1$, the MS's SNIR was on average 13.5 dB below its SNR_{MRC} value. This applied constant factor though could be tuned in order for the SNIR Predictor method to be applied in other scenarios.

Notice that considering a minimum threshold for packet reception on the lowest MCS of 5 dB together with a SNR drop of 13.5 dB results in a required SNR of 18.5 dB, which is the region for the highest MCS in 802.16e. Therefore, only high SNR MSs are likely to cope with strong interference without drop-

ping below the minimum threshold for packet reception; e.g. a MS u having a $SNR_{\mathrm{TxBf}}(u) < 18.5$ and $\chi > 1$ is considered to be in outage.
In addition, for a given frequency resource block b, the term $\tau^s_{u,u'}$ in Equation 5 is computed as:

$$\tau_{u,u'} = (p_{u,u'})^2 \tag{6}$$

where given the channels h_u and $h_{u'}$ of MSs u and u', respectively, p represents the spatial correlation between these two MSs on a given frequency resource block and is given by the maximum normalized scalar product [Cal04]:

$$p_{u,u'} = \frac{|h_u^{\mathrm{H}} h_{u'}|}{||h_u||_2 ||h_{u'}||_2} \tag{7}$$

Complexity

The proposed SNIR prediction algorithm, as simple as it seems, holds the potential for being equally complex as any arbitrary precoding technique. This means the most important part of the proposed method is, in fact, the symmetric properties of the correlation function which reduces the computational cost severely. Through applying basic caching techniques any redundant operations can be prevented and each SDMA group evaluation is reduced to a simple sum of scalars.

An arbitrary MIMO beam-former technique that optimizes beamforming weights under a certain criteria does so by solving a system of N linear equations. Using Gaussian elimination for this purpose results in an arithmetic complexity of $O(N^3)$ [Far88], where N is the number of antennas in the BS. The SNIR predictor proposed can be just as complex as $O(N^3)$. However, by using caching techniques as well as the nature of iterative groupers the complexity can be reduced to $O(N)$.

The complexity of computing the term τ itself is linear on the number of antennas in the BS, i.e. $O(N)$. In addition, within the context of an SDMA group evaluation, τ must be calculated for each MS to every other interfering MS. This means that for a given group size g there will be a total of $g(g-1)$ evaluations of the term $\tau_{u,u'}$, hence resulting in an overall complexity of $O(M)g(g-1) = O(M \cdot g^2)$. Obviously, g can be at most N, in which case the complexity would be $O(N^3)$, at worst.

However, the term $\tau_{u,u'}$ can be pre-computed and stored in a lookup table. Pre-generating the values for every MS pair reduces the complexity of evaluating an SDMA group to $g(g-1)$ simple lookups, and, given that $g \leq N$, the overall worst case complexity would be $O(N^2)$. Finally, for any SDMA grouper that evaluates SDMA groups iteratively through sequentially adding MSs (e.g. BFA or our CBA grouper) the complexity could be reduced even further. The reason is that every time a new MS is added to a group, every MS within the group only needs to add its correlation to the added MS, whereas the newly added MS sums the correlation to every other MS, hence resulting in $(g-1+g-1)$ lookups and a complexity of $O(g)$, or at worst $O(N)$.

3.2.4 Adaptive Modulation and Coding (AMC) Module

The Adaptive Modulation and Coding (AMC) module is used for adapting the coding-scheme on a burst-by-burst basis per link, depending on the current channel conditions. In order to do so a set of SDMA groups is given by the SDMA grouping module, each associated to a given sub-band, and an accurate SNIR estimate is used to assign an MCS to each burst for a given target block error rate. In case the SDMA grouping was performed using the low complex SNIR prediction algorithm additional measures are undertaken to compensate its error. This section describes the general MCS selection, as well as these extra measures.

In order to derive the MCS to be assigned, the AMC module computes the actual pre-coding matrices per frequency block within the used sub-band for each MS in an SDMA group. Once the pre-coding matrices are computed, the AMC module can obtain an accurate SNIR estimate for each MS and frequency block by means of Equation 9. Then, the AMC module collapses the set of per-frequency block SNIR values, $\{\gamma_{u,b}\}$, into a single equivalent SNIR value for each MS making use of the EESM mapping technique. Finally, a single scalar SNIR value is available for each MS and given a target block error rate (1% in this case), the AMC module selects the proper MCS by doing a simple look up operation on a pre-computed *Block Error Rate-SNR* table.

In addition, as mentioned, the AMC module compensates introduced errors by the low complex SNIR prediction algorithm, in case it has been applied for the SDMA group formation. The low complex SNIR prediction, as will be shown later, tends to be optimistic especially as the number of antennas grows, which means some groups might have MSs that can not be served and must be regrouped accordingly. This situation is dealt with in Algorithm 1, where for each SDMA group G having MSs in outage, the worst MS u (largest $SNR_{\mathrm{MRC}}(u) - \gamma(u, G)$), is removed from the group. Then, a new precoding matrix is calculated for the remaining MSs left in the group. The removed MSs are collected and handed back to the SDMA grouping module that re-computes SDMA groups for the in-outage MSs (see Figure 5). Also, every time the SDMA grouping module has to re-group a set of MSs, the maximum SDMA group size is reduces by one. This is done out of two reasons: for one, the computational cost for SNIR prediction used by the grouper reduces as the maximum SDMA size decreases, and for another the likelihood of outage MSs to be in outage again decreases as well, since the number of potential interferer in an SDMA group is reduced.

Complexity

Given a set of SDMA groups, the number of precoding calculations performed by the AMC module corresponds to the product between the number of groups, $|\mathcal{G}|$, and the number of frequency blocks, B. However, when using the low complex SNIR prediction, additional precoding calculations may be required if there are MSs in outage. Assume that in the worst case the SNIR prediction leads to $|\mathcal{G}|$ groups; each consisting of N highly correlated MSs that cannot be served together. In that case, in each iteration of Algorithm 1 each group is reduced until its group size reaches one, resulting in a total of $O(|\mathcal{G}|N^2B)$ precoding calculations. Therefore, if $|\mathcal{G}| = \lceil K/N \rceil$ the worst case time-complexity regarding the number of precoding calculations is $O(BNK)$. In case of multiple sub-bands the number of groups $\mathcal{G}$ is increased with equal magnitude as the frequency blocks per group is reduced, which effects the complexity in the same manner as the grouping described in Section 3.2.2 and, therefore, is neglectable.

Algorithm 1 Required preprocessing step in the AMC module when SNIR prediction algorithm is used for SDMA grouping.

Require:
K - Number of MSs; N - Number of BS antennas
procedure ProcessOutageMSs
  $U \leftarrow 1 : K$ ▷ Set of all MSs.
  $\mathcal{G}^* \leftarrow \emptyset$ ▷ Final grouping configuration.
  for $maxGroupSz = N$ downto 2 **do**
    $\mathcal{G} \leftarrow$ SDMAGroupingModule($U, maxGroupSz$) ▷ Using SNIR predictor
    $U \leftarrow \emptyset$ ▷ MSs to be regrouped
    for $G_i \in \mathcal{G}$ **do**
      while $O =$ getOutageMSsAfterPrecoding$(G_i) \neq \emptyset$ **do**
        $u \leftarrow \arg\max_{u \in O}\{SNR_{\mathrm{MRC}}(u) - \gamma(u, G_i)\}$ ▷ The MS which suffered most
        $G_i \leftarrow G_i \backslash u$ ▷ Remove u from group G_i
        $U \leftarrow U \cup u$ ▷ Add u to MSs to be regrouped.
      end while
      $\mathcal{G}^* \leftarrow \mathcal{G}^* \cup G_i$ ▷ Valid group G_i is added to final grouping
    end for
    if $U == \emptyset$ **then** ▷ No more MSs in outage; terminate
      break
    end if
  end for
return $\mathcal{G}$ ▷ Return grouping
end procedure

3.2.5 Frame Construction

The Frame Construction algorithm constructs an OFDMA frame from multiple per-sub-band available SDMA groups and their packets with the main goal to maximize the total utility carried by an OFDMA frame[4]. Additionally it supports a variable number of sub-bands to be used with the goal to achieve a FSS gain. The higher the total number of sub-bands available, the higher the probability to find a sub-band with favorable channel conditions for an SDMA group.

Given a set of SDMA groups per sub-band, the algorithm selects a subset of these groups to be allocated within the current SDMA-OFDMA frame depending on the utility of corresponding packets. In addition the space needed for signaling the DL-MAP, which varies depending on the selected groups, is allocated as well. Algorithm 2 describes the algorithm consisting of two phases: the *extension* phase (outer loop) and the *selection* phase (inner loop). The *extension* phase allocates the frame in the time dimension in a stepwise fashion, whereas the selection phase chooses the best groups over all sub-bands as to achieve a FSS gain.

In the *extension* phase a certain area is made available for packing. In the *selection* phase the best SDMA group over all sub-bands, which maximizes the sum utility of the frame using the available packing area is scheduled with its given size. For each next iteration within the *selection* phase the best group of all *remaining* sub-bands is scheduled. After all sub-bands have been allocated the *extension* phase increases the total packing area in width, int the time domain, and the *selection* process is repeated, either increasing the area of an

[4]The packets within an SDMA layer are greedily allocated through sequentially adding packets until the available capacity has been reached. This packing algorithm can be replaced by a more efficient solution like the Greedy Scheduling Algorithm from [ZDMXCF10].

already scheduled group or adding a new one to the corresponding sub-band. The algorithm terminates in case there is no more space left within the frame for the *extension* phase to increase the packing area for any given SDMA group.

It should be noted that the SDMA groups of different sub-bands are obviously overlapping, which means that the same packets can be allocated in different groups/sub-bands. The *frame* structure used in Algorithm 2 was used for simplification and should provide the necessary functionality as to freeze and unfreeze packets depending on whether or not they have already been scheduled on some sub-band. In other words, once an SDMA group is added to the frame (line 18) the selected packets that fit into the current space must be prohibited from being allocated on another sub-band (group). Only the group which has allocated packets in a previous iteration is allowed to reuse, or unfreeze them for others. Furthermore, the amount of packets which can be packed in the packing area of a SDMA group is determined by the space available in the MAP.

Algorithm 2 The SDMA-OFDMA packing algorithm partitions a frame into different areas each representing the packing area of given SDMA group.

Require:

1: $\mathcal{G}$ - set of SDMA groups over all sub-bands; P - list of packets destined to MSs in G sorted according to utility per slot;
2: ofdmaFrame() - Creates a frame object which handles the frame layout and any packing related features.
3: SCSB - Number of subchannels per sub-band.
4: SB - Total number of sub-bands.
5: **procedure** framePart(G, P)
6: stepSize $\leftarrow \lceil \frac{\text{minSlotSize}(G,P)}{\text{SCSB}} \rceil \times \text{SCSB}$
7: vLimit $\leftarrow \max(\text{initSz} \times \text{SCSB}, \text{stepSize})$
8: usedSpace $\leftarrow [0 \cdots 0_{\text{SB}}]$ ▷ Packed space per sub-band
9: frame $\leftarrow$ ofdmaFrame(P) ▷ Initialize structure representing the frame
10: **while** $vLimit \times \text{SB} + \text{frame.MapSizeSlots}() < \text{frame.Size}$ **do**
11: **for** $G_i^j \in G$ **do** ▷ For each group on every sub-band
12: $G_i^j \leftarrow \text{vLimit} - \text{usedSpace}(j) + \text{frame.getSize}(G_i^j)$ ▷ Set the available size
13: **end for**
14: $J \leftarrow \{1 \cdots \text{SB}\}$;
15: **while** $|J| \neq 0$ **do**
16: $G_i^j \leftarrow \arg\max_{G_i^j} \text{util}(\text{frame} \cup G_i^j), j \in J, G_i^j \in G$
17: **if** $\text{util}(\text{frame} \cup G_i^j) > \text{util}(\text{frame})$ **then** ▷ Ensure gain in terms of total utility
18: frame.set(G_i^j) ▷ Add or update the allocated space for the found group
19: $J \leftarrow J \setminus j$ ▷ Remove all groups of sub-band j
20: usedSpace(j) $\leftarrow$ usedSpace(j) + stepSize
21: **else**
22: $J \leftarrow \emptyset$ ▷ Stop searching if we cannot increase the total utility
23: **end if**
24: **end while**
25: stepSize $\leftarrow \min(\max(\text{frame.freeCols}, 1) \times \text{SCSB}, \text{stepSize})$
26: vLimit $\leftarrow$ vLimit + stepSize ▷ Adjust the vertical limit
27: **end while**
28: **return** (frame)
29: **end procedure**

Example:

Figure 8 illustrates a detailed example of the packing algorithm. As can be seen the extension phase moves a vertical limit for the burst packing area from right

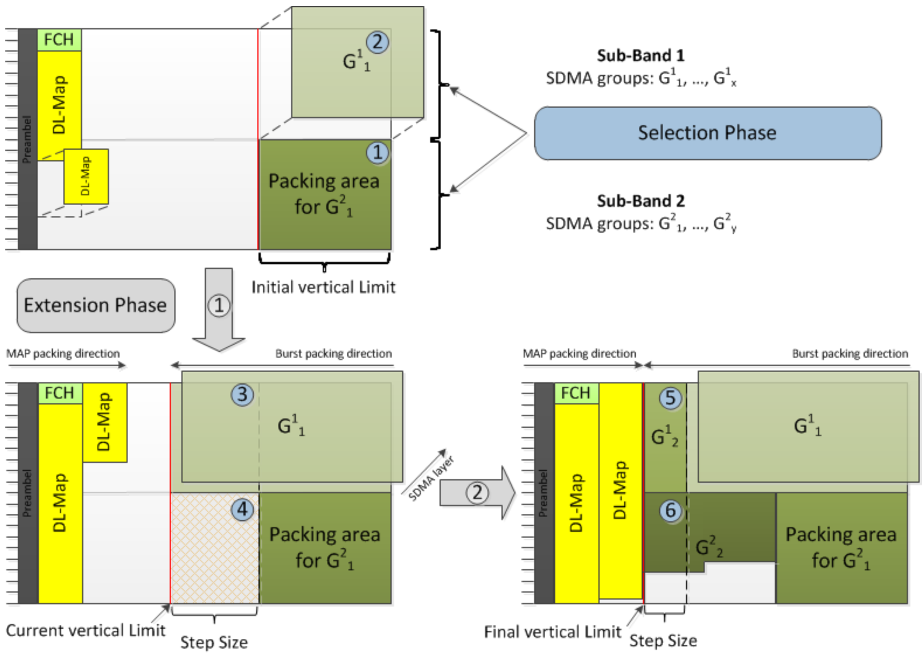

Figure 8: SDMA-OFDMA packing example. The selection phase schedules SDMA groups per sub-band based on the maximum increase of the total utility carried by the frame. In case (4) the additional space is not sufficient to pack additional packets, stays therefore unallocated and can be scheduled during the next extension phase (6). The packing area gradually approaches the MAP area until it becomes a limiting factor for the last added SDMA groups.

to left towards a column wise growing MAP growing from left to right. The initial vertical limit sets the largest available space for the selection phase, after which it is incremented by a predetermined *step size* that eventually reduces to one single column until MAP and packed area cover all columns of the whole frame. In the last step the last scheduled SDMA groups are limited by the lack of available space for the MAP (ref. to group G_2^2 in Figure 8).

Additional details:

The vertical limit was a necessary design choice due to the column wise growing MAP. It ensures that the packing area of all sub-bands grow more evenly in the end so that no single sub-band blocks the MAP to early due to favorable channel conditions. The initial vertical limit as well as the step wise incrementation are an essential part of that process. The *initial* vertical limit should be as large as possible in favor of the frequency-selective scheduling gain, as well as to minimize the eventually needed steps that are necessary for a gradual approach between the MAP and the packed area. The *initial* vertical limit in full columns is calculated as follows:

$$initSz = \left\lceil \frac{(DL_{sl} - 1) \times SC/MSB - Map^*_{sz} \times N}{SC} \times SB \right\rceil, \tag{8}$$

where N is the number of antennas at BS, DL_{sl} the total number of slots (i.e. columns) and SC the total number of subchannels (i.e. rows) in the OFDMA frame. Map^*_{sz} is the predicted MAP size in slots per SDMA layer and SB the number of sub-bands used. Map^*_{sz} was estimated based on the total number of 40 Byte packets that can be merged and signaled into SC many slots on an average MCS. The idea behind this concept is that the initial vertical limit allows the same number of slots for any number of sub-band with MSB being the maximum and that the pessimistically estimated large MAP allows the rest of the frame to be used for an SDMA group. This value is set in Algorithm 2 line 7. The step size should be as small as possible incrementing the packing area by one column per step. In case the packet queue is filled with very large packets a small step size might be to small for additional packets, therefore, it is overridden by the maximum minimum burst size encountered in all MS packet queues rounded to full columns (line 6). As the incrementation of the packet area approaches the map area the step size will be limited by the remaining available space (line 25).

Complexity:

There are K MSs per sub-band each corresponding to a spatial layer, independent of SDMA group affiliation. This means the expression line 16 has to estimate the utility gain of $K \times SB$ many layers. The selection phase is performed SB times and the extension phase at most DL_{sl}. The worst case total number of utility estimations performed is, therefore, $DL_{sl} \times SB[K \times SB] = O(SB^2)$.

4 System-level Simulations

The performance of the proposed scheduling algorithm is analyzed in this work by means of simulations. For this purpose a simulation framework was implemented in Matlab focusing on a cellular multi-user OFDMA-MIMO system on the downlink. Table 1 summarizes the most important parameters.

Parameter	Value
System bandwidth	5/10/20 MHz
Subcarrier bandwidth	10.9375 kHz
FFT size	512/1024/2048
Center frequency	2.5 GHz
Frequency reuse pattern	3x1x1
Transmit power	46 dBm
MS noise density (dBm/Hz)	-167 dBm/Hz
Cell radius	$\sim 288\,m$
WIM scenario	C2 (urban macro-cell, LOS)
Number of antennas at BS (M)	1-5 omni elements separated by half wavelength
MSs' speed	2 km/h
OFDMA frame duration	5 ms
Tx Precoding algorithm	MinMSE
WiMAX permutation scheme	AMC
Packet buffer size	$K \times 12.96$ KiB

Table 1: Simulation Parameters.

All BS are equipped with a linear antenna array of N antenna elements placed half a wavelength apart. Figure 9 depicts the complete cellular network including neighboring cells. Each BS performs SDMA without cooperation, causing inter-cell interference for every MS. The MSs are uniformly distributed within a cell, equipped with single antennas and are associated to each BS depending on a minimum SNR requirement plus an additional outage margin. The transmit power of the BS is adjusted so that the mean SNR of the MSs placed at the cell edge is equal to the mentioned minimum requirement using a single transmit antenna.

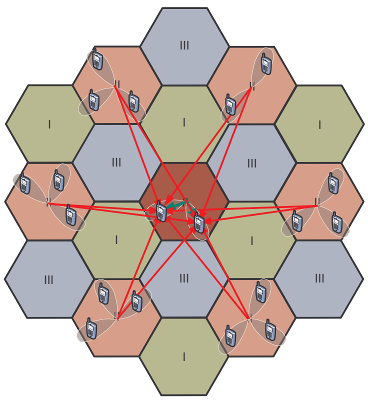

Figure 9: Inter-cell interference is modeled based on a frequency reuse pattern of 3x1x1.

In each simulation an 802.16 OFDMA frame is considered having a duration of 5ms, for which the channel response is assumed to be flat. This assumption is based on low MS mobility and is well within the boundaries of a typical 802.16 use-case, where the coherence time for a carrier frequency of 2.5 GHz and a mobility speed of 2 km/h is roughly 200 ms [AGM07]. The OFDMA frame is divided into DL and UL sub-frames with a 35/12 DL/UL ratio. In order to exploit the frequency-selective-channel and multi-user diversity adjacent sub-carriers are grouped into logical subchannels (AMC). An equal transmit power and the same average noise power is assumed for all sub-carriers.

The resource allocation is delivered in MAP messages at the beginning of each frame and is preceded by the Frame Control Header (FCH). FCH provides the necessary information required to decode the subsequent MAP message, e.g. MAP length and coding. For robustness the MAP is send omni-directionally with QPSK 1/2. The size of the DL-MAP varies depending on the allocated resources on a frame by frame basis. Table 2 depicts the different information elements and their sizes.

Variable	**Value**
F_0	fixed_dl_map_overhead (72 bits)
F_1	dl_map_ie_fixed_overhead (44 bits)
F_2	dl_map_ie_cid_size (16 bits)
SC	num_subcarriers_slot (48)

Table 2: Parameter used for the calculation of the DL-MAP size.

Channel state information is know at the BS and is acquired through UL sounding, and/or through channel reciprocity. For each MS the BS knows the channel transfer coefficients of every 28th subcarrier for every antenna element. Depending on the channel bandwidth of 5/10/20 MHz there are a total of $N \times 15$, $N \times 30$, or $N \times 60$ coefficients per MS, respectively. Furthermore, the BS has full CSI without estimation errors.

4.1 Link Level

The MIMO channel and its coefficients where generated with the WIM simulator of the WINNER-Phase II project (Wireless World Initiative New Radio). An implementation based on the WINNER II model, referred to as WIM, has been implemented and is available under GPL license [KMH07b]. This includes a universal system model capable of adapting to a wide range of mobile communication scenarios, including various propagation models for a frequency range of 2 to 6 GHz [KMH07a].

In Figure 10 a single MIMO link is depicted. The spatial characteristics of the environment are geometrically modeled through clusters of scatters, and are influenced by the antenna fields on receiver and transmitter side. A total of 20 scatters results in 20 sub-paths per cluster, independent of the scenario. The number of clusters varies from 8 to 24 and are scenario dependent. The spatial and temporal characteristics of each link between MSs and BS where modeled according to the LoS/NLoS C1 Suburban macro-cell scenario.

The BS uses the knowledge of the channel response to serve multiple MSs separated in space. For each MS $u (u = 1, \dots, K)$, associated with the BS, the frequency dependent channel response is represented by a vector $H_u =$

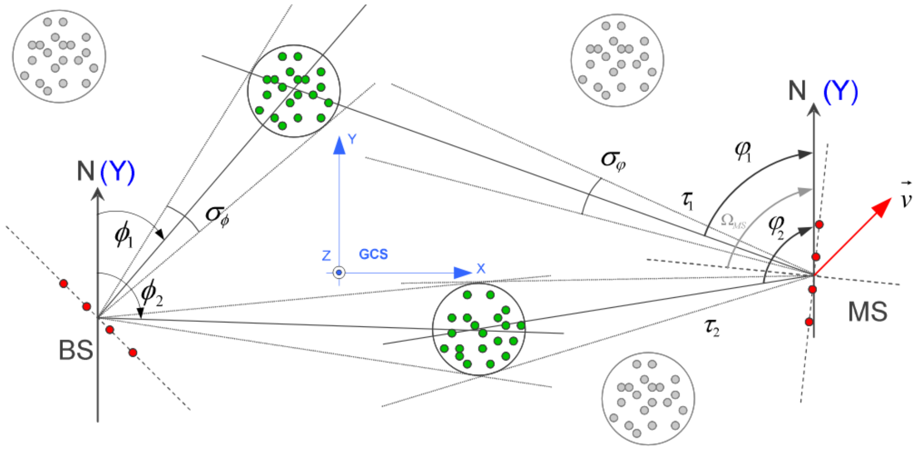

Figure 10: Link Level

$[h_{1,u}, h_{2,u}, \ldots, h_{N,u}]^T$, where each coefficient $h_{i,u}$ holds all spatial correlations and multi-path effects of the channel between the BS antenna element i and the receiver antenna of the $u-$th mobile station as depicted in Figure 11. The BS uses a precoding vector to adjust the transmitted signal in order to mitigate the propagation effects of the channel and to control the interference among MSs within an SDMA group. Accordingly, each MS u within an SDMA group is given a normalized precoding weight vector $W_u = [w_{1,u}, w_{2,u}, \ldots, w_{N,u}]^T$.

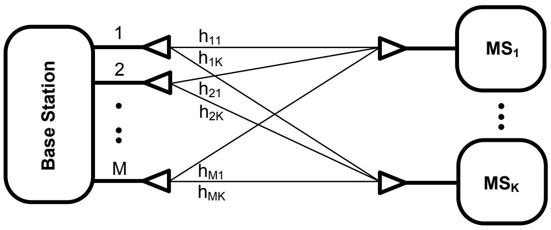

Figure 11: System model: Downlink (DL) of a single BS equipped with an Antenna Array and multiple single-antenna MSs.

In OFDMA the total channel bandwidth is separated into S orthogonal sub-carriers of which, in this case, the channel response of every 28-th sub-carrier is known at the BS, each relationship represented by the index $b = 1, \cdots, \frac{S}{28}$. The effective channel gain at the BS can be computed as $||W_{u,b}^H \cdot H_{u,b}||_2^2$, where the operator $(\cdot)^H$ stands for the complex conjugate transpose of a vector/matrix. Notice, that multiple MSs within an SDMA group sharing the same time-frequency resource cause interference upon each other. The resulting SNIR for a MS u can be computed as [Mac08]:

$$\gamma_{u,b} = \frac{P_{u,b} \cdot ||W_{u,b}^{\mathrm{H}} \cdot H_{u,b}||_2^2}{\sigma^2 + \sum_{v=1,u\neq v}^{M} P_{u,b} \cdot ||W_{v,b}^{\mathrm{H}} \cdot H_{u,b}||_2^2} \tag{9}$$

where $P_{u,b}$ represents the average received power at MS u on sub-carrier b.

Based on the prior defined vector representation of the channel, the BS implements two beamforming techniques. In SDMA mode the sub-streams of multiple MSs are separated through weights given by the *Minimum Mean Square Error* technique (MinMSE) [Gro05]. Whereas, in case of a single MS the weights are simply aligned to the given channel according to the *Maximum Ratio Combining* (MRC) technique also known as transmit beamforming [OC07].

MinMSE works in the following way. For a given sub-carrier b the matrix $\mathbf{H}_b$ contains in each u-th column the channel coefficients ($H_{u,b}$) of the u-th MS. Let $\mathbf{R_{ss}} = \mathbf{H}_b \cdot \mathbf{H}_b^H$ be the autocorrelation matrix of all MSs within an SDMA group, and $\mathbf{R_{nn}}$ the noise correlation matrix. Then, the weight matrix $\mathbf{W_b}$ containing the beamforming weights for every u-th MS, can be computed as:

$$\mathbf{W}_b = (\mathbf{R_{ss}} + \mathbf{R_{nn}})^{-1} \cdot \mathbf{H}_b, \tag{10}$$

where the u-th column in $\mathbf{W}_b$ contains the weights for the u-th MS for the given frequency resource unit b.

For the single MS case the full antenna diversity gain is achieved through applying MRC weights, calculated as:

$$W_{u,b}^{\mathrm{H}} = \frac{H_{u,b}^{\mathrm{H}}}{||H_{u,b}||_2^2}. \tag{11}$$

The achieved SNR of a single MS u is also given by Equation 9, using the MRC beamforming weights and ignoring the intra-cell interference term. The denominator of Equation 11 normalizes the weights to unity so that the average total transmit energy remains the same. The same normalization is performed for the weights given by the MinMSE technique.

4.2 Error Model

The main focus of this work lies on the performance of the SDMA-OFDMA MAC scheduler in terms of sevice throughput, which depends on the successfully decoded data on receiver side. A common approach is to predict the decoded PDU error rate given a set of SNIR values for each sub-carrier used by the PDU, and its MCS [For08]. For an SDMA-OFDMA system there are three main factors that contribute to the actual SNIR. Namely, the accuracy of the estimated channel state (CSI), the interference between spatially multiplexed MSs within the cell (intra-cell), and the interference of neighboring cells (inter-cell).

The accuracy of the estimated CSI heavily depends on the applied technique. Since these techniques lie outside of the scope of this work, as mentioned, perfect CSI is assumed. As a result, the intra-cell interference of MSs can be accurately computed by equation 9. This leaves the inter-cell interference as main contributer to the error between the assumed SNIR at the BS and the actual SNIR.

The effects of the inter-cell interference where modeled as follows. Based on a typical 3x1x1 deployment, as depicted in Figure 9, the tier-one interfering cells where simulated, and performed the same SDMA-OFDMA MAC scheduling as the center cell. Consequently, creating random interference on the target cell unknown prior to transmission. For each burst packed by the center cell in an

Percentage	MPDU Size (bytes)
18.89%	40
12.09%	1500
6.14%	62
4.67%	1420
4.60%	52
53.61%	uniform(40,1500)

Table 3: MPDU Size Distribution

OFDMA frame the real SNIR for each used sub-carrier, including the actual intra-cell as well as inter-cell interference, is computed as:

$$\tilde{\gamma}_{u,b} = \frac{P_{u,b} \cdot ||W_{u,b}^{\mathrm{H}} \cdot H_{u,b}||_2^2}{\delta_{u,b} + \sum_{c=1}^{\#\mathrm{BSIntf}} P_{u_c,b} \cdot ||H_{u_c,b}||_2^2}, \tag{12}$$

where $\delta_{u,b}$ represents the intra-cell interference plus noise given by the denominator term of Equation 9 and #BSIntf is the number of tier-one interfering BSs. The set of obtained SNIR values is collapsed into an equivalent SNR over an AWGN channel using the EESM mapping. Given the equivalent SNR and the used MCS, a block error rate is obtained from a pre-computed table look up (BuER).

Finally, this error is included into the system performance evaluation through multiplying the utility/bytes of each carried burst by (1-BuER) to account for retransmissions and thus getting the service throughput.

4.3 Traffic Model

The performance of a SDMA system, not only depends on the applied MIMO techniques, but on the amount of data traffic as well. There is no need for spatial multiplexing of multiple MSs in case they have no data for transmission. Nevertheless, in a realistic environment situations like this occur, where the available data traffic is not equally balanced amongst MSs.

The model used within this work assumes an unbalanced situation where 50% of the MSs generate 80% of the traffic at MAC layer. Furthermore, the packets where generated with random sizes according to the packet distribution given in Table 3. This distribution was derived from a data collection campaign by SPRINT[5]. In favor of SDMA, $K \times 12.95$ KB of data are generated and assigned to different MSs to ensure a certain degree of minimum traffic, unless otherwise stated. This value is based on the total frame capacity achievable using the highest MCS at a 10 MHz bandwidth, minus one DL-slot for map-overhead.

4.4 Quality-of-Service Scheduler Model

The QoS scheduler controls the QoS requirements of every users traffic, as well as maintaining fairness across data flows. Its task is to select a candidate list of packets for each flows buffer to be transmitted in the next DL sub-frame. In addition, the QoS scheduler tags each individual packet P_i with a utility value u_i. In order to achieve fairness across flows the *Proportional Fair Utility (PF)*

[5] https://research.sprintlabs.com/packstat/

function was used to derive the utility value u_i for each packet within the buffer for the next transmission.
The Proportional Fair Utility (PF) [LSM08], $u_{\mathrm{pf}}()$ achieves fairness through assigning each packet its size in bytes weighted by the average throughput of prior sessions. This means that, on one hand, efficient packets with a large payload in bytes attain a high priority, thus, maximizing the throughput, and on the other hand are reduced in priority in case the associated flow has a history of high throughput, therefore, allowing less effective packets to be eventually prioritized. The utility U_{packet} assigned to each packet of size B in Bytes is calculated as $U_{packet} = B/T$, where T is an Exponentially Weighted Moving Average (EWMA) of consumed resources in Bytes that have been scheduled for the flow of this packet in previous frames. It is defined as $T(i) = (1-\alpha) \cdot T(i-1) + \alpha \cdot N(i-1)$, where $N(i-1)$ is the actual number of Bytes of the flow in frame $i-1$, with $N(i-1) = 0$ in case no data was send . [6]

[6] The value for α in the EWMA filter is set to 0.2

5 Evaluation

In this section the results of the evaluation are presented. First, in Section 5.1 the quality of the low complexity SNIR prediction algorithm and its margin of error will be evaluated for the single-carrier, and multi-carrier (EESM) case. In Section 5.2 the effects of the proposed SNIR algorithm on system-level performance, without *frequency-selective scheduling* will be evaluated for different SDMA grouping algorithm, whereas the best performing SDMA grouper and its results will be compared to an exhaustive search. Furthermore, in Section 5.3 the option of the proposed scheduler to perform *frequency selective scheduling* with varying degree of granularity and its effects on performance will be analyzed.

5.1 SNIR Prediction

The SNIR Prediction module, described in Section 3.2.3 provides two distinct ways for estimating the SNIR values per subcarrier. The most accurate technique is based on the actual precoding weights used for SDMA. Due to the high computational costs for this method a suboptimal low complex alternative was proposed/included as well. In this section the margin of error of the low complex alternative is evaluated for the single carrier, the multi carrier case (EESM), as well as its performance with different beamforming algorithms.

5.1.1 Single Carrier/Multi Carrier

The reduced computational cost of the SNIR Prediction method comes at a price. Depending on the SDMA group size and the total number of available antennas, the predicted SNIR differs from the actual SNIR. Figure 12 shows the CDF plot of the error between the predicted SNIR and the actual SNIR for all possible SDMA group sizes, and random user combinations, averaged over 101 seeds. Most noticeably, the error is almost non existent for two MSs, but as the group size approaches the maximum possible, the error increases. The reason for this behavior lies in the nature of the prediction method, which only considers the correlation of every MS pair, instead in the context of the whole group.

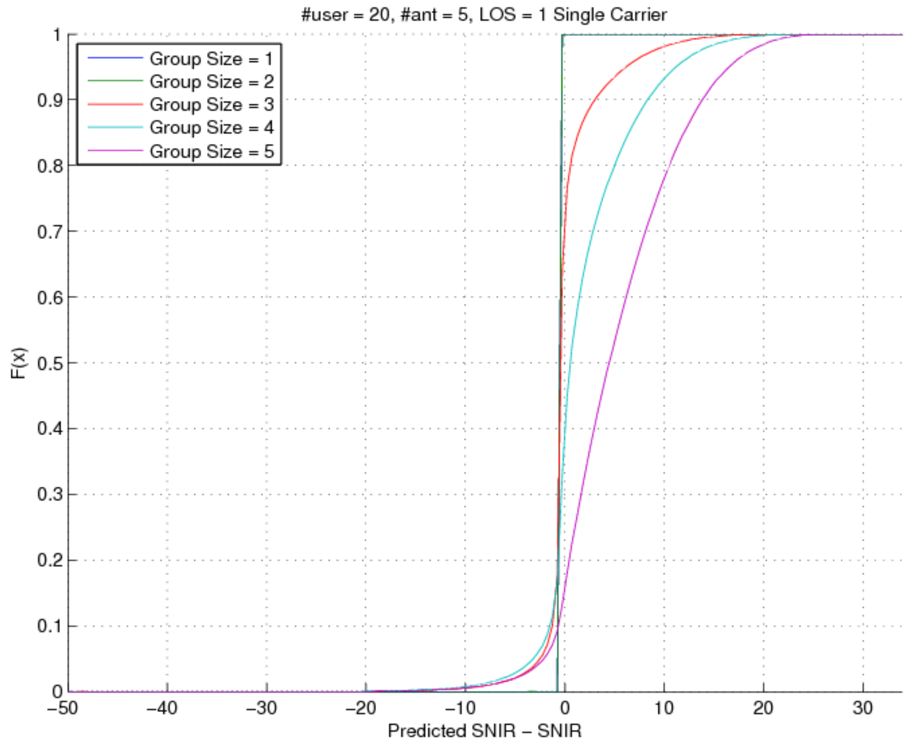

Figure 12: The Cumulative Distribution Function of the error between predicted SNIR and actual SNIR for all possible group sizes for $N = 5$.

In Figure 13 the mean error of the predictor is depicted for different number of antennas. Similar to the CDF plot it shows that the error increases as it reaches the maximum possible group size, which seems valid for any number of antennas. For only two MSs within a group the predictor underestimates the SNIR, but only with one db on average which could be a result of rounding errors. In general the predictor seldomly underestimates the per subcarrier SNIR, but far more often, overestimates it, especially as the number of antennas grows. What is also noteworthy is the fact that for all sizes of N the SNIR for each MS within an SDMA group, containing half the amount of MSs as there are degrees of freedom (N), can be accurately estimated.

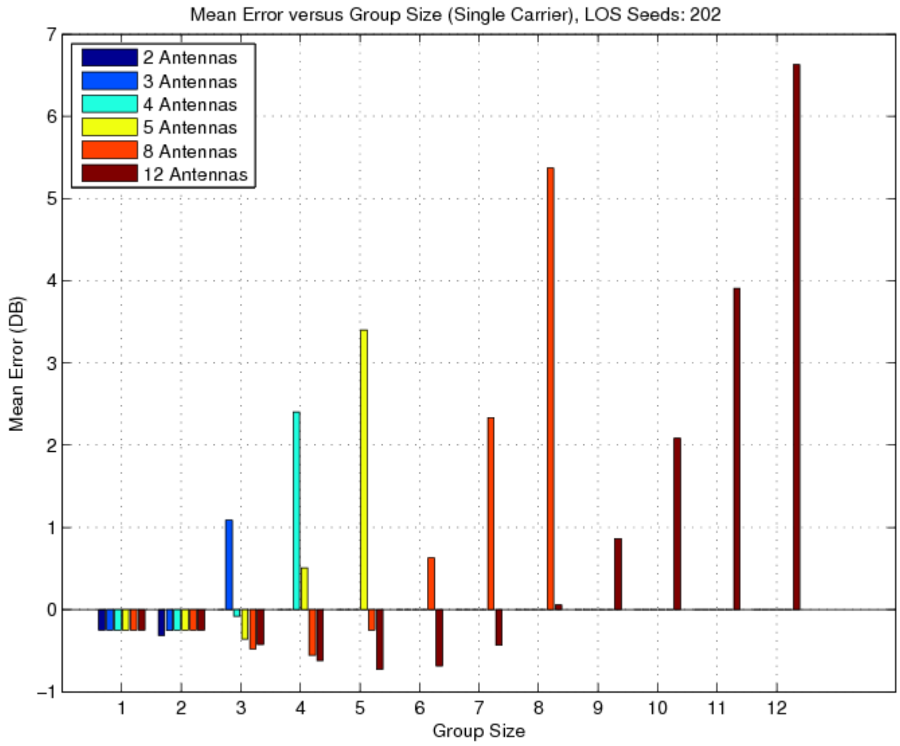

Figure 13: Mean SNIR prediction Error per subcarrier, for various number of antennas and group sizes. The number of users where 20 for 25 antennas, and 15, 14 for antennas 8 and 12 respectively.

As described in Section 3.2.2, the SDMA grouping metric collapses multiple SNIR values into a scalar using the EESM compression method. Figure 14 depicts the SNIR prediction error for 30 SNIR values spread over a bandwidth of 10 MHz collapsed into one SNIR representation. As can be seen the error slightly increases by roughly one dB, in cases where the SDMA group size reaches N, in comparison to the single carrier case. Overall, the average error behaves in a similar fashion with multiple subcarriers as with a single carrier.

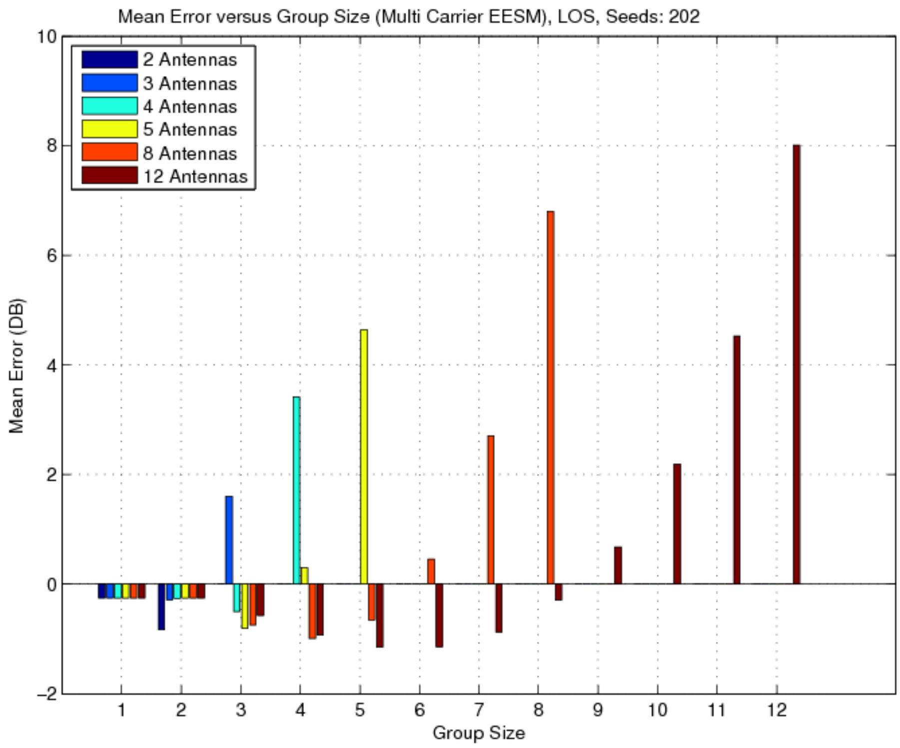

Figure 14: Mean SNIR prediction Error for 30 predicted subcarriers collapsed into one scalar, various number of antennas and group sizes. The number of users where 20 for 25 antennas, and 15, 14 for antennas 8 and 12 respectively.

5.1.2 Influence of Beamforming techniques

The SNIR prediction method proposed is based on the channel correlation amongst MSs. It is based on the assumption that the correlation reflects the amount of interference that needs to be dealt with by any MIMO technique used. In Figure 15 the mean error between the estimated SNIR based on precoding and the predicted SNIR is depicted for common linear MIMO techniques. These include the Maximum SIR, Minimum Mean Square Error, Minimum Variance Distortion-less Response [Gro05],and the Zero-Forcer [KBF$^+$er]. The first three techniques are know to be optimal in the sense that they optimize the SIR. As can be seen they all lead to the same mean error which increases as the group size approaches N. Only the Zero-Forcer technique holds a higher mean error as it sacrifice more in order to enforce the nulls towards each interferer. As the Figure shows, the proposed method is fairly independent of the MIMO techniques applied and even holds a lower error for techniques considered to be optimal.

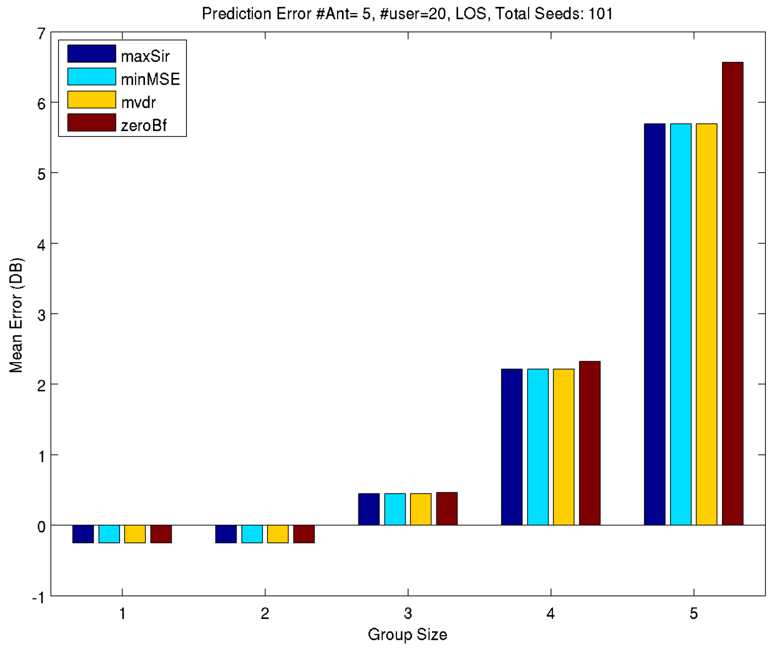

Figure 15: Mean SNIR prediction error for various MIMO techniques.

5.2 SDMA System Performance

The greatest problem faced by any practical SDMA-OFDMA scheduling solution is the limiting factor of complexity. As the degrees of freedom increase, with multiple antennas and multiple subcarriers, system performance can be further enhanced at the expense of exponentially-increasing scheduling complexity. This means that the main objective for any scheduler should be to maximize performance whilst minimizing the involved computational costs.

In Section 5.2.1 the benefits of the proposed low complex SNIR Prediction method will be analyzed in terms of performance versus complexity. It will be shown that this method has the capabilities of reducing the computational costs of a highly complex SDMA grouper, like the CBA, to that of the most simple state-of-the-art SDMA grouper, the FFA. The performance is measured in terms of system throughput, whereas the measurement of complexity was chosen as the total number of executed floating point operations (FLOP) [Mac08]. Next, the different SDMA grouping algorithms will be evaluated more thoroughly in Section 5.2.2 using the SNIR Prediction method. Finally, in Section 5.2.3 the overall performance of the low complex SNIR Prediction method combined with the power of the CBA SDMA grouper will be compared against an exhaustive search grouper using the precoding based SNIR estimation under different system parameters.

5.2.1 SDMA Metric

One of the main contributor to the complexity of SDMA grouping is the grouping metric which is used to compare different SDMA groups. A common metric is the group capacity metric based on the achievable SNIR values per MS. The most accurate SNIR estimation is given by the actual precoding weights. Its computational cost is rather high, therefore, low complex alternatives are promising techniques to reduce the complexity. In the following the two SNIR estimation methods where compared: i) the SNIR Precoding, and ii) the SNIR Prediction from Section 3.2.3.

Figure 16 illustrates, for a scenario with $N = 5$ antennas at the BS and $K = 20$ MSs within the cell, the achieved service throughput for different SDMA grouping algorithms (CBA, BFA, FFA) and the two methods for calculating the SNIR. Even though the SNIR Prediction holds a certain margin of error (Section 5.1), it has only a small impact on performance. The CBA and BFA both suffer a negligible loss in throughput. The FFA is even better with the SNIR Prediction method, than with the precoding derived SNIR. A simple explanation would be that since the SNIR Prediction overestimates the SNIR, which results in malformed but larger SDMA groups, the additional processing step performed by the AMC module creates an unexpected bonus. The FFA is very simple, so the additional logic of the AMC post processing seems to contribute to the performance.

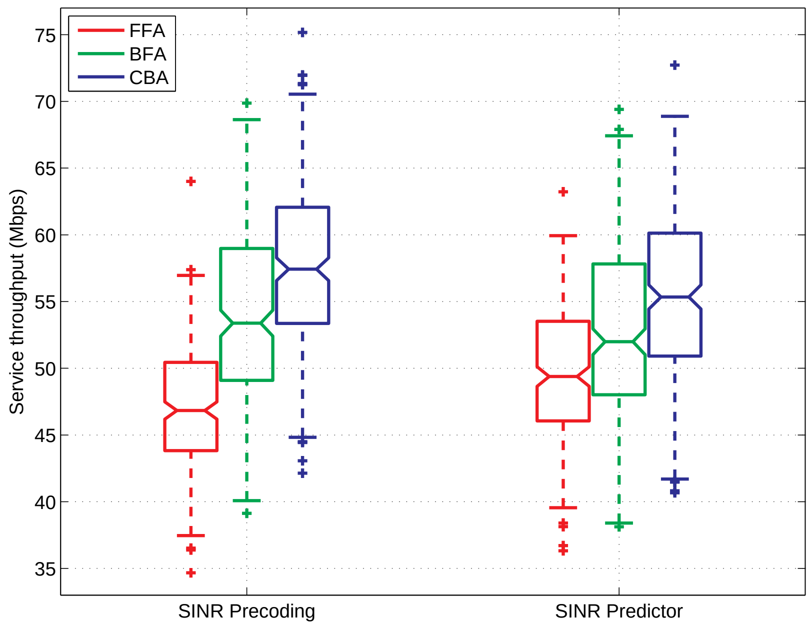

Figure 16: Service throughput.

In order to evaluate the reduction in complexity by the SNIR Prediction in comparison to the precoding based method Figure 17 depicts the number of overall required FLOPs [Mac08]. These values include the initial costs for generating the spatial correlations between every MS pair, the number of reevaluations of SDMA groups, as well as the precoding calculations of the AMC module if needed. As clearly observed in the figure, the reduction in the number of FLOPs achieved by the SNIR Prediction is very significant for all grouping algorithms. In case of the most complex grouper (CBA) the total number of FLOPs is reduced by almost two orders of magnitude. Note, that the computational cost for the CBA is even cheaper, using the SNIR Prediction, than the simple FFA using the precoding based method. Overall a low complex metric like the SNIR Prediction is able to compensate the additional computational costs of a complex grouper, such as the CBA, to a significant degree.

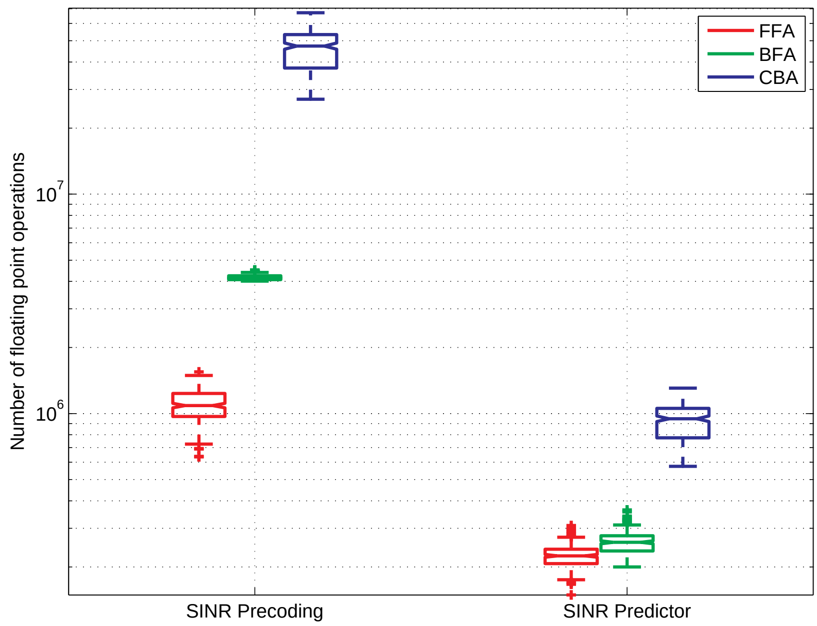

Figure 17: Number of floating point operations.

5.2.2 SDMA Grouping Algorithm

As the previous section has shown, the low complex per-MS SNIR computation method used to derive the group capacity metric can reduce the complexity of an SDMA-OFDMA scheduler significantly, for any SDMA grouping algorithm. In this section the performance/complexity trade-off as a result of the applied SDMA grouping algorithms is investigated further.

Based on the results of the prior section the low complex SNIR Prediction metric is applied here, as it has proven its usefulness. Therefore, all the SDMA grouping algorithms use the same low-complex comparison metric, and the only differences in computational costs will be governed by the number of group comparisons performed.

The key value for any SDMA grouper and its complexity is the number of available MSs K. As the BFA has a worst-case complexity of $O(K^2)$, the CBA comes with $O(K^4)$ [ZMDMXC12]. Only a higher performance would sanction such a burden. Figure 18 illustrates the service throughput of the algorithms under study, again with $N = 5$ antennas, and different numbers of

MSs $K = 5, 10, 20, 30$. As can be seen the CBA justifies its complexity through the achieved performance gain for any number of MSs, followed by the BFA and FFA, respectively. Even at $K = 30$ the CBA still holds a gain of nearly 10% in comparison to the BFA. Overall it can be seen that the gain increases for all groupers as the number of MSs increases, where it seems to saturated at $K = 30$. The CBA not only gains more from more MSs, but also seems to be able to improve its gain better at $K = 30$ than the BFA/FFA can.

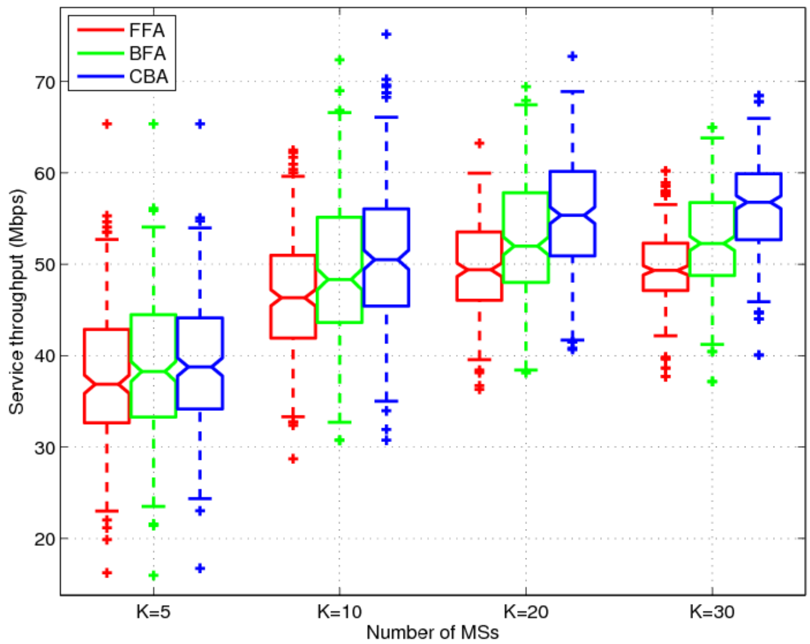

Figure 18: Service throughput.

Similar as in the prior section, the computational complexity of the different SDMA grouping algorithms were evaluated in terms of FLOPs, depicted in Figure 19. As could be expected the CBA is the most expensive algorithm, followed by the BFA and FFA, with its computational costs increasing as the number of MSs increase. The most dominant reason being the number of evaluated SDMA groups, which is shown in Figure 20. The performance gain of each grouping algorithm is a direct consequents of the total number of evaluated SDMA groups, independent of the used grouping metric. Overall, the BFA/CBA provide a complementary performance versus complexity trade-off, both performing better than the FFA. The CBA would be the best choice which maximizes the overall performance, but it comes with the highest computational cost. The FFA would be the least computational intensive algorithm, with the worst performance, whereas the BFA would be a good mix of both worlds.

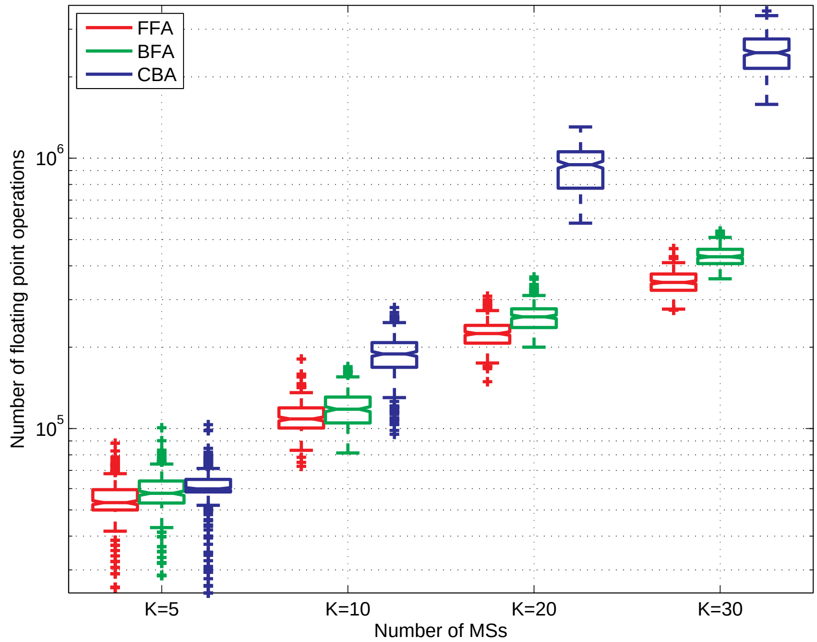

Figure 19: Number of floating point operations.

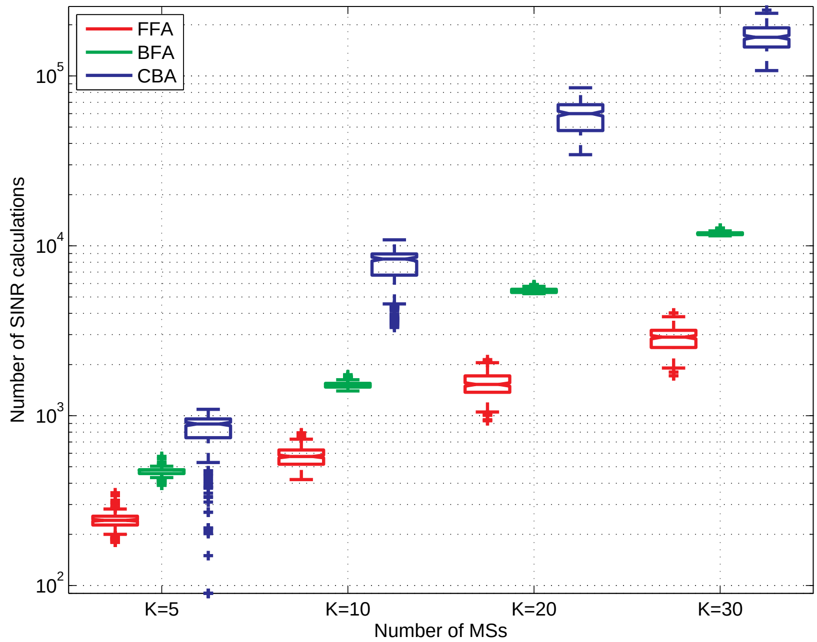

Figure 20: Number of evaluated SDMA groups.

5.2.3 Overall Performance

The last two sections have shown that the SNIR Prediction method grants a substantial reduction in computational costs, and the CBA offers the highest performance in throughput. In this section the performance of the combined strength of both is compared with three alternative algorithms:

1.) OPT+SDMA: This algorithm solves the same problems, but uses exhaustive search for SDMA grouping, and perfect knowledge of the per-MS SNIR used for SDMA group evaluation, as well as brute force to maximize the utility for the OFDMA frame.

2.) sGSA+TxBF: In this case the number of antennas are used purely for beamforming to boost the signal of each MS in a single time-frequency region, without the interference of additional MSs (no SDMA).

3.) GSA: The GSA is an OFDMA scheduling solution [ZDMXCF10] that uses only the time-frequency dimensions for allocations and is, therefore, only applicable for $N = 1$.

The combination of the low complex SNIR Prediction method and the CBA algorithm will be referred to as sGSA. The QoS scheduler uses the Proportional Fair utility to guarantee fairness amongst MSs.

In general, the performance of an SDMA-OFDMA system scales with the number of available antennas at the BS. But in particular, it is the amount of offered load that effects the performance just as strongly (Section 4.3). Both of these aspects will be evaluated next.

Impact of Number of Antennas

The number of antennas at the BS increase the transmit diversity, which translates into a SNR improvement for traditional transmit beamforming, and increases the degrees of freedom for SDMA. Figure 21 depicts the achieved cell throughput in the LOS scenario with $K = 10$ active MSs for the various algorithms, and different N. As could be expected the sGSA+SDMA solution outperforms sGSA+TxBf, and comes very close to Opt+SDMA.

The sGSA+TxBf performance improves only slightly with additional antennas. This is due to the fact that the cell deployment is adequate for $N = 1$ antennas, therefore, additional antenna diversity quickly reaches the highest, or more efficient MCS. Also, with $K = 10$ and the PF utility the Multi-User Diversity gain is already high enough to compensate the effects of unfavorable channel conditions of those few MSs suffering under it.

Now, in case of sGSA+SDMA the additional antenna diversity leads to a more dramatic gain in system performance. Even in comparison to the optimal Opt+SDMA solution, the suboptimal sGSA+SDMA performance very well. Surely, the difference seems marginal and barely increase with additional antennas, but it would be wrong to assume this is the general case. Since, with $K = 10$ and a maximum of $N = 5$ the search space for optimal SDMA groups is fairly limited. For larger K the computational requirements quickly grows, in which case it should be expected that the difference in performance of the sGSA+SDMA versus the optimal solution increases. But, no matter how large this difference might become, any other approach would suffer the same fate as the sGSA+SDMA, which still contains a very promising SDMA grouping solution (ref. 5.2.2).

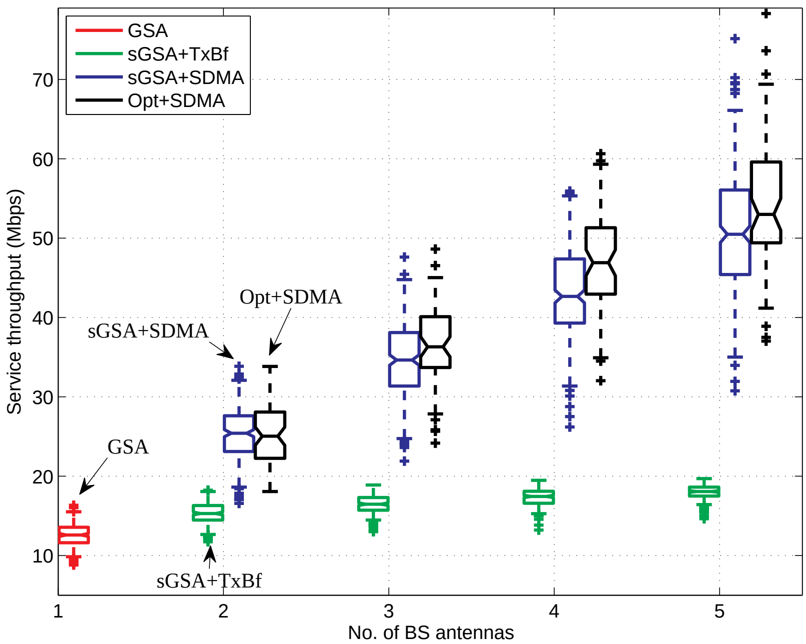

Figure 21: Service throughput.

The results indicate the expected performance on the MAC layer. Since the simulations included not only the MAC overhead due to signaling (DL-MAP), but also draw-backs due to packing, like additional padding as well as unused slots in the frame, a deeper insight into the results can be achieved through considering these effects as well.

In order to grasp the impact of packing inefficiencies, and what could be gained through more efficient ways the spatial gain was evaluated invariant to the used MCS. Figure 22 depicts the spatial multiplexing gain which was calculated as follows: all used slots on each spatial layer were summed up and divided by the total number of slots available in a frame. Note, that a spatial gain of N is not reachable in practice as long as there are omni-directional transmissions for the MAP and MSs that can not be fully separated in space. Overall, the SDMA solutions (sGSA+SDMA, Opt+SDMA) deliver a linear spatial gain slightly below N. For example, the spatial gain of the sGSA+SDMA for $N = 5$ is below ca. 3.2 in 75% of all cases. The Opt+SDMA, which can be assumed to be optimal in the sense of SDMA grouping, offers 25% of all cases a maximum gain of 3, and up to 3.8 for 75%. As these values are very close to N, it can be concluded that for small $N \leq 5$ there is only little room left to gain by applying a more pack efficient solution, independent of the SDMA grouping problem.

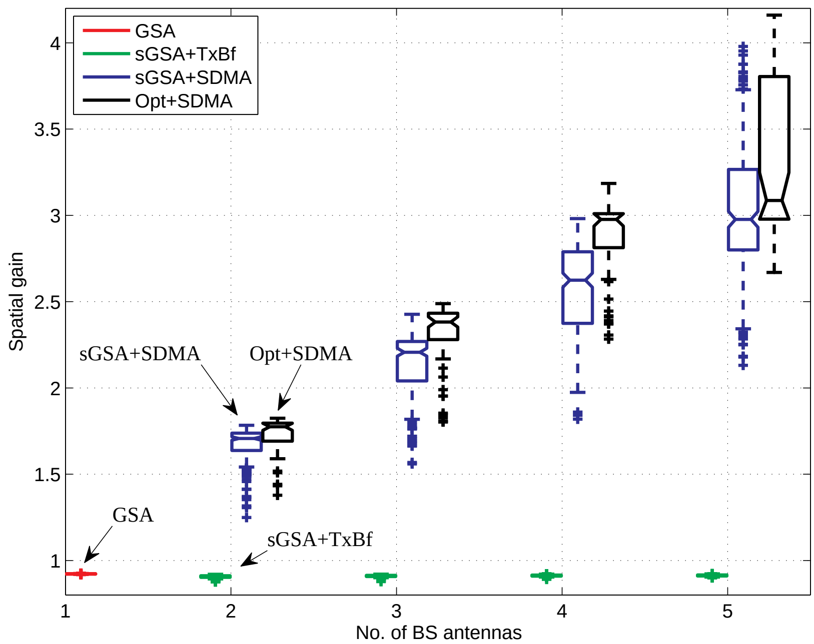

Figure 22: Spatial gain.

Finally, an overview of the needed space for the signaling overhead is given in Figure 23. The algorithms using SDMA, obviously, fabricate a higher signaling overhead that grows with N. The more spatial resources are available the more MSs can be packed, and must be referenced. Note how the GSA has a higher overhead than the sGSA+TxBF, this is a result of the heavy use of burst concatenation for TxBF since the MSs where forced to span the whole frequency block more consecutive space per MS was available.

The highest overhead was reached by the sGSA+SDMA, which allocates up to 7.6% of the total frame space for the DL-MAP when using 5 antennas at the BS. In order to assess the potential gain of a more efficient signaling approach, let us assume the DL-MAP could be compressed to half its size. This would translate in a throughput increase of 0.3 to 2 Mbps for 1 and 5 antennas respectively. It can therefore be conclude that the benefit from a more efficient MAP signaling is much higher for a SDMA scheme with a large number of antennas, but negligible for only a few.

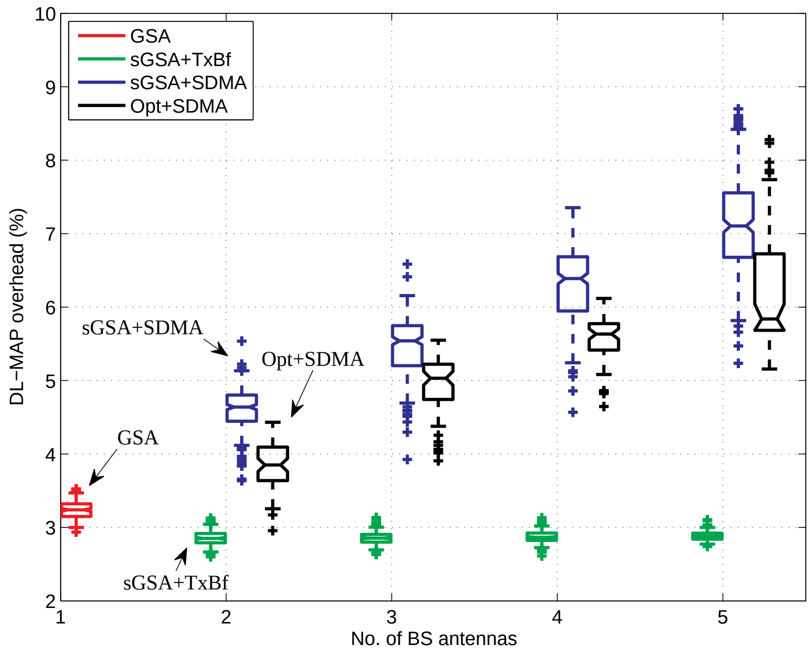

Figure 23: DL-MAP overhead.

Impact of Offered Load

In order to exploit the additional spatial resources to their full extent the quantity of data traffic is crucial, which is effected by the total number of MSs, as well as the amount of data at hand per-MS. The higher the number of MSs the higher the probability of finding spatially orthogonal channels, which improves the quality of the average SDMA resource. Additionally, the more data traffic is available the more of these resources can be employed.

For this purpose the system throughput was evaluated for a fixed $K = 10$ MSs with different per-MS buffer sizes (Figure 24), and for a fixed buffer size of $B = 12.66$ KB with different numbers of MSs (Figure 25). The algorithms included for evaluation are the GSA with $N = 1$, sGSA+TxBf with $N = 5$, and the sGSA+SDMA for $N = 2$ to $N = 5$ antennas at the BS.

As can be seen the performance of all algorithms is effected most by the per-MS buffer size, than by the total number of MSs. Obviously, additional resources increase capacity, but the extent to its potential output can only be achieved by an appropriate input. Most noticeably the capacity improvement through diversity techniques like the sGSA+TxBf with $N = 5$ can be easily achieved through spatial multiplexing with only $N = 2$ (sGSA+SDMA). Looking at Figure 25, it can be seen that increasing the number of MSs in the cell also increases capacity, with the greatest gain already achieved with K twice the number of antennas N, which beyond that quickly saturates for any N.

It is also interesting to see how the number of MSs effects the sGSA+TxBf with $N = 5$ in comparison to sGSA+SDMA with $N = 2$. From the available buffer size of $B = 12.66$ KB both algorithms have a well saturated capacity in conclusion to Figure 24. Nevertheless, it can be seen that the additional user diversity improves the sGSA+SDMA capacity, unlike is the case for the sGSA+TxBf, which is due to the increased probability of finding spatially uncorrelated MSs.

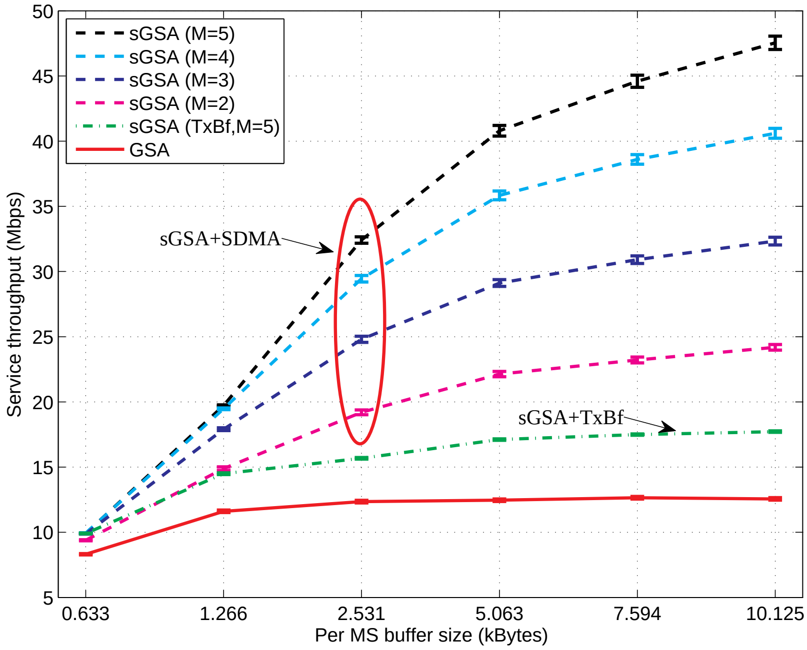

Figure 24: Impact of buffer size.

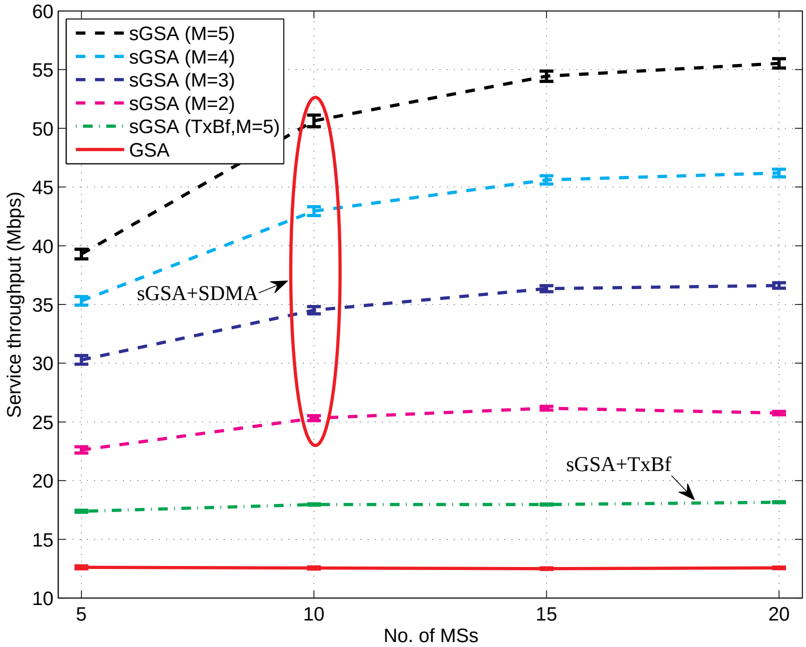

Figure 25: Impact of number of MSs.

5.3 Frequency Selective Scheduling

In this section the impact of a frequency-selective channel will be evaluated, in terms of achievable performance through the use of Frequency-Selective Scheduling (FSS) performed by the proposed SDMA-OFDMA scheduling solution. Due to frequency selectivity the channel transfer matrix is different for different subcarriers, which directly influences the spatial compatibility of different MSs within an SDMA group. Therefore, to ensure that the SDMA grouping algorithm estimates the underlying channel perfectly the SNIR is estimated using the actual precoding. The SDMA grouping is performed on each sub-band by the BFA. The total number of sub-bands, dictate the granularity of the frequency resource.

In the following section the number of sub-bands will be varied from one to six in the interest of illustrating how an increased granularity in the frequency domain could improve performance. The range of correlated frequencies over which a channel can be considered flat depends on the encountered channel conditions. Therefore, the impact of changes in bandwidth will be evaluated for fixed channel conditions, as well as how a fixed bandwidth is affected by changes in channel conditions (LOS versus NLOS). At the end, a final evaluation is made in how far multi-user diversity complements the performance gain of FSS.

5.3.1 Influence of Bandwidth

Frequency selective, or flat fading is not a property of the channel alone, it is a result of the relationship between the bandwidth BW and the channel's coherence bandwidth BW_C, which is reciprocal to the multi-path delay spread [TV05]. The WIM C2 LOS (urban macro-cell) scenario has a delay spread of 41 ns, that translates into a coherence bandwidth of roughly 12 MHz.

In Figure 26 the channel bandwidth of 5, 10, and 20 MHz where evaluated for different number of BS antennas and a fixed number of MSs $K = 12$. It depicts for each parameter pair (BW, N) the service throughput for an increasing number of sub-bands from one to six. As could be expected there is no FSS gain when using a channel bandwidth, e.g 5 MHz, that is substantially smaller than the coherence bandwidth of 12 MHz. Even worse, increasing the number of sub-bands can have a negative effect on throughput for large number of antennas. In case of 8 antennas and 6 sub-bands the throughput decreases by one-fifth compared to the baseline of one sub-band. Similar effects are seen for a bandwidth of 10 MHz, even though there is a small FSS gain of +7.2% using 2 antennas, the FSS gain decreases with additional antennas and increasing granularity until an actual loss of up to −4.1% occurs. Only a bandwidth of 20 MHz, which is well beyond the coherence bandwidth of the channel, can achieve a very small FSS gain with 8 antennas and 6 sub-bands. This severe loss of throughput when increasing the FSS granularity can only be explained by the increasing MAC layer overhead, i.e. increased MAP signaling overhead.

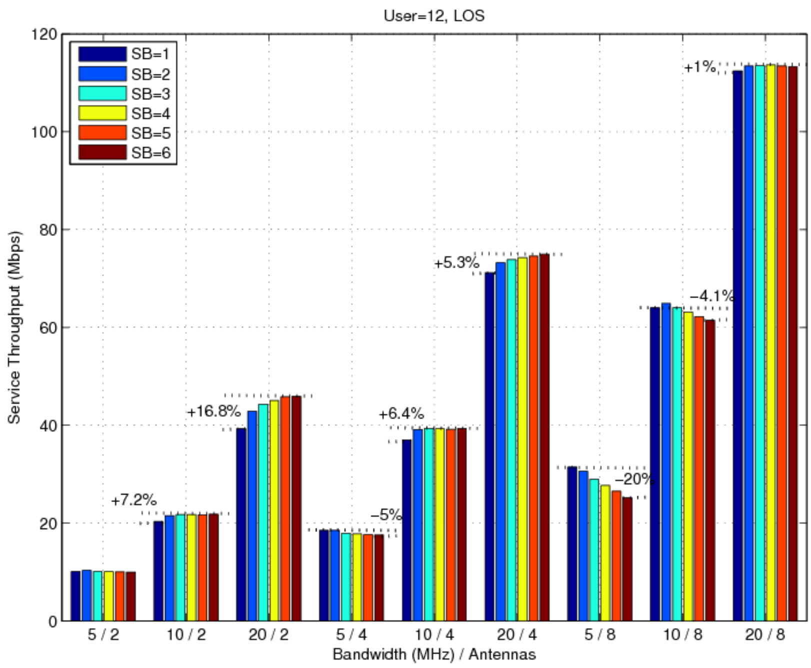

Figure 26: Impact of Bandwidth on Throughput.

Figure 26 depicts the overall DL-MAP overhead in terms of allocated frame space. The substantial performance loss for a bandwidth of 5 MHz and $N =$ 8 antennas and 6 sub-bands seems to be a direct consequence of a two-third increase in signaling overhead.

In conclusion, the highest FSS gain of 16.8% was achieved with a high granularity of 6 sub-bands, only 2 antennas and a bandwidth of 20 MHz at roughly twice the coherence bandwidth of the channel. But, there is a substantial drawback due to increased signaling overhead as the number of additional SDMA resources increases. This means that systems, with bandwidths close to or below the coherence bandwidth and/or a high number of antennas, will forfeit baseline performance through increasing FSS granularity.

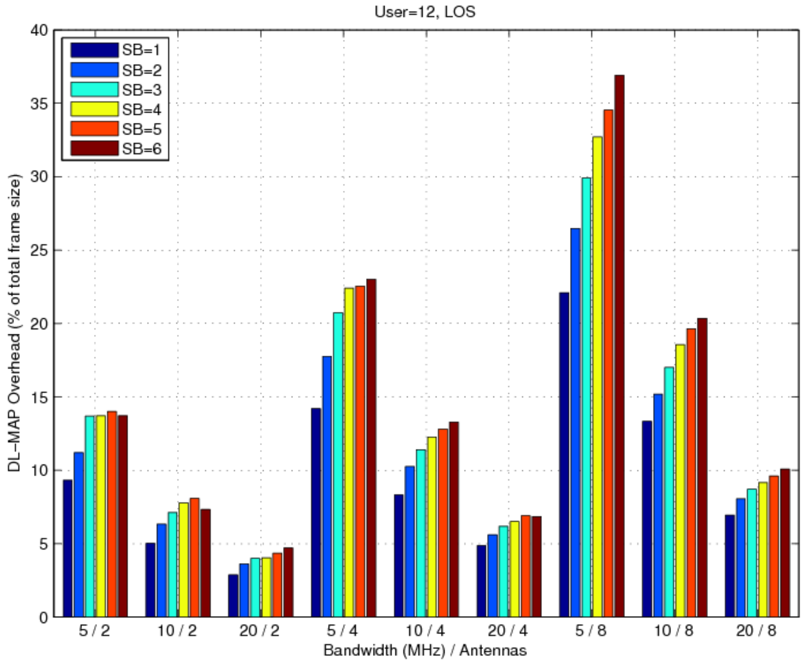

Figure 27: Impact of Bandwidth on DL-MAP Overhead.

5.3.2 LOS/NLOS

In Figure 28 the FSS gain is compared between different channel conditions of an LOS and NLOS environment. It should be noted that these conditions are not simply limited to changes in the multi-path delay spread, but also include things like angular spread, scatters and so on, which directly influences MS correlation. From the Figure 28 it can be seen that for 4 antennas the gain from FSS slightly increases from 6.4% (LOS) to 8.5% (NLOS). This is an even higher gain than was achieved by increasing the bandwidth from 10 to 20 MHz within the LOS environment. This was expected, since the coherence bandwidth of the NLOS environment is roughly 5 times smaller than the 10 MHz bandwidth, in contrast to the 12 MHz coherence bandwidth of the LOS environment.

Similar to earlier results, the signaling overhead increases as the number of BS antennas increase and there is a trade-off between FSS granularity and performance. The NLOS situation does not prevent this trade-off from occurring, even though it seems to be less severe.

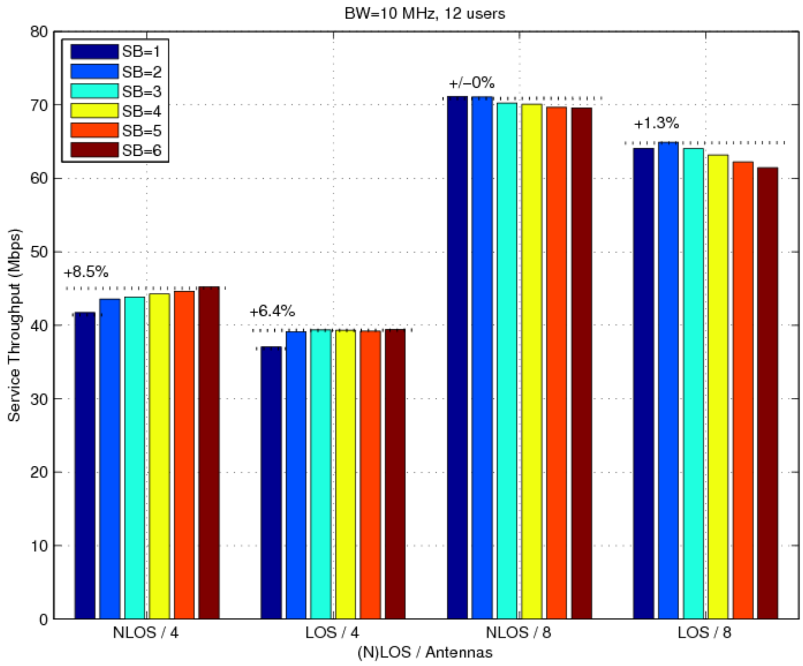

Figure 28: Impact of LOS

5.3.3 Multi-User Diversity

Finally, the impact of available MSs, K, on the achievable FSS gain will be evaluated for the LOS scenario, using a fixed 10 MHz bandwidth and multiple BS antennas. Figure 29 depicts the results. Overall, increasing the FSS granularity increases signaling overhead, and results in the same trade-off as has been seen before. The increasing number of MSs increases the gain from FSS only slightly. E.g. with $N = 8$ the gain can be increased from only 1.3% to 5.4%, when having a network of 36 instead of 12 MSs and using at most 2 sub-bands. Additionally it can be seen that in some cases the capacity increase through multi-user diversity is substantially higher than the FSS gain. For example, in case of 8 antennas, 12 MSs with the best performing number of sub-bands could not reach the same capacity as 24 MSs without FSS (1-sub-band).

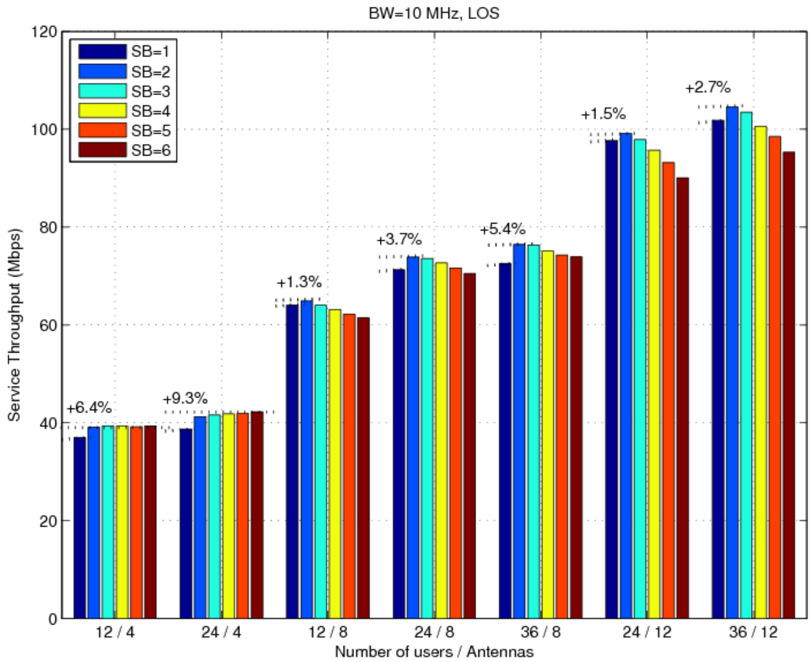

Figure 29: Impact of User Number.

6 Conclusion

6.1 Summary

In this work a generic SDMA-OFDMA MAC scheduler, that integrates current state-of-the-art suboptimal SDMA grouping strategies with varying complexity, and a low complex SNIR predictor have been proposed. Additionally the proposed solution is capable of frequency selective scheduling through splitting the entire frequency band into multiple sub-bands upon which SDMA groups are generated and the most favorable ones scheduled. The algorithm was evaluated in terms of computational cost and performance through system-level simulation, based on the real world use-case of the WiMAX IEEE 802.16-2009 standard.

The evaluation consisted of three main parts: analysis of the performance of the proposed low complexity SNIR prediction algorithm, investigation of the system level performance of the proposed solution without frequency selective scheduling, and finally evaluation of a possible gain provided by frequency selective scheduling. The SNIR prediction algorithm and its margin of error where evaluated for single-carrier, multi-carrier systems, and different SDMA-MIMO techniques. In general the SNIR prediction overestimates the actual SNIR as the number of users within an SDMA group approach the number of available antennas. Nevertheless, it provides a fair indication of suitable SDMA groups at low computational costs. The SNIR prediction algorithm and its benefits on system level where analyzed and compared using different SDMA grouping algorithms. It was shown that the corresponding reduction in complexity, in some cases, reached an improvement of two orders of magnitude. It was shown that a complex grouper like the CBA[ZMDMXC12], in combination with the SNIR prediction algorithm, could reduce its computational cost to the same level comparable to that of the very simple First Fit Algorithm, whilst providing higher capacity gains than the more complex Best Fit Algorithm. Based on these observations the overall performance of the low complex SNIR Prediction algorithm combined with the power of the CBA SDMA grouper where compared against an exhaustive search grouper using the precoding based SNIR estimation under different system parameters. The results show that the proposed OFDMA-SDMA MAC scheduler in combination with the CBA grouper and the low complex SNIR prediction algorithm, allow the highest gains in capacity compared to other SDMA grouping algorithms. As the proposed SDMA-OFDMA MAC scheduler neglects the frequency dimension in order to turn a complex three dimensional allocation into a two dimensional one, the final evaluation within this work focused on the possible gain of frequency-selective scheduling. By explicitly simulating the MAC layer overhead (i.e. DL-MAP), which is required to signal every packed data burst in the OFDMA frame, it was shown that besides the complexity versus performance trade-off dominated by the SDMA grouping problem, there exists an additional trade-off between performance and signaling overhead that can invalidate the advantages of frequency-selective scheduling. However, under specific conditions, like reduced antenna diversity, or a large channel bandwidth, a significant gain of up to 16.8% was achieved from FSS.

6.2 Future Work

There are a number of possibilities for practical application that arise through the combination of MIMO and OFDMA techniques that are yet to be explored. The proposed scheduling solution, for example, used a very simple approach for allocating multiple packets within an SDMA layer. More complex packing algorithms, that have been proposed for OFDMA, could be applied. As the performance of simple packing algorithms might be significantly outperformed by more complex ones [Xav08], the combination of a simple SDMA grouper with a complex OFDMA packing algorithm might prove to be an efficient solution.

In case of SDMA techniques, we have seen in this work that a reduction in MAP signaling overhead is a promising direction for further capacity improvements, especially for a large number of antennas. The impact of channel estimation error on system performance, particularly on the FSS gain should be analyzed. One could assume that allocating one SDMA group across the whole frequency spectrum is more robust against estimation errors then performing FSS, if so FSS in SDMA-OFDMA systems could be neglected.

In case of multiple receive antennas, multiple receive antennas at the MS can increase signal robustness as well as allowing multiple data-streams to the same MS. Therefore, it could be interesting to evaluate the possiblities of extending SDMA scheduling techniques to serve not only multiple-users but optionaly multiple-streams as well.

Additionally, multiple receive antennas at the BS could be used for interference reduction amongst cells. A joint scheduling approach shared by all neighboring cells could sacrifice one or more degrees of freedom of each BS for interference alignment through nulling certain directions. In an interference limited system this could prove to hold significant capacity gains.

References

[80209] *IEEE Standard for Local and Metropolitan Area Networks. Part 16: Air Interface for Broadband Wireless Access Systems.* IEEE 802.16-2009 Standard, May 2009

[AGM07] Andrews, Jeffrey G. ; Ghosh, Arunabha ; Muhamed, Rias: *Fundamentals of WiMAX: Understanding Broadband Wireless Networking.* 1. Prentice Hall, 2007. – ISBN 0132225522

[Ala98] Alamouti, Siavash M.: A simple transmit diversity technique for wireless communications. In: *IEEE Journal on Selected Areas in Communications* 16 (1998), S. 1451–1458. http://dx.doi.org/10.1109/49.730453. – DOI 10.1109/49.730453

[Cal04] Calvo, Diego B.: *Fairness Analysis of Wireless Beamforming Schedulers*, Iniversitat Politecnica de Catalunya (UPC), Diss., 2004

[Cos83] Costa, M.: Writing on dirty paper (Corresp.). In: *Information Theory, IEEE Transactions on* 29 (1983), Mai, Nr. 3, 439–441. http://dx.doi.org/10.1109/TIT.1983.1056659. – DOI 10.1109/TIT.1983.1056659. – ISSN 0018–9448

[Dra10] DragonWave: *Take Microwave Spectral Efficiency to a New Level.* http://www.wcai.com/whitepapers-microwave-spectral-efficiency-dragonwave.html. Version: 2010

[EES05] *CINR measurements using the EESM method.* IEEE 802.16 Broadband Wireless Access Working Group. http://www.ieee802.org/16/tge/contrib/C80216e-05_141r3.pdf. Version: March 2005

[Far88] Farebrother, R.W.: *Linear Least Squares Computations (Statistics: A Series of Textbooks and Monographs).* 1. CRC Press, 1988. – ISBN 0824776615

[FN96] Farsakh, C. ; Nossek, J. A.: A real time downlink channel allocation scheme for an SDMA mobile radio system. In: *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications* 3 (1996), Oktober, 1216–1220 vol.3. http://dx.doi.org/10.1109/PIMRC.1996.568477. – DOI 10.1109/PIMRC.1996.568477. ISBN 0–7803–3692–5

[For08] Forum, Wimax: WiMAX System Evaluation Methodology. In: *Forum American Bar Association* (2008), Nr. Version 2.1, S. 1–209

[Gol05] Goldsmith, A.: *Wireless Communications.* Cambridge University Press, 2005 http://books.google.de/books?id=b9GeSQAACAAJ. – ISBN 9780521684545

[GPK09] Gollakota, Shyamnath ; Perli, Samuel D. ; Katabi, Dina: Interference alignment and cancellation. In: *SIGCOMM Comput. Commun. Rev.* 39 (2009), August, Nr. 4, 159–170. http://dx.doi.org/10.1145/1594977.1592588. – DOI 10.1145/1594977.1592588. – ISSN 0146–4833

[Gro05] Gross, Frank: *Smart Antennas for Wireless Communications: With MATLAB (Professional Engineering)*. 1. McGraw-Hill Professional, 2005. – ISBN 007144789X

[HCCW08] Huang, Ching Y. ; Chang, Chieh-Yao ; Chen, Po-Han ; Wu, Ming-Hsien: Performance Evaluation of SDMA Based Mobile WiMAX Systems. In: *International Wireless Communications and Mobile Computing Conference* (2008), August, 1042–1046. http://dx.doi.org/10.1109/IWCMC.2008.181. – DOI 10.1109/IWCMC.2008.181. ISBN 978–1–4244–2201–2

[KBF+er] Kaiser, T. ; Bourdoux, A. ; Fonollosa, J. R. ; Boche, H. ; Andersen, J. B. ; Utschick, W.: *Smart Antennas - State of the Art*. Bd. 3. EURASIP Book Series on Signal Processing and Communications, Hindawi Publishing Corporation, December. – 1 – 876 S.

[KMH07a] Kyösti, Pekka ; Meinilä, Juha ; Hentilä, Lassi: WINNER II Channel Models. In: *IST-4-027756 WINNER II* D1.1.2 V1.1 (2007). http://www.ist-winner.org/WINNER2-Deliverables/D1.1.2v1.1.pdf

[KMH07b] Kyösti, Pekka ; Meinilä, Juha ; Hentilä, Lassi: *WINNER Phase II Implementation*. 2007. – "Avaiable: http://www.ist-winner.org/phase_2_model.html"

[Lee07] Lee, Seung J.: Trade-Off Between Frequency Diversity Gain and Frequency-Selective Scheduling Gain in OFDMA Systems with Spatial Diversity. In: *IEEE Communications Letters* 11 (2007), Nr. 6

[LHSH04] Love, David J. ; Heath, Robert W. ; Santipach, Wiroonsak ; Honig, Michael L.: What is the value of limited feedback for MIMO channels. In: *IEEE Commun. Mag* 42 (2004), S. 54–59

[LSM08] Lakkakorpi, J. ; Sayenko, A. ; Moilanen, J.: Comparison of Different Scheduling Algorithms for WiMAX Base Station: Deficit Round-Robin vs. Proportional Fair vs. Weighted Deficit Round-Robin. In: *IEEE Wireless Communications and Networking Conference* (2008), März, 1991–1996. http://dx.doi.org/10.1109/WCNC.2008.354. – DOI 10.1109/WCNC.2008.354. – ISBN 978–1–4244–1997–5

[Mac08] Maciel, Tarcisio F.: *Suboptimal Resource Allocation for Multi-User MIMO-OFDMA Systems*, Technical University of Darmstadt, Diss., 2008

[NR10] Nascimento, Alberto ; Rodriguez, Jonathan: A joint utility scheduler and SDMA resource allocation for mobile WiMAX networks. In: *Proceedings of the 5th IEEE international conference on Wireless pervasive computing*. Piscataway, NJ, USA : IEEE Press, 2010 (ISWPC'10). – ISBN 978–1–4244–6855–3, 589–594

[OC07] Oestges, C. ; Clerckx, B.: *MIMO wireless communications: from real-world propagation to space-time code design*. Academic Press, 2007 (MIMO Wireless Communications: From Real-world Propagation to Space-time Code Design). – ISBN 9780123725356

[PFE08] Pérez-Fontán, F. ; Espiñeira, P.M.: *Modeling the wireless propagation channel: a simulation approach with MATLAB*. Wiley, 2008 (Wiley series on wireless communications and mobile computing). http://books.google.de/books?id=N-mdDXvs0kAC. – ISBN 9780470727850

[PHS05] Peel, C. B. ; Hochwald, B. M. ; Swindlehurst, A. L.: A vector-perturbation technique for near-capacity multiantenna multiuser communication-part I: channel inversion and regularization. In: *Communications, IEEE Transactions on* 53 (2005), Januar, Nr. 1, 195–202. http://dx.doi.org/10.1109/TCOMM.2004.840638. – DOI 10.1109/TCOMM.2004.840638. – ISSN 0090–6778

[RRTW11] R. Ramjee, S. K. K. ; Thuel, S. ; Wang, X.: *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 20102015*. CISCO White Paper. 'http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html'. Version: February 2011

[SCA+06] Shen, Zukang ; Chen, Runhua ; Andrews, Jeffrey G. ; Heath, Robert W. ; Evans, Brian L.: Sum capacity of multiuser MIMO broadcast channels with block diagonalization. In: *IEEE Trans. Wireless Comm*, 2006, S. 886–890

[STKL01] Shad, F. ; Todd, T. D. ; Kezys, V. ; Litva, J.: Dynamic Slot Allocation (DSA) in Indoor SDMA/TDMA Using a Smart Antenna Basestation. In: *IEEE/ACM Transactions on Networking* 9 (2001)

[TJC99] Tarokh, Vahid ; Jafarkhani, Hamid ; Calderbank, A. R.: Space-time block codes from orthogonal designs. In: *IEEE Trans. Inform. Theory* 45 (1999), S. 1456–1467

[TSMC98] Tarokh, Vahid ; Seshadri, Nambi ; Member, Senior ; Calderbank, A. R.: Space-time codes for high data rate wireless communication: Performance criterion and code construction. In: *IEEE Trans. Inform. Theory* 44 (1998), S. 744–765

[Tso06] Tsoulos, G.V.: *MIMO system technology for wireless communications*. CRC/Taylor & Francis, 2006 (The electrical engineering and applied signal processing series Bd. 0). http://books.google.de/books?id=bcXO1U77ijAC. – ISBN 9780849341908

[TV05] Tse, D. ; Viswanath, P.: *Fundamentals of wireless communication*. Cambridge University Press, 2005 http://books.google.de/books?id=66XBb5tZX6EC. – ISBN 9780521845274

[WC98] Wesel, R. D. ; Cioffi, J. M.: Achievable rates for Tomlinson-Harashima precoding. In: *Information Theory, IEEE Transactions on* 44 (1998), März, Nr. 2, 824–831. http://dx.doi.org/10.1109/18.661530. – DOI 10.1109/18.661530. – ISSN 0018–9448

[WGSB07] Wang, Fan ; Ghosh, Amitava ; Sankaran, Chandy ; Benes, S.: WiMAX System Performance with Multiple Transmit and Multiple Receive Antennas. In: *VTC Spring*, IEEE, 2007, 2807-2811

[Xav08] *On the Challenges for the Maximization of Radio Resources Usage in WiMAX Networks*. 2008

[YG06] Yoo, Taesang ; Goldsmith, A.: On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming. In: *Selected Areas in Communications, IEEE Journal on* 24 (2006), März, Nr. 3, 528–541. http://dx.doi.org/10.1109/JSAC.2005.862421. – DOI 10.1109/JSAC.2005.862421. – ISSN 0733–8716

[ZDMXCF10] Zubow, A. ; D.Camps-Mur ; X.Pérez-Costa ; Favaro, P.: *Greedy Scheduling Algorithm (GSA) - Design and Evaluation of an Efficient and Flexible WiMAX OFDMA Scheduling Solution*. Elsevier Computer Networks Journal, Volume 54, Issue 10, July 2010

[ZM12] Zubow, Anatolij ; Marotzke, Johannes: On the Frequency-Selective Scheduling Gain in SDMA-OFDMA Systems. In: *CoRR* abs/1201.6282 (2012)

[ZMDMXC12] Zubow, A. ; Marotzke, J. ; D.Camps-Mur ; X.Pérez-Costa: sGSA: An SDFMA-OFDMA Scheduling Solution / Humboldt University Berlin. 2012. – Forschungsbericht

Selbständigkeitserklärung

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbständig verfasst und nur unter Verwendung der angegebenen Quellen und Hilfsmittel angefertigt habe. Weiterhin erkläre ich, eine Diplomarbeit in diesem Studiengebiet erstmalig einzureichen.

Berlin, den 10. Juli 2012 Johannes Marotzke

Statement of authorship

I declare that I completed this thesis on my own and that information which has been directly or indirectly taken from other sources has been noted as such. Neither this nor a similar work has been presented to an examination committee.

Berlin, July 10, 2012 Johannes Marotzke