TECHNISCHE UNIVERSITÄT DRESDEN

Spectrally and Energy Efficient Radio Resource Management for Multi-Operator Shared Networks

Osman Aydin

Fakultät Elektrotechnik und Informationstechnik
der Technischen Universität Dresden

zur Erlangung des akademischen Grades

D O K T O R I N G E N I E U R
(Dr.-Ing.)

genehmigte Dissertation

Vorsitzende:	Prof. Dr.-Ing. habil. Frank Ellinger
Gutachter:	Prof. Dr.-Ing. Eduard A. Jorswieck
	Prof. Dr. Stephan ten Brink
	Dr. Alessio Zappone
Tag der Einreichung:	14. 10. 2019
Tag der Verteidigung:	13. 03. 2020

Abstract

Commercial mobile communication systems are mainly based on licensed frequency spectrum, and the license is very expensive as the spectrum is a sparse wireless resource. Therefore, sharing this wireless resource is an essential requirement not only at the present but also in the future considering trends like connectivity for everybody and everything. In this thesis, we study the sharing of wireless resources with different approaches for realizing fair, efficient, and predictable sharing solutions in a controlled manner.
The efficient use of wireless channel resources is an important target to reduce the costs of network operation and deployment. To achieve this, we need practical scheduling algorithms for wireless resources, out of which several of them will be presented and analyzed in this work. Different optimization frameworks for the spectral efficiency utility are presented, with an individual focus on guaranteeing resource or rate fairness among the operators in a network with shared radio resources. Thus, the presented proposals will help the mobile network operators to overcome the issues of losing network control and traceability of used wireless resources in a shared environment.
Besides this, emerging vertical industries, such as automotive, healthcare, industry 4.0, internet of things (IoT) industries will put a certain burden on the wireless networks asking for guaranteed service level requirement from the mobile network operators.
In this regard, this thesis provides the necessary methods addressing these challenges with the help of scheduling methods which are based on the joint optimization of spectral and energy efficiency. Thus, wireless networks will be enabled as a service function in a controlled and scalable way for new emerging markets. Furthermore, the presented solutions fit well with the requirements of fifth generation (5G) network slicing.

Zusammenfassung

Kommerzielle Mobilfunksysteme nutzen hauptsächlich lizenzierte Frequenzspektren, deren Lizenzen sehr teuer sind, da die Spektren eine rare Funkressource sind. Die gemeinschaftliche Nutzung dieser Funkressourcen ist daher eine wesentliche Notwendigkeit, nicht nur in der Gegenwart, sondern auch in der Zukunft unter Berücksichtigung von Entwicklungen wie der Bereitstellung von Konnektivität für Mensch und Maschine. In dieser Dissertation untersuchen wir die gemeinsame Nutzung von Funkressourcen mit unterschiedlichen Ansätzen zur Realisierung von fairen, effizienten und vorhersagbaren Ressourcenvergabe Lösungen, die auch regelbar sind.

Die effiziente Nutzung von Funkkanalressourcen ist ein wichtiges Ziel zur Reduzierung der Betriebs- und Bereitstellungskosten der Netzwerke. Um dies zu erreichen, brauchen wir praktische Schedulingalgorithmen für Funkressourcen. In dieser Dissertation werden mehrere dieser Algorithmen vorgestellt und analysiert. Verschiedene Rahmenbedingungen für die Optimierung der spektralen Effizienzziele werden dargestellt mit einem individuellen Fokus auf die Gewährleistung einer gerechten Ressourcen- oder Datenratenverteilung zwischen den Netzwerkbetreibern, die diese gemeinsamen Funkressourcen teilen. Somit helfen wir mit den präsentierten Lösungen, die Vorbehalte der Netzbetreiber bezüglich dem Verlust der Netzwerkkontrolle und der Nachvollziehbarkeit der gemeinschaftlich genutzten Funkressourcen abzubauen. Auf der anderen Seite werden neue vertikale Industriemärkte wie die Automobilindustrie, das Gesundheitswesen, die Industrie 4.0, das Internet der Dinge die Mobilfunknetzwerke belasten und Forderungen in Bezug auf garantierte Dienstequalität an die Mobilfunknetzbetreiber stellen.

In dieser Hinsicht liefert diese Dissertation die notwendigen Methoden, um die gezeigten Herausforderungen mit der Hilfe von Schedulingmethoden anzugehen, die auf der gemeinsamen Optimierung von spektraler und Energie Effizienz basieren. Somit werden Mobilfunknetzwerke mit kontrollierbaren und skalierbaren Funktionen befähigt als Dienstleister für neu aufkommende Kundenmärkte zu operieren. Darüber hinaus sind die vorgestellten Lösungen auf die Anforderungen der fünften Mobilfunkgeneration (5G) inklusive der Network Slicing Technologie gut anwendbar.

Acknowledgements

I would not be able to start and complete this work presented in this thesis without the support of many people around me during my life and Ph.D. studies.

Therefore, first of all, I show extensive gratitude to my father inspired me to become an engineer, my mother who is supporting me my whole life, my thoughtful sister, my wonderful wife, my dearest daughters, my parents-in-law and my sister-in-law.

My sincere gratitudes towards Prof. Dr.-Ing. Eduard A. Jorswieck for providing me the opportunity for this Ph.D. and steered me through the studies with his knowledge. I am grateful to Prof. Dr. Stephan ten Brink and to Dr. Alessio Zappone for being the second and the third reviewer.

A very special gratitude goes out to my former Bell Labs managers Tod Sizer, Ulrich Barth, Dr. Hans-Peter Mayer and Wolfgang König for providing me the framework for this work.

I show extensive gratitude to Dr. Danish Aziz for his generic motivation and support.

I would like to acknowledge Dr. Ilaria Malanchini, Prof. Dr. Stefan Valentin, Dr. Andreas Weber, Wahid Jamil, Dr. Bho Matthiesen and Dr. Stefan Wesemann for their valuable discussions and work together. I am grateful to Dr. Johannes Richter for providing me the Latex infrastructure to finalize this work in an excellent format.

Finally, I would also like to thank all people which are not listed but supported this work.

Contents

List of Figures

List of Tables

Abbreviations

3GPP	3rd Generation Partnership Project
BS	base station
CDMA	code division multiple access
C-RAN	Cloud-RAN
CSI	channel state information
EE	energy efficiency
eMBB	enhanced mobile broad band
FDMA	frequency division multiple access
GDA	generalized dinkelbach algorithm
GEE	global energy efficiency
GPS	generalized processor sharing
GRS	generalized resource sharing
GSM	Global System for Mobile Communications
IoT	internet of things
LTE	Long Term Evolution
MAC	media access control
MAX	maximum rate
MBB	mobile broad band
MIMO	multiple-input multiple-output
mMTC	massive machine type communications
MNO	mobile network operator
MOCN	multiple operator core network
MO-CRRM	multi-operator common radio resource management
MOMU-MIMO	multi-operator multi-user MIMO
MOP	multi-objective program

MORAN	multi operator radio access network
MU-MIMO	multi-user MIMO
OFDM	orthogonal frequency division multiplexing
OFDMA	orthogonal frequency division multiple access
PF	proportional fair
QoS	quality of service
RAN	radio access network
RAT	radio access technology
RR	round robin
RRM	radio resource management
SDMA	space division multiple access
SE	spectral efficiency
SINR	signal-to-interference-plus-noise ratio
SLA	service level agreement
SNR	signal-to-noise ratio
TDM	time-division multiplexing
TDMA	time-division multiple-access
TP	transmission point
TSS	two step scheduler
TTI	transmission time interval
UE	user equipment
UMTS	Universal Mobile Telecommunication System
URLLC	ultra-reliable low-latency communications
V2N	vehicular-to-network
V2V	vehicular-to-vehicular

VoIP	voice-over-IP
WF	water-filling
ZF	zero-forcing

Part I

Preliminaries

Chapter 1

Introduction

Services provided by wireless mobile communications have become an essential part of our lives. Long Term Evolution (LTE) [1] also generally known as fourth generation (4G) mobile communication system of 3rd Generation Partnership Project (3GPP) based radio standard is considered to be the first generation that has truly provided the first mobile broad band (MBB) services. One main target of 4G was to achieve 100 Mbps downlink and 50 Mbps uplink throughput rates within a 20 MHz spectrum allocation [2]. The targets for the upcoming fifth generation (5G) [3] standard come along with more challenging requirements. The most relevant ones dealing with the content of this thesis are given in the following:

- 5G High Performance Data Rate: For the enhanced mobile broad band (eMBB) service 300 Mbps downlink and 50 Mbps uplink data rates are given as a target for outdoor users in dense urban areas [4].
- 5G High Performance Services: Still one has to remark that the allocated spectrum has to be divided between further high performance services like massive machine type communications (mMTC) and ultra-reliable low-latency communications (URLLC). Corresponding measures are already taken to solve efficiently the 5G spectrum scheduling with physical layer methods. This is done by the introduction of different numerologies in the frequency dimension or in the time domain with different lengths of scheduling time slots, namely transmission time interval (TTI), or symbol lengths. The corresponding discussions and illustrations are given in [5] and [6].
- 5G High Spectrum Demand: To cope with high traffic densities like 750 Gbps/km2 in dense urban areas [4], the 5G system considers more spectrum bandwidth ,e.g., by refarming the LTE spectrum, new spectrum allocations as given in the appendix of [7] and possibly by using unlicensed spectrum [8].

The aim of this thesis is to provide radio resource management (RRM) solutions in the case the frequency spectrum, time and transmit power resources are shared among different network operators. This is called the multi-operator sharing scenario. Hence, concepts and solutions are provided which meet the high demands of, e.g., 5G and beyond mobile communication systems.

1.1 Motivation

The capacity of wireless channel has been defined by the well-known Shannon-Hartley theorem [9]. It states that the capacity is directly proportional to the available frequency spectrum bandwidth that could be used on a channel. Availability of frequency spectrum for each mobile network operator (MNO) is limited. Hence, spectrum pooling among multiple MNOs working in a partnership can provide each operator eventually a higher system capacity. This will require efficient sharing of frequency resources. Moreover, for this case, [10] shows a sharing gain already with two sharing operators.
From an economical perspective, the constantly increasing data traffic demand forces MNOs to densify their network nodes to satisfy the demand. In addition to such densification, this is achieved with higher frequency spectrum reuse. Wireless network deployment costs become dramatically higher for operators, due to increasing traffic demand, in contrast with the revenues per subscriber as shown in Figure 1.1 and analyzed in [11].

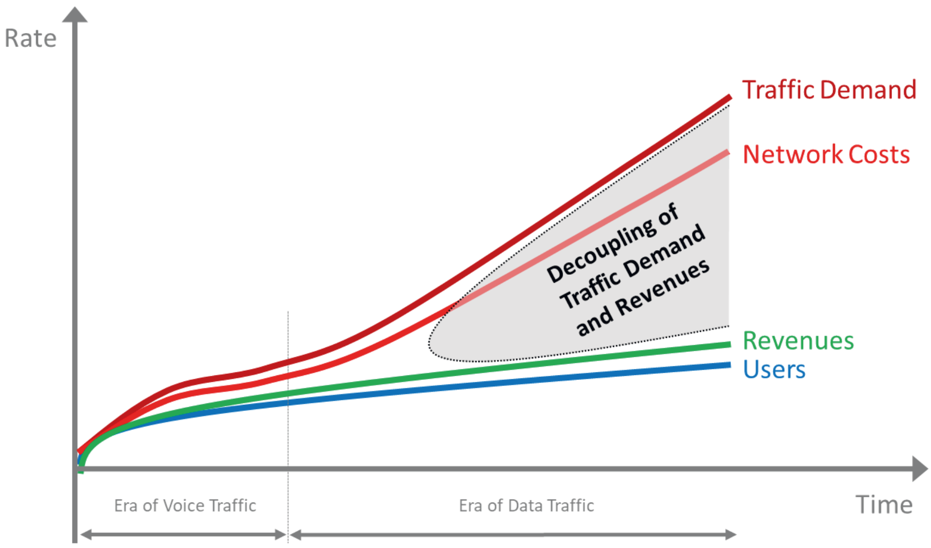

Figure 1.1: Data traffic demand versus operator’s revenue adapted from [11]

Sharing frequency resources to provide more capacity would be a more cost effective way instead of continuous network densification [12] in anyhow over-

lapping coverage areas with other MNOs possible with different business portfolios [13].
When doing a cost analysis, on top of the deployment investments also the operational costs have to be taken care of. One of the cost factors, there is the operational cost for consumed energy in a base station (BS) [14]. In [15], the relation between traffic demanded and the cost for the energy in a 4G mobile network is illustrated with realistic pricing. The energy used for data transmission is highlighted in [16] as cost per bit which should be pursued as a requirement for 5G. We will define this ratio of spent energy per data bit as energy efficiency (EE).

The main impact from this thesis is to present BS scheduler solutions for multi-operator environments showing with the set-up capabilities the bunch of possible different business relationship realizations in a shared environment. Therefore, in the core part of this thesis, different approaches and results are presented guaranteeing fairness, data rate assurance for the MNOs, possibilities how to tune spectrally or energy efficient operation for the overall shared wireless system on radio access network (RAN) level according to MNOs targets. Thus, the presented achievements will help the MNOs to remove fears of losing network control and traceability of used wireless resources in an shared environment.

1.2 State of the Art

Sharing of resources among users is the intrinsic topic of wireless resource allocation methods. The first widespread mobile communication systems, e.g. Global System for Mobile Communications (GSM) [17] used the time dimension, namely time-division multiple-access (TDMA), to share the frequency spectrum resources among the users.
On top of that multiplexing scheme in Universal Mobile Telecommunication System (UMTS) [18] the wireless resources were shared using the code division multiple access (CDMA) approach. With LTE based on multiple-input multiple-output (MIMO) [19, 20], a specific space division multiple access (SDMA) approach has become popular for sharing the wireless resources among multiple users known as multi-user MIMO (MU-MIMO) [21]. In wireless networks the different multiplexing schemes were initially applied for spectrally efficient scheduling of the under the subscription of same MNO. With similar reasons these multiplexing schemes were also studied for wireless resource sharing

among MNOs. Furthermore, in [22], the spectrum pooling aspect is highlighted as further extension in which the sharing operators jointly use their licensed frequency spectrum. Regarding radio interference management, when multiple operators co-exist and share a common pool of spectrum, [23] proposes a dynamic spectrum allocation algorithm with satisfactory level of quality of service (QoS) for the users. The study in [24] shows that pooled spectrum can be distributed with an auction mechanism based scheme for joint RRM. Also the approach there, goes a step further and provides the spare radio resources to be fully utilized by different radio access technology (RAT)s of the MNOs being short of radio resource. Similar approach is presented in [25] which considers a pricing for the different parts of shared spectrum as a scaling factor in the objective function. On the other hand, [26] presents game theory based approaches maximizing the sum rate of the sharing operators.
A different sharing approach is given in [27] and [28]. Here, single shared radio network infrastructures such as relays or small cells are commonly used by sharing operators whereas the surrounding infrastructure is still realized as independent single operator networks. Interference between both infrastructures is suppressed by MIMO techniques or by signaling of utilized spectrum. Both approaches show sharing gains in the assumed system for the data throughput. In comparison to the given approaches, earlier described sharing approaches analyse and optimize the voice traffic distribution and the resulting specific call blocking rates [29–32] in multi-operator scenarios.
Today, the focus of multi-operator sharing scenarios are not only limited to the opimization of the spectral efficiency (SE), but extending the focus to fair distribution of the available resources to the sharing operators [33] and extending the utility also to EE [34, 35] of the shared wireless network resources as also described and presented in the following chapters.
For 5G, since the focus is on service-specific sharing concepts, e.g., for network slicing [34] or virtualization, IoT [36] and Cloud-RAN (C-RAN) [35, 37, 38], different sharing approaches are introduced.

1.3 Thesis Outline and Contribution

The optimization of wireless channel resources management is an effective approach to reduce the costs of network operation and deployment. To achieve this, we need practical scheduling algorithms for wireless resources, several of them will be presented and analyzed in the following sections. Different optimization frameworks for the spectral efficiency utility are presented, with an individual focus on guaranteeing resource or rate fairness among the operators in a network with shared radio resources. Prior to this, the first part gives insights to the challenges of future mobile networks and how sharing mechanisms can support the future wireless traffic demands. This part highlights explicitly the targets of 5G of mobile communications regarding QoS and EE. The main requirements are the needs for more enhanced and adjustable resource scheduling mechanisms. Different services will have different requirements where each service has to be guaranteed and efficiently operated within wireless networks. Figure 1.2 summarizes the thesis outline in a tabular form per part and chapter according to the presented sharing approaches. For each sharing approach, it lists up the information about the domain of shared resources, namely time, frequency and power resource. Further, the constraints, describing the fairness dimension to be achieved between sharing operators, like the resource or rate fairness, e.g. in a resource fairness constraints the utility is achieved in the limits of tuned wireless resource distribution per operator. The utilities column consists of maximizing spectral efficiency by data rate maximization or in a more advanced solution jointly maximizing rate and EE.

		Sharing Approach	Shared Resources	Constraints	Utilities
Part II	Chapter 4	TDMA Two Step Scheduler	Time	Resource Fairness	Rate Maximization
	Chapter 5	Generalized Processor Sharing	Rate Matrix: Frequency & Time	Rate Fairness	Rate Maximization
	Chapter 6	Generalized Resource Sharing	Frequency	Resource Fairness	Rate Maximization
Part III	Chapter 7	MOMU-MIMO I	Power	Resource Fairness Rate Fairness	Rate Maximization
	Chapter 8	MOMU-MIMO II	Power	Rate Fairness	Rate Maximization
Part IV	Chapter 9	MOMU-MIMO III	Power	Rate Fairness Energy Efficiency Fairness	Rate Maximization Energy Efficiency Max.

Figure 1.2: Overview of presented sharing approaches

The main component of RRM solutions is set-up on scheduling algorithms which fulfill the basic service requirements and also provide fairness between the sharing partners, namely the mobile network operators. This will be shown in the second part, which includes the analysis of the solutions from Chapter 4-6, as given in Figure 1.2. The scheduling algorithms are based on the time and frequency domains of wireless networks. The resulting limits and penalties attended with these approaches will be closely discussed.

In the third part, the RRM focus is on the spatial multiplexing of wireless resources as a further sharing dimension to be exploited for multi-operator scenarios. This proposed sharing paradigm will be discussed and evaluated in the light of the shortcomings of former ones together with its provided performances in the light of spectral efficiency and fairness of distributed resources among sharing mobile network operators.

The fourth part optimizes different approaches that were presented in second and third parts. The transmission power used for transmitting radio channels is added and evaluated as a sharing dimension into the sharing model. And consequently EE requirements are introduced as a further utility besides SE. Both utilities are investigated with a multi-objective implementation in different sharing case studies. For our case studies, the problem formulations for the generalized dinkelbach algorithm (GDA) are kept practical. Together with the performance results achieved for SE and EE it is shown that the assumptions and results have promising applicability for the upcoming mobile network generations.

Chapter 2

Sharing Framework in Mobile Communications

In this chapter, an overview of the most relevant and applied methods for sharing resources in mobile networks is given. The case of national network roaming [39], which could also be considered as network sharing, is not further discussed. Figure 2.1 presents from left to right the increasing number of shared dimensions, beginning with the no sharing case. Next, with site sharing, the site will be used jointly by the BSs of the operator, which is indicated with the green coloring. In the multi operator radio access network (MORAN) case, beside the site sharing, the BSs and antennas are marked in green to highlight the next level of sharing. Finally on the most right setup, the multiple operator core network (MOCN) is illustrated with the further sharing dimension: the frequency spectrum. So the frequencies of the operators are pooled together, which is indicated by the green markings.

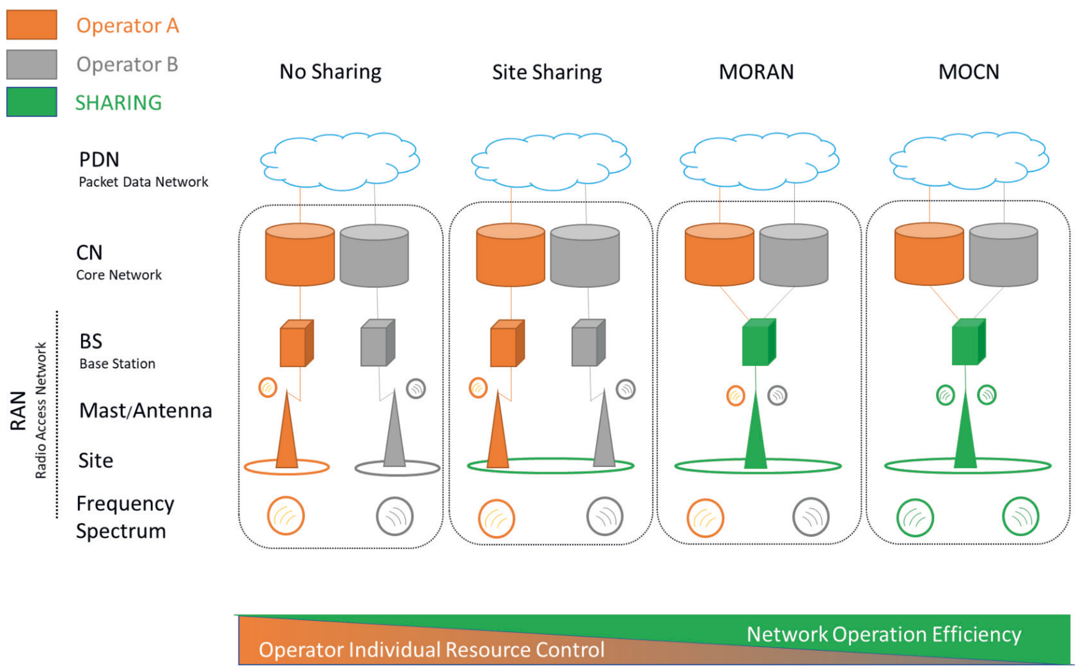

Figure 2.1: Mobile network sharing options

2.1 Site and Infrastructure Sharing

The site sharing as illustrated in Figure 2.1 is mostly meant regarding sharing of the BS or antenna tower ground. This helps different operators to share the rental costs and to reduce overall network operation costs. The shared objects can be extended to sharing of further so called passive elements like the radio tower itself, the antenna, the BS cabinet and its cooling system, which will be called infrastructure sharing. Such passive hardware resources have already been shared between multiple mobile network operators. The interested reader is referred to [40] for further details about site and infrastructure sharing.

2.2 Multi Operator Radio Access Network

As given in Figure 2.1, MORAN sharing contains the sharing of the passive elements as given in Section 2.1 and additionally the sharing of active elements in a BS such as transceivers. It is not standardized in the standardization body 3GPP. More insights regarding MORAN and its benefits are given in [41].

2.3 Multiple Operator Core Network

In MOCN sharing, as given in Figure 2.1, all elements of RAN, passive and active, including the wireless spectrum are shared. Therefore, we can list here the main advantage as the efficiency gained by pooling of the wireless spectrum. The overall pooling gain could be transformed to an individual gain to the sharing partner, i.e., MNOs, if a service provisioning per partner is guaranteed according an agreed SLA. This can be achieved by different means of sharing the wireless resource as given in the next chapter.

Further details of MOCN regarding service aspects and requirements for network sharing are given in [39] and details of architecture and functional description for network sharing are provided in [42].

2.4 Sharing of Radio Resources

The wireless resources of an operator are already provided to its users as a shared media. The multiplexing of the users [43, 44] can be realized by TDMA, frequency division multiple access (FDMA), CDMA or spatial multiplexing, i.e., by using MIMO concepts, see Figure 2.2.

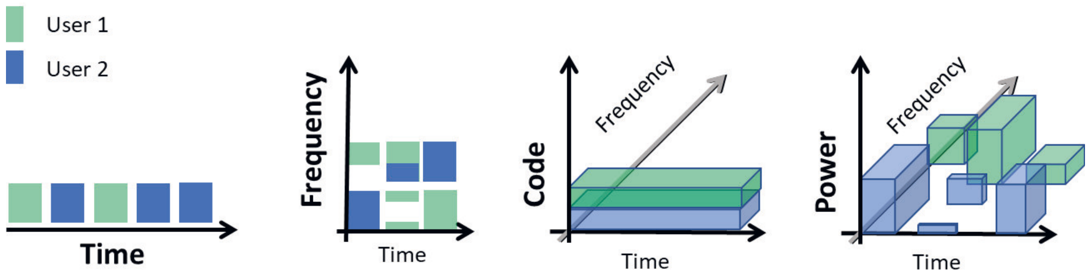

Figure 2.2: Multiplexing mechanisms in wireless networks

In this thesis, the application of multiplexing mechanisms is extended from the user domain to the operator domain providing efficient radio resource scheduling mechanisms, as listed in Figure 1.2, for different multi-operator sharing scenarios.

2.5 Service Level Agreements

The efficient operation of shared wireless network requires efficient centralized scheduling mechanisms in the case of multi operators. We call it here, multi-operator common radio resource management (MO-CRRM). Generally, terms and conditions about the service parameters related to MO-CRRM are pre-agreed between operators with the help of SLAs [45, 46]. Therefore, details of such SLAs agreements are captured in contracts which are not public. In 5G networks, SLA-type of service level requirements could be realized with the network slicing feature. Vertical industries, e.g. the automotive industry or public safety authorities can define own QoS parameters to be realized in an independent end-to-end logical network namely the network slice with reserved access to it [47]. Automotive use cases, including traffic safety relevant use cases and their service level requirements can be found in [48].

Part II

Fairness Triggered Algorithms

Chapter 3

Two Step Scheduler

3.1 Introduction

Sharing the wireless channel resources and BS hardware, is an effective approach to reduce the costs of operation and deployment. Sharing allows multiple operators to utilize the BS and bandwidth more efficiently. To do so, operators and infrastructure providers agree on a fixed sharing ratio that has to be guaranteed during operation. In traditional approaches, each operator receives a constant fraction of the overall bandwidth that is proportional to the sharing ratio. At a first glance, such bandwidth splitting is appealing. It is simple, and the allocated resource fraction is guaranteed. However, allocating fixed sub-bands leads to a mismatch between sharing ratio and operator's cell capacity in frequency selective propagation scenarios [49]. Moreover, due to the fixed bandwidth limit, the cell capacity of one operator may be too low to support the current traffic while sufficient bandwidth is available in the sub-band of another operator. This artificially created bottleneck wastes resources and reduces QoS in bursty traffic scenarios. In order to avoid such shortcomings, we present an approach which cyclically allocates the whole bandwidth among the operators. The first multi-operator scheduler presented in this thesis is a simple and practical scheduler based on time-division multiplexing (TDM). It is a TSS [50] that allocates time slots among operators in a round robin (RR) manner and allocates subsequently resources to the users of each operator, see Figure 3.1 and specified in *Two-Step Scheduler (TSS)* in Appendix A. This two-step approach guarantees the agreed resource shares in the first step while allowing each operator to execute its own multi-user scheduling policy in the second step. This flexibility is a significant benefit of the TSS design. The simulation results demonstrate an insignificant increase in scheduling delay at high spectral efficiency and resource fairness.

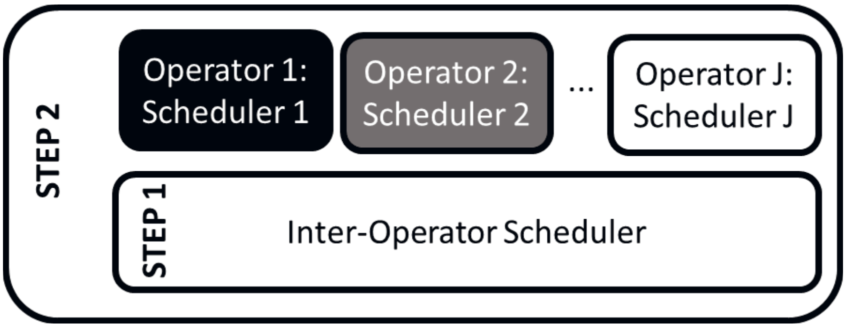

Figure 3.1: Two-Step Scheduler (TSS) architecture from [50]

Key functions of TSS:

- Practical architecture: Our new TSS architecture enables dynamic resource sharing among multiple operators and integrates easily into existing BS designs and 3GPP media access control (MAC) specifications [51].
- Lightweight scheduling algorithm: Our scheduling algorithm allocates resources among multiple operators without exceeding the complexity order of single operator scheduling. Although resource sharing in time always comes at the cost of delay, the latency increase of TSS is insignificant.
- Efficient and fair resource allocation: Our simulation results show that our scheduling approach reaches the spectral efficiency of proportional fair (PF) scheduling. It is fair among users, by reaching PF's fairness and it is fair among operators by fulfiling the guaranteed sharing ratio.

3.2 System Model

We study the downlink of a single cell covered by one BS that is shared among $J = |\mathcal{J}|$ network operators. We denote the set of operators by $\mathcal{J}$ and the set of active users in the cell as $\mathcal{K}$. The number of active users is $K = |\mathcal{K}|$. The system bandwidth is denoted as B and shared among the users of a single operator. The duration of a scheduling period, within which the users of all J operators are scheduled is given by T. The sharing ratio g_J within T defines the time resource allocated to an operator J. The multi-operator scheduling operation is exemplified in Figure 3.2.

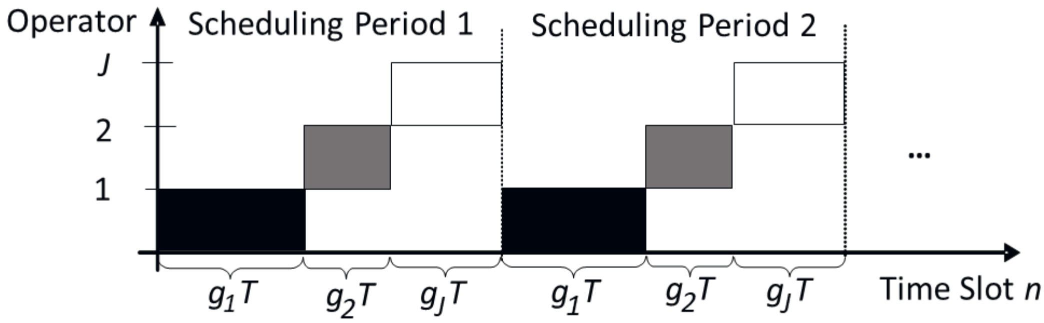

Figure 3.2: Two-Step Scheduler exemplary scheduler operation from [50]

For the quasi-static block flat-fading channel and traffic models, we assume a simple fading model where the channel coefficient $h_k[n]$ of an arbitrary user k and time slot n is a circularly symmetric complex Gaussian random variable. This random variable is independent and identically distributed (i.i.d.) among time slots and users. This classic Rayleigh fading model [43, Section 2.4.2] leads to the i.i.d. exponentially distributed instantaneous signal-to-noise ratio (SNR)

$$\gamma_k[n] = |h_k[n]|^2 \,\mathrm{SNR}_k, \tag{3.1}$$

where SNR_k is the average SNR given by the user's path loss and Shadowing. We calculate the physical layer rate an arbitrary user k achieves per time slot using Shannon's equation as

$$r_k = \log_2(1 + \gamma_k), \tag{3.2}$$

To focus on scheduler operation, we exclude further dynamics from our study. To this end, we assume the full buffer traffic model. Here, the downlink transmission at the BS is filled during the complete simulation time for each user in $\mathcal{K}$. This simple traffic model corresponds to the download of a large file and creates a constant load. We vary K to study our system for different load.
Referring the scheduling assumptions, from the set of active users the user set $\mathcal{K}^* \subseteq \mathcal{K}$ is scheduled. We denote the number of scheduled users as $k_{max} = |\mathcal{K}|$ and assume that each scheduled user receives the fraction B/k_{max} of bandwidth B. An important building block for our inter-operator scheduler is the RR policy, where the resource allocation is exclusive and alternates among the requesting entities. We employ this simple strategy for inter-operator scheduling in the first stage of our algorithm and study it as one possible option for

inter-user scheduling. Several scheduling policies would be possible. Here, we focus on the widely-used class of PF schedulers, which assigns resources to those users k_{max} with the highest weight

$$\lambda_k(n) = \frac{r_k(n)^\alpha}{R_k(n)^\beta} \tag{3.3}$$

in the current time slot. In (3.3), $\alpha \in [0, \infty[$ and $\beta \in [0, \infty[$ are constant parameters to adjust the effect of the corresponding terms. In the extreme cases with $\alpha = 0$ and $\beta = 1$ the term will be independent of the current channel conditions of the users, so users will be scheduled in a RR style. On the other hand, if the settings $\alpha = 1$ and $\beta = 0$ are used the scheduling will be a maximum rate (MAX) scheduler, as the users with best channel conditions are favored. The quantity R_k is the moving average over time, which is calculated as

$$R_k(n) = \begin{cases} (1-\theta)R_k(n-1) + \theta r_k, & k \in \mathcal{K}^* \\ (1-\theta)R_k(n-1), & k \notin \mathcal{K}^* \end{cases} \tag{3.4}$$

where the constant forgetting factor θ allows to trade off average and update. In addition to PF scheduling, we compare our TSS multi-operator scheduler to MAX scheduling for users of a single operator. With this scheduling policy, weight given in (3.3) reduces to $\lambda_k(n) = r_k(n)$, i.e., the first k_{max} users with the highest instantaneous physical layer rate are scheduled.

3.3 Results and Analysis

For the data rate results, the scatter plot in Figure 3.3 provides a first impression on the fairness and average sum rate. The latter metric is the time-average of the aggregated downlink rates over all users in the system and points to the average throughput of the system. Fairness is measured as the 5% quantile of the data rates, which expresses the throughput of users at the cell edge. As shown, the performance of TSS is similar to applying the PF policy to users of a single operator. Constant performance values are achieved also with changing network conditions. In summary, TSS can be used in a multi-operator environment without losing data rate compared to single-operator PF scheduling.

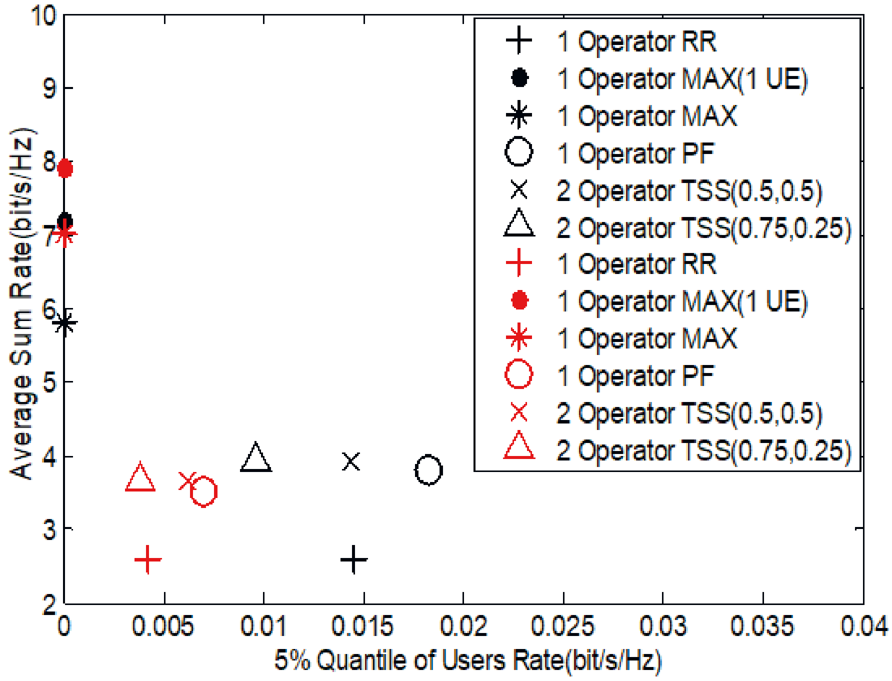

Figure 3.3: TSS: Average sum rate and 5% quantile with 30 users (red values) and 100 users (black values) from [50]

To validate that TSS reaches the agreed and configured operator sharing ratios, we study the fraction of resources allocated to the operators. The results in Table 3.1 validate that the agreed sharing ratios are achieved for 30 and 100 users scheduled in the BS.

Table 3.1: Achieved resource sharing ratios with TSS from [50]

	Configured sharing ratio g		Achieved sharing ratio	
	$Operator1$	$Operator2$	$Operator1$	$Operator2$
30 Users	0.50	0.50	0.4983	0.5017
100 Users	0.50	0.50	0.5016	0.4984
30 Users	0.75	0.25	0.7487	0.2513
100 Users	0.75	0.25	0.7509	0.2491

Since TSS realizes operator sharing in time we need to also study scheduling delay as a performance indicator for the applied schedulers. As the scheduling delay will influence the users perception in using delay sensitive application (e.g., gaming applications), we study the user's scheduling delay for different network loads assuming 1 ms for the subframe duration for scheduling. For 30 active users, TSS shows no effect on the mean of the scheduling delay in Table

3.2. However, the standard deviation σ of the delay is affected by TSS. Since TSS has one further scheduling step compared to the single operator schedulers. Thus, at a sharing ratio of 0.5, TSS increases the standard deviation of the delay of single-operator PF by 1.16 ms. Our results show that 50% of all users are served with an average delay smaller than 20 ms. This is sufficient to support even delay-sensitive applications such as voice-over-IP (VoIP). Note that, in many BS designs, this delay would be further reduced by traffic-specific prioritization, which is not considered in this study. All scheduling delay performance results are summarized in Table 3.2.

Table 3.2: Achieved scheduling delay in milliseconds with TSS from [50]

	30 Users		100 Users	
	μ	$\mu + \sigma$	μ	$\mu + \sigma$
RR	5.00	5.01	16.67	17.15
PF	5.00	6.63	16.67	18.42
TSS (0.5;0.5)	5.00	7.79	16.67	20.09
TSS (0.75;0.25)	5.00	9.34	16.67	27.08

3.4 Conclusion

In this chapter, we presented a simple TSS for sharing wireless channel resources among the users of multiple operators. Studying the sharing ratio, it is shown that TSS reliably allocates the resources among the operators according to the agreed shares. From studying sum rate, fairness, and delay, we can conclude that our TSS approach shows no considerable drawback compared to common single-user schedulers as PF. Although sharing resources in time always comes at the cost of delay, the average delay is not increased and the increase in delay variance is insignificant.

Chapter 4

Generalized Processor Sharing

4.1 Introduction

A different multi-operator scheduler approach is followed with the GPS [52] than the TSS which is based on TDM. By applying GPS and Bin Packing, the channel capacity is distributed among the operators respectively to their users. Here, the channel capacity corresponds to an achievable data rate per user. A scheduling approach close to the presented one is proposed in [53]. Therein, the author follows the idea of GPS to allocate resources to users of multiple operators. Although GPS is algorithmically promising, it requires the instantaneous cell capacity to be known a priori. To provide this value, a constant cell capacity is assumed in [53]. This assumption is impractical for mobile communication systems where the channel gains vary in time. Our algorithm, however, is not based on such unrealistic assumption. In particular, the scheduling algorithm assigns channel capacity based on the users' instantaneous physical layer rates. A second difference to [53] is that our algorithm aims to maximize the number of users that are scheduled per slot. By applying a Bin Packing heuristic, our algorithm minimizes cut-off and, thus, reaches a substantially higher spectral efficiency. And similarly as in TSS the flexible scheduler design of GPS allows each operator to execute its own multi-user scheduling policy.

The details of the GPS algorithm is illustrated in Figure 4.1 and specified in the algorithm given in Appendix B. For a clear presentation, we structure the algorithm in three phases and apply PF for all operators. The algorithm is, thus, called GPS-PF. In its first phase, the data rate of the user with the highest channel gain is calculated. This rate is used as an estimate for the instantaneous cell capacity in a second phase. Therein, GPS is applied to calculate rate c_j that is granted to an arbitrary operator J. Then, the users are weighted according to the operator's scheduling policy, and all users are selected that can be served with the operator's rate. This selection process can be seen as a filter that removes users with unfeasible rates $r_{jk} > c_j$ from the scheduling decision of the current time slot. Figure 4.1 illustrates this filtering process, which results in a feasible rate matrix f for all users and all operators. In the third phase, users with feasible rates are scheduled for each operator according to their weights.

Following the highest-weight-first policy, the users of an arbitrary operator J are selected such that their rates are packed inside c_j. This is performed by a Bin Packing heuristic and assures that (i) the operator's capacity is efficiently used and (ii) its weighting policy is applied.

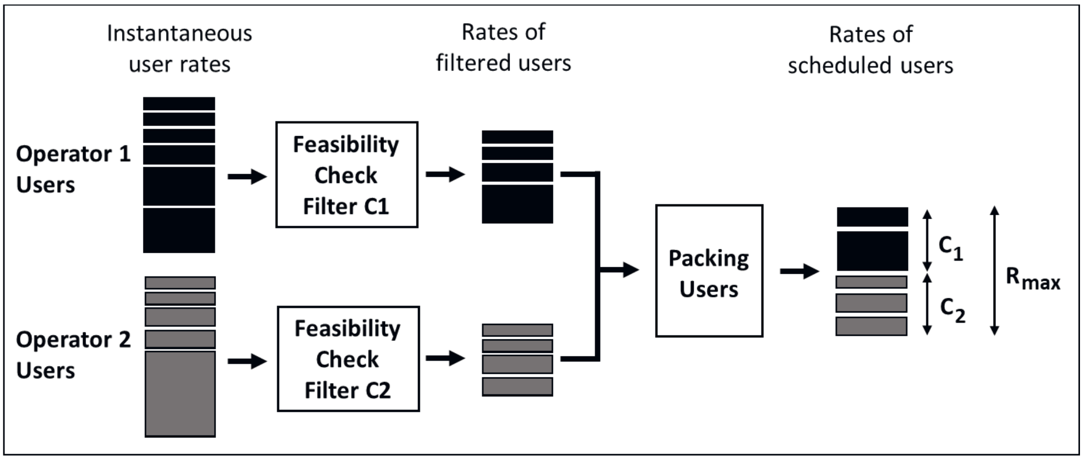

Figure 4.1: GPS scheduler architecture from [52]

To follow the time-variant channel gain, the GPS-PF algorithm is executed every time slot, which has a duration T in the millisecond regime. The algorithm returns matrix S that defines the allocated rates for all users of all operators for the current time slot. In S , a zero rate indicates that a user is not scheduled. Note that due the feasibility test in Phase 2 and packing in Phase 3, the number of scheduled users may vary from cycle to cycle. As users may be excluded for several time slots, their communication delay can be increased. This aspect is studied in the following passages.

4.2 System Model

The system model, for which GPS-PF is applied, is identical to the system model in Section 3.2. The downlink of a single cell covered by one BS is assumed which is shared among X network operators. We denote the set of operators by $\mathcal{J}$ and the set of active users in the cell as $\mathcal{K}$. The number of active users is $K = |\mathcal{K}|$. Also the same equations for the channel, data traffic assumptions and scheduling policies as given by (3.1), (3.2), (3.3) and (3.4) are used.

4.3 Results and Analysis

In this section, we will study how the multi-operator scheduler performs compared to common single-operator schedulers such PF, MAX, and RR. After comparing data rate and fairness, we will focus on the specifics of multi-operator scheduling. In particular, we study if GPS-PF keeps the agreed sharing ratios and by how much multi-operator scheduling in time increases the users' scheduling delay.
We study data rate, sharing ratio and delay for varying signal-to-interference-plus-noise ratio (SINR). We choose a high number of $K = 100$ users to study the system under full load. With $J = 2$ operators, we focus on a simple example where each operator serves 50 users. All schedulers are studied under equal assumptions on parameters, load, and channel statistics. We study performance as the time average of the aggregated downlink rates over all users in the system, called average sum rate. Fairness is measured as the 5% quantile of the data rates, which is a widely-used metric to expresses the throughput of users at the cell edge.

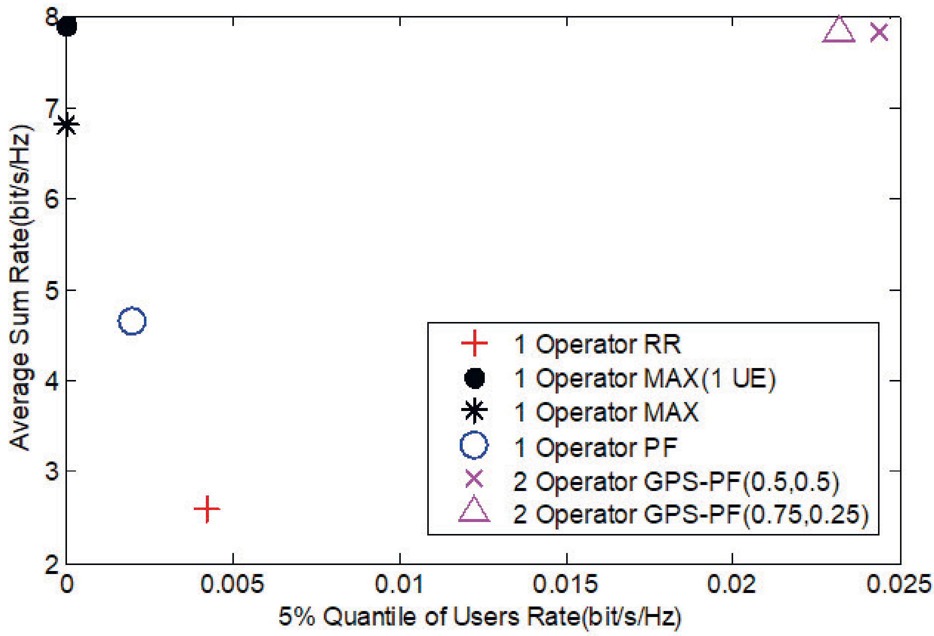

Figure 4.2: GPS: Average Sum Rate and 5% Quantile with 100 users from [52]

The scatter plot in Figure 4.2 provides a first impression on the fairness and average sum rate reached by the studied schedulers. The average sum rate of GPS-PF shows a very efficient use of radio resources. As shown in Figure 4.2, almost the performance of MAX is achieved, which is a very high efficiency for multi-operator scheduling. Our GPS-PF scheduler reaches this outstanding performance by distributing the complete spectrum to the users via the Bin Packing strategy. Only a small fraction may remain unused if no further users with sufficient rates can be found. Note that this is a substantial efficiency benefit compared to other multi-operator schedulers [2,3,8]. Moreover, GPS-PF outperforms the other schedulers even in terms of fairness. This is shown by the 5% quantile axis in Figure 4.2. The high fairness is a result of the Bin Packing approach in the GPS-PF algorithm, which favors users with low rates to fill-up the remaining bandwidth. To validate that GPS-PF achieves the agreed operator sharing ratio g, we study the fraction of resources that is actually allocated to the operators. The results in Table 4.1 shows that this fraction is very close to the values defined in g. Thus, the agreed sharing ratios are achieved sufficiently close.

Table 4.1: Achieved resource sharing ratios with GPS from [52]

	Configured sharing ratio g		Achieved sharing ratio	
	$Operator 1$	$Operator 2$	$Operator 1$	$Operator 2$
100 Users	0.50	0.50	0.4999	0.5001
100 Users	0.75	0.25	0.7526	0.2474

Scheduling delay: Since GPS-PF may not schedule users with a capacity higher than the share of an operator, the queuing delay of such users may increase at the BS. Such increase in scheduling delay can penalize delay-sensitive applications such as voice, video telephony, or gaming and has to be carefully studied. Table 4.2 summarizes the simulation results for the scheduling delay at the base station assuming 1 ms as subframe duration for scheduling. Our GPS-PF algorithm increases the average scheduling delay μ by approximately 0.6 ms with a slight dependency on the sharing ratio. Also the standard deviation of the scheduling delay σ increases. We can conclude that the high throughput and fairness of GPS-PF come at the cost of scheduling delay. However, we have to note that the slight increase is insignificant for most delay-sensitive applications. Our results show that 50% of all users can be served with an average delay smaller than 20 ms. This is sufficient to support even delay-sensitive applications as VoIP. Note that, in many base station designs, this delay would be further reduced by traffic-specific prioritization.

Table 4.2: Achieved scheduling delay in milliseconds with GPS from [52]

	μ	$\mu + \sigma$
RR	12.50	13.00
PF	12.50	14.22
GPS (0.5;0.5)	13.07	20.67
GPS (0.75;0.25)	13.09	22.76

4.4 Conclusion

We presented a scheduling algorithm that allocates transmission time to users of multiple operators. Based on Generalized Processor Sharing (GPS) and Bin Packing, our algorithm reaches the high spectral efficiency of Max-Rate scheduling [43] without sacrificing fairness. Our scheduler keeps high throughput for users at the cell edge and provides the agreed sharing ratios to the operators. The additional functionality and high gains come at a slightly increased scheduling delay. However, this increase is insignificant for real-time services such as VoIP. Moreover, the presented algorithm is the first feasible application of GPS for multi-operator scheduling.

Chapter 5

Generalized Resource Sharing

5.1 Introduction

This approach represents the first application of a utility function for a multi-operator scheduler in this thesis. It is targeting to maximize the single utility spectral efficiency by guaranteeing rate fairness for the sharing operators based on TDM. While it is based on generalized resource sharing (GRS) which is basically introduced previously in [54], here, the focus is on the mathematical formulation of multi-operator scheduling as published in [55]. The presented scheduling policy allows to trade-off sharing guarantees versus spectral efficiency, covering current fixed and dynamic approaches as special cases. Analyzing this general scheduling policy leads to a profound understanding of its most important parameters and of their effect on rate-dependent utilities. The presented proofs are confirmed by simulation results that hold for insightful scenarios. Sharing guarantees among operators and infrastructure providers must be held at a minimum reduction of spectral efficiency. We proposed a centralized multi-operator scheduler in [54] that allows to trade off sharing guarantees versus spectral efficiency. This dynamic sharing profits from fluctuations in traffic demands and channel quality by deviating from the agreed sharing guarantees in a controlled manner. Although this approach promises high gains, a rigorous understanding based on a solid mathematical formulation is required before operators can accept this new technology. In [56], we further extended mathematical analysis with mathematical proofs to provide a consistent mathematical understanding of the convex optimization problem. From an algorithmic point of view, [10] provided early performance results for sharing orthogonal resource blocks, which have been extended to non-orthogonal, distributed techniques in [49]. In the following analysis, the focus is on orthogonal spectrum sharing through a centralized scheduler. At large time scale, [57] presents so-called Spectrum Leasing based on convex optimization. With this approach, operators grant access to each other's resources at a much larger time scale than considered in this chapter. A negative side effect of such long-term reservation is a reduction of statistical multiplexing gains, as variations in traffic and channel gains cannot be exploited. This is possible by scheduling-based

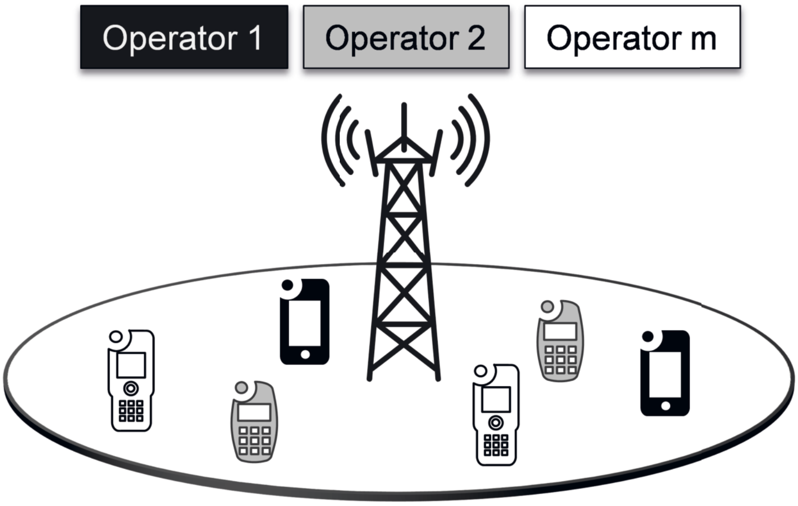

Figure 5.1: GRS: Sharing resources of a single base station among multiple operators from [55]

approaches that operate in the regime of several milliseconds, which are studied in [52, 53] and in this scheduling approach. The approaches in [52, 53] allocate constant shares of wireless channel resources to the operators. Although we cover such static sharing as a special case, our dynamic formulation allows to vary the resource allocation over time by allocating more or less than the agreed resource shares to the operators. Such dynamic sharing was proposed in [54] and a formal analysis is provided below.

5.2 System Model

We consider the common case of an independent scheduler at a BS, where neighboring BS are covered by inter-cell interference in the SINR parameter. Focusing on one such BS, we assume that its wireless channel resources are shared by the users of multiple operators, see Figure 5.1. Let $\mathcal{J}$ be the set of operators and $|\mathcal{J}|$ their number. Likewise, the overall set of $|\mathcal{K}|$ active users served by the BS is denoted by $\mathcal{K}$ with $\mathcal{K}_j$ being the subset of users associated with an arbitrary operator j. We consider bandwidth and transmit power to be constant and time to be discretized in slots, with n indexing an arbitrary time slot.

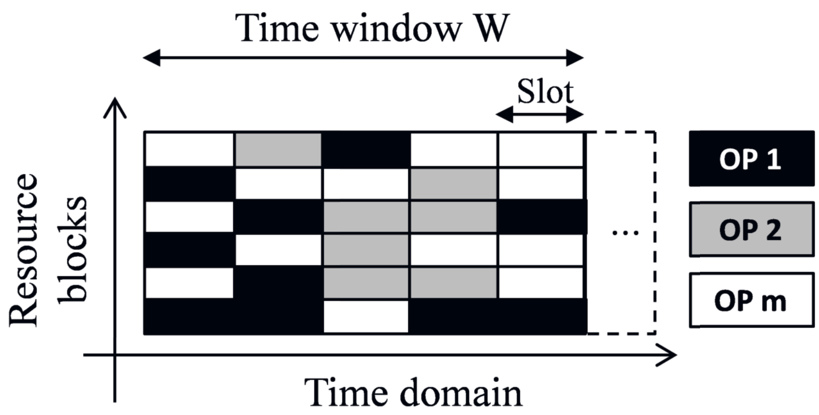

Figure 5.2: GRS: Resource allocation over time with TTI=1ms interval from [55]

5.2.1 Channel Model

We assume the downlink from BS to an arbitrary user k to be a time-variant, frequency-flat block fading channel with i.i.d. Rayleigh channel coefficients h_k. This model yields exponentially distributed random channel gains $|h_k|^2$ and provides the instantaneous SINR per time slot n as $\gamma_k[n] = |h_k[n]|^2$ SINR$_k$. Here, the average SINR of user k is given as SINR$_k = Pd_k^{-\alpha}/(\sigma^2 + I_0)$, assuming the simplified Okumura-Hata [58, 59] propagation model where P denotes the constant transmission power in Watts, d_k the distance between user k and the BS in meters, α the path-loss exponent, σ^2 the average thermal noise power, and I_0 the average power of interference. Using this model for instantaneous SINR, allows to calculate the spectral efficiency in bits/s/Hz of an arbitrary user k per time slot n as $r_k[n] = \log_2(1 + \gamma_k[n])$ in bits/s/Hz.

5.2.2 Scheduling Assumptions

At every time slot, the scheduler decides how to allocate the wireless channel resources (e.g., time-frequency blocks) among the operators. This dynamic resource allocation is illustrated in Figure 5.2. We assume that infrastructure provider and operators agree on constant sharing ratios a priori, e.g., within a service level agreement. We define the sharing ratios for each arbitrary operator j to be a continuous variable $\tilde{S}_j \in (0, 1]$. Although practical sharing ratios are usually a discrete fraction of resource blocks, this continuous model assures simplicity and comes at no loss in generality when bandwidth is large [60].

Assuming dynamic sharing, we allow the scheduler to deviate from the agreed sharing ratios for a certain time so the utility can converge to ergodic capacity [61]. For an arbitrary time slot n, we use $s_j[n] \in [0, 1]$ to denote the instantaneous sharing ratio assigned to operator j and $x_k[n] \in [0, 1]$ to denote the instantaneous sharing ratio assigned to user k. To control the degree of dynamic sharing, we propose to limit the deviation from the agreed sharing ratio to an interval, where $\Delta \in [0, 1]$ defines the *maximum deviation*. This parameter defines the interval size within which the allocated resources may vary. Thus, choosing D allows GRS to trade off sharing guarantees versus spectral efficiency. The resulting interval within which a moving average of $s_j[n]$ may vary is $[\max(\tilde{S}_j - \Delta, 0), \min(\tilde{S}_j + \Delta, 1)]$. This moving average evaluated over a *window size* of $W \in \mathbb{N}_+$ slots. This parameter W defines the time window within which deviation from the agreed sharing ratio is allowed.

Finally, we assume that the scheduler aims at maximizing a sum of the continuous concave utility function $f : \mathbb{R}_+ \times [0, 1] \to \mathbb{R}_+$. We assume this utility function to depend on the spectral efficiency $r_k[n] \in \mathbb{R}_+$ of the users, which is defined as above. This general model is common in wireless resource allocation and covers important scheduling policies such as Proportional-Fair or Maximum-Rate as special cases.

Using the previous definitions and notation, we now introduce GRS as the multi-operator scheduling problem

$$
\begin{aligned}
\max_{\mathbf{x}} \quad & \sum_{k \in \mathcal{K}} f(r_k[n], x_k[n]) && && \text{(5.1a)} \\
\text{s.t.} \quad & \sum_{k \in \mathcal{K}} x_k[n] = 1 && && \text{(5.1b)} \\
& \sum_{k \in \mathcal{K}_j} x_k[n] = s_j[n] && \forall\, j \in \mathcal{J} && \text{(5.1c)} \\
& \varepsilon_j[n] = \left(\frac{1}{W} \sum_{i=n-W+1}^{n} s_j[i] \right) - \tilde{S}_j && \forall\, j \in \mathcal{J} && \text{(5.1d)} \\
& -\Delta \le \varepsilon_j[n] \le \Delta && \forall\, j \in \mathcal{J} && \text{(5.1e)} \\
& x_k[n] \ge 0. && \forall\, k \in \mathcal{K} && \text{(5.1f)}
\end{aligned}
$$

According to this formulation, the scheduler aims to maximize the sum of a continuous concave utility function f by allocating the resource fraction $\mathbf{x} = (x_1, \dots, x_{|\mathcal{K}|})$ per time slot n. Auxiliary variables are $\mathbf{s} = (s_1, \dots, s_{|\mathcal{J}|})$ and

$\varepsilon = (\varepsilon_1, \ldots, \varepsilon_{|\mathcal{J}|})$, which depend on $\mathbf{x}$. Due to the concavity of f and due to the linear constraints, the optimization problem is concave and can be efficiently solved with standard algorithms [62].

5.2.3 Constraints of GRS

Let us now discuss the constraints in detail. Constraint (5.1b) assures full use of the wireless channel resources by assuring that all components of $\mathbf{x}$ aggregate to one. Constraint (5.1c) ensures that all users of arbitrary operator j receive a resource fraction that matches the instantaneous sharing ratio $s_j[n]$.

Constraint (5.1d) sets $\varepsilon_j[n]$ equal to the difference between the average sharing ratio of operator j over time window W and the agreed sharing ratio $\tilde{S}_j$. For $\varepsilon_j[n] > 0$, the users served by operator j receive more than $\tilde{S}_j$. For $\varepsilon_j[n] < 0$, the users receive less resources than agreed. Note that this constraint is defined recursively, since the moving average $\varepsilon_j[n]$ depends on the previous values $s_j[i]$ within time slots $i = n - W + 1, \ldots, n - 1$.

By defining lower and upper bound of $\varepsilon_j[n]$, constraint X (5.1e) allows to control the deviation from the agreed sharing ratio through parameter Δ. Finally, constraint (5.1f) imposes the sharing values to be non-negative.

5.2.4 Effect of Maximum Deviation Δ on GRS

Assuming W to be constant, the proposed GRS formulation allows us to implement a variety of multi-operator scheduling policies by choosing the maximum deviation Δ. Two relevant special cases are called Fixed Sharing and Free Sharing.

Fixed Sharing refers to multi-operator scheduling policies with a constant instantaneous sharing ratio. Keeping the sharing ratio fixed assures that the operator receives the agreed fraction of resources for the complete operation time. In (5.1a), this is achieved by choosing $\Delta = 0$ which forces $\varepsilon_j[n] = 0$ for every operator j in every time slot n. Since now the scheduler has to exactly meet the agreed sharing ratio at every point in time, users have to be served even if their low channel gain reduces the overall sum utility. Thus, the strict guarantees of Fixes Sharing result into a high cost for overall efficiency. We will now proof that the rate-dependent utility reached by Fixed Sharing is, in fact, the lower bound for the utility of all alternative GRS-scheduling policies.

Proposition 5.2.4.1. "Fixed Sharing" is a lower bound for GRS. $\triangleleft$

Proof. Let $F := \{\mathbf{x}, \mathbf{s}, \varepsilon \text{ s.t. } (5.1\text{b}), (5.1\text{c}), (5.1\text{d}), (5.1\text{e}), (5.1\text{f})\}$ be the set of constraints of formulation (5.1). Since, in the fixed sharing scheduling, $\Delta = 0$, the inequality constraints (5.1e) converge to

$$\varepsilon_j[n] = 0 \qquad \forall\, j \in \mathcal{J}. \tag{5.2}$$

□

Then, the fixed sharing problem is equivalent to

$$\max\{f(\mathbf{r}, \mathbf{x}) : \mathbf{x} \in F'\}$$

where $F' := \{\mathbf{x}, \mathbf{s}, \varepsilon \text{ s.t. } (5.1\text{b}), (5.1\text{c}), (5.1\text{d}), (5.1\text{f}), (5.2)\}$. Since set F' is a subset of set F, i.e., $F' \subseteq F$,

$$\max\{f(\mathbf{r}, \mathbf{x}) : \mathbf{x} \in F'\} \leq \max\{f(\mathbf{r}, \mathbf{x}) : \mathbf{x} \in F\}.$$

Free Sharing refers to a scheduler choosing any sharing ratio that maximizes the objective function in (5.1). This case is realized by choosing $\Delta \geq \tilde{\Delta} := \max_{j \in \mathcal{J}}\{\tilde{S}_j, 1 - \tilde{S}_j\}$. This choice of Δ imposes the least restrictions on the scheduler and, thus, represents the optimal parametrization in terms of utility. However, it provides no guarantees on the sharing ratio at a certain instant in time. Before we prove that the sum utility reached by Free Sharing upper bounds the utility of all other GRS-scheduling policies, we introduce two remarks.

Lemma 5.2.4.2. Variable $\varepsilon_j[n]$ is by definition in $[-\tilde{S}_j, 1 - \tilde{S}_j]$.

Proof. Substituting constraint (5.1c) in constraint (5.1d), one can write $\varepsilon_j[n]$ as:

$$\varepsilon_j[n] = \left(\frac{1}{W} \sum_{i=n-W+1}^{n} \sum_{k \in \mathcal{K}_j} x_k[i] \right) - \tilde{S}_j$$

The term in the round brackets is the average over W time slots of a sum of W terms, such that $x_k[i] \leq 1$, $\forall k, i$, and its value is in $[0, 1]$. Therefore, the minimum value of $\varepsilon_j[n]$ is $-\tilde{S}_j$ and its maximum value is $1 - \tilde{S}_j$.

Lemma 5.2.4.3. Constraints (5.1e) are irrelevant for GRS when $\Delta \geq \tilde{\Delta}$.

Proof. This follows directly from Remark 5.2.4.2. Indeed, $\varepsilon_j[n] \leq 1 - \tilde{S}_j \leq \max_{j\in\mathcal{J}}\{\tilde{S}_j, 1-\tilde{S}_j\} \leq \Delta$ and $\varepsilon_j[n] \geq -\tilde{S}_j \geq -\max_{j\in\mathcal{J}}\{\tilde{S}_j, 1-\tilde{S}_j\} \geq -\Delta$. In other words, this means that constraints (5.1e) define a range for $\varepsilon_j[n]$ which is larger than the range in which this variable lays by definition. □

Proposition 5.2.4.4. "Free Sharing" is an upper bound for GRS.

Proof. As shown in Remark 5.2.4.3, constraints in (5.1e) are irrelevant for GRS when $\Delta \geq \tilde{\Delta}$. Since variable $\varepsilon_j[n]$ is not needed any longer, also auxiliary constraints (5.1c) and (5.1d) can be removed. Therefore, the free sharing problem is equivalent to

$$\max\{f(\mathbf{r}, \mathbf{x}) : \mathbf{x} \in F''\}$$

where $F'' := \{\mathbf{x} \text{ s.t. constraints (5.1b), (5.1f)}\}$. Since F'' is a superset of F, i.e., $F'' \supseteq F$,

$$\max\{f(\mathbf{r}, \mathbf{x}) : \mathbf{x} \in F''\} \geq \max\{f(\mathbf{r}, \mathbf{x}) : \mathbf{x} \in F\}.$$ □

Corollary 5.2.4.5. The system utility in the fixed sharing and in the free sharing cases is independent of W.

Proof. For the fixed sharing, this comes directly from the fact that the resource assignment is fixed, independently of the window size. For the free sharing, observe that the constraints in set F'' (proof of Prop. 5.2.4.4) do not contain W. □

5.2.5 Effect of Window Size W on GRS

In this section, we will analyze how choosing the window size W for the moving average (5.1d) affects the utility of GRS.

We will prove that the utility function f is non-monotonic w.r.t. W, making it not straightforward to characterize a deterministic dependency of the utility on W. Consequently, we choose a probabilistic approach where we will characterize the probability that the utility increases with W. Our proof holds for the cases $W = 1$ and $W = 2$, which will be indicated as superscripts. In order to compare both cases, we will calculate $U_{TOT} = U[n] + U[n+1]$, assume $\mathcal{J} = \{1, 2\}$, and the sum rate utility $U[n] = f(r_1[n], r_2[n], s_1[n], s_2[n]) = r_1[n]s_1[n] + r_2[n]s_2[n]$. Therein, the random variable $i[n]$ refers to the rate of the user served by operator i at time slot n and is defined as in 5.2. We assume $\tilde{S}_1 \leq \frac{\tilde{S}_2}{2}$ and $s_j[n-1] = \tilde{S}_j$.

Theorem 5.2.5.1. When $0 < \Delta < \tilde{S}_1$, the objective function (5.1a) is not decreasing with W (i.e., not decreasing with $W = 2$, w.r.t. $W = 1$), with probability

$$\Pr(\mathcal{U}) = \Pr(\mathcal{E}_1)\Pr(\mathcal{A})^2 + \\ + 2\Pr(\mathcal{A})(1 - \Pr(\mathcal{A})) + \Pr(\mathcal{E}_2)(1 - \Pr(\mathcal{A}))^2 \quad (5.3)$$

◁

where

$$\mathcal{U} = U_{TOT}^{W=2} \geq U_{TOT}^{W=1}$$

$$\mathcal{A} = r_1[n] \geq r_2[n]\} \ , \ \bar{\mathcal{A}} = \{r_1[n] < r_2[n]$$

$$\mathcal{E}_1 = r_1[n] - r_2[n] \geq r_1[n+1] - r_2[n+1] | \mathcal{A}[n], \mathcal{A}[n+1]$$

$$\mathcal{E}_2 = r_1[n] - r_2[n] < r_1[n+1] - r_2[n+1] | \bar{\mathcal{A}}[n], \bar{\mathcal{A}}[n+1].$$

Theorem 5.2.5.2. When $\tilde{S}_1 \leq \Delta < \tilde{S}_2$, the objective function (5.1a) is not decreasing with W with probability

$$\Pr(\mathcal{U}) = 1 + \Pr(\mathcal{E}_1)\Pr(\mathcal{A})^2 - \Pr(\mathcal{A})^2. \quad (5.4)$$

◁

Theorem 5.2.5.3. The objective function (5.1a) is not decreasing with W with probability 1 when $\Delta \geq \tilde{S}_2$.

The detailed proof of Theorem. 5.2.5.1 is given in the appendix of [55] and here only the rationale for the proofs of Theorems 5.2.5.2 and 5.2.5.3 are added which follow the same principle.

Corollary 5.2.5.4. The probability, $\Pr(\mathcal{U})$, that system Utility with $U[n] = r_1[n]s_1[n] + r_2[n]s_2[n]$, increases with $W = 2$, w.r.t $W = 1$, when $\mathcal{J} = \{1, 2\}$ and $s_j[n-1] = \tilde{S}_j$ is independent of Δ.

We can conclude that, in general, the utility is non-monotonic in W. Even if a larger W increases the degree of freedom of the scheduling problem (5.1), an increase of the sum utility cannot be guaranteed. This operational insight can be explained by the temporal correlation of the resource allocation. Using its complete fraction of resources at time slot n may prevent an operator to use any resources at the following time slot $n + 1$. In this extreme case, a larger value of $W > 1$ reduces spectral efficiency, i.e., utility.

5.3 Results and Analysis

In this section, we will discuss some numerical results that show how the parameters Δ and W affect the utility of (5.1).

5.3.1 Parameters and Studied Scenarios

We consider the downlink of a single BS shared among $|\mathcal{J}| = 2$ operators. We assume that users are uniformly distributed over the covered area. Simulations are repeated for 5000 independent, random user distributions. We assume that the scheduler maximizes the utility function $f(r_k[n], x_k[n]) = w_k r_k[n] x_k[n]$ and select the QoS weight $w_k = 1 \; \forall \; k \in \mathcal{K}$, which reflects the Maximum-Rate scheduling [63].

In the first studied scenario, all users have equal average SINR. This reflects the widely-used ring scenario [43] where all users are 100 m away from the BS (i.e., $d_1 = d_2 =$100). In our simulations, we set the average SINR$_k$ = 24 dB $\forall \; k \in \mathcal{K}$ to reflect this case. In the following scenarios, we keep the SINR for the users of operator 1 but reduce the SINR for the users of operator 2. This reflects the users being placed on two concentric rings around the base station with radius $d_1 < d_2$. In our second scenario, choosing $d_2 =$150 m reduces the SINR to SINR$_k$ = 17.8 dB $\forall \; k \in \mathcal{K}_2$ for the users of operator 2. In the third scenario, we choose $d_2 =$500 m which further reduces the SINR to SINR$_k = 0$ dB $\forall \; k \in \mathcal{K}_2$.

5.3.2 Performance Results

Figure 5.3 shows the performance of the proposed GRS as the CDF of the aggregate spectral efficiency for different choices of Δ. We assume $\tilde{S}_1 = 0.25$ and $\tilde{S}_2 = 0.75$ and $W = 1$. The results show that the proposed GRS increases the system performance with respect to fixed sharing ($\Delta = 0$). Comparing the results for $d_2 =$100 m to those for $d_2 =$500 m shows that the gain increases in scenarios with non-uniform SINR.

Figure 5.4 shows the average aggregate spectral efficiency for the three scenarios described above. The figure illustrates a clear trend for this utility measure w.r.t. the parameters Δ and W. As expected, minimum utility is reached for $\Delta = 0$, while the maximum value is reached for any $\Delta \geq \tilde{\Delta} = 0.75$.

The results for the scenarios $d_2 = 100$ or $d_2 = 150$ also indicate that higher W increase the spectral efficiency. However, this increase is not shown for

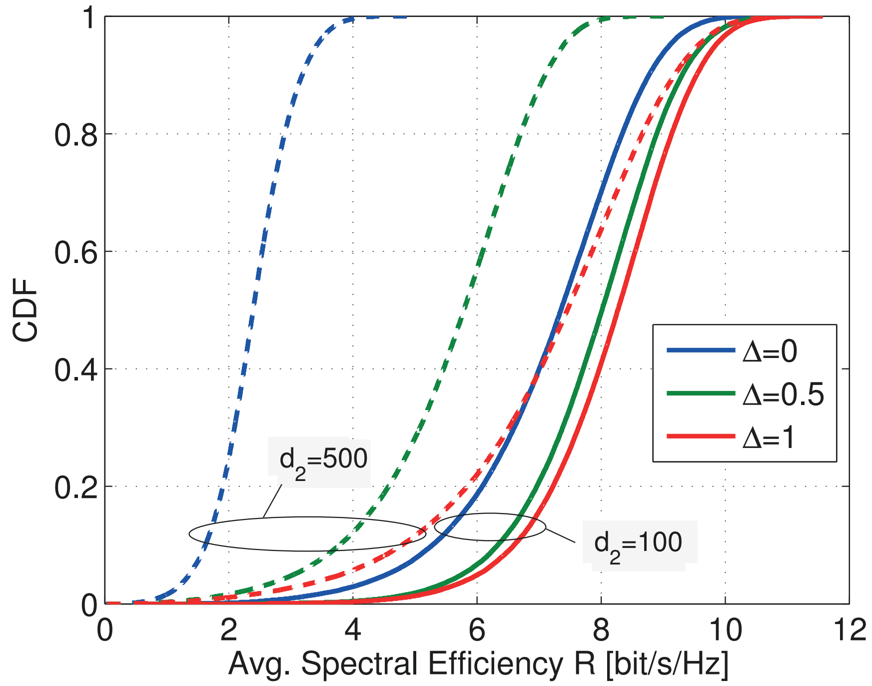

Figure 5.3: GRS: CDF of the aggregate spectral efficiency for two different placements for the users of operator 2, $W = 1$, and $\tilde{S}_1 = 0.25$ from [55]

$d_2 = 500$ where the spectral efficiency is not affected by the window size.

To explain this effect, Figure 5.5 presents some numerical results on the probability of the event $\mathcal{U}$ considering the scenario from Theorems 5.2.5.1, 5.2.5.2, and 5.2.5.3. The first plot of this figure compares the three different placements for the users of operator 2 for $\tilde{S}_1 = 0.25$ and $\tilde{S}_2 = 0.75$. The lower plot compare different values of $\tilde{S}_1$ when all users are equally placed using $d_1 = d_2 =$100 m radius.

We conclude that theoretical and numerical values coincide in confirmation to our proofs. As expected from the theoretical results, Figure 5.5 shows a stepwise probability function $\Pr(\mathcal{U})$ that changes its value at $\Delta = \tilde{S}_1$ and $\Delta = \tilde{S}_2$. We observe further that this probability is always greater than 0.5 when the SINR of the users for operator 2 is not too small (i.e., $d_2 = 100$ and $d_2 = 150$ m). This probability value explains the average utility gain when increasing W to 2 for the respective scenarios, as illustrated in Figure 5.4. In contrast, the probability is exactly 0.5 for the third scenario ($d_2 =$500 m) and for $\Delta < \tilde{\Delta}$. Consequently, all the curves for these cases coincide in Figure 5.4.

Finally, when Free Sharing is chosen by $\Delta \geq \tilde{S}_2$, $\Pr(\mathcal{U})$ is always one (cp.

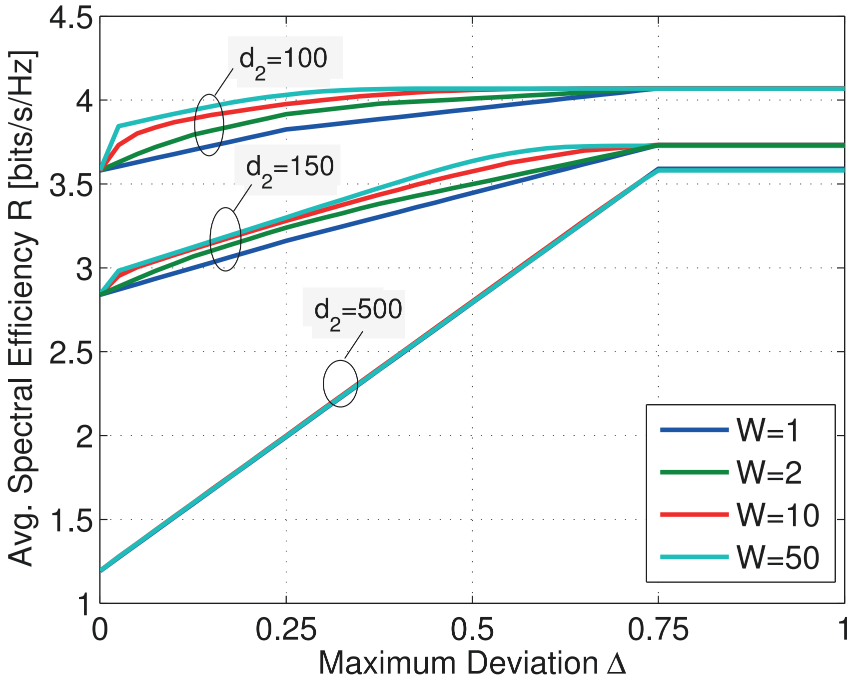

Figure 5.4: GRS:Average aggregate spectral efficiency for three different placements for users of operator 2, different W, and $\tilde{S}_1 = 0.25$ from [55]

Theorem 5.2.5.3). This observation corresponds to the fact that the utility does neither decrease nor increase with W, which leads to coinciding curves for $\Delta \geq \tilde{\Delta} = 0.75$ in Figure 5.4.

5.4 Conclusion

With GRS a new scheduling approach is analyzed which shares wireless channel resources among the users of multiple operators. By controlling the maximum deviation parameter Δ, this approach allows to trade-off sharing guarantees versus performance.

It is proved that controlling Δ allows to realize a variety of scheduling policies. Fixed Sharing guarantees the agreed sharing ratio to the operators but represents the lower bound for spectral efficiency while Free Sharing provides maximum spectral efficiency at no guarantees. The analysis of further important effects of Δ is complemented with a probabilistic analysis on the time span W over which sharing is performed.

The result is a profound understanding of how multi-operator sharing af-

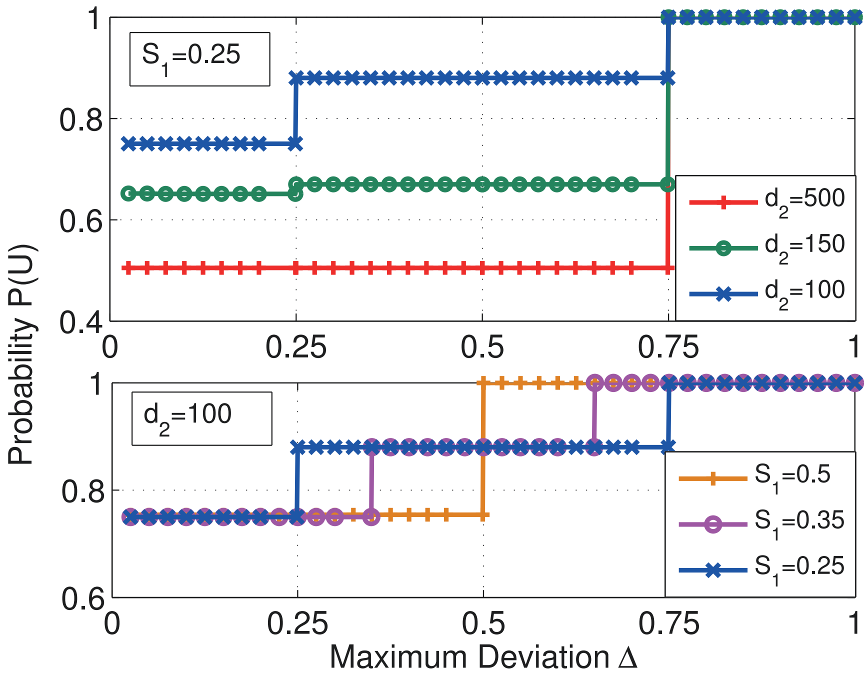

Figure 5.5: GRS: Probability of the event $\mathcal{U}$ for different values of $\tilde{S}_1$ and different distribution of the rates for users of operator 2 from [55]

fects spectral efficiency or similar utility functions. The fact that such operational understanding can be given as rigorous mathematical proof shows the full strength of the presented framework for multi-operator scheduling.

Part III

Optimized Fairness Triggered Algorithms

Chapter 6

Multi-Operator Multi-User MIMO (Single Objective)

6.1 Introduction

Spectrum sharing is seen as one key element to achieve significantly higher spectral efficiency to satisfy the ambitious goals of 5G wireless systems. In the previous chapters, we have presented scheduling approaches for resource sharing to improve spectral efficiency and enhance network coverage with smaller operational and investment costs. One of the other key service parameters in shared environments is to provide the agreed fairness between the partners. In this chapter, a new approach is proposed which guarantees the achievement of agreed fairness criteria to the smallest available granularity of time-frequency dimensions. This solution is based on the multi-user MIMO approach for wireless network sharing. We use multiple antennas for spatial multiplexing of multiple operators [64]. Our approach is novel as we use the spatial domain multiplexing of multiple operators which has never been exploited for MO-CRRM in one base station which is also shared by the operators. We call this approach as the multi-operator multi-user MIMO (MOMU-MIMO). Therefore, our proposal is applicable to all possible radio access interfaces, e.g., TDMA, FDMA, CDMA and orthogonal frequency division multiple access (OFDMA). The operation is well illustrated in Figure 6.1.

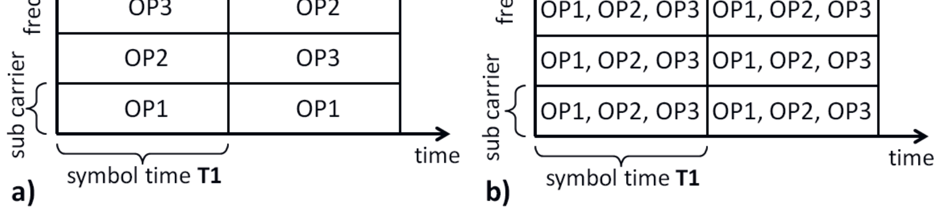

Figure 6.1: a) Current sharing methods for multiple operators b) MOMU-MIMO based sharing approach from [64]

6.2 System Model

The new concept for MO-CRRM presented in this chapter is based on the spatial dimension of the radio channel in combination with time, frequency and power domain. The primary objective is to consider SLAs that equally share the time and frequency resources among the operators and provide fairness using the BS transmit power. In [65], a spectrum sharing approach is presented for co-located transmitters of multiple operators. The transmitters of different operators do not share the user data; they rather only share the channel state information (CSI). Using the shared CSI, a coordinated beamforming is designed to mitigate the inter operator interference. In contrast to [65], a full sharing scenario is considered, in which BS hardware and spectrum are shared. Therefore, it is assumed that the transmitter has access to the user data (i.e., CSI and SNR) of all the shared operators.

We consider an orthogonal frequency division multiplexing (OFDM) based closed loop multi-user MIMO downlink system operated by a neutral service provider. The service area is covered by a BS equipped with $(M_t > 1)$ transmit antennas which serves single antenna users with $(D_t \leq M_t)$ independent data streams over a single OFDM time-frequency resource element. Let $\mathcal{L}$ be the set of indices of all the active user equipment (UE)s in the coverage area of the BS such that $(|\mathcal{L}| \geq M_t)$. The BS performs user selection, with $\mathcal{S}$ be the set which contains the indices of the selected users such that $(\mathcal{S} \subseteq \mathcal{L}, |\mathcal{S}| = K, K \leq M_t)$. Assume that the users in the coverage area of the BS are associated with different cellular network operator entities. Let $\mathcal{J}$ be the set which contains the indices of the network operators $(1 < |\mathcal{J}| \leq I, I \leq M_t)$. The signal received by the mth UE associated with the operator $i \in \mathcal{J}$ served by the BS can be represented by $y_m \in \mathbb{C}$. Using the narrow-band OFDM assumption it can be written in discrete time form as follows:

$$y_m = \mathbf{h}_m^T \mathbf{x} + n_m \tag{6.1}$$

The symbol $\mathbf{h}_m \in \mathbb{C}^{M_t \times 1}$ represents the channel vector between the BS and the mth UE and is given by, $\mathbf{h}_m^T = [h_{m,1} \quad h_{m,2} \quad h_{m,3} \quad \dots \quad h_{m,M_t}]$. The entries of $\mathbf{h}_m$ are modeled as identically and independently distributed zero-mean circularly symmetric complex Gaussian variables. We assume a quasi-static block flat-fading channel over a symbol transmission time on a subcarrier. The time and frequency indices are dropped due to the notational convenience. We assume perfect channel estimation by the user and a perfect feedback of the channel vector from each active user to the BS. The symbol $\mathbf{x}$ is the transmit signal vector which can be written as the linear combination of all the symbols transmitted by the BS to the users in $\mathcal{S}$ and is given by,

$$\mathbf{x} = \sum_{m \in \mathcal{S}, m=1}^{K} \mathbf{b}_m \sqrt{p_m} s_m, \tag{6.2}$$

where, $\mathbf{b}_m \in \mathbb{C}^{M_t \times 1}$ is the unit norm precoding vector designed for the mth UE, p_m is the allocated power and s_m is the information symbol for the respective UE. We assume a BS power constraint ,i.e., $E[\mathbf{x}\mathbf{x}^*] \leq P_t$. The noise at the receive antenna of the mth UE is represented by n_m and is modelled as white Gaussian with zero mean and variance σ^2. The substitution of (6.2) in (6.1) gives us the following form for received signal:

$$y_m = (\mathbf{h}_m^T \mathbf{b}_m) \sqrt{p_m} s_m + \sum_{n \neq m, \forall n \in \mathcal{S}} (\mathbf{h}_m^T \mathbf{b}_n) \sqrt{p_n} s_n + n_m \tag{6.3}$$

The first term in (6.3) is the desired signal by the mth UE but the second term is the undesired multi-user interference. We use the well-known zero-forcing (ZF) precoding to get rid of this interference. The ZF precoding vector can be design to fulfill the following constraints:

$$(\mathbf{h}_m^T \mathbf{b}_m) = \delta_m, \delta_m \neq 0, m \in \mathcal{S} \tag{6.4}$$

$$(\mathbf{h}_m^T \mathbf{b}_n) = 0, \forall n \in \mathcal{S} \tag{6.5}$$

The constraint in (6.5) can be achieved when $\mathbf{b}_m^*$ lies in the orthogonal compliment of the subspace spanned by all $\mathbf{h}_n, \forall n \in \mathcal{S}, n \neq m$. For this purpose, the BS uses the system channel matrix $\mathbf{H} \in \mathbb{C}^{M_t \times K}$ mth row of which consists of the row channel vector received by the BS from the mth selected user, it can be written as $\mathbf{H} = [\mathbf{h}_1^T \;\; \mathbf{h}_m^T \;\; \ldots \;\; \mathbf{h}_K^T]^T \in \mathbb{C}^{K \times M_t}$. The pseudo inverse of $\mathbf{H}$ is represented by $\tilde{\mathbf{H}} \in \mathbb{C}^{M_t \times K}$, the mth column of $\tilde{\mathbf{H}}$ represents the ZF precoding vector that fulfills the constraints in (6.4) and (6.5).

$$\tilde{\mathbf{H}} = \mathbf{H}^H (\mathbf{H}\mathbf{H}^H)^{-1} = [\mathbf{b}_1 \;\; \mathbf{b}_m \;\; \ldots \;\; \mathbf{b}_K] \tag{6.6}$$

Now using (6.4) and (6.5) the received signal from (6.3) can be written as:

$$y_m = \delta_m \sqrt{p_m} s_m + n_m. \tag{6.7}$$

If the maximization of system spectral efficiency is the only objective, then the selection of $\mathcal{S}$ and allocation of p_m are jointly optimized with the help of system sum-rate maximization and water-filling. However, in this case, the agreed fairness among the UEs and their associated operators are completely neglected. In the next section, we explain how do we achieve the agreed fairness among the operators by the selection of the UEs in the set $\mathcal{S}$ and the allocation of power.

6.3 Multi-Operator Multi-User MIMO

As explained in Section (6.2), the active users in the service area of the BS are associated with multiple operators. So the main objective and first priority of the service provider in the coverage area is to fulfill the fairness criteria and comply with the SLA. We divide the fairness aspect of SLA into two parts.

6.3.1 MOMU-MIMO based user grouping to fulfil SLA

The first part of SLA is based on user grouping which ensures that all the operators are simultaneously served with the help of spatial multiplexing. We create user candidate sets out of the super set $\mathcal{L}$. Let $\mathcal{L}_i$ be the set which contains the indices of the active UEs associated with the ith operator such that $(\mathcal{L}_i \subset \mathcal{L})$. Let $\mathcal{S}_c$, $(|\mathcal{S}_c| = K)$ be one of the candidate set of users which fulfils the fairness criteria and it is given by:

$$\begin{aligned} \mathcal{S}_c &= \{\mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3 \cup \ldots\ \mathcal{S}_I\} \\ \text{for any } \{i,j\} \in \mathcal{J} &\ : (\mathcal{S}_i \cap \mathcal{S}_j) \end{aligned} \tag{6.8}$$

Where $\mathcal{S}_i, \mathcal{S}_j$ correspond to the operators i and j, respectively, such that $(\mathcal{S}_i \subseteq \mathcal{L}_i)$ and $(\mathcal{S}_j \subseteq \mathcal{L}_j)$. Each set in the union for $\mathcal{S}_c$ is independent or mutually exclusive with any other set. The constraint in (6.8) nothing but ensures that all the operators are simultaneously served.

6.3.2 SLA based utility optimization in MOMU-MIMO

The second part is based on the allocation of the power to the users in the candidate sets under different SLA fairness constraints. The fairness constraints differ in their agreed fairness criteria and correspondingly in the shared power resource for spatially multiplexed users of different operators. The details of different fairness criteria and accordingly the power allocation are presented in Section 6.3.4.

6.3.3 User group selection for higher system spectral efficiency

After fulfiling the SLA fairness aspects, we want to reach the highest spectral efficiency. Therefore, we select one of the candidate set for the transmission such that the system sum rate is maximized to reach the ergodic capacity [61]. The power had been already allocated to each user candidate set according to the second part of the SLA.

$$R(\mathcal{S}) = \max_{\forall \mathcal{S}_c} R(\mathcal{S}_c), \tag{6.9}$$

where $\mathcal{S}$ represents the set of users selected for transmission.

6.3.4 Algorithms to achieve different criteria

In the following, we describe our algorithms for power allocation to the candidate users sets. Without loss of generality and for the convenience of notation, from now on we assume that $M_t = 2, I = 2, \mathcal{S} = \{m, n\}, m \in \mathcal{L}_1, n \in \mathcal{L}_2, |\mathcal{S}| = K = 2$. Let g_1 and g_2 be the two input factors which represent the agreed service share of operator 1 and 2 respectively such that $(g_1 \leq g_2)$ and $(g_1 + g_2) = 1$. The sum rate for the selected set can be written as,

$$R(\mathcal{S}) = R_m + R_n, \tag{6.10}$$

where, R_m=$\log_2$ $(1 + \frac{\delta_m}{\sigma^2} p_m)$ and R_n=$\log_2$ $(1 + \frac{\delta_n}{\sigma^2} p_n)$ under the following constraints:

$$(p_m + p_n) \leq P_t \tag{6.11}$$

$$p_m, p_n \geq 0 \tag{6.12}$$

Where, p_m and p_n represent the allocated powers and $\frac{\delta_m}{\sigma^2}$ and $\frac{\delta_n}{\sigma^2}$ are the signal to noise ratios respectively for the users m and n.

Criteria for Agreed Rate Fairness

The first agreement that we consider is on the achieved rate level. The state of the art [53, 57] fulfill rate fairness criterion which is based on a minimum rate demand. However, we consider an agreement where the operators have a guaranteed rate demand for each time-frequency resource element. With the service shares g_1 and g_2 for operator 1 and operator 2 the agreed rate conditions can be given as:

$$R_m = g_1 R(\mathcal{S}) \tag{6.13}$$

$$R_n = g_2 R(\mathcal{S}) \tag{6.14}$$

A relaxation of the service share can be achieved introducing a Tolerance Value λ with a value greater than zero, e.g., as seen later in the *Algorithm MOMU-MIMO* in step 13) given in Appendix C. As the rate of the UEs is a result of the logarithmic function of output signal to noise ratio, a linear conversion of the given operator service share to the allocated power share is not possible. Therefore, we propose an iterative power allocation algorithm based on incremental power to achieve the desired rate fairness as in (6.13) and (6.14) presented in Appendix C.

Criteria for Agreed Power Fairness

Here, we consider the SLA where the operators share the power as the resource according to the given service share over a shared time-frequency sample. The agreement conditions can be written as:

$$p_m = g_1 P_t \tag{6.15}$$

$$p_n = g_2 P_t \tag{6.16}$$

The algorithm to achieve this agreement is straight forward. We follow the **Step 1** (creation of candidate sets) and **Step 4** (selection of set $\mathcal{S}$ according to (6.9)) as in *Algorithm MOMU-MIMO*, see Appendix C. In **Step 3**, we allocate the power to all the possible candidate sets such that (6.15) and (6.16) are satisfied.

Criteria for System Spectral Efficiency (Water Filling)

Here we consider the case where the SLA is applied only on the time-frequency resource sharing among the operators. Therefore, the service provider is free to allocate power. Hence, we use the well-known classical water-filling (WF) based power allocation in **Step 3** which is optimum for system spectral efficiency. The **Step 1** and **Step 4** are as in *Algorithm MOMU-MIMO* given in Appendix C.

6.4 Results and Analysis

We consider three different SLAs between the shared operators and the service provider. These agreements differ with respect to the pre-defined system utility optimization. The utilities we consider are the achievement of fairness among operators and the optimization of system spectral efficiency. So, we propose new algorithms for the achievement of these utilities. The primary system utility (fairness) that we consider is the achievement of agreed service share between the operators. This utility helps us to show the feasibility and proof of MOMU-MIMO basic operation. For the operators it provides a transparent fairness metric. It works as an essential input for decision making processes to adopt network sharing based solutions. The secondary utility is the system spectral efficiency. For the service provider, it is an important performance metric as it shows the efficient use of the invested resources. In this way, we assess the complete system performance and provide key performance indicators for all the concerned parties.

6.4.1 Assumptions and Parameters

We model a coverage area of a BS in a cellular network where the effect of average inter cell interference is included in Gaussian noise as in (6.1). We consider a scenario where two operators are sharing the service in the coverage area of the BS. Each operator has equal number of active users such that $|\mathcal{L}_1| = |\mathcal{L}_2|$. The users are randomly dropped with a uniform distribution in the coverage area. An average input SNR of 10 dB is used for all the users to model the average path-loss. This value represents the nominal average SNR in a universal reuse cellular network. We assume full buffer traffic. As explained before, we employ Rayleigh fading channel. We assume a perfect channel estimation and perfect error-free feedback from all the users. Further salient parameters are summarized in Table 6.1.

In the following, we focus on the operator sharing ratios from 50% to 90% as the other ratios are the complementary one from 50% down to 10% in the case of two operators in the simulated system setup.

Table 6.1: Simulation parameters from [64]

Parameter	**Value**
Channel Realizations	5000
UEs per Operator	1, 5, 10, 25, 50
Service Shares (g_1/g_2)	0.1/0.9, 0.2/0.8, 0.3/0.7, 0.4/0.6, 0.5/0.5
Rate Fair Service Share Tolerance Value $\lambda(\%)$	0, 2, 4, 8

6.4.2 Results for Agreed Fairness as Utility

It is trivial that the basic fairness between the operators over time and frequency resources is perfectly guaranteed through the first step of user selection, as this fairness characteristic is inherited from MOMU-MIMO based sharing through spatial multiplexing. This is well illustrated on the right side of Figure 6.1.

On top of the time and frequency resources, we have considered transmit power as a shared resource. In Section 6.3, we have defined further fairness criteria in terms of SLAs based on the allocation of power between different operators. Two SLAs in Subsection 6.3.2 are based on the data rate fairness and power (resource) fairness. Figure 6.2 shows the achievement of service shares as agreed in the SLA. For the resource fairness, achieving these agreed service shares is straight forward. As given in (6.15) and (6.16), the shares can be linearly transformed into the allocated transmit powers p_m and p_n. The Y-axis in Figure 6.2 presents these achieved shares against the agreed shares on the X-axis. In Figure 6.2, the achieved shares for rate fairness (*Algorithm MOMU-MIMO* given in Appendix C) are also presented with different tolerance values λ. The *Step 3* in *Algorithm MOMU-MIMO* performs the allocation of power to achieve agreed service share of each operator for the rate fairness. We can see from Figure 6.2 that the agreed shares for rate fairness are also perfectly achieved in case the tolerance value for achieved service share is 0%. In fact, the results are overlapping with the resource fairness. If we increase the tolerance value λ, we can see that the achieved service share is less than the agreed one. In exchange the spectral efficiency can be increased for the case $\lambda > 0$ as we will see in the next sub chapter. Further note that the absolute difference in achieved and agreed rate fairness is in the same order as of the value λ for service shares greater than (50%+λ). We can summarize for Figure 6.2 both types of fairness

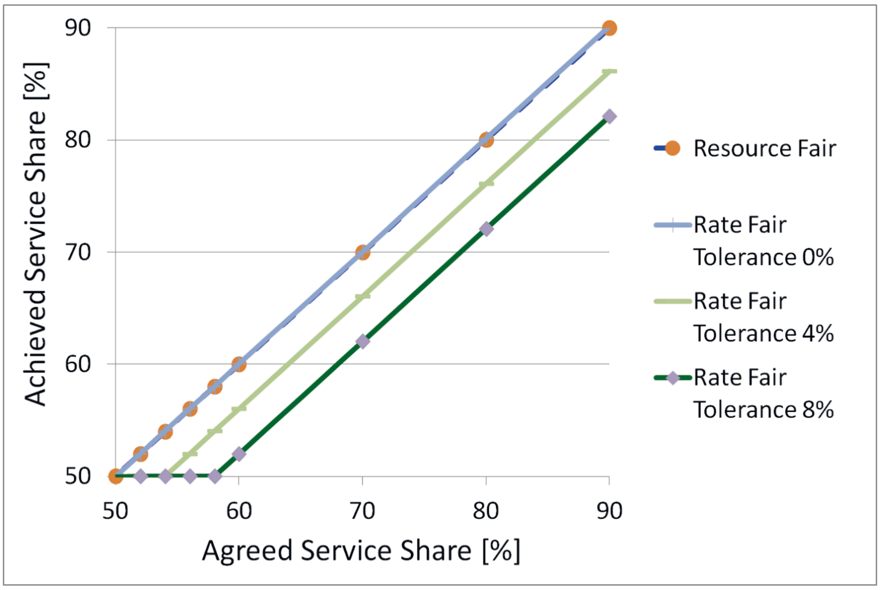

Figure 6.2: Spectral efficiency for different service shares for rate fairness (0% tolerance) and water filling based SLAs from [64]

criteria are supported by the presented MOMU-MIMO approach.

6.4.3 Results for Spectral Efficiency as Utility

The third SLA for power allocation (as in Subsection 6.3.2) is dedicated to optimize the achievable spectral efficiency in a shared MOMU-MIMO system. Therefore, we have used water-filling based power allocation between the operators. This method is optimal in terms of achievable spectral efficiency. Hence, we used it as the baseline to assess the performance of the other two SLAs based on rate and resource fairness.

Figure 6.3 shows the spectral efficiency for the optimal power allocation and for the rate fairness based power allocation. We have considered different service shares for rate fairness with a tolerance value for the agreed service share of 0%. We observe in Figure 6.3 that with the decrease in the difference in service share of operators, the difference in the spectral efficiency between WF (optimal) and the rate fairness algorithm also decreases. The lowest performance can be seen with the share pair $(g_1 = 10\%, g_2 = 90\%)$. In this case with 100 UEs, we have a loss of 45% system spectral efficiency. However, with equal service share $(g_1 = g_2 = 50\%)$ this loss is reduced to 21%. In Figure 6.4 we show the performance results with the tolerance value of 8%. If we com-

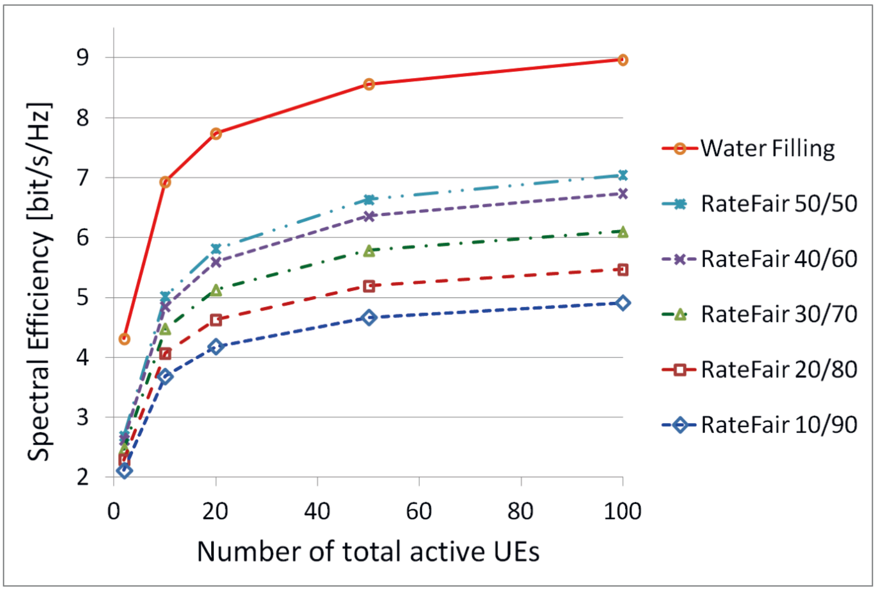

Figure 6.3: Spectral efficiency for different service shares for rate fairness (0% tolerance) and water filling based SLAs from [64]

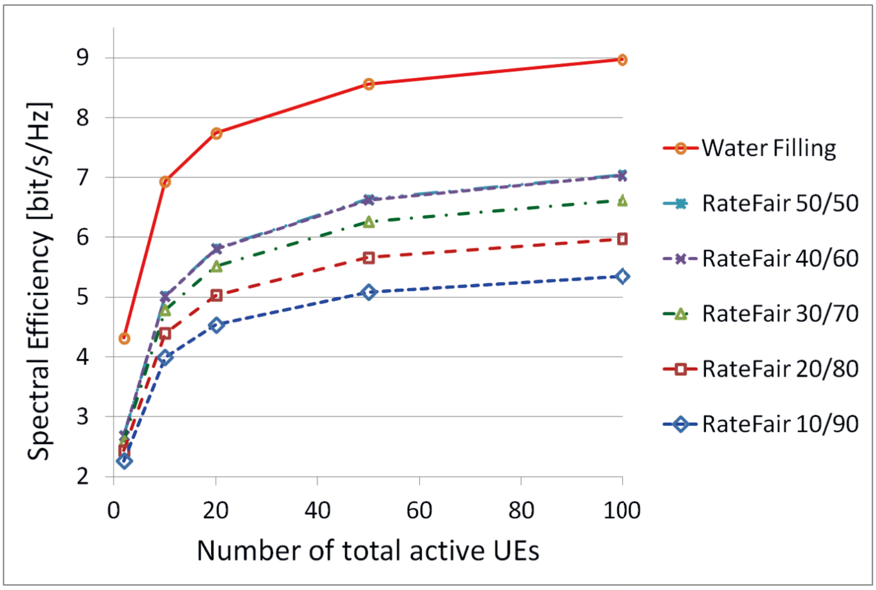

Figure 6.4: Spectral efficiency for different service shares for rate fairness (8% tolerance) and water filling based SLAs from [64]

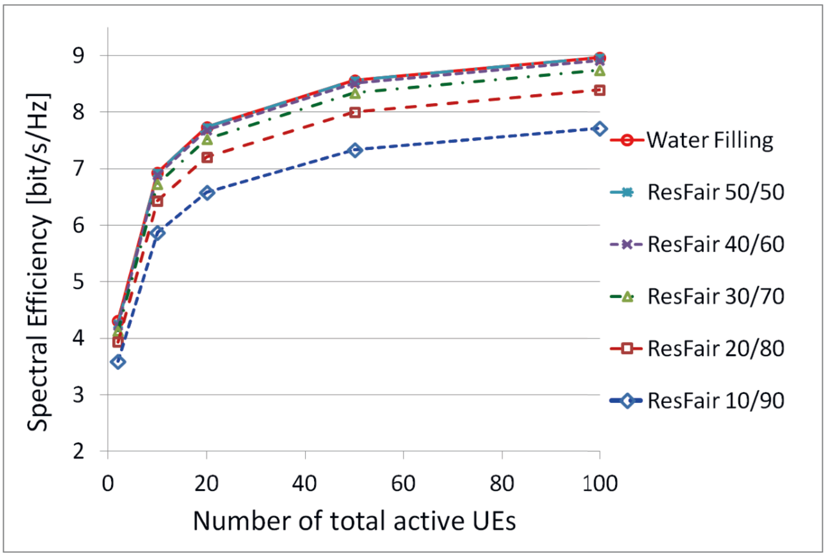

Figure 6.5: Spectral efficiency for different service shares for resource fairness and water filling based SLAs from [64]

pare these results with Figure 6.3, we can observe that with $\lambda = 8\%$ there is an increase in spectral efficiency for all service shares except $RateFair$ $50/50$. This is because the increase in the tolerance of service agreement compromises the achieved service share which is already shown in Figure 6.2. Let us focus on the service share of $RateFair$ $10/90$ with 100 active UEs, the increase in spectral efficiency as compared to Figure 6.3 is of the order of the tolerance value $\lambda = 8\%$. This result is also compliant with the results shown in Figure 6.2. Generally one can observe that the spectral efficiency performances converge to the results of the $RateFair$ $50/50$ service share result. This is the maximum rate fairness condition which is also given in Step 20) of *Algorithm MOMU-MIMO*.

Figure 6.5 presents the spectral efficiency for resource fairness based power allocation. We can see that also here, the spectral efficiency has the same behaviour for the difference in service share as in rate fairness approach. However, the comparative losses are much lesser than rate fairness because the service shares are linearly coupled with the resource which is power. We see that for $(g_1 = 10\%, g_2 = 90\%)$ and 100 UEs the performance loss is only 13%. This is the maximum loss against the maximum difference in service share. If we look at $(g_1 = 30\%, g_2 = 70\%)$ service share, it is almost overlapping the optimal spectral efficiency without any compromises to the achieved service share as shown in Figure 6.2 for the resource fair case.

6.5 Conclusion

In this chapter a fundamentally new approach for the sharing of network resources among multiple operators is proposed and analyzed. The concept is based on the spatial multiplexing of the users associated with different operators in a multi user MIMO system. With the help of multi user selection the perfect sharing of network resources between the operators is ensured. This approach enables the achievement of agreed service shares (fairness) with in the smallest time-frequency resource element. In addition to the time and frequency, transmit power resource can also be shared. With the help of two different SLAs, it is demonstrated that the rate and resource fairness can be achieved over a time-frequency resource element. Such fairness is an important system requirement for shared networks [66]. Spectral efficiency is another system utility which is considered in this analysis for this scheduling approach. Based on the optimal power allocation (water-filling), it is shown that even in a perfectly fair sharing scenario, high spectral efficiency can be achieved. So the present algorithms for the MOMU-MIMO approach fulfils the SLAs based on rate and resource fairness criteria.

Chapter 7

Improved Multi-Operator Multi-User MIMO (Single Objective)

7.1 Introduction

In this chapter, the focus is on two problems related to the rate fairness based SLAs in an environment where multiple operators are sharing the spectrum [67]. Based on the MOMU-MIMO concept for sharing first introduced in [64] it is analytically shown that achieving rate fairness among arbitrary number of operators on the smallest time-frequency radio resource (e.g., one OFDM resource element) is feasible. Then the problem of rate fairness will be transformed to a power allocation problem and an analytical proof is provided that shows that an unique solution is possible for multiple operators. In addition to the analytical method, also an algorithm is provided which is capable of guaranteeing rate fairness for any arbitrary number of operators. The algorithm used, is based on the Newton-Raphson numerical approximation [68] which is well known for its fast convergence and simplicity. In the following also simulation results are presented for the assessment of the proposed approach and algorithm. The results show that rate fairness based SLAs can also achieve system spectral efficiency as high as the resource fairness based SLAs. This provides an extra degree of freedom to the 5G service providers in terms of rate based QoS provisioning.

In the EU project called METIS [69], which was designated to lay the foundation of 5G wireless communications system, related work was performed in [70, 71], discussing coordination of inter-operator spectrum sharing and implementation approach for adaptive spectrum sharing among co-located distributed RANs of different operators. One early work on spectrum sharing and resource allocation for rate fairness is reported in [72] but the results are based on non-cooperative and cooperative game theory. Furthermore for fulfilling the agreed fairness between the operators such approach require a high amount of radio resources (e.g. time and frequency) compared to our proposal in this chapter.

For 5G wireless systems one of the challenges is to efficiently use the available resources with respect to hardware infrastructure and available frequency spectrum. Frequency spectrum sharing is one of the measures which can offer efficient use of the scarce spectrum resource. For such a multi-operator sharing environment efficient scheduling approaches with controllable performance regarding QoS are needed. Provisioning resource fairness among sharing operator is simple as the resources in wireless networks are quantifiable, e.g., time resource= 1 ms, frequency resource= 1 Hz, power resource= 1 mW. For the case of providing QoS in terms of rate fairness to sharing operators, enhanced schedulers have to be implemented. The proposed MU-MIMO based scheduler can guarantee such QoS parameters for the operators as they are used by QoS definitions for their subscribers. Such a scheduler providing rate fairness for the sharing operators has to cope with a non-linear mathematical problem for each of its scheduling decisions. This non-linearity is based on the relation between the performance of spectrum efficiency and (amongst others to) the logarithm of the SNR which is described in the Shannon/Hartley theorem. Accordingly the mathematical problem formulation of rate fairness scheduling is a non-linear feasibility problem. This non-linear mathematical problem of rate fairness can be solved sub-optimally with a heuristic approach as shown in previously in [64] for two operators, with analytic methods for any number of operators within this presented approach.

7.2 System Model

An OFDM based closed-loop multi-user MIMO downlink system is considered which is operated by a neutral service provider as described in [64]. For the fine details on multi user MIMO one can refer to [64]. Here, the attributes of the considered system are explained that are necessary to understand the proposed approach and the results.

- $\mathcal{J}$: Set of indices representing the operators engaged in spectrum sharing, such that $\{|\mathcal{J}| = I, 1 \leq I \leq M_t\}$ where, $(M_t > 1)$ is the number of transmit antennas at the BS.
- $\mathcal{L}$: Set of indices of all the active UEs in the coverage area of the BS, we assume that $(|\mathcal{L}| \geq M_t)$.
- $\mathcal{L}_i$: Set which contains the indices of the active UEs associated with the operator i $(i \in \mathcal{J})$, such that $(\mathcal{L}_i \subset \mathcal{L})$.
- $\mathcal{S}_t$: Set of indices of the users selected by the BS for transmission on an OFDM resource element, such that $\{\mathcal{S}_t \subseteq \mathcal{L}, 1 \leq |\mathcal{S}_t| \leq M_t\}$.

- P_t: Maximum transmit power available at the BS.
- p_i: Power allocated to UE i associated with ith operator such that $(0 < p_i \leq P_t)$.
- δ_i: Channel gain at the antenna of the UE (single antenna) after the transmit beamforming. We assume ZF beamforming with perfect channel estimation by the user and a perfect feedback of the channel vector from each active user to the BS as in [64].
- σ^2: Variance of the noise at the receive antenna of the UE modelled as white Gaussian with zero mean.

7.3 Rate Fairness for Multiple Operators

The objective is to design an algorithm which fulfils a SLA that guarantees the rate fairness with high system spectral efficiency among arbitrary number of operators which are sharing spectrum in a given coverage area. The spectrum sharing environment is based on MOMU-MIMO concept [64]. Therefore, the essentials of the concept are also integrated here and explained to some extent for better understanding. Figure 7.1 represents our algorithm in terms of main building blocks. The functions of these blocks are described in the following.

7.3.1 User Groups Formation

As part of the user selection problem, at first we create the candidate sets of users. While creating all the possible candidate sets, we ensure that at least one UE from each operator is present in each candidate set. Assuming $I = M_t$ and $\mathcal{S}_i$ is the set that contains the indices of the UEs corresponding to operator i such that $\{(\mathcal{S}_i \subset \mathcal{L}_i), |\mathcal{S}_i| = 1)\}$. Let $\mathcal{S}_c(|\mathcal{S}_c| = M_t)$ be one of the candidate set of users that has to be considered for rate fairness criteria and it is given by:

$$\begin{gathered} \mathcal{S}_c = \{\mathcal{S}_1 \cup \mathcal{S}_2 \ldots \cup \mathcal{S}_i \cup \mathcal{S}_j \cup \ \ldots \ \cup \mathcal{S}_I\} \\ \forall (i,j), i \neq j \ : \{\mathcal{S}_i \cap \mathcal{S}_j\} = \emptyset \end{gathered} \tag{7.1}$$

There will be $|\mathcal{L}_i|^I$ candidate sets if we assume the same number of UEs per operator i.

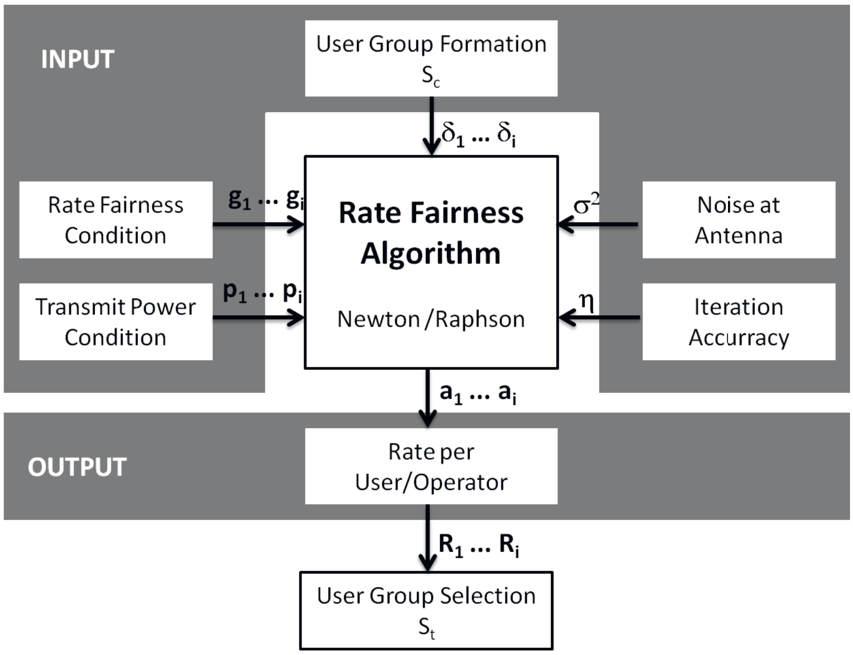

Figure 7.1: Rate fairness calculation based on presented analytical algorithm from [67]

7.3.2 Defining System Constraints

In this subsection we formulate the constraints that we have with respect to the SLA (rate fairness constraint) and service provider (transmit power constraint).

Rate-based SLA

Let g_i be the rate share for operator i that has to be guaranteed by the service provider, as agreed in the SLA. Using this service share, the rate based SLA for operator i can be defined as,

$$R_i = g_i R(\mathcal{S}_t) \tag{7.2}$$

where, R_i is the rate that the operator i will get and $R(\mathcal{S}_t)$ represents the total sum rate of all the operators including operator i. In the following, we define the share for each operator such that the sum of the service shares from all the operators is given as:

$$\sum_{i \in \mathcal{J}, i=1}^{I} g_i = 1 \tag{7.3}$$

Exploitation of Transmission Power

To ensure the use of complete available radio resources and to maximize the system spectral efficiency in the interest of the service provider, we impose the following constraint for the allocated power to all the selected users represented by the set $\mathcal{S}_t$:

$$\sum_{i \in \mathcal{J}, i=1}^{I} p_i = P_t \tag{7.4}$$

7.3.3 Feasibility Problem for Rate Fairness

In this section, we formulate the problem of achieving rate fairness with maximum spectral efficiency for any number of operators involved in spectrum sharing. Moreover, we also provide the feasibility proof for the agreed rate fairness, as written in (7.2). Using the well known Shannon-Hartley theorem and the definitions given in Section 7.2, we can write the expression for R_i as follows:

$$R_i = \log_2 \left(1 + \frac{\delta_i}{\sigma^2} p_i\right) \tag{7.5}$$

The objective problem is to find $p_i, \forall i \in \mathcal{J}$ such that (7.2) is fulfilled for each i. For the ease of further analysis we perform the following substitution,

$$(1 + \frac{\delta_i}{\sigma^2} p_i) = a_i \tag{7.6}$$

and re-write (7.5) as:,

$$R_i = log_2(a_i) \tag{7.7}$$

With the help of (7.7) we have transformed our problem from p_i to a_i.
Transformation of the Rate Constraint:
Using (7.2), the relation between the rates of any two operators $(i, j) \in \mathcal{J}$ can be written as:

$$R_i = \tfrac{g_i}{g_j} R_j \quad , i \neq j \tag{7.8}$$

Without loss of generality, in the following we assume $(j = 1)$. Using (7.7) we can re-write (7.8) as follows:

$$log_2(a_i) = \tfrac{g_i}{g_1} log_2(a_1) \tag{7.9}$$

The above expression can be further simplified as:

$$a_i = a_1^{\frac{g_i}{g_1}} \tag{7.10}$$

Transformation of the Power Constraint:
With $j = 1$ we can re-write (7.4) as,

$$p_1 + \sum_{i \neq j, i=2}^{I} p_i = P_t \tag{7.11}$$

where, p_i can be written in terms of a_i with the help of (7.6) as follows:

$$p_i = \frac{(a_i - 1)\sigma^2}{\delta_i} \tag{7.12}$$

Using (7.11) and (7.12) we can write,

$$\frac{(a_1 - 1)\sigma^2}{\delta_1} + \sum_{i \neq j, i=2}^{I} \frac{(a_i - 1)\sigma^2}{\delta_i} = P_t \tag{7.13}$$

By reordering the terms according to a_i we get:

$$a_1 + \delta_1 \sum_{i \neq j, i=2}^{I} \frac{1}{\delta_i}(a_i - 1) = 1 + P_t \frac{\delta_1}{\sigma^2} \tag{7.14}$$

Substituting (7.10) in (7.14) we get:

$$a_1 + \delta_1 \sum_{i \neq j, i=2}^{I} \frac{1}{\delta_i}(a_1^{\frac{g_i}{g_1}} - 1) = 1 + P_t \frac{\delta_1}{\sigma^2} \tag{7.15}$$

Proposition 7.3.3.1. The power allocation per operator which satisfies rate fairness constraint (7.2) and the total power constraint (7.4) is uniquely determined by the following K equalities $\triangleleft$

$$\begin{cases} a_1 + \delta_1 \sum_{i=2}^{I} \frac{1}{\delta_i}(a_1^{\frac{g_i}{g_1}} - 1) = 1 + P_t \frac{\delta_1}{\sigma^2} \\ a_i = a_1^{\frac{g_i}{g_1}} \qquad\qquad\qquad\qquad\qquad\qquad (i \in \mathcal{J}, i \neq j) \end{cases} \tag{7.16}$$

Proof. The proof of the Proposition 7.3.3.1 is partitioned in two parts. The first part is showing the *equivalence* of its equalities regarding (7.2) and (7.4). The second part of the proof shows the *uniqueness* of the results of the Proposition 7.3.3.1. as given in [(7.16)].

1. Part: We can assume a bijective transformation from (7.2) and (7.4) to (7.16), as each element of the relation of (7.2) and (7.4) is paired at least with one element out of (7.16). And no element of the relation of (7.2) and (7.4) can be paired with more than one element out of (7.16) as shown in the presented derivation of (7.16). So for a bijective transformation [(7.2),(7.4)] $\rightarrow$ [(7.16)] we can assume equivalence in the transformation direction [(7.2),(7.4)] $\leftarrow$ (7.16).

2. Part: A strictly monotonic increase of the left hand side (LHS) of the first part of (7.16) can be seen. Therefore, there is a unique intersection with the constant right hand side (RHS) of the equation. With the other equation as given in (7.16) unique components a_i (with $i = 2...K$) are described. From this it follows that the equation provides uniqueness for the rate fairness condition.□

7.3.4 Finding a_1 by Newton-Raphson Method

Let us define a function $f(a_i)$ from (7.16) such that:

$$f(a_1) = a_1 + \delta_1 \sum_{i \neq j, i=2}^{I} \frac{1}{\delta_i}(a_1^{\frac{g_i}{g_1}} - 1) - (1 + P_t \frac{\delta_1}{\sigma^2}) \tag{7.17}$$

It can be seen that $f(a_1)$ is a non-linear function. As a standard practice the Newton-Raphson method is used to calculate by linear approximation the roots of the non-linear function $f(a_i)$. The general form of known approximation method is given by

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} \tag{7.18}$$

where x_0 is an assumed starting value for the root of $f(x_0)$ and x_1 a better root approximation calculated with the help of the function $f(x_0)$ and its first derivative $f'(x_0)$. In the next iteration x_0 takes the value of x_1 and a new x_1 is calculated until the difference between x_0 and x_1 becomes smaller than a given iteration accuracy value η.

In our case, the first derivative $f'(a_1)$ of the feasibility problem function $f(a_i)$ is given by:

$$f'(a_1) = 1 + \frac{\delta_1}{g_1} \sum_{i=2}^{i \in \mathcal{J}} \frac{g_i}{\delta_i} a_1^{(\frac{g_i}{g_1} - 1)} \tag{7.19}$$

Finally to summarize, we find a_1 by Newton-Raphson and then we find the remaining a_i by using the second part of (7.16).

7.3.5 User Group Selection

As described in Section 7.3.4, the BS computes power allocation for all the operators in all the candidate sets (which are formed in Section 7.3.1). After the computation of power, the BS selects the set $\mathcal{S}_t$ for transmission which maximizes the system spectral efficiency as given in the following:

$$R(\mathcal{S}_t) = \max_{\forall \mathcal{S}_c} R(\mathcal{S}_c) \tag{7.20}$$

7.4 Simulation Results

In this section, we show with the help of simulation results that we can achieve rate fairness among multiple operators with system spectral efficiency close to optimum. For the sake of fair assessment of our approach, we compare the results with two reference approaches. The next subsections describe these approaches, the simulation assumptions and the analysis of the results. Note that the fulfilment of the SLA has already been proven analytically by proposition 7.3.3.1. The focus of the simulations is to show how the optimum spectral efficiency can be achieved by using our algorithm for rate fairness among any number of operators.

7.4.1 Reference Approaches for Computing p_i

Water-Filling Approach for Maximum Spectral Efficiency

Here, we consider an SLA which is limited only to the user group formation for each time-frequency resource element. Hence, the service provider is free to allocate power independent of SLA. Therefore, classical WF based power allocation approach can be used which is known for achieving optimum for system spectral efficiency.

Heuristic Approach for Rate Fairness

In [64], we proposed an iterative power allocation algorithm based on incremental adjustment of the transmission power until the desired rate fairness as

Table 7.1: Simulation parameters from [67]

Parameter	Value
Channel Realizations	1000
UEs per Operator	1, 5, 10, 25, 50
Number of Transmission Antennas	2OP: 2 3OP: 3
Service Shares (2OP: g_1/g_2) (3OP: $g_1/g_2/g_3$)	2OP: 0.1/0.9, 0.2/0.8, 0.3/0.7, 0.4/0.6, 0.5/0.5 3OP: 0.05/0.05/0.9, 0.1/0.1/0.8, 0.15/0.15/0.7, 0.25/0.25/0.5, 0.3/0.3/0.4, 0.35/0.35/0.3, 0.4/0.4/0.2, 0.45/0.45/0.1
Newton/Raphson: Iteration Accuracy η	10^{-3}

in (7.2) is achieved. However, the power constraint in (7.4) was relaxed due to which the approach suffered the loss in achievable spectral efficiency. Moreover, that approach was limited to the sharing among two operators only.

7.4.2 Assumptions and Parameters

We model a coverage area of a BS in downlink where the users are randomly dropped with a uniform distribution. The users are associated with different operators and the network has full buffer traffic for each user. In every TTI, each operator has equal number of active users. The users are facing a Rayleigh fading channel with an average SNR of 10 dB. The effect of average inter cell interference is assumed to be included in Gaussian noise. Here, perfect CSI at the transmitter is considered and any feedback or scheduling delay is neglected. Further parameters are summarized in Table 7.1. The simulations are carried out for the cases when two and three operators are sharing the spectrum.

7.5 Results and Analysis

Figure 7.2 presents the spectral efficiency results against the number of active users for the two operators in agreement with the service provider through a rate fairness SLA with different service shares. The figure presents a comparison of methods for the computation of power by our analytic approach (based

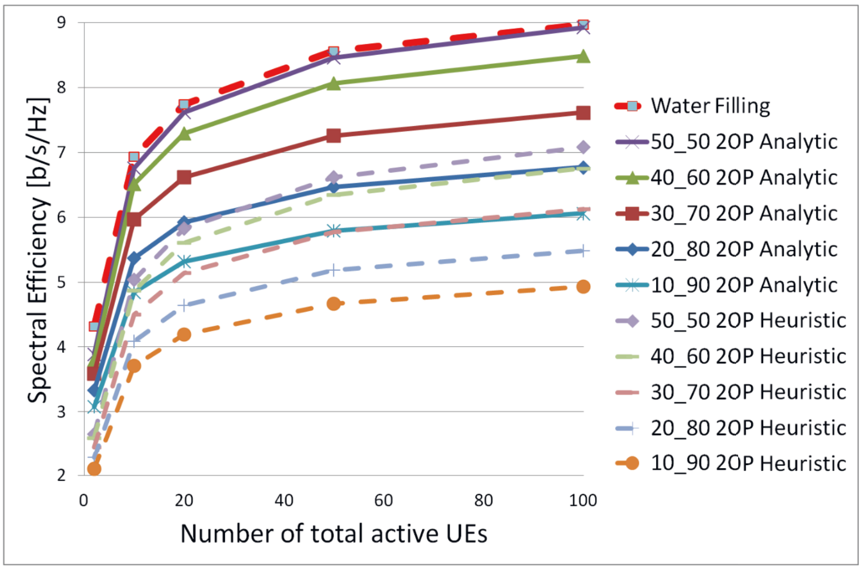

Figure 7.2: 2 Operators: Spectral efficiency for different service shares for rate fairness with a heuristic and analytic approach and WF based SLAs from [67]

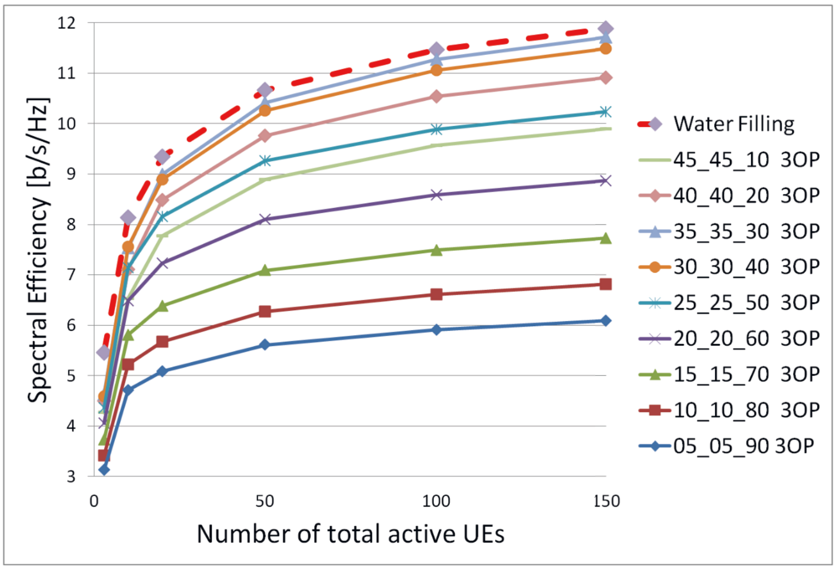

Figure 7.3: 3 Operators: Spectral efficiency for different service shares for rate fairness with analytic approach and WF based SLAs from [67]

on Newton-Raphson approximation as proposed in Section 7.3), by the heuristic method in [64] and by the WF method. The focus is to show that for any given service share, the analytic method outperforms the heuristic method. This high performance of analytic method is due to the tight power constraint which is introduced in (7.4). For the case of $(g_1 = 50\%, g_2 = 50\%)$, the analytic method for rate fairness provides the same performance as the WF method as presented for the resource fairness approach in [64]. However, for the same set of service shares, the heuristic approach is way beyond the WF (e.g., 22% loss for 100 UEs). If we increase the difference in the service share of the two operators, the performance of rate fairness methods decreases. However, the analytic approach remains closer to the WF. This implies that the service provider can afford an SLA with higher variance and still achieve a better system performance. This flexibility allows the service provider to cover the operators with wide variety of use cases. For example, some operator may have only low data rate users in that area and require only a smaller service share while the other may have broadband users and require a bigger service share.

In Figure 7.3, the focus is to prove that the analytic rate fair approach can be applied for any number of operators. Figure 7.3 presents the exemplary results with three operators bearing different service shares. Please note that with three operators we have $M_t = 3$ which provides us an absolute gain in spectral efficiency as compared to the two operators case where $M_t = 2$. When the variance in service shares is low, we can observe the same trend in Figure 7.3 as in Figure 7.2, i.e., the performance is close to WF. However, there is a higher loss in performance when the variance is higher. For example, compare the case of high variance with three and two operators for 100 UEs. The loss in three operators case $(g_1 = 5\%, g_2 = 5\%, g_3 = 90\%,)$ as compared to WF is 91% whereas the loss in two operators case $(g_1 = 10\%, g_2 = 90\%)$ is only 50%. This implies that it is beneficial for the service provider to provide service to higher number of operators however with low variance in agreed service shares.

7.6 Conclusion

In this chapter, it is shown that rate fairness based SLAs can also achieve system spectral efficiency as high as resource fairness based SLAs in an environment where multiple operators are sharing the spectrum. Moreover, we showed that rate fairness approach is also scalable to any number of operators. We transformed the problem of achieving rate fairness among operators to a power allocation problem. We proved analytically that for each transmission, the power for each operator can be uniquely determined which satisfies the total rate constraint of the system. We used Newton-Raphson method to find the power values and devise an algorithm that ensures the agreed rate fairness and achieves high spectral efficiency. We also provided simulation results to assess the performance of our algorithm as well as compared it with other methods. The assessment of results endorse the usefulness of our multi-operator scheduling algorithm. Moreover, it also implies that resource fairness (as in most of the literature work) is not the only metric which supports the interest of the service providers.

Part IV

Optimized Efficiency Triggered Algorithms

Chapter 8

Multi-Operator Multi-User MIMO (Multi-Objective)

8.1 Introduction

Along with spectral efficiency, energy efficiency is a key performance metric for the design of 5G wireless networks. At the same time, infrastructure sharing among multiple operators has also emerged as a new trend in wireless communications networks. The approach in this chapter [73] presents an optimization framework for energy efficiency and spectral efficiency maximization, in a network where radio resources are shared among multiple operators. We define a heterogeneous SLA framework for a shared network, in which the constraints of different operators are handled by two different multi-objective optimization approaches namely the utility profile and scalarization methods. Pareto-optimal solutions are obtained by merging these approaches with the theory of generalized fractional programming. The approach applies to both noise-limited and interference-limited systems, with single-carrier or multi-carrier transmissions. Extensive numerical results illustrate the effect of the operator-specific SLA requirements on the global spectral and energy efficiencies. Three network scenarios are considered in the numerical results, each one corresponding to a different SLA, with different operator-specific energy efficiency and spectral efficiency constraints.
Besides the 1000x increase of mobile data volume per geographical area, EE in bit/J is one of the central requirements for the definition of 5G networks [74]. One technique to reduce energy consumption and provide ultra-high broadband access cost efficiently is to share or pool the available network resources in a multi-operator environment. The contributions in [75] and [76] describe in detail the partnership-based business models among multiple operators for infrastructure and spectrum based sharing. However, efficient scheduling approaches with controllable performance regarding QoS provisioning are still in their evolutionary phase. It is well known that smart scheduling algorithms are required in cellular networks to achieve certain levels of user fairness and fulfill the QoS requirements. Likewise, in a shared network with multiple operators, fairness of resource usage among sharing operators has to be provided along

with the guaranteed QoS. Resource fairness can be easily achieved among the sharing operators, by simply dividing the available resources (e.g., time, frequency and power) between them in a fair manner. The shared resource will be linearly partitioned according the required resource fairness ratios. However, it is more challenging to guarantee QoS requirements (i.e., minimum throughput rates) of the sharing operators. The beforehand presented scheduling approaches focus on the maximization of the SE with either guaranteed resource fairness or rate fairness in multi-operator networks [64, 67]. Here, we focus on the maximization of SE along with the global energy efficiency (GEE) while guaranteeing individual operator constraints, e.g., QoS and energy consumption.

8.1.1 Related Work

There has been considerable amount of studies on infrastructure sharing under 3rd Generation Partnership Project (3GPP) for future wireless networks. Specifically, the technical report in [66] presents the results of a study on Radio Access Network (RAN) sharing enhancements. At first, the report describes the structure of a shared RAN, by considering it as a set of access points which are used by the *participating operators*, labeled as *sharing operators* hereafter. Several use cases and scenarios are described which encompass different techno-economic needs and targets of the sharing operators and RAN providers. The types of operators span from Mobile Virtual Network Operator (MVNO) to incumbent operators having own infrastructure and radio resources and also taking into account the operators merely hosting the shared RAN. The presented approaches are applicable to most of the presented use-cases, and it provides a technical framework which supports the efficient realization of these scenarios.

Next Generation Mobile Networks (NGMN) consortium has defined multi-operator network sharing also as a key technology building block for efficient 5G network resource usage [74]. The set of prime benefits envisioned by the consortium for individual operators includes the minimization of capital and operational expenditure to provide the required network coverage and capacity. Enabling technologies to achieve these benefits should be based on efficient resource management, minimizing power consumption and maximizing SE. These targets can be achieved by means of enhanced radio resource management in shared RAN environments. The scheduler for a shared RAN should further ensure fairness and fulfilment of operator policies by considering SLA conditions.

In the previous approaches, as given in [64] and [67], the spatial dimension of the channel is considered as a radio resource which can be shared among multiple operators. This allows us to guarantee resource and rate fairness instantaneously (e.g., on the smallest time-frequency resource element in an OFDM-based RAN). So, it is shown that multi-antenna BSs allow multiple operators to share a given time-frequency resource. This requires efficient allocation of the transmit power in order to achieve high SE and to guarantee resource or rate fairness. This novel approach is named as Multioperator-Multiuser-MIMO (MOMU-MIMO). In contrast to [24, 50, 52, 53, 56, 57, 77], MOMU-MIMO allows us to guarantee SLA conditions at each resource element. The scheduling approach in [64] is focused more on the development of heuristic algorithms to provide proof-of-concept for MOMU-MIMO. Based on the optimal power allocation, it is shown that we can achieve high SE along with the fulfilment of SLAs based on rate and resource fairness criteria even on each resource element. The approach in [64] is further extended in [67] with the emphasis on the achievement of near optimal SE with guaranteed rate fairness, regardless of the number of participating operators. Using the Newton-Raphson, method an algorithm is provided that ensures the agreed rate fairness and achieves high SE.

While the previously presented literature only focuses on the traditional performance metrics of throughputs and rates, ecological and sustainable growth concerns are bringing forth the need of designing energy-efficient networks. Indeed, over the last years EE has been acknowledged as a key performance metric for 5G networks, for which the goal has been set to increase the data rate by a factor 1000, while at the same time halving the energy consumption. This requires a 2000 fold increase of the EE, given as bit-per-Joule in [74]. A recent survey of the energy-efficient 5G networks was written in [78]. Therein, resource allocation was observed to be one of the key approaches to improve network EE. In line with these observations and the requirements from NGMN [74], the scheduling algorithm in this chapter is strengthened further in this direction by considering not only the SE but also the EE. In this chapter, a multi-objective solution is presented for which each objective (EE or SE utility) will have diverse constraints coming from different sharing operators simply participating and/or hosting the shared RAN.

In the context of EE maximization, fractional programming theory appears the most suited tool, and indeed it has been successfully applied for energy-efficient resource allocation in many different wireless communication scenarios [79]. Closest to the scenario studied in this chapter are the works on energy-efficient resource allocation in multi-carrier and multi-antenna systems. Recent studies on EE maximization by fractional programming in Orthogonal Frequency Division Multiple Access (OFDMA) systems are [80–86]. [80, 83–85] consider energy-efficient resource allocation subject to QoS constraints, [81] studies energy-per-bit minimization, while in [82] the EE is optimized subject to proportional fairness constraints. The work in [86] focuses on joint power and beamforming weights optimization for an energy efficient system. The trade-off between EE and SE in OFDMA networks is studied in [87, 88]. The papers [89–94] consider energy-efficient resource allocations for cooperative system in which clusters of base stations are coordinated via backhaul connections. Different energy-efficient performance metrics are considered, including the GEE of the cluster [89, 91], the sum of the individual EEs [90, 91], the product of the individual EEs [91], and the minimum of the EEs [85, 92]. Many recent works study the EE of multi-antenna wireless links. Relevant to the current scheduling approach are works on EE and SE trade-offs in [95–98]. [95] studies the problem of minimizing the energy consumption per transmitted bit, whereas [96] investigates the problem of energy and outage probability optimization. The energy-spectral efficiency trade-off is discussed in [97], whereas [98] provides energy-efficient precoding algorithms for a single-user MIMO system.

8.1.2 Novel Contributions

The novelty of this enhanced scheduling approach compared to the references given above, it provides a framework to model new heterogeneous SLAs for 5G networks, where multiple operators share the networks, each one having multiple QoS and fairness constraints. The objective is to achieve operator-specific EE and SE targets while keeping global spectral efficiency and global energy efficiency at maximum in a shared system. According investigated literature, the combined problem of global system utilities and operator specific multiple objectives in shared networks has not been previously investigated.

The problem is formulated as a multi-objective optimization, in which the GEE and SE are simultaneously maximized. The Pareto boundary of this multi-objective optimization problem is characterized by using the multi-objective methods of utility profile and of scalarization [99], together with the theory of generalized fractional programming [100], applied to EE [79] and SE. The proposed optimization framework for this MOMU-MIMO system has the following attractive features:

1. It globally maximizes the GEE and SE in the shared environment subject to sharing constraints.
2. It provides a baseline for heterogeneous SLAs with multiple constraints for each operator, that are tailored according to the specific and heterogeneous service types of the different operators.
3. The framework applies to a network shared by an arbitrary number of operators, each serving multiple UEs.
4. It optimally fulfils the utility targets and operator constraints at each allocable time-frequency resource element (e.g., one symbol duration on a sub-carrier in OFDM based radio interface).

8.2 System Model

An OFDM-based closed-loop multi-user MIMO downlink transmission is considered for a MOMU-MIMO system as given in [64]. The focus will be on the coverage area of a single transmission point (TP) equipped with M transmit antennas and connected to a BS serving I active users in a cellular network. The TP is shared by O operators sharing the available spectrum, and each active user is associated to a possibly different operator. Let $\mathcal{U}_o$ be the set of users associated to operator o. Thus, we have $\{1, 2, \ldots, I\} = \mathcal{U}_1 \bigoplus \mathcal{U}_2 \ldots \bigoplus \mathcal{U}_O$. Each mobile device has a single antenna and hence, the downlink channel is formally equivalent to an $I \times M$ MIMO system, with channel matrix $\mathbf{H} = [\mathbf{h}_1 \;\; \mathbf{h}_2 \;\; \mathbf{h}_i \;\; \ldots \;\; \mathbf{h}_I]^H \in \mathbb{C}^{I \times M}$, with $\mathbf{h}_i$ the $M \times 1$ channel vector between the TP and the i-th mobile user. Then, the received signal at mobile i can be expressed as

$$y_i = \sqrt{p_i}\mathbf{h}_i^H \mathbf{b}_i s_i + \sum_{j \neq i}^{I} \sqrt{p_j}\mathbf{h}_i^H \mathbf{b}_j s_j + z_i \;, \tag{8.1}$$

wherein p_i, $\mathbf{b}_i$, and s_i are the transmit power, unit-norm beamforming vector, and information symbol intended for user i, while z_i is the noise term, modeled as a realization of a zero-mean Gaussian random variable with variance σ^2. Moreover, let us define the equivalent channel gain $\delta_i = |\mathbf{h}_i^H \mathbf{b}_i|^2$, for all i. Then, the SINR enjoyed by user i is expressed as

$$\gamma_i = \frac{p_i \delta_i}{\sigma^2 + \sum_{j \neq i} p_j \delta_j} \;. \tag{8.2}$$

Based on (8.2), the spectral efficiency of operator o, and the system spectral efficiency are defined as

$$\mathrm{SE}_o = \sum_{i \in \mathcal{U}_o} \log_2(1+\gamma_i) \tag{8.3}$$

$$\mathrm{SE} = \sum_{i=1}^{I} \log_2(1+\gamma_i) \tag{8.4}$$

Similarly, the bit-per-Joule energy efficiency of operator o, and also the system global energy efficiency (GEE), are defined as [79]

$$\mathrm{EE}_o = \frac{\sum_{i \in \mathcal{U}_o} \log_2(1+\gamma_i)}{\sum_{i \in \mathcal{U}_o} (p_i + p_{0,i})} \tag{8.5}$$

$$\mathrm{GEE} = \frac{\sum_{i=1}^{I} \log_2(1+\gamma_i)}{p_0 + \sum_{i=1}^{I} p_i} \tag{8.6}$$

with $p_{0,i}$ being the static circuit power dissipated in all hardware blocks required to operate the communication with user i (e.g., DA/AD converters, baseband filters, cooling equipment, digital signal processing operations), and $p_0 = \sum_{i=1}^{I} p_{0,i}$. In general, the GEE and the SE are two conflicting performance measures, since the maximization of one can lead to a decrease of the other. The goal will be to analyze this trade-off, designing power control algorithms for multi-operator MIMO system, that are able to strike the optimal balance between SE and GEE, subject to total power constraints, per-operator minimum SE and EE constraints, and system GEE constraints. Two main scenarios will be addressed.

- Noise-limited scenario. By setting $\mathbf{b}_i = \frac{\tilde{\mathbf{b}}_i}{||\tilde{\mathbf{b}}_i||}$, where, $[\tilde{\mathbf{b}}_1 \ \tilde{\mathbf{b}}_2 \ \tilde{\mathbf{b}}_i \ \ldots \ \tilde{\mathbf{b}}_K] = \mathbf{H}^H(\mathbf{H}\mathbf{H}^H)^{-1}$, multi-user interference is completely suppressed. This leads to better performance and simpler resource allocation algorithms. However, it requires a sufficient number of antennas, i.e., $M \geq I$, as well as perfect CSI. The noise-limited scenario will be addressed in Section 8.3
- Interference-limited scenario. In this case, no assumption on M is done, and interference is explicitly accounted for in the power allocation procedure, for fixed beamforming vectors $\{\mathbf{b}_i\}_{i=1}^{I}$. This might be necessary for example when $M < I$. This scenario is addressed in [73, Section IV].

Before concluding this section, it should be observed that (8.1) does not explicitly include the presence of out-of-cell interference, which however can be regarded as included in the noise term. The detailed study on the impact of out-of-cell interference can be considered as possible extension. Moreover, while the analysis refers to a single-carrier system, the results and optimization framework carry over to the multi-carrier scenario with minor modifications. More details on this point will be provided in Chapter 8.3.

8.3 Power control in the noise-limited scenario

In the noise-limited scenario, the SINR in (8.2) simplifies to $\gamma_i = p_i\delta_i/\sigma^2$, and the metrics introduced in (8.3), (8.4), (8.5), and (8.6) simplify accordingly. Thus, applying the variable change

$$x_i = 1 + \frac{\delta_i}{\sigma^2}p_i \quad \Longleftrightarrow \quad p_i = \frac{\sigma^2}{\delta_i}(x_i - 1), \tag{8.7}$$

and defining $\mathbf{x} = \{x_i\}_{i=1}^{I}$, the optimization problem to be solved is stated as

$$\max_{\mathbf{x}} \left[\mathrm{GEE}(\mathbf{x}), \mathrm{SE}(\mathbf{x})\right] \qquad \text{s.t. } \mathbf{x} \in \mathcal{X},$$

wherein $\mathcal{X}$ is the problem feasible set, defined by the following constraints:

$$x_i \quad > 1 \;,\; \forall\, i = 1, \ldots, I \tag{8.8a}$$

$$\sum_{i \in \mathcal{U}_o} \log_2 x_i - \mathrm{EE}_o^* \left(\sum_{i \in \mathcal{U}_o} (x_i - 1)\frac{\sigma^2}{\delta_i} \right) \geq \mathrm{EE}_o^* \sum_{i \in \mathcal{U}_o} p_{0,i} \;,\; \forall\, o = 1, \ldots, O \tag{8.8b}$$

$$\sum_{i=1}^{I} \log_2 x_i - \mathrm{GEE}^* \sum_{i=1}^{I} (x_i - 1)\frac{\sigma^2}{\delta_i} \quad \geq \mathrm{GEE}^* p_0 \tag{8.8c}$$

$$\sum_{i \in \mathcal{U}_o} \log x_i \quad \geq \underline{R}_o \;,\; \forall\, o = 1, \ldots, O \tag{8.8d}$$

$$\sum_{i=1}^{I}(x_i - 1)\frac{\sigma^2}{\delta_i} \quad \leq P_t. \tag{8.8e}$$

The expression in (8.8a) sets a general requirement that each operator has to be scheduled in each scheduling period with a transmission power p_i greater than zero. The constraints in (8.8b) and (8.8c) take into account the SLA regarding individual EE levels to be achieved (follows from (8.5)) for a specific operator or the GEE level for the network provider (follows from (8.6)), whereas the per-operator constraints in (8.8d) define the minimum target spectral efficiency $\underline{R}_o$ to be guaranteed to an operator as defined in the SLA. Finally, (8.8e) ensures that the maximum transmit power is below the maximum feasible power P_t. Note that with the help of constraints (8.8b)-(8.8d) we can define new SLAs circumscribing the requirements for operators as well as neutral site owners involved in a shared environment.

Problem is a multi-objective program (MOP) [99], i.e., an optimization problem in which the objective function is a vector, whose components are to be optimized simultaneously. In our case we have a bi-objective problem. The most widely accepted solution concept when dealing with multi-objective optimization is to obtain Pareto-optimal points. Specifically, for the case at hand, the Pareto region is defined as the region of all feasible pairs $(\text{GEE}(\mathbf{x}), \text{SE}(\mathbf{x}))$, namely

$$\mathcal{P} = \{(\text{GEE}(\mathbf{x}), \text{SE}(\mathbf{x})) \ , \ \mathbf{x} \in \mathcal{X}\} \ . \tag{8.9}$$

The outer frontier of $\mathcal{P}$ is the Pareto boundary of the problem and its points are called Pareto-optimal and represent the optimal set of points regarding as far as solving (8.8). All Pareto-optimal points have the property that it is not possible to further increase one of the two objectives, without decreasing the other. Multi-objective optimization theory provides several approaches to convert the vector objective of a multi-objective problem, into a scalar objective whose maximization results in a Pareto-optimal point. The two most widely used approaches are the utility profile method and the scalarization method. Both approaches will be considered in the sequel.

8.3.1 Utility profile approach

According to the utility profile approach, a ratio of GEE and SE is to be maximized. This leads to the single-objective problem:

$$\max_{\mathbf{x}\in\mathcal{X}} \ t \qquad \text{s.t. GEE} \geq wt \qquad \text{and SE} \geq (1-w)t \ , \tag{8.10}$$

wherein the weighting factor w can be tuned to give different priority to the two utilities. Equal prioritization of the utilities is achieved by setting $w = 0.5$. It is also important to remark that for every choice of w, a different point on the Pareto-frontier is obtained, and that the whole Pareto frontier can be characterized by sweeping the weight w.

Now, problem (8.10) is equivalently reformulated as

$$\max_{\mathbf{x}\in\mathcal{X}} \min \left(\frac{1}{w} \frac{\sum_{i=1}^{I} \log x_i}{p_0 + \sum_{i=1}^{I} (x_i - 1) \left(\frac{\sigma^2}{\delta_i}\right)}, \frac{1}{1-w} \sum_{i=1}^{I} \log x_i \right). \tag{8.11}$$

Problem (8.11) can be cast into the framework of generalized fractional programming [79, Section 3.3], [101]. In particular, the GDA is able to maximize the minimum of a family of ratios, provided each ratio has a concave numerator and a convex denominator. In addition, the constraint set must be convex. These assumptions are clearly fulfilled for problem (8.11). In the following, we describe how the GDA can be applied to (8.11), showing that this requires a polynomial complexity.

GDA for solving (8.11): In order to apply the theory of generalized fractional programming to (8.11), the fundamental step is to introduce the following auxiliary function [101]:

$$F(\lambda) = \max_{\mathbf{x}\in\mathcal{X}} \min_{k} \left(f_k(\mathbf{x}) - \lambda g_k(\mathbf{x}) \right), \tag{8.12}$$

with $k = \{\text{GEE}, \text{SE}\}$, $f_{\text{SE}}(\mathbf{x}) = \frac{1}{1-w} \sum_i \log x_i$, $g_{\text{SE}} = 1$, as well as $f_{\text{GEE}}(\mathbf{x}) = \frac{1}{w} \sum_i \log x_i$, and $g_{\text{GEE}}(\mathbf{x}) = p_0 + \sum_i (x_i - 1)\frac{\sigma^2}{\delta_i}$.
A fundamental result of generalized fractional programming establishes that solving (8.11) is equivalent to finding the zero of the function (8.12). In other words, $F(\lambda)$ has a unique zero, which we denote by λ^*. Then, $\mathbf{x}^*$ is a solution of (8.11) if and only if $\min_k \left(f_k(\mathbf{x}^*) - \lambda^* g_k(\mathbf{x}^*)\right) = 0$. As a consequence, in order to solve (8.11) it suffices to determine the zero of $F(\lambda)$. This is accomplished by the GDA, whose pseudo-code is reported below here.

It is known that the convergence rate of Algorithm 1 is linear [101]. In addition, it can be seen that since f_k and g_k are respectively concave and convex

Algorithm 1 GDA

$\varepsilon > 0, n = 0, \lambda_n > 0$
while $F(\lambda_n) > \epsilon$ **do**
  $\mathbf{x}_n^* = \arg\max_{x \in X} \left\{ \min_k \left\{ f_k(\mathbf{x}_n) - \lambda_n g_k(\mathbf{x}_n) \right\} \right\};$
  $F(\lambda_n) = \min_k \left\{ f_k(\mathbf{x}_n^*) - \lambda_n g_k(\mathbf{x}_n^*) \right\};$
  $\lambda_{n+1} = \min_k \frac{f_k(\mathbf{x}_n^*)}{g_k(\mathbf{x}_n^*)}; n = n + 1;$
end while
Optimal solution in $\mathbf{x}_{n-1}^*$.

for each k, and $\mathcal{X}$ is a convex set, Algorithm 1 requires to solve a convex problem in each iteration, which can be accomplished with polynomial complexity. Finally, the value of λ output by the algorithm can be seen to be the optimal value of the objective of (8.11).

8.3.2 Scalarization approach

The scalarization approach is an alternative approach to solve a MOP and consists of maximizing a convex combination of the objectives of problem (8.8), which leads to the problem:

$$\max_{\mathbf{x} \in \mathcal{X}} \quad (1 - w)\text{SE} + w\text{GEE}, \tag{8.13}$$

subject to the constraints in (8.8). Unlike (8.11), problem (8.13) does not fall into the framework of generalized fractional programming, but instead belongs to the class of sums-of-fraction optimization problems, for which available solution methods exhibit exponential complexity in the number of variables [102].

However, for the case at hand it is possible to globally solve (8.13) with affordable complexity. To this end, let us introduce the auxiliary variable y defined as $y = \sum_i \log x_i$, which represents the system total spectral efficiency. Then, for fixed y, the problem in (8.13) reduces to[1]

$$\max_{\mathbf{x} \in \mathcal{X}} \quad w \frac{y}{p_0 + \sum_{i=1}^{I} (x_i - 1) \frac{\sigma^2}{\delta_i}} + (1 - w) y,$$

[1]Note that the equality constraint $y = \sum_{i=1}^{I} \log x_i$ can be relaxed to an inequality constraint, because the objective is increasing in y.

$$\text{s.t.} \quad \sum_{i=1}^{I} \log x_i \geq y \tag{8.14}$$

It should be observed that since y is fixed, problem (8.14) is equivalent to minimizing the linear function $\sum_{i=1}^{I}(x_i - 1)\frac{\sigma^2}{\delta_i}$ at the denominator, subject to the convex constraints which describe the set $\mathcal{X}$ and to the additional constraint $\sum_{i=1}^{I} \log x_i \geq y$. Therefore, problem (8.14) can be solved by standard convex optimization methods [62]. Then, the optimal solution of (8.13) can be obtained by performing a line search over y, solving (8.14) for each fixed y. In order to implement an efficient line search on y, we can bound the range in which y must lie by observing that:

$$y_{lb} = \sum_{o=1}^{O} \underline{R}_o \leq \sum_{i=1}^{I} \log x_i \leq \sum_{i=1}^{I} \log\left(1 + \frac{\delta_i}{\sigma^2} P_t\right) = y_{ub}.$$

The resulting overall algorithm is illustrated in Algorithm 2, which is able to determine the global solution of (8.13) with a predefined accuracy which depends on the number of steps η.

Algorithm 2 Weighted GEE/SE Algorithm

$\eta > 0$, $\tau = \frac{y_{ub} - y_{lb}}{\eta}$

for $y \in [y_{lb} : \tau : y_{ub}]$ **do**

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} \sum_{i=1}^{I}(x_i - 1)\frac{\sigma^2}{\delta_i}\ , \ \text{s.t.} \sum_{i=1}^{I} \log x_i \geq y\ ,\ \mathbf{x} \in \mathcal{X}$$

end for

Optimal solution in $\mathbf{x}^*$.

Finally, we make the following remark.

Remark 8.3.2.1. The analysis in this section has been performed assuming a single carrier. However, it can be extended to multi-carrier systems, too. Indeed, if $N > 1$ carriers are available, the per-operator spectral efficiency and the system sum spectral efficiency modify to

$$SE_o = \sum_{i\in\mathcal{U}_o}\sum_{n=1}^{N}\log_2\left(1+\frac{p_{i,n}\delta_{i,n}}{\sigma^2}\right) \tag{8.15}$$

$$SE = \sum_{i=1}^{I}\sum_{n=1}^{N}\log_2\left(1+\frac{p_{i,n}\delta_{i,n}}{\sigma^2}\right) \tag{8.16}$$

and the per-operator energy efficiency and the GEE modify to

$$EE_o = \frac{\sum_{i\in\mathcal{U}_o}\sum_{n=1}^{N}\log_2\left(1+\frac{p_{i,n}\delta_{i,n}}{\sigma^2}\right)}{\sum_{i\in\mathcal{U}_o}p_{o,i}+\sum_{n=1}^{N}p_{i,n}} \tag{8.17}$$

$$GEE = \frac{\sum_{i=1}^{I}\sum_{n=1}^{N}\log_2\left(1+\frac{p_{i,n}\delta_{i,n}}{\sigma^2}\right)}{\sum_{i=1}^{I}p_{o,i}+\sum_{n=1}^{N}p_{i,n}} \tag{8.18}$$

It is seen that, with respect to the transmit powers $\{p_{i,n}\}_{i,n}$, (8.15) and (8.16) are still concave functions, while (8.17) and (8.18) have concave numerators and affine denominators. Hence, all the requirements to employ the methods from Section 8.3 still hold. ◁

8.4 Results and Analysis

Numerical system-level simulations are shown here to assess the performance of the proposed algorithms. The general use-case considered for the simulation is described in the coming subsection.

8.4.1 Sharing Scenarios with Heterogeneous SLAs

Our motivation for studying heterogeneous SLAs stems from the possibility of diverse services and their requirements offered by current and future networks. For example, an operator focusing on massive machine type communication service with short data packets may be interested more in EE than in achieving a required date rate. Another operator may focus on ultra-broadband service and therefore, would like to have a service quality based on strict data rate requirements.

As an example, the system model of Section 8.2 applies to the following use case. Consider the indoor scenario with a large building such as a big shopping

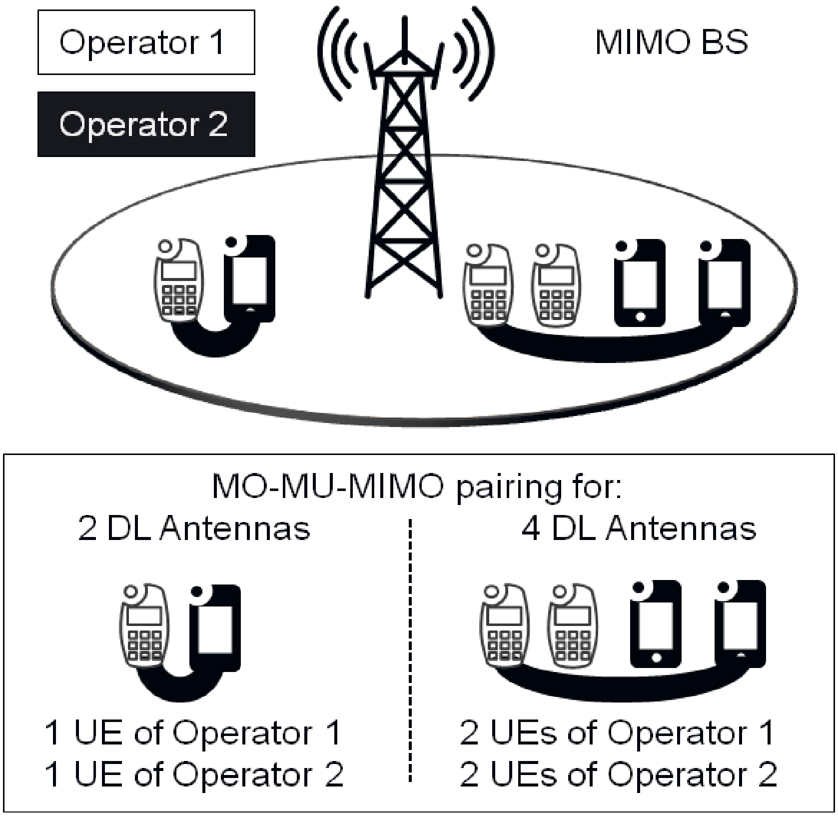

Figure 8.1: Network set-up case-1: with 2 operators with 2 UEs or 4 UEs in total scheduled according number of MIMO antennas at BS owned by one of the sharing operators from [73]

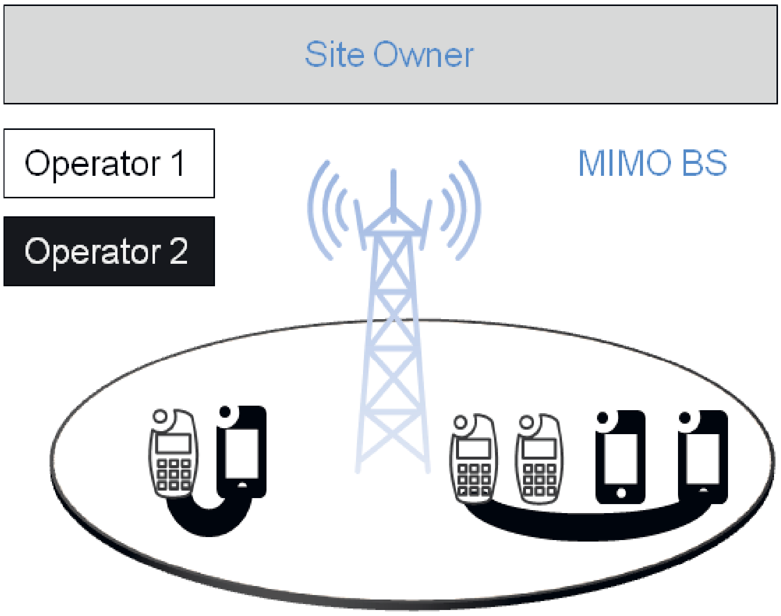

Figure 8.2: Network set-up case 2: with 2 operators sharing resources of a BS controlled by a neutral site owner from [73]

mall, a hotel or an airport. On the one hand it is quite difficult for an operator to provide indoor coverage from an outdoor site. The problem is even more challenging for underground floors. On the other hand it is difficult for the building owner to allow site installations for all operators. This problem will be even more challenging for deployments with millimeter waves. Such business scenarios and their possible architectural options are discussed in the two 3GPP documents [103] and [104]. In line with the mentioned business cases, we consider in our numerical results a site (consisting of BS and TP) in a big shopping mall. The site is either owned by one of the operators or is set-up and controlled by a neutral site owner (host of the shared RAN). The frequency spectrum used by the operators is either pooled spectrum of both or one of the operators provides his frequency resources to the other operator, while the neutral site owner could either rent operators frequency resources and/or focus on unlicensed frequency resources. The neutral site owner offers the shared site as a service provider for the operators and targets an energy-efficient operation of it. However, sharing operators may have different requirements with respect to EE or QoS. In the following, we have considered three different SLAs corresponding to three different case-studies employing this network deployment. The EE and data rate constraints in these three cases are summarized in Table

8.1, with R_o denoting the rate of operator o in bit/s.

Table 8.1: SLA cases with different EE and data rate constraints from [73]

SLA case-study		EE Constraint	Data Rate Constraint
Case-1	Operator 1	$\text{EE}_1^* \geq \{10,15,..40\}$ bit/J	-
	Operator 2	-	$R_2 \geq 5$ bit/s
Case-2	Operator 1	-	$R_1 \geq 2$ bit/s
	Operator 2	-	$R_2 \geq 3$ bit/s
	Site Owner	$\text{GEE}^* \geq \{10,15,..30\}$ bit/J	-
Case-3	Operator 1	-	$R_1 \geq 2$ bit/s
	Operator 2	-	$R_2 \geq 3$ bit/s
	Operator 3	$\text{EE}_3^* \geq \{10,15,..40\}$ bit/J	-

Case-1 reflects the scenario where there is no global system constraint. In this case, Constraints (8.8a), (8.8b), (8.8d) and (8.8e) will be active in the multi-objective optimization. One global system constraint is considered in case-2 where the site owner has defined a GEE constraint for the operation of its TP, either to control the operational cost or to limit the interference level in a cellular network. In addition to that, both operators target minimum required QoS. For this case, the multi-objective optimization is carried out with the constraints (8.8a), (8.8c), (8.8d) and (8.8e). In the third case we consider a similar scenario as in case 1, but with three operators instead of two.

We also assess the performance of the optimization framework with $M = 2$, $M = 3$ and $M = 4$ transmit antennas. The purpose of studying the impact of transmit antennas is to show that the proposed scheme scales with the number of served users or with the number of antennas for MIMO systems.

A baseline case is also considered for the sake of comparison. Specifically, the baseline solves the problem in (8.11) only with (8.8a) and (8.8e). Other constraints on EE, GEE and individual data rates are not considered. The comparison of this (almost) unconstrained programming problem with the constrained cases in Table 8.1 shows the impact of individual EE or rate constraints on the GEE. Hence, this comparison is useful in understanding the impact of heterogeneous SLAs on the overall GEE of the multi-operator shared system.

Table 8.2: Simulation parameters from [73]

Parameter	**Value**
Channel Realizations	1000
Number of Scheduled UEs	$I = 2$, $I = 4$ (SLA case-1,SLA case-2); $I = 3$ (SLA case-3)
Number of Transmission Antennas	$M = 2$, $M = 4$ (SLA case-1,SLA case-2); $M = 3$ (SLA case-3)
Iteration Accuracy for DBA	$\epsilon = 10^{-3}$
Weighting Factor w	$\{0.1, 0.2, ..0.9\}$

8.4.2 Simulation scenario and parameters

The users are randomly dropped with a uniform distribution in the coverage area of the TP. The scheduled users are associated with different operators and the network has full buffer traffic for each user. In every TTI, each operator has equal number of active users. For any user i, the channel from the TP has been modeled as $\mathbf{h}_i = \alpha \tilde{\mathbf{h}}_i$, wherein α models path-loss and other possible slow-fading effects (such as shadowing), while $\tilde{\mathbf{h}}_i$ models fast fading effects and is generated as a zero-mean complex Gaussian vector with identity covariance matrix. In our simulations, we set the channel-to-noise-ratio $\mathbb{E}[\delta_i/\sigma^2] = \alpha^2/\sigma^2 = 20\,\text{dB}$ for any $i = 1, \ldots, I$. Average inter-cell interference is included in σ^2 and treated as noise.

Each simulation is carried out for a given set of constraints, and the same set of independent channel realizations is used to obtain average results. The channel realizations are selected in such a way that the optimization framework always provides a solution (the constraints are feasible)[2]. We focus on the dynamic power consumption for the transmitted data as in [86] with a maximum transmit power P_t (normalized to 1 Watt). The other part of consumed TP power corresponding to, e.g., circuit power and cooling has been considered by p_0. Further simulation parameters are given in Table 8.2.

[2] We observe that the presence at the same time of maximum power constraints and QoS constraints might lead to an empty problem feasible set.

8.4.3 Results and Analysis

In the following, we present simulation results for the cases given in Table 8.1. The EE constraints associated to individual operators have been considered in case-1 and case-3. These constraints span from 10 bit/J to 40 bit/J with a granularity of 5 bit/J. A GEE constraint has been instead considered in case-2.

SLA case-study 1

The results in Figures 8.3 and 8.4 are based on case-study 1 from Table 8.1 with 2 operators and in total 4 UEs as shown on the right side of the network set-up in Figure 8.1. The GEE utility against the weighting factor w is shown in Figure 8.3 for different EE constraints of operator 1. When $w < 0.5$, the GEE is constant and at its minimum level as the maximum transmission power is consumed increase the SE. This can be seen in Figure 8.4 where SE achieves the largest value when $w < 0.5$. As w increases, the the GEE is prioritized compared to the SE and therefore, the GEE value increases for $w > 0.5$. Further, increasing w until $w \geq 0.9$ has a negligible impact on the GEE value.

The influence on the SE of the fixed rate constraint of Operator 2 is seen in Figure 8.4. For $w < 0.5$ only the lowest EE constraint values of 10 bit/J and 15 bit/J can approach the maximum of SE set by the baseline at 14.7 bit/s/Hz. The softer the EE restrictions of Operator 1 are, the higher the SE is (as expected). Still, even when the weighting coefficient is around $w = 0.9$, which prioritizes the GEE, the SE keeps a high rate value (e.g., 7.8 bit/s/Hz for OP1_EE_40) if compared to the baseline scheme, which achieves a rate of 5.2 bit/s/Hz, since it enforces no rate constraint.

Figures 8.5 and 8.6 show that the operator requirements on EE and data rate given in Table 8.1 are indeed met. Specifically, the EE constraint of Operator 1 from 10 bit/J up to 40 bit/J is satisfied as seen in Figure 8.5. For $w > 0.7$, as the multi objective formulation in (8.11) prioritizes the GEE, the EE constraints of Operator 1 from 10 bit/J up to 30 bit/J are fulfiled with strict inequalities. In analogy to the EE constraint, also the rate constraint of Operator 2 is fulfiled as seen in Figure 8.6. For $w > 0.7$ the data rate results are slightly above or equal to 5 bit/s, whereas for $w < 0.7$ the rate performance is clearly above the given data rate constraint.

Figure 8.7 considers the constraint OP1_EE_40 and illustrates the ratio of the transmit power p_i compared to the transmit power that is obtained by the baseline scheme. For $w \leq 0.5$, the baseline case consumes full power as the

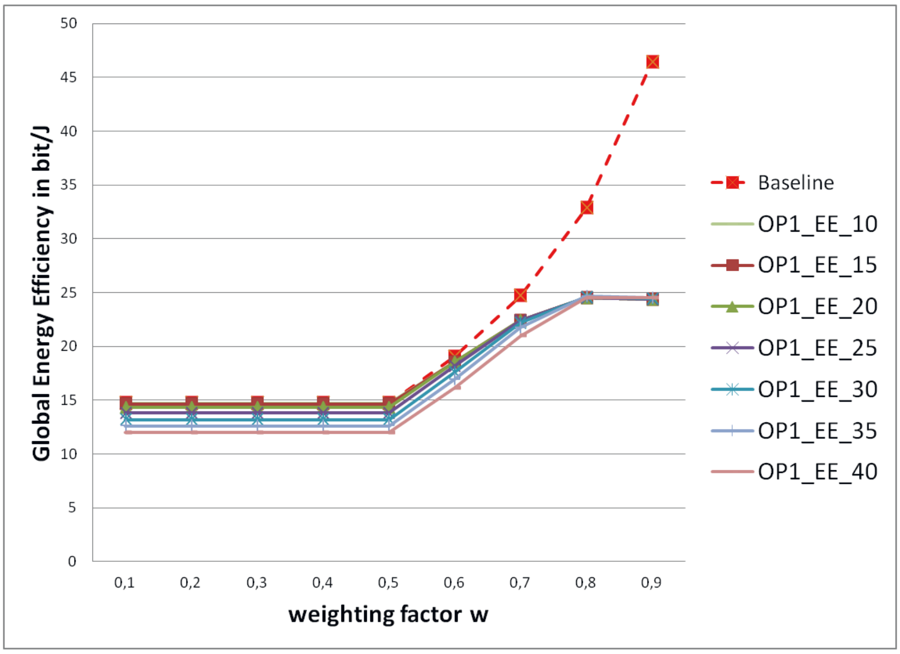

Figure 8.3: Averaged GEE for case-1 with 2 operators and in total 4 UEs from [73]

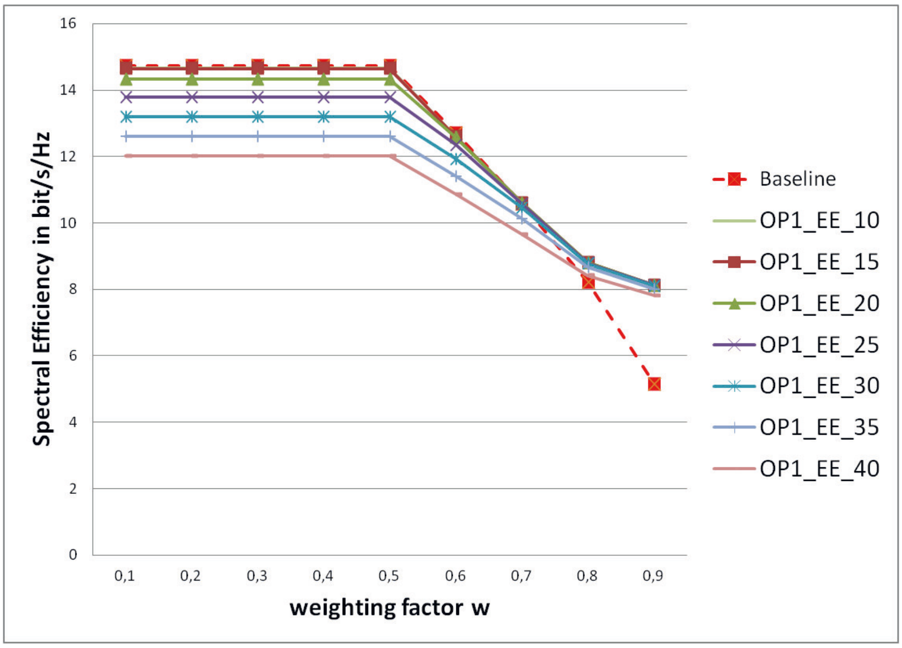

Figure 8.4: Averaged SE for case-1 with 2 operators and in total 4 UEs from [73]

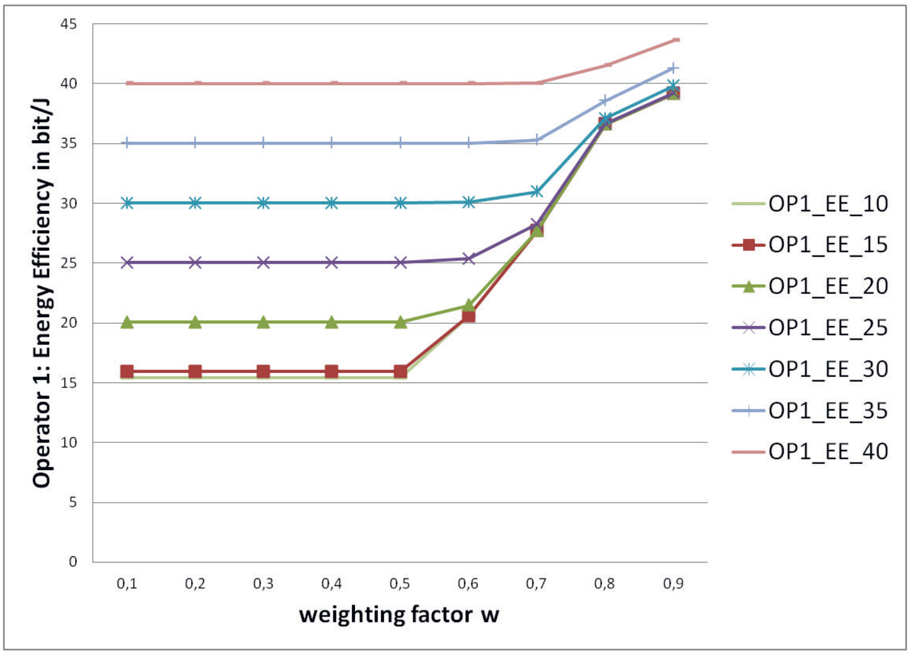

Figure 8.5: Achieved average EE for operator 1 in case-1 with 2 operators and in total 4 UEs from [73]

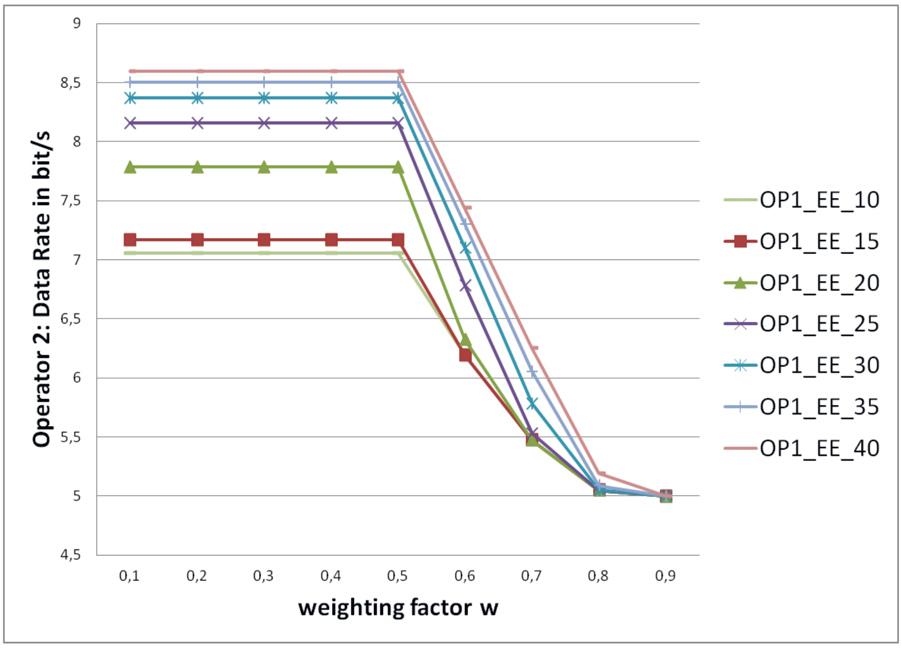

Figure 8.6: Achieved average data rate for operator 2 in case-1 with 2 operators and in total 4 UEs from [73]

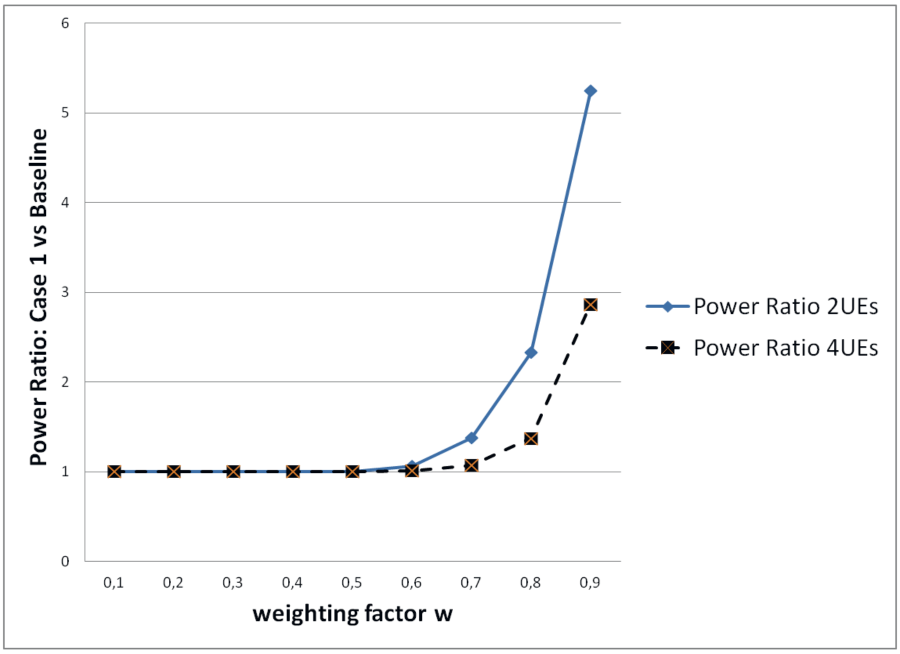

Figure 8.7: Power ratio of the baseline and case-1 with 2 UEs or 4 UEs with an EE constraint for operator 1 ($\mathrm{EE}_1 \geq 40$ bit/J) from [73]

utility function in (8.11) prioritizes the spectral efficiency. For $w \geq 0.5$ the baseline yields the minimum value of transmit power, because it enforces no spectral efficiency constraint. In order to analyze the impact of the number of UEs, both the cases of 2 UEs and 4 UEs are illustrated in the figure. The results show that the case of 4 UEs consumes less power as compared to the case of 2 UEs. This is because the spatial multiplexing due to the higher antenna number at the transmitter grants higher rate values without requiring to increase the transmit power. Indeed, we recall that the number of antennas is increasing with the number of scheduled UEs, as shown in Table 8.2.

The transmission power analysis results show that an optimal constrained solution comes with a price to be paid. In practice, a trade-off between the SLA constraints and the global system performance has to be found and agreed according to the resulting transmit power to be used at the shared TP. The weighting factor w can be used to tune the operation point from spectral-efficient to energy-efficient operation and to control the transmission power.

SLA case-study 2

Figure 8.8 presents the GEE and SE for case 2 of Table 8.1, with 2 UEs or 4 UEs. The site owner wants to operate his wireless cell ensuring a GEE level

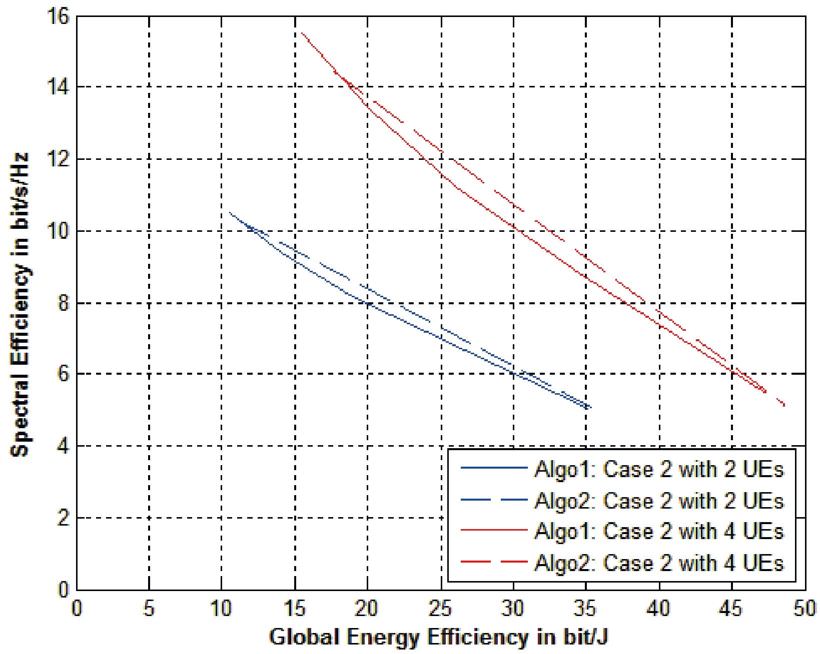

Figure 8.8: GEE and SE for case-2 with Algorithm 1 and Algorithm 2 with a GEE constraint for the site owner (GEE$\geq$ 10 bit/J) and two data rate constraints for operator 1 and 2 with ($R_1 \geq 2$ bit/s, $R_2 \geq 3$ bit/s) from [73]

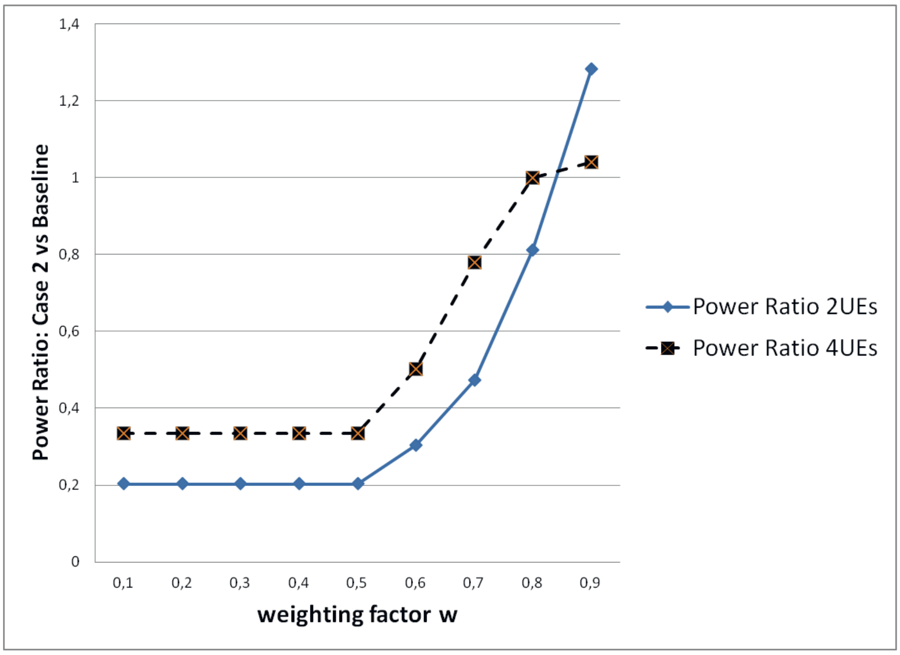

Figure 8.9: Power ratio of the baseline and case-2 with 2 UEs or 4 UEs with a GEE constraint of the site owner (GEE $\geq$ 30 bit/J) from [73]

larger than 10 bit/J, whereas the 2 sharing operators demand rates larger than 2 bit/s and 3 bit/s. The Pareto boundary with 2 UEs has an operation range in SE from 10.5 bit/s/Hz down to 5.0 bit/s/Hz and for the GEE between 10.5 bit/J and 35.2 bit/J. In the same configuration, when 4 UEs are considered, the Pareto boundary is enlarged spanning from 15.5 bit/s/Hz down to 5.5 bit/s/Hz for the SE, and between 15.5 bit/J and 47.3 bit/J for the GEE.

In Figure 8.8, we can further compare the operating points achieved with Algorithm 1 and Algorithm 2. It is important to recall that while Algorithm 1 is able to achieve all points on the Pareto boundary of the GEE-SE Pareto region, Algorithm 2 can only describe the convex hull of the Pareto region. Figure 8.8 confirms this point, since the true boundary computed by Algorithm 1 shows that the Pareto region is not convex, while the boundary obtained by Algorithm 2 defines the convex hull of the region, connecting the boundary end points determined by Algorithm 1.

In Figure 8.9 we compare the power ratios of 2 UEs and 4 UEs for SLA case-2. One can observe for $w < 0.8$, the power ratio is ≤ 1. For $w \leq 0.5$ we have an additional constraint of GEE therefore, in both cases less power is consumed than for the baseline. Now for $w \geq 0.5$, as the power consumed by the baseline decreases, the power ratio increases. For higher values of w, the ratio is larger than one implying higher power consumption by SLA cases. Hence, the range $0.1 \leq w < 0.8$ provides the possibility to operate a shared system with higher energy efficiency at the same time fulfiling the rate requirements from individual operators.

SLA Case-Study 3

The results in Figures 8.10 and 8.11 are based on case-study 3 from Table 8.1, where 3 operators and 3 UEs (1 UE per operator) are considered. This particular SLA case emphasizes the fact that the presented multi-operator algorithmic framework is scalable, and indeed similar results have been obtained as in case-study 1. In Figures 8.10 and 8.11, for $w < 0.5$, the GEE and SE are close to the baseline for the different EE values of Operator 3. Instead, for $w > 0.5$ the GEE increases and the SE decreases as expected. In any case, the SE is still above the baseline guaranteeing that Operator 1 and 2 meet their data rate requirements $R_1 \geq 2$ bit/s or $R_2 \geq 3$ bit/s.

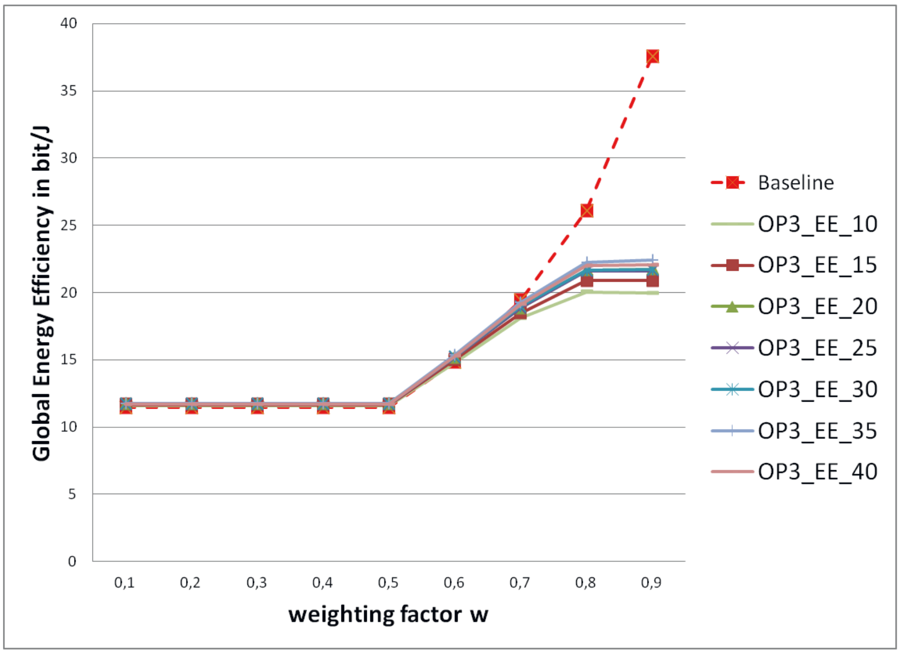

Figure 8.10: Averaged GEE for case-3 with 3 operators and in total 3 UEs from [73]

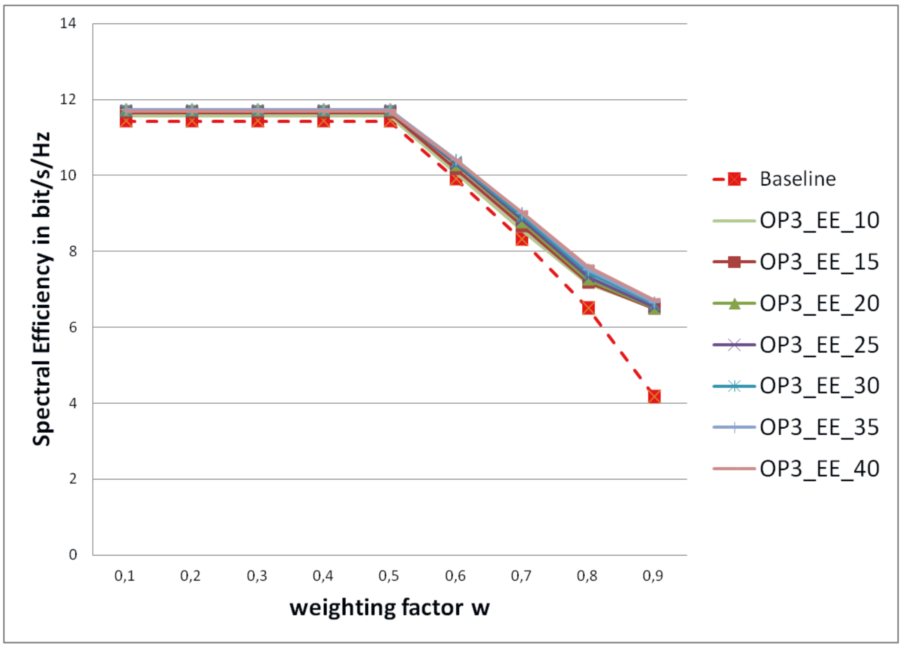

Figure 8.11: Averaged SE for case-3 with 3 operators and in total 3 UEs from [73]

8.5 Conclusion

This scheduling approach has provided a framework to optimize heterogeneous SLAs serving multiple operators, subject to heterogeneous constraints. Two resource allocation algorithm have been presented, considering both noise-limited and interference-limited scenarios. In the noise-limited case, multi-objective optimization theory has been merged with fractional programming theory, which has led to determining the globally optimal resource allocation policy, thus, fully characterizing the system Pareto-boundary. In the next step we have applied the same approach to the interference limited scenario in a multi-carrier system. This case appeared to be more challenging, leading in general to NP-hard resource allocation problems. In order to solve this problem, the tool of sequential optimization has been employed together with multi-objective optimization and fractional programming. The resulting optimization algorithms, although not being proved globally optimal, enjoy strong optimality properties and require the solution only of convex problems.

Part V

Conclusion

Chapter 9

Conclusion and Outlook

Sharing of network resources is important to offer the data capacity and performance needed for the services provided in mobile and wireless communications. The challenge is to provide the most efficient technical sharing solution in different environments, e.g., with different number of sharing MNOs, geographically limited sharing at hot-spot areas and with constraints and requirements driven by regulatory or business aspects. It is therefore, the objective of the present thesis to analyze and develop new sharing solutions which could be implemented for future mobile communication network environments, e.g. 5G.

Accordingly, in this thesis, a comprehensive set of several network sharing approaches for different multi-operator sharing scenarios have been proposed and analyzed with extensive numerical results along with theoretical analyses. They are mainly based on exploiting the effects of multiplexing different radio resources, namely, time, frequency and combination of power and spatial multiplexing resources, i.e., MIMO. Efficient sharing of resources among multiple MNOs has been achieved by using effective scheduling algorithms which exploit possible multiplexing gains.

Guaranteed fairness among the sharing partners was one of the main objectives in almost all the presented approaches. In addition to that, spectral and energy efficiencies were also considered. In each approach, multiple objectives were set for different constraints to keep differentiation factors between operators alive although using same network resource pool. With the help of simulations and analytical methods, solutions were presented to show how the objectives in each approach could be achieved.

In the beginning, in Part II, sharing solutions are provided which are based on simple and known scheduling approaches. They are adapted in the respect that they can provide reliable sharing guarantees among the operators for rate and resource fairness. The shared resources were time and frequency spectrum resources.

As presented in Part III, a new sharing solution is introduced based on spatial multiplexing. It provides fairness jointly based on achieved data rate and used network resources among the MNOs. Important to mention is that fairness is achieved over a time-frequency resource element, which corresponds in 5G to time fractions of 1 ms and smaller.

The highlights of the novel developed scheduling approaches in this thesis are given in Part IV. The main achievement of these sharing solutions is that they show that even a Pareto optimization of SE and EE is achieved in different operator sharing scenarios. This optimization was investigated under different business cases including a business case with a neutral site owner providing the service for the sharing operators guaranteeing data rate fairness and EE fairness in terms of consumed transmission power. The key for solving the Pareto Optimization was the underlying Dinkelbach algorithm.
These deeply analyzed approaches provide an insight to future applications of this algorithm for Pareto optimization of SE and EE e.g. in a network slicing scenario which is one of the main 5G features as presented in [105]. In a follow up study [106], we have highlighted the applicability of the present results of this thesis for vehicular-to-vehicular (V2V) and vehicular-to-network (V2N) use cases with their associated QoS requirements which can be mainly guaranteed by the realization of wireless resource sharing either in a multi-operator scenario or based on a network slicing solution. In both future scenarios, the presented thesis with its sharing models and achieved results can deliver solutions for implementation of multi-operator based service provisioning.

Part VI

Appendix

Appendix A

Algorithm: Two-Step Scheduler (Chapter 3)

In:
Set of sharing operators: $\mathcal{J}$
Set of active users: $\mathcal{K}$
Operator's sharing ratio: $g := g_1, .., g_J$
Duration of scheduling period: T
Users' instantaneous SNR: $\gamma := \gamma_1, .., \gamma_K$
Proportional Fairness parameters: α, β
Number of scheduled users per time frame: k_{max}
Bandwidth: B
Out:
Allocated rate matrix: $\mathcal{S} \in \mathbb{R}^{JxK}$

//**Phase 1:** Distribute transmission time among operators
for each operator: $j \in \mathcal{J}$
    $\Delta n_j = |g_j T|$
end //for each operator

//**Phase 2:** Distribute channel resources among users
for each operator: $j \in \mathcal{J}$
    //choose operator to schedule in n via Round Robin
    $j^* := j \mod J + 1$
    for time slot $n \leq \Delta n_j$**:**
        for each user $k \in \mathcal{K}$**:**

```
22.             //calculate instantaneous rate
23.             r_k := log_2(1 + γ_k)
24.             // PF weight calculation
25.             w_k = r_k^α / R_k^β
26.         end //for each user
27.         r'_k = sort(r, with highest w_k first),
28.         // Assign rate to the first users
29.         for user u ≤ k_max
30.             S_ju = (B / k_max) r'_u
31.         end //for k_max users
32.     end //for each time slot
33. end //for each operator
34. return S
```

In the first scheduling phase, we grant each operator a fraction g of the scheduling period duration T. Applying the floor function in line 14 provides the number of time slots Δn_j for an arbitrary operator J. In the second phase, this number of slots is allocated to each operator following the Round Robin principle. Within each time slot, an operator performs its own policy to schedule its users. In Algorithm TSS, we chose the Proportional Fair policy as a widely-used example. The algorithm returns the matrix S, which includes non-zero rates for the scheduled users of all operators. The algorithm operation is illustrated in Figure 3.2. Per scheduling period, each operator $j \in \mathcal{J}$ receives Δn_j time slots. After J of such allocations, a new scheduling period starts. Two of such periods are illustrated in Figure 3.2.

Appendix B

Algorithm: Generalized Processor Sharing (Chapter 4)

In:
Set of sharing operators: $\mathcal{J}$
Set of active users: $\mathcal{K}$
Operator's sharing ratio: $g := g_1, .., g_J$
Users' instantaneous SINR: $\gamma := \gamma_1, .., \gamma_K$
Proportional Fairness parameters: α, β
Initial set of feasible rates: $f \in \mathbb{R}^{JxK}, \forall j \in \mathcal{J}, \forall k \in \mathcal{K} : f_{jk} = 0$
Out:
Allocated rate matrix: $\mathcal{S} \in \mathbb{R}^{JxK}$

//**Phase 1:** Calculate maximum system data rate
for each operator: $j \in \mathcal{J}$
    $k'_j := argmax_{j \in \mathcal{J}}(\gamma_{jk})$
end //for each operator
    $k^* := max_{j \in \mathcal{J}}(k'_j)$
    $R_{max} := \log_2(1 + \gamma_{k*})$

//Phase 2: Construct feasible rate matrix f
for each operator: $j \in \mathcal{J}$
    //Share R_{max} among operators
    $c_j := g_j R_{max}$
        for each user $k \in \mathcal{K}$**:**

//calculate instantaneous rate

$r_{jk} := \log_2(1 + \gamma_{jk})$

// PF weight calculation

$w_{jk} = \frac{r_{jk}^{\alpha}}{R_{jk}^{\beta}}$

//Select users with feasible rates

if $r_{jk} \leq c_j$

$f_{jk} = r_{jk}$

end //if

end //for each user

end //for each operator

//Phase 3: Scheduling users

for each operator: $j \in \mathcal{J}$

$f'_{j,\forall k \in \mathcal{K}} = sort(f_{j,\forall k \in \mathcal{K}}$, with highest w_{jk} first)

//Pack users in remaining capacity Δc_j

$\Delta c_j := c_j$

for each user $k \in \mathcal{K}$**:**

if $f'_{j,k} \leq c_j$

//Allocate rate to scheduled user

$S_{jk} = f'_{j,k}$

$\Delta c_j = \Delta c_j - f'_{j,k}$

else $S_{jk} = 0$

end //if

end //for each user

end //for each operator

return S

Appendix C

Algorithm: MOMU-MIMO (Chapter 6)

In:

Set of active users from each operator: $\mathcal{L}_1$ and $\mathcal{L}_2$

Service shares: (g_1, g_2),$(g_1 \leq g_2)$ and $(g_1 + g_2) = 1$

Total system power: P_t

Input SNRs

Incremental Power Step: Δ_p

Out:

Set of Selected Users: $\mathcal{S} = \{m, n\}$

Power Allocation: p_m and p_n

for each resource element

  //**Step 1:** Create UE pairs (all possible candidate sets) according to (6.8)

  //**Step 2:** Determining calculation interval c_i

    ci=min(λ, min(g_1, g_2)

    Lower threshold: lt=g_1-ci

    Upper threshold: ut=g_1+ci

  //**Step 3:** Transmit power allocation

  for each pair of UEs in candidate set $\mathcal{S}_c$= $\{u, v\}$, **where** $u \in \mathcal{L}_1$ **and** $v \in \mathcal{L}_2$

    //**Step a:** Initialization $(l = 0)$

    // l is the number of iteration step

    Compute ${}^{l=0}p_u$ and ${}^{l=0}p_v$ such that $R_u = R_v$

    //**Step b:** /Check for the rate condition

do{

$p_u(l) = p_u(l-1) - \Delta_p$

$p_v(l) = p_v(l-1) + \Delta_p$

$R_u(l) = log_2(1 + \frac{\delta_u}{\sigma^2} p_u(l))$

$R_v(l) = log_2(1 + \frac{\delta_v}{\sigma^2} p_v(l))$

$r_u(l) = \frac{R_u(l)}{R_u(l)+R_v(l)}, r_v(l) = \frac{R_v(l)}{R_u(l)+R_v(l)}$

}**while** ($(lt{<}^l\ r_u < ut)${

$l = l + 1$

}

//**Step c:** $R_u \leftarrow R_u^l$ and $R_v \leftarrow R_v^l$

//**Step d:** Calculate: $R(\mathcal{S}_c) = R_u + R_v$

end //each UE pair

//**Step 4:** // Select the set $\mathcal{S}$ out of candidate sets according to (6.9)

return $\mathcal{S} = \{m, n\}, p_m, p_n$

end // for each resource element

Part VII

Bibliography

Introduction

[1] S. Parkvall, A. Furuskar, and E. Dahlman. *Evolution of LTE Toward IMT-advanced*. Feb. 2011.

[2] 3GPP. *Long Term Evolution of the 3GPP radio technology*. Nov. 2004. URL: ftp.3gpp.org/Inbox/2008_web_files/LTA_Paper.pdf.

[3] P. Marsch, Ö. Bulakci, O. Queseth, and M. Boldi. *5G System Design: Architectural and Functional Considerations and Long Term Research*. 2018.

[4] 3GPP. *Service requirements for the 5G system; Stage 1 (Release 16), TS 22.261 , V16.1.0*. TS. 3rd Generation Partnership Project (3GPP), Sept. 2017.

[5] M. Tesanovic et al. *Design framework and suitability assessment proposal for 5G air interface candidates*. Oct. 2016.

[6] 3GPP. *Study on New Radio (NR) Access Technology; Physical Layer Aspects (Release 14), TR 38.802 V2.0.0*. Tech. rep. 3rd Generation Partnership Project (3GPP), Mar. 2017.

[7] 3GPP. *Work Item on New Radio (NR) Access Technology, RP-172115*. WI. 3rd Generation Partnership Project (3GPP), Sept. 2017.

[8] 3GPP. *Study on NR-based Access to Unlicensed Spectrum, RP-171496*. SI. 3rd Generation Partnership Project (3GPP), June 2017.

[9] C. E. Shannon. *Communication in the Presence of Noise*. Jan. 1949.

[10] L. Badia et al. *Resource Sharing among Wireless Network Operators: Spectrum, Infrastructure, and Full Radio Access Network Sharing*. Mar. 2013. URL: http://www.saphyre.eu/publications/index.html.

[11] S. Banerjee, T. Loozen, M. Fanno, and A. Saksena. *The Future Communication Service Provider*. 2015. URL: http://www.accenture.com.

[12] M. A. Marsan and M. Meo. *Network sharing and its energy benefits: A study of European mobile network operators*. Dec. 2013.

[13] I. Sugathapala, I. Kovacevic, B. Lorenzo, S. Glisic, and Y. Fang. *Quantifying Benefits in a Business Portfolio for Multi-Operator Spectrum Sharing*. Dec. 2015.

[14] V. Mancuso and S. Alouf. *Reducing Costs and Pollution in Cellular Networks*. Aug. 2011.

[15] Z. Salman. *The Economic and Sustainability Future of Cellular Networks.* Aug. 2013. URL: `http://arxiv.org/abs/1308.0795`.

[16] Ericsson White Paper. *5G Radio Access.* Apr. 2016. URL: `http://www.ericsson.com`.

[17] M. Mouly and M. B. Pautet. *The GSM System for Mobile Communications.* 1992.

[18] H. Holma and A. Toskala. *WCDMA for UMTS Radio Access for Third Generation Mobile Communication.* 2000.

[19] G. Foschini and M. Gans. *On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas.* Mar. 1998. URL: `https://doi.org/10.1023/A:1008889222784`.

[20] D. Gesbert and J. Akhtar. *Breaking the barriers of Shannon's capacity: An Overview Of Mimo wireless systems.* 2002.

[21] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath. *Capacity limits of MIMO channels.* June 2003.

[22] T. A. Weiss and F. K. Jondral. *Spectrum pooling: an innovative strategy for the enhancement of spectrum efficiency.* Mar. 2004.

[23] G. Alnwaimi, K. Arshad, and K. Moessner. *Dynamic Spectrum Allocation Algorithm with Interference Management in Co-Existing Networks.* Sept. 2011.

[24] X. Zeng, Z. Feng, V. Le, Y. Xue, and Y. Lin. *An Auction Based Joint Radio Resource Management Scheme and Architecture in a Multi-Operator Scenario.* May 2008.

[25] A. Padaki, A. Sengupta, M. Abdelbar, J. H. Reed, and W. H. Tranter. *An Orthogonal Spectrum Sharing Scheme for Cognitive LTE Networks.* Mar. 2014.

[26] T. Sanguanpuak, S. Guruacharya, N. Rajatheva, M. Bennis, and M. Latva. *Multi-Operator Spectrum Sharing for Small Cell Networks: A Matching Game Perspective.* June 2017.

[27] J. Zhan, F. Roemer, and M. Haardt. *Relay Assisted Physical Resource Sharing: Projection Based Separation Multiple Operators (ProBaSeMO) for Two-Way Relaying With MIMO Amplify and Forward Relays.* Sept. 2012.

[28] P. Luoto, P. Pirinen, M. Bennis, S. Samarakoon, S. Scott, and M. Latva-aho. *Co-Primary Multi-Operator Resource Sharing for Small Cell Networks*. June 2015.

[29] M. K. Pereirasamy, J. Luo, M. Dillinger, and C. Hartmann. *Dynamic inter-operator spectrum sharing for UMTS FDD with displaced cellular networks*. Mar. 2005.

[30] B. Aazhang, J. Lilleberg, and G. Middleton. *Spectrum sharing in a cellular system*. Jan. 2004.

[31] G. Salami, A. U. Quddus, D. Thilakawardana, and R. Tafazolli. *Nonpool based spectrum sharing for two UMTS operators in the UMTS extension band*. Sept. 2008.

[32] G. Salami, D. Thilakawardana, and R. Tafazolli. *Dynamic Spectrum Sharing Algorithm Between Two UMTS Operators in the UMTS Extension Band*. June 2009.

[33] Y. Lin, H. Chien, H. Chang, and C. Lai. *Multi-operator Fairness in Transparent RAN Sharing*. Mar. 2018.

[34] J. Opadere, Q. Liu, T. Han, and N. Ansari. *Energy-Efficient Virtual Radio Access Networks for Multi-Operators Cooperative Cellular Networks*. Sept. 2019.

[35] S. Farhat, B. H. Hassan, and A. Ellatif Samhat. *Energy Efficient Resource Sharing in Multi-Operator Heterogeneous Cloud RAN*. Oct. 2018.

[36] Y. Xiao, M. Krunz, and T. Shu. *Multi-Operator Network Sharing for Massive IoT*. Apr. 2019.

[37] S. Park, O. Simeone, and S. Shamai. *Multi-Tenant C-RAN With Spectrum Pooling: Downlink Optimization Under Privacy Constraints*. Nov. 2018.

[38] S. Park, O. Simeone, and S. Shamai Shitz. *Optimizing Spectrum Pooling for Multi-Tenant C-RAN Under Privacy Constraints*. June 2018.

Sharing Models

[39] 3GPP. *Service aspects and requirements for network sharing*. Technical Report TR 22.951 V15.0.0. 3GPP, June 2018.

[40] GSMA. *Mobile Infrastructure Sharing*. Sept. 2012. URL: `http://www.gsma.com/publicpolicy/wp-content/uploads/2012/09/Mobile-Infrastructure-sharing.pdf/`.

[41] Nokia. *Network Sharing: Delivering mobile broadband more efficiently and at lower cost.* URL: https://onestore.nokia.com/asset/200192/.

[42] 3GPP. *Network Sharing;Architecture and functional description.* Technical Report TR 23.251 V15.1.0. 3GPP, Sept. 2018.

[43] D. Tse and P. Viswanath. "Fundamentals of Wireless Communication". In: Cambridge University Press, 2005.

[44] T. Rappaport. *Wireless communications: Principles and practice.* 1996.

[45] J. Hultell, K. Johansson, and J. Markendahl. *Business models and resource management for shared wireless networks.* Sept. 2004.

[46] A. Akbari-Moghanjoughi, J. Amazonas, G. Boada, and J. Solé-Pareta. *Service Level Agreements for Communication Networks: A Survey.* June 2019.

[47] GSMA. *Network Slicing Use Case Requirements.* Apr. 2018. URL: https://www.gsma.com/.

[48] 5GAA. *White paper: C-V2X Use Cases, Methodolgy, Examples and Service Level Requirements.* June 2019. URL: https://5gaa.org/.

Two Step Scheduler

[49] E. Jorswieck, L. Badia, T. Fahldieck, E. Karipidis, and J. Luo. *Spectrum sharing improves the network efficiency for cellular operators.* Mar. 2014.

[50] O. Aydin, W. Jamil, and S. Valentin. "A Two-Step Scheduler for the Dynamic Sharing of Wireless Channel Resources Among Operators". In: *Proc. Vehicular Technology Conf. VTC-Spring.* June 2013.

[51] 3GPP. *3GPP TS 36.321 V15.3.0 Medium Access Control (MAC) protocol specification (Release 15).* Sept. 2018.

Generalized Processor Sharing

[52] S. Valentin, W. Jamil, and O. Aydin. *Extending Generalized Processor Sharing for multi-operator scheduling in cellular networks.* July 2013.

[53] S. A. AlQahtani. *Adaptive rate scheduling for 3G networks with shared resources using the generalized processor sharing performance model.* Jan. 2008.

Generalized Resource Sharing

[54] I. Malanchini, S. Valentin, and O. Aydin. "Generalized resource sharing for multiple operators in cellular wireless networks". In: *2014 International Wireless Communications and Mobile Computing Conference (IWCMC)*. Aug. 2014, pp. 803–808.

[55] I. Malanchini, S. Valentin, and O. Aydin. "An analysis of generalized resource sharing for multiple operators in cellular networks". In: *2014 IEEE 25th Annual International Symposium on Personal, Indoor, and Mobile Radio Communication (PIMRC)*. Sept. 2014, pp. 1157–1162.

[56] I. Malanchini, S. Valentin, and O. Aydin. "Wireless resource sharing for multiple operators: Generalization, fairness, and the value of prediction". In: *Computer Networks* 100 (2016), pp. 110–123.

[57] G. Salami and R. Tafazolli. "Inter-operator Dynamic Spectrum Sharing: (Analysis, Cost and Implications)". In: *International Journal of Computer Networks*. Mar. 2010.

[58] Y. Okumura. "Field strength and its variability in VHF and UHF land-mobile radio service". In: *Rev. Electr. Commun. Lab.* 16 (1968), pp. 825–873. URL: `https://ci.nii.ac.jp/naid/10010001461/en/`.

[59] M. Hata. "Empirical formula for propagation loss in land mobile radio services". In: *IEEE Transactions on Vehicular Technology* 29.3 (Aug. 1980), pp. 317–325.

[60] S. Stanczak, M. Wiczanowski, and H. Boche. *Fundamentals of Resource Allocation in Wireless Networks*. Vol. 3. Springer, Berlin, 2009.

[61] A. Goldsmith. *Wireless Communications*. Cambridge Univ. Press, 2005.

[62] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[63] P. Viswanath, D. Tse, and R. Laroia. "Opportunistic beamforming using dumb antennas". In: 48.6 (June 2002), pp. 1277–1294.

Multi-Operator Multi-User MIMO (Single Objective)

[64] O. Aydin, D. Aziz, and E. Jorswieck. *Radio Resource Sharing among Operators through MIMO based Spatial Multiplexing in 5G Systems*. Dec. 2014.

[65] R. Gangula, D. Gesbert, J. Lindblom, and E. Larsson. *On the Value of Spectrum Sharing among Operators in Multicell Networks*. June 2013.

[66] 3GPP. *Study on Radio Access Network sharing enhancements*. Technical Report TR 22.852 V13.1.0. 3GPP, Sept. 2014.

Improved Multi-Operator Multi-User MIMO (Single Objective)

[67] O. Aydin, E. Jorswieck, and D. Aziz. *Rate Fairness Based QoS Provisioning for Operators in 5G Shared Networks*. May 2015.

[68] E. W. Weisstein. *Newton's Method*.

[69] METIS. "Mobile and wireless communications Enablers for the Twenty twenty Information Society". In: 2012. URL: https://metis2020.com.

[70] B. Singh, K. Koufos, and O. Tirkkonen. *Coordination protocol for inter-operator spectrum sharing based on spectrum usage favors*. June 2014. URL: https://metis2020.com.

[71] S. Hailu, A. A. Dowhuszko, O. Tirkkonen, and L. Wei. *One-shot games for spectrum sharing among co-located radio access networks*. Nov. 2014.

[72] R. Etkin, A. Parekh, and D. Tse. *Spectrum sharing for unlicensed bands*. Nov. 2005.

Multi-Operator Multi-User MIMO (Multi-Objective)

[73] O. Aydin, E. Jorswieck, D. Aziz, and A. Zappone. *Energy-Spectral Efficiency Tradeoffs in 5G Multi-Operator Networks With Heterogeneous Constraints*. Sept. 2017.

[74] NGMN Alliance. *NGMN 5G White Paper*. Mar. 2015.

[75] L. Badia and et all. *Resource Sharing between Operators*. Mar. 2013. URL: http://www.saphyre.eu/publications/index.html/.

[76] 5G Infrastructure Public Private Partnership. *5G Vision*. Feb. 2015. URL: https://5g-ppp.eu/wp-content/uploads/2015/02/5G-Vision-Brochure-v1.pdf.

[77] M. A. Khan and et all. *Cooperative game theoretic approach to integrated bandwidth sharing and allocation*. Istanbul, Turkey, May 2009.

[78] S. Buzzi, C.-L. I, T. E. Klein, H. V. Poor, C. Yang, and A. Zappone. *A Survey of Energy-Efficient Techniques for 5G Networks and Challenges Ahead*. 2016.

[79] A. Zappone and E. Jorswieck. *Energy Efficiency in Wireless Networks via Fractional Programming Theory*. 2015.

[80] X. Xiao, X. Tao, and J. Lu. *QoS-Aware Energy-Efficient Radio Resource Scheduling in Multi-User OFDMA Systems*. Jan. 2013.

[81] S. Kim, B. G. Lee, and D. Park. *Energy-Per-Bit Minimized Radio Resource Allocation in Heterogeneous Networks*. Apr. 2014.

[82] C. He, G. Y. Li, F. C. Zheng, and X. You. *Energy-Efficient Resource Allocation in OFDM Systems With Distributed Antennas*. Mar. 2014.

[83] Q. Xu, X. Li, H. Ji, and X. Du. *Energy-Efficient Resource Allocation for Heterogeneous Services in OFDMA Downlink Networks: Systematic Perspective*. June 2014.

[84] R. ArabLoodaricheh, S. Mallick, and V. K. Bhargava. *Energy-Efficient Resource Allocation for OFDMA Cellular Networks With User Cooperation and QoS Provisioning*. Nov. 2014.

[85] A. Zappone, L. Sanguinetti, G. Bacci, E. Jorswieck, and M. Debbah. *Energy Efficient Power Control: A Look at 5G Wireless Technologies*. Apr. 2016.

[86] O. Tervo, L. Tran, and M. Juntti. *Optimal Energy Efficient Transmit Beamforming for Multi-User MISO Downlink*. Oct. 2015.

[87] C. Xiong, G. Y. Li, S. Zhang, Y. Chen, and S. Xu. *Energy- and Spectral-Efficiency Tradeoff in Downlink OFDMA Networks*. Nov. 2011.

[88] X. Chen and S. Ouyang. *Energy- and Spectral-Efficiency Trade-Off in OFDMA-Based Cooperative Cognitive Radio Networks*. 2014.

[89] D. W. K. Ng, E. S. Lo, and R. Schober. *Energy-Efficient Resource Allocation in Multi-Cell OFDMA Systems with Limited Backhaul Capacity*. Oct. 2012.

[90] S. He, Y. Huang, S. Jin, and L. Yang. *Coordinated Beamforming for Energy Efficient Transmission in Multicell Multiuser Systems*. Dec. 2013.

[91] L. Venturino, A. Zappone, C. Risi, and S. Buzzi. *Energy Efficient Scheduling and Power Allocation in Downlink OFDMA Networks With Base Station Coordination.* Jan. 2015.

[92] B. Du, C. Pan, W. Zhang, and M. Chen. *Distributed Energy-Efficient Power Optimization for CoMP Systems With Max-Min Fairness.* June 2014.

[93] K. G. Nguyen, L. N. Tran, O. Tervo, Q. Vu, and M. Juntti. *Achieving Energy Efficiency Fairness in Multicell MISO Downlink.* Aug. 2015.

[94] O. Tervo, A. Tolli, M. Juntti, and L. Tran. *Energy-Efficient Beam Coordination Strategies with Rate Dependent Processing Power.* 2016. URL: http://arxiv.org/abs/1610.02307.

[95] X. Chen, X. Xu, and X. Tao. *Energy Efficient Power Allocation in Generalized Distributed Antenna System.* July 2012.

[96] G. Brante, I. Stupia, R. Souza, and L. Vandendorpe. *Outage Probability and Energy Efficiency of Cooperative MIMO with Antenna Selection.* Nov. 2013.

[97] O. Onireti, F. Heliot, and M. Imran. *On the Energy Efficiency-Spectral Efficiency Trade-Off of Distributed MIMO Systems.* Sept. 2013.

[98] V. Varma, S. Lasaulce, M. Debbah, and S. Elayoubi. *An Energy-Efficient Framework for the Analysis of MIMO Slow Fading Channels.* May 2013.

[99] K. Miettinen. *Nonlinear Multiobjective Optimization.* Springer, 1999.

[100] W. Dinkelbach. *On nonlinear fractional programming.* Mar. 1967.

[101] J. P. Crouzeix, J. A. Ferland, and S. Schaible. *An algorithm for generalized fractional programs.* 1985. URL: http://dx.doi.org/10.1007/BF00941314.

[102] J. E. Falk and S. W. Palocsay. *Recent Advances in Global Optimization.* Ed. by C. A. Floudas and P. M. Pardalos. Princeton, NJ, USA, 1992. URL: http://dl.acm.org/citation.cfm?id=137305.137325.

[103] 3GPP. *Service aspects and requirements for network sharing.* Technical Report 22.951 V 13.0.0. 3GPP, Dec. 2015.

[104] 3GPP. *Network Sharing; Architecture and functional description.* Technical Specification 23.251 V 13.2.0. 3GPP, June 2016.

Conclusion

[105] 5GAA. *Cellular V2X: Conclusions based on Evaluation of Available Architectural Options*. Feb. 2019. URL: `http://5gaa.org/5gaa-in-motion/news/#white-papers/`.

[106] B. Matthiesen, O. Aydin, and E. Jorswieck. "Throughput and Energy-Efficient Network Slicing". In: *WSA 2018; 22nd International ITG Workshop on Smart Antennas*. Mar. 2018, pp. 1–6.