# Suboptimal Resource Allocation for Multi-User MIMO-OFDMA Systems

Vom Fachbereich 18
Elektrotechnik und Informationstechnik
der Technischen Universität Darmstadt
zur Erlangung der Würde eines
Doktor-Ingenieurs (Dr.-Ing.)
genehmigte Dissertation

von
M.Sc. Tarcisio Ferreira Maciel
geboren am 15.05.1977 in Fortaleza, Brasilien

| | |
|---|---|
| Referent: | Prof. Dr.-Ing. Anja Klein |
| Korreferent: | Prof. Dr.-Ing. Martin Haardt |
| Tag der Einreichung: | 14. April 2008 |
| Tag der mündlichen Prüfung: | 22. September 2008 |

D 17
Darmstädter Dissertation
Darmstadt 2008

# Acknowledgments

This thesis was prepared in the period of time between April 2005 and September 2008, during which I have been with the Communications Engineering Lab at the Institute of Telecommunications of the Technische Universität Darmstadt.

I would like to thank Prof. Dr.-Ing. Anja Klein for her support, incentive, guidance, and valuable suggestions and comments during the supervision of my studies, which have decisively contributed to the elaboration of this work.

I also thank Prof. Dr.-Ing. Martin Haardt from the Technishe Universität Ilmenau for taking the time to be on my dissertation committee and evaluate my work.

Moreover, I thank Prof. Dr.-Ing. Dr. h.c. Willmut Zschunke from the Technische Universität Darmstadt for his support and Prof. Dr.-Ing. habil. Dr.-Ing. E.h. Paul Walter Baier from Technische Universität Kaiserslautern for reviewing parts of this thesis. My acknowledgments also to the colleagues and to the administrative staff at the Communications Engineering Lab.

I would like to acknowledge the scholarship suport of DAAD and CAPES-Brazil.

Finally, I would like to thank my family for their love and support. I especially thank my wife Patrícia, who stood by me at all times.

Darmstadt, September 2008.

Tarcisio F. Maciel

# Kurzfassung

Von zukünftigen drahtlosen Kommunikationssystemen wird erwartet, dass sie verschiedenste Datendienste zuverlässig zur Verfügung stellen, wobei diese Dienste Raten im Bereich von wenigen kbit/s bis zu mehreren Mbit/s fordern. Wegen der hohen Kosten für Funkfrequenzen müssen diese Systeme außerdem besonders effizient bezüglich der Spektrumsnutzung sein. Die Anwendung von Orthogonal Frequency Division Multiple Access (OFDMA) und Multiple Input Multiple Output (MIMO) basierten Verfahren wird als besonders vielversprechend angesehen, um diesen Anforderungen zu genügen.

Auf der einen Seite sind MIMO-OFDMA Systeme sehr flexibel und besitzen eine hohe spektrale Effizienz. Auf der anderen Seite ist die Zuweisung der Funkressourcen aufgrund der erheblichen Anzahl von Sub-Trägern und der Berücksichtigung der räumlichen Komponente besonders komplex. Die optimale Zuweisung der Funkressourcen, die die Summenrate des Systems maximiert, ist meist zu komplex für praktische Anwendungen. Daher werden suboptimale und effiziente Verfahren zur Funkressourcenzuweisung mit geringer Komplexität benötigt, die den Mobilstationen die verfügbaren Frequenz-, Zeit- und Raumressourcen des Systems zuteilen.

Diese Arbeit befasst sich mit suboptimalen Verfahren zur Funkressourcenzuweisung mit dem Ziel, die Summenrate des Systems zu maximieren. Um ein effizientes Verfahren zur Funkressourcenzuweisung mit akzeptabler Komplexität zu entwerfen, wird das ursprüngliche Problem der Summenratenmaximierung des Systems neu formuliert, wobei es in vier Unterprobleme zerlegt wird. Diese sind das Space Division Multiple Access (SDMA)-Gruppierungsproblem, das Vorkodierungsproblem, das Leistungszuweisungsproblem und das Ressourcenvergabeproblem. Für jedes dieser Unterprobleme werden verschiedene existierende und neu vorgeschlagene Algorithmen angewendet, die alle die Maximierung der Summenrate des Systems zum Ziel haben. Durch die Kombination dieser Algorithmen entstehen suboptimale Verfahren zur Funkressourcenzuweisung, die jedoch äußerst effizient arbeiten.

Für das SDMA-Gruppierungsproblem werden vier neue SDMA-Algorithmen vorgestellt: ein Algorithmus basiert auf konvexer Optimierung und drei Greedy-Algorithmen basieren auf einfachen heuristischen Ansätzen. Die vorgeschlagenen Algorithmen erzeugen die SDMA-Gruppen anhand von räumlichen Korrelationseigenschaften und Kanalgewinnen der Mobilstation und benötigen keine Vorkodierung oder Leistungszuweisung. Es wird gezeigt, dass die vorgestellten Algorithmen bezüglich der mittleren Summenrate genauso gute Ergebnisse liefern wie einige existierende SDMA-Algorithmen, wobei sie beachtlich niedrigeren Rechenaufwand als die existieren Verfahren benötigen.

Zur Lösung des Vorkodierungsproblems werden zwei existierende Algorithmen angewendet. Für das Leistungszuweisungsproblem wird ein neuer iterativer Soft Dropping Algorithm (SDA) vorgeschlagen, der nachträglich mit Generalized Eigen-Precoding (GEP) kombiniert wird und zu einem neuen Algorithmus führt, der

Vorkodierung und Leistungszuweisung vereint. Des Weiteren wird die Konvergenz dieses neuen Algorithmus gezeigt. Eine besondere Eigenschaft ist, dass der SDA und damit auch der Algorithmus, der Vorkodierung und Leistungszuweisung vereint, entweder zur Maximierung der Summenrate oder zur Sicherstellung von Dienstgütekriterien (QoS-Kriterien) an der Mobilstation mit Hilfe einfacher Parametereinstellungen genutzt werden können. Des Weiteren wird bei den Algorithmen zur Vorkodierung und Leistungszuweisung ein neuer Sequential Removal Algorithm (SRA) vorgeschlagen, der es ermöglicht, Mobilstationen aus SDMA-Gruppen zu entfernen, wenn dies die Summenrate erhöht.

Um das Ressourcenvergabeproblem zu lösen, werden Algorithmen vorgestellt und verglichen, die entweder getrennte oder gemeinsame SDMA-Gruppierung und Ressourcenvergabe verwenden. Es wird gezeigt, dass die getrennte Vergabe der Ressourcen zu den SDMA-Gruppen genauso gute Ergebnisse bezüglich der Maximierung der Summenrate des Systems liefert wie die gemeinsame Verarbeitung der SDMA-Gruppierung und der Ressourcenvergabe, wobei der getrennte Ansatz deutlich einfacher ist. Des Weiteren werden durch die Algorithmen zur Ressourcenvergabe verschiedene Kriterien zur Priorisierung von Mobilstationen oder SDMA-Gruppen berücksichtigt, und es wird gezeigt, dass durch geschickte Anpassung der Prioritätskriterien die Fairness zwischen den Mobilstationen bezüglich ihres Durchsatzes maßgeblich erhöht werden kann, ohne dabei die Summenrate des Systems nennenswert zu reduzieren.

Es zeigt sich, dass die neuen suboptimalen Verfahren zur Funkressourcenzuweisung, die durch die Kombination der vorgeschlagenen Algorithmen entstehen, einen erheblichen Teil der maximal erreichbaren Summenrate des Systems erzielen, wobei ihr Rechenaufwand beachtlich niedriger ist als der eines optimalen Verfahrens. Die vorgestellten Verfahren zur Funkressourcenzuweisung ereichen über 90% der mittleren Summenrate, die mit einem Exhaustive Search Verfahren für die SDMA-Gruppierung, das die Summenrate maximiert, erzielt wird.

# Abstract

Future wireless communication systems are expected to reliably provide data services with rate requirements ranging from a few kbit/s up to some Mbits/s and, due to the high costs of frequency spectrum, these systems also need to be extremely efficient in terms of the spectrum usage. In particular, the application of transmission schemes based on Orthogonal Frequency Division Multiple Access (OFDMA) and on Multiple Input Multiple Output (MIMO) is considered as a promising solution to meet these requirements.

On the one hand, MIMO-OFDMA systems are flexible and spectrally efficient. On the other hand, the considerably large number of subcarriers and the inclusion of the space dimension make the Resource Allocation (RA) in such systems very complex. In fact, the optimum RA that maximizes the sum rate of the system is often too complex for practical application and, consequently, suboptimal rather efficient and low-complexity RA strategies are required in order to allocate the frequency, time, and space resources of the system to the Mobile Stations (MSs).

This thesis deals with suboptimal RA strategies in the downlink of MIMO-OFDMA systems aiming at the maximization of the sum rate. In order to solve the problem of maximizing the sum rate with affordable complexity, a new formulation for the problem is proposed which divides it into four subproblems, namely the Space Division Multiple Access (SDMA) grouping problem, the precoding problem, the power allocation problem, and the resource assignment problem. For each subproblem, several existing or newly proposed algorithms are applied, which are also oriented towards the maximization of the sum rate of the system. Through the combination of these algorithms, new suboptimal rather efficient RA strategies are obtained.

For the SDMA grouping problem, four new SDMA algorithms are proposed: one algorithm based on convex optimization and three greedy algorithms based on simple heuristics. The proposed algorithms build the SDMA groups based only on the spatial correlation and channel gain of the MSs and depend neither on precoding nor on power allocation. The proposed algorithms are shown to perform as good as some existing SDMA algorithms in terms of the achieved average sum rate and to have considerably lower computational complexity than the existing ones.

For the precoding problem, two existing algorithms are adopted. For the power allocation problem, a new iterative Soft Dropping Algorithm (SDA) is proposed, which is subsequently combined with Generalized Eigen-Precoding (GEP) into a new iterative joint precoding and power allocation algorithm. Moreover, the proof for the convergence of the joint precoding and power allocation algorithm is provided. In particular, the SDA and, consequently, the joint precoding and power allocation algorithm can pursue either the maximization of the sum rate or the provision of Quality of Service (QoS) to the MS by means of a simple parameter setting. Also as part of the precoding and power allocation algorithm, a new Sequential Removal Algorithm (SRA) is proposed, which might remove MSs from SDMA groups as to enhance the sum rate.

For the resource assignment problem, algorithms performing either separated or joint SDMA grouping and resource assignment are proposed and compared. For the maximization of the sum rate of the system, it is shown that a sequential assignment of resources to SDMA groups performs as good as assignment algorithms performing joint SDMA grouping and resource assignment while being considerably more simple. Moreover, different criteria to prioritize MSs or SDMA groups are considered by the assignment algorithms, and it is shown that by adopting a suitable priority criterion the throughput fairness among the MSs can be considerably improved at the expense of only small reductions of the average sum rate of the system.

The new suboptimal RA strategies that result from the combination of the proposed algorithms are shown to obtain a considerable fraction of the maximum achievable sum rate of the system with computational costs considerably lower than that of an optimum RA. Indeed, the proposed RA strategies achieve over 90% of the average sum rate obtained by an RA strategy performing an Exhaustive Search (ES) for the SDMA group that maximizes the sum rate.

# Contents

# Chapter 1

# Introduction

## 1.1 Resource Allocation in MIMO-OFDMA Systems

### 1.1.1 Key Technologies

Future wireless communication systems are expected to reliably provide data services with rate requirements ranging from a few kbit/s up to some Mbits/s and, due to the high costs of frequency spectrum, these systems also need to be extremely efficient in terms of the spectrum usage. In particular, the application of transmission schemes based on Orthogonal Frequency Division Multiple Access (OFDMA) and on Multiple Input Multiple Output (MIMO) is considered by most of the proposals for future wireless communication systems as a promising solution to meet these requirements [Tel95, FG98, NP00, LL05, DSK+06].

OFDMA employs Orthogonal Frequency Division Multiplexing (OFDM) as transmission technique, which divides a wideband channel into a number of orthogonal narrowband subchannels, termed subcarriers. Simple OFDM-based transceivers can be efficiently implemented using the Fast Fourier Transform (FFT) and, because the channel transfer function of the subcarriers is non-frequency-selective, sophisticated equalization structures are not needed [NP00, LL05, FKKC06]. Moreover, by allocating a variable number of subcarriers to a given radio link, the system can accommodate services with different rate requirements and provide the required rate flexibility [LL05].

MIMO communication employs multiple transmit and receive antennas to enhance the system performance by exploiting spatial diversity to improve communication reliability and/or spatial multiplexing to improve throughput [PNG03]. Over the same radio channel, spatial multiplexing can be used to simultaneously transmit multiple data streams separated in space, thus enabling to obtain huge system throughput gains [Tel95, FG98]. In this way, high spectral efficiency values can be achieved without requiring additional frequency resources [FDH05, MK06, DSK+06]. Since the data streams transmitted by a Base Station (BS) might be intended for different Mobile Stations (MSs) sharing the same radio resource in space, a spatial component in the multiple access scheme is obtained, namely Space Division Multiple Access (SDMA) [LR99, Van02, PNG03]. The multiple antenna systems employed at the BS and MSs are usually termed Antenna Arrays (AAs) and the individual antennas of the AAs are termed AA elements [LR99, Van02].

Further on, a system combining MIMO and OFDMA transmission schemes will be termed a MIMO-OFDMA system, which in fact combines Frequency Division Multiple

Access (FDMA), Time Division Multiple Access (TDMA), and Space Division Multiple Access (SDMA) in its multiple access scheme as to exploit their best properties.

### 1.1.2 Scenario Description

In this section, the scenario considered in this thesis is described. The Downlink (DL) of a Multi-User (MU) MIMO-OFDMA system considering a single sectorized cell of the cellular system is assumed. Fig. 1.1 illustrates the considered scenario, in which there is a BS located at the corner and the MSs are uniformly distributed within the cell.

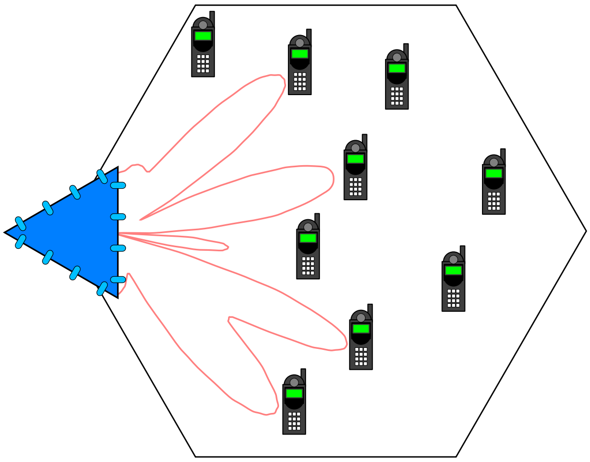


Figure 1.1. General scenario.

In Fig. 1.1, the BS is equipped with an AA and the MSs are equipped with a single antenna each. The BS employs its AA to adaptively form beams towards a group of MSs that share the same subcarrier in space and this subcarrier is said to be allocated to the MSs. In the DL, adaptive beamforming weights the signals transmitted by each AA element as to improve the power of the received signal at the MS and/or mitigate interfering signals [LR99, Van02, PNG03]. At the transmitter side, adaptive beamforming is also termed precoding and the weights applied to a transmitted signal are usually organized in a vector termed the precoding vector [PNG03, SSH04, MBQ04, JUN05, BHV06]. Thus, the BS sends signals to each of these MSs on the same subcarrier while separating the signals intended for the MSs in space through precoding.

### 1.1.3 Resource Allocation Problem

In this section, the Resource Allocation (RA) problem in frequency, time, and space is discussed in non-mathematical terms. On the one hand, MIMO-OFDMA systems are flexible and spectrally efficient. On the other hand, the considerably large number of subcarriers and the inclusion of the space dimension make the RA in such systems very complex. In general terms, the radio resource units in such a system are elements organized in a three-dimensional structure with a frequency, a time, and a space axis, as illustrated at the left side of Fig. 1.2. How units are defined in each axis direction depends on the system design. Frequency resource units can be represented by subcarriers, as in OFDMA [Rap99, LL05], time resource units by Time-Slots (TSs), as in TDMA [Rap99, HRM02], and space resource units by spatial layers/dimensions [DSK+06]. Each of the resource units illustrated in Fig. 1.2 may be assigned to a different MS and, therefore, a huge number of possible assignments exists even for a relatively small number of resource units and MSs.

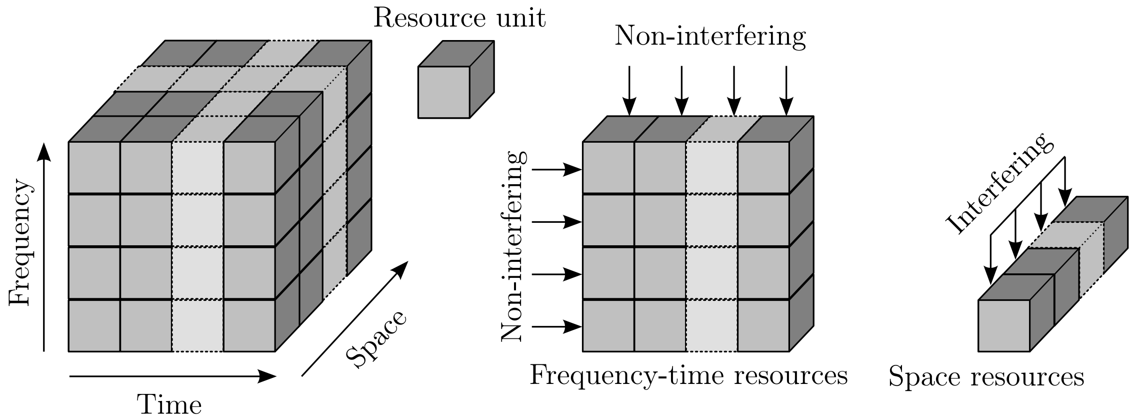



Figure 1.2. Radio resource units in time, frequency, and space.

Through adequate frequency and time synchronization, frequency and time resources can be effectively made orthogonal by design. This orthogonality simplifies the RA in frequency and time, since an MS allocated on a resource does not interfere with another MS allocated on a different resource. From the point of view of the RA, such resources can be considered as relatively independent of each other, since they are not coupled through interference, as illustrated in the middle of Fig. 1.2. This does not hold for space resources and obtaining such orthogonality is relatively more complicated in this case. In fact, space resources result from the reuse of a same resource in frequency and time by a group of MSs, so that MSs allocated on space resources associated with the same frequency-time resource will essentially interfere with each other. From the point of view of the RA, such resources are dependent on each other since they are coupled through interference, as illustrated at the right side of Fig. 1.2. Due to this characteristic, frequency-time resources can be interpreted as the real radio resources in the system, which are in fact shared in space by several MSs through SDMA. Further on, a group of MSs sharing the same frequency-time resource in space is termed an SDMA group.

The performance of the system depends on the following aspects:

- The composition of the SDMA groups to which resources are allocated.
- The precoding technique employed to spatially separate the signals intended for the different MSs of an SDMA group.

- The distribution of the available power among the resources and the MSs to which the resources are allocated.
- The selection of the frequency-time resource which is allocated to each SDMA group.

In the following, the dependency of the performance of the system on the above four aspects is shortly discussed. If the channels of the MSs are highly spatially uncorrelated, these MSs can be efficiently separated in space through precoding. However, if their channels are highly spatially correlated, these MSs cannot be efficiently separated in space through precoding. MSs that can be efficiently separated in space are said to be spatially compatible [Cal04, SS04a, FDH05, MK06]. In order to obtain SDMA gains and improve the sum rate of the system, SDMA groups have to be composed of spatially compatible MSs. Otherwise, the MSs in the SDMA group strongly interfere with each other and the sum rate of the system is compromised [SS04a, FDH05, MK06, MK07a]. Consequently, the composition of the SDMA groups affects the performance of the system.

The spatial separation through precoding of the signals intended for the MSs is simplified whenever the MSs are spatially compatible. There are different precoding techniques which might, e.g., suppress interference among MSs totally, in part, or ignore it [SSH04, MBQ04, JUN05]. Therefore, such precoding techniques present different performances and the selection of one of them influences the performance of the system.

The interference between MSs sharing a given resource through SDMA is a function of the power distribution among the MSs in the SDMA group, as well as, of the power distribution among resources. Considering a certain amount of power available for an SDMA group, allocating more power to one MS enhances the quality of its signal, e.g., in terms of Signal-to-Interference plus Noise Ratio (SINR), but also reduces the SINR of the other MSs in the SDMA group [ZKAQ01]. Analogously, allocating more power to a certain resource enhances the SINR of the MSs sharing this resource, but reduces the SINR of the MSs allocated on the other resources [ZKAQ01]. Consequently, an efficient distribution of the power among the MSs and resources has to be performed in order to improve the performance of the system, e.g., in terms of sum rate.

Finally, the spatial compatibility among MSs is resource-dependent, i.e., MSs that are spatially compatible on a given resource might be incompatible on another resource [MK06, LZ06, MK07a, MK07b]. Consequently, the selection of the resources which are allocated to the SDMA groups also influences the performance of the system.

In fact, these four aspects can be associated with the four following problems:

- The **SDMA grouping problem**, which corresponds to building groups of MSs as spatially compatible as possible on each resource.
- The **precoding problem**, which corresponds to determining precoding vectors able to spatially separate the signals intended for the MSs efficiently.

- The **power allocation problem**, which corresponds to allocating power to MSs and resources efficiently.
- The **resource assignment problem**, which corresponds to assigning resources to the best SDMA groups, e.g., the groups achieving the highest rate on each resource.

The SDMA grouping problem and the resource assignment problem are essentially integer problems, since only whole resources can be assigned to whole MSs. The precoding problem and the power allocation problem are not integer problems, as long as the precoding vectors and allocated powers are not discretized. Integer optimization problems are usually hard to solve due to its combinatorial nature [NW99].

From the previous discussion, it can be noted that the RA problem in MU MIMO-OFDMA systems has in fact the four problems above as subproblems. The previous discussion also reveals some interdependencies among the referred subproblems, such as the efficiency of the precoding technique which is conditioned on the SDMA group composition and the dependency between the performance of an SDMA group and the power distribution among resources. Due to such interdependencies, the RA in frequency, time, and space in MU MIMO-OFDMA systems becomes complicated. Indeed, an optimal RA strategy would have to jointly solve all the four subproblems and, consequently, becomes a complex combinatorial and non-convex optimization problem [LZ06]. The computational complexity of such an optimal RA strategy rapidly increases with the number of MSs, resources, and transmit antennas, and becomes unfeasible for most practical cases. Therefore, suboptimal rather efficient low-complexity RA strategies are required [LZ06].

In MU MIMO-OFDMA systems, different RA objectives can be pursued, such as the maximization of the sum rate of the system [FN96, STKL01, KT02, Cal04, TC04a, Lau04, ZL05, CC05, TUBN06, MSLT07], the maximization of minimum SINR in the system [Cal04, SBO06], or the minimization of the total transmit power under rate constraints [ZTC06]. In this work, the maximization of the sum rate of the system is considered as the RA objective.

## 1.2 Framework for Suboptimal Resource Allocation Strategies

In this section, a new framework for suboptimal RA strategies is introduced. Herein, the model illustrated in Fig. 1.3 is proposed as a framework for suboptimal RA strategies, which divides the RA problem into the four subproblems mentioned in Section 1.1.3. Then, existing or new algorithms oriented towards the maximization of the sum rate can be applied to each subproblem and combined to form suboptimal rather efficient RA strategies.

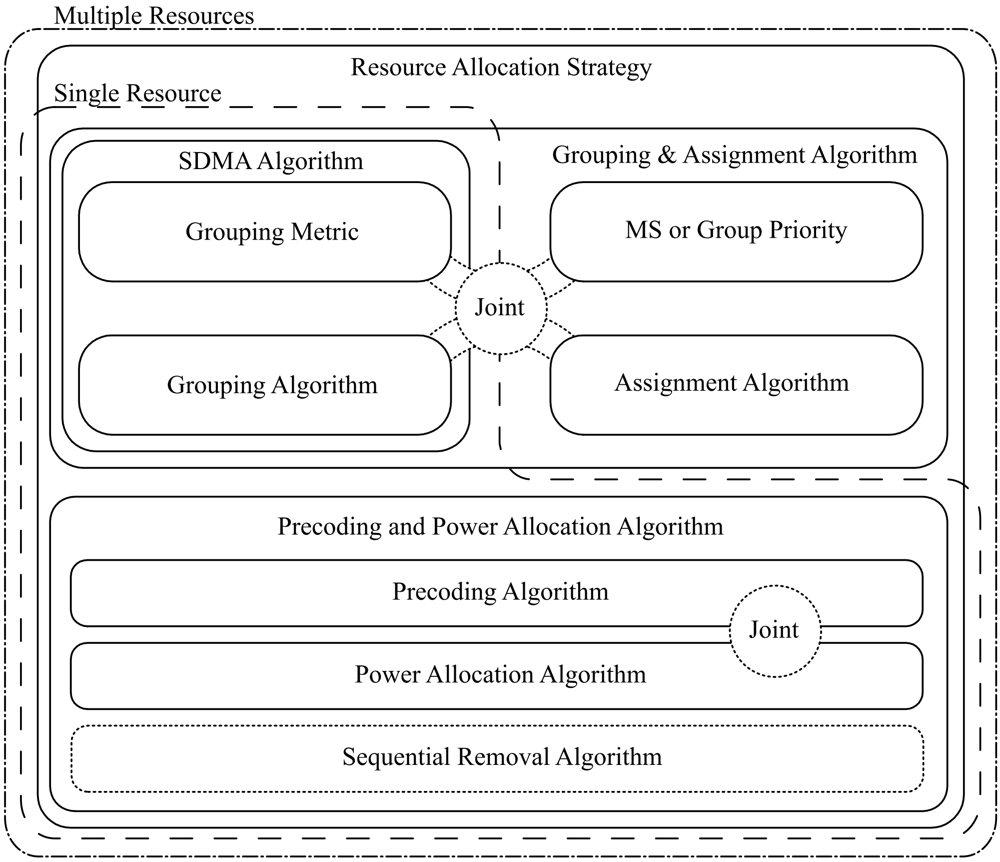



Figure 1.3. Framework for suboptimal RA strategies for MIMO-OFDMA systems.

In Fig. 1.3, two cases are considered regarding the number of resources taken into account by the RA strategy: a case considering a single resource, which is indicated by the dashed line, and another case considering multiple resources, which is indicated by the dot-dashed line.

The single-resource case does not mean that the system has only a single resource, but that resources are considered individually by the RA strategies. Since in this case, for the RA strategy there exist not multiple resources to be assigned, the resource assignment problem has no importance. In the multiple-resource case, the RA strategy takes all the multiple resources into account and the resource assignment problem has to be considered.

In the following, the algorithms employed by the RA strategies to solve each of the RA subproblems are discussed.

In order to solve the SDMA grouping problem, an SDMA algorithm is employed. However, the SDMA grouping problem is known to be a Non-deterministic Polynomial time Complete (NP-C) problem [GJ79, STKL01, Cal04]. NP-C problems are problems that cannot be solved in polynomial time or, in other words, whose complexity might exponentially increase with the problem dimensions [GJ79, NW99, Leu04]. Thus, in order to find the SDMA group that maximizes the sum rate of the system, an Exhaustive Search (ES) over all the possible SDMA groups is required [FDH05, MK06, MK07a].

Therefore, it is important to investigate how MSs can be efficiently organized into SDMA groups with low complexity, i.e., suboptimal rather efficient SDMA algorithms are required. Most suboptimal SDMA algorithms share a common structure which is composed of two main elements, namely:

1. A grouping metric, which measures the spatial compatibility among MSs in an SDMA group, i.e., which quantifies how efficiently the MSs in an SDMA group can be separated in space. Additionally, the grouping metric can also be used to compare different SDMA groups.
2. A grouping algorithm, which, based on the grouping metric, builds and compares SDMA groups composed of spatially compatible MSs without needing to perform an ES.

In this work, these two elements are considered to define an SDMA algorithm, as illustrated in Fig. 1.3.

In the multiple-resource case, it is required to solve the resource assignment problem and decide which resource to allocate to which SDMA group. For this purpose, the two following elements are usually considered, namely:

1. An MS or group priority, which measures the efficiency of allocating a given resource to an MS or an SDMA group, respectively.
2. An assignment algorithm, which based on the MS or group priority assigns the system resources to an MS or an SDMA group, respectively.

The SDMA algorithm, the MS or group priority, and the assignment algorithm are the elements composing the Grouping & Assignment (GA) algorithm, as illustrated in Fig. 1.3.

The GA algorithm involves both the SDMA grouping and the resource assignment problems. The resource assignment problem becomes almost irrelevant whenever a single resource is considered, while the SDMA grouping problem becomes irrelevant if SDMA groups containing only a single MS are considered. The division of the GA algorithm in Fig. 1.3 is similar to that considered by some RA strategies designed for Multi-User Single Input Single Output (SISO)-OFDMA systems, which split the GA algorithm into smaller steps, as in [BGWM07].

Because the GA algorithm involves both the SDMA grouping and the resource assignment problems, different designs are possible. The designs of the GA algorithm may differ, for example, in the assignment of resources to single MSs [KT02, TUBN06] or to whole SDMA groups [MK07b, MK08]. They may also differ by performing Separated Grouping and Assignment (SGA), in which first SDMA groups are built and afterwards some of them are assigned to the resources [MK07b], or Joint Grouping and Assignment (JGA), in which SDMA groups are simultaneously built on the multiple resources [KT02]. Considering JGA, the ideas of grouping metric and MS or group priority merge into each other and one can talk of a global priority which at the same time involves RA aspects in frequency, time, and space. The same happens with the grouping and assignment algorithms, which based on the new global priority performs in parallel SDMA grouping and resource assignment on the different system resources. The GA algorithm and its elements, as well as the possibility of performing JGA are also indicated in Fig. 1.3.

In order to solve the precoding and the power allocation problems, a precoding algorithm and a power allocation algorithm are employed, respectively. The precoding algorithm and the power allocation algorithm are illustrated in Fig. 1.3 as two separated blocks. However, precoding and power allocation do not need to be performed necessarily separated from each other, so that joint precoding and power allocation algorithms can also be considered, as indicated in Fig. 1.3. In this case, precoding vectors and power allocation are usually optimized alternately and iteratively [BO02, CTRF02, SB04, MK07c].

Despite the fact that some blocks in Fig. 1.3 appear isolated, the exchange of information among them is not excluded in this work. On the one hand, the separation between the GA algorithm and the precoding and power allocation algorithm helps to simplify the RA strategy. For example, whenever the SDMA grouping and resource assignment problems, i.e., the combinatorial subproblems of the RA problem, are solved by the GA algorithm, the precoding and power allocation are considerably simplified. On the other hand, if the GA algorithm, and especially the SDMA algorithm, is aware of the actual precoding and power allocation, the performance of the RA may be improved but at the expense of increased complexity. For example, if the SDMA algorithms employs a grouping metric that depends on the actual precoding and power allocation, it may be possible to accurately estimate the performance of an SDMA group, e.g., in terms of the group capacity [FDH05, MK06] or of the SINRs perceived by the MSs [STSK98, STKL01, KTSK01] and, consequently, avoid putting MSs into the group that, e.g., do not contribute to enhance the group capacity. However, the complexity of the SDMA algorithm increases because precoding vectors and power allocation have to be recomputed whenever the composition of SDMA group changes. The contrary occurs with grouping metrics that do not depend on the actual precoding and power allocation, which may have low complexity but may not be able to quantify the spatial compatibility as accurately as the former ones.

Therefore, if due to the RA strategy design the SDMA algorithm is made unaware of the precoding and power allocation, resources might be unadvisedly allocated to MSs that do not contribute to improve the sum rate of the system. In this case, some SDMA groups have more MSs than they should contain and the SDMA group size has to be adjusted by removing some MSs from the group. Consequently, the determination of the size of the SDMA group is also a problem that has to be taken into account by the RA strategy. In order to remove MSs and adjust the size of the SDMA group, a Sequential Removal Algorithm (SRA) is employed, as illustrated in

Fig. 1.3. The SRA selects the MSs to remove from the SDMA group based on a given removal criterion [Cal04, SBO06, MK06, MK07a, MK07c]. Because a more reliable decision about which MS to remove from the group can be made considering the actual precoding and power allocation, the SRA is introduced as part of the precoding and power allocation algorithm, as shown in Fig. 1.3. Anyway, an SRA is only needed if the SDMA algorithm is unaware of the precoding and power allocation. The precoding algorithm, the power allocation algorithm, and the SRA are the elements composing the precoding and power allocation algorithm, as shown in Fig. 1.3.

## 1.3 State-of-the-Art

In this section, a literature review on suboptimal RA strategies is provided. A joint solution of the four subproblems described in Section 1.1.3 is too complex for application in a realistic scenario, so that RA strategies structured as shown in Fig. 1.3 are usually employed to solve the RA problem. However, there are some algorithms that could be used to maximize the sum rate of the system by jointly solving the four subproblems described in Section 1.1.3, e.g. in [CS03, JRV+05].

Differently from Single-User (SU) SISO-OFDMA and SU MIMO-OFDMA systems, the direct maximization of the sum rate in the DL of MU MIMO-OFDMA systems is a hard optimization problem [LL05, JRV+05, AHSB+06, LZ06, CC05, CC07]. In the SU cases, the sum rate of the system is maximized by transmitting to the MSs with highest channel gain on each resource, employing Singular Value Decomposition (SVD) precoding [Tel95, FG98, PNG03], and the Water Filling Algorithm (WFA) over all the resources [PNG03, PF05]. However, for MU MIMO-OFDMA systems, such a simple policy for the RA does not exist [LL05, JRV+05, AHSB+06, LZ06, CC05, CC07]. In [CS03], it has been shown that the maximum sum rate of the system, which reaches Sato's bound [Sat78], can be achieved through the application of non-linear Dirty Paper Coding (DPC) techniques [Cos83], such as Tomlinson-Harashima precoding [Joh04, SH05] or vector-perturbation [PHS05, HPS05]. However, maximizing the sum rate of a MIMO-OFDMA system using DPC renders a computationally complex non-convex optimization problem [JRV+05]. In [JRV+05], the duality between the MIMO Multiple Access Channel (MAC) and the MIMO Broadcast Channel (BC) is exploited to derive an iterative algorithm that achieves the maximum sum rate of the BC channel by solving a convex dual MAC problem. Although simpler than the DPC-based solutions, the algorithm in [JRV+05] is still relatively complex. It is an iterative approach that relies on the alternate optimization of the instantaneous covariance matrix and of the power of the transmit signals. The algorithm requires to perform Eigenvalue Decompositions (EVDs) of the Gram dual MAC effective channel matrices of all the MSs on each resource, as well as to apply the WFA on the eigenvalues obtained at each iteration [JRV+05, PF05, TUBN06]. Due to its potentially unaffordable computational complexity, the algorithms in [CS03, JRV+05] may not be suitable for application in a more realistic scenario and only draw a theoretical upper bound for the achievable sum rate of the system.

In a more realistic situation, a fixed precoding algorithm or a fixed set of precoding algorithms is employed, e.g., in order to fulfill some real system constraint such as to limit by design the complexity of the system. In these cases, the optimal solution of [JRV+05] does not apply anymore. For a fixed precoding algorithm, the problem solution does not necessarily become more simple. In fact, in this case the selection of the MSs to share a given resource in space, i.e., on different spatial layers of the same resource, determines the achievable rate on the resource. This problem corresponds to the SDMA grouping problem, which is NP-C [GJ79, STKL01, Cal04], cf. Section 1.2, and an ES over all the possible groups on each resource would be required in order to find the optimum RA. Therefore, suboptimal RA strategies able to approximate the maximum sum rate of the system without incurring prohibitively high computational efforts remain an important research topic nowadays [LZ06].

In the following, RA strategies fitting into the framework illustrated by Fig. 1.3 are discussed. In order to overcome complexity problems, several works concentrated on suboptimal RA strategies and divided the RA problem into smaller subproblems. Other works concentrated directly only on one or on a subset of the four subproblems of Section 1.1.3, e.g., by considering a single resource. The framework illustrated in Fig. 1.3 encompasses a large number of suboptimal RA strategies available in the literature. Table 1.1 gives a set of RA strategies and its classification into the framework of Section 1.2. The columns of Table 1.1 are associated with the elements of the proposed framework and appear in the same order in which the elements have been discussed in Section 1.2. Because most of the works in Table 1.1 concentrated on the single-resource case and, consequently, did not regard the resource assignment problem, a column for the assignment algorithm is not included. A few others works considered the presence of multiple resources and also proposed an assignment algorithm and will be addressed later in this section. Moreover, a column indicating whether the GA algorithm is aware of the actual precoding and power allocation is also included in Table 1.1.

It can be seen in Table 1.1, that SDMA algorithms considering a large variety of grouping metrics and grouping algorithms have already been proposed. They also have been combined with a considerable number of precoding and power allocation algorithms. One can note that the Best Fit Algorithm (BFA) [STSK98, STKL01, KTSK01] (or algorithms like the BFA) has been often employed as grouping algorithm. The BFA is part of several SDMA algorithms studied in this thesis and it will be described in more details in Section 3.2.2

In the following, the previous works listed in Table 1.1 are shortly discussed in the order in which they are referenced in the last column of Table 1.1.

In [FN96, STSK98, STKL01, KTSK01, SBO06], SDMA algorithms incorporating Quality of Service (QoS) constraints, e.g., in terms of a target SINR, have been investigated.

Table 1.1. RA strategies – State-of-the-art – Single resource.

| **Grouping & Assignment Algorithm** | | | **Precoding and Power Allocation** | | | **Ref.** |
|---|---|---|---|---|---|---|
| **SDMA algorithm** | | **Aware** | **Precoding algorithm** | **Power allocation algorithm** | **SRA criterion** | |
| **Grouping metric** | **Grouping algorithm** | | | | | |
| Max. subspace separation | Best Fit Algorithm (BFA) | No | GEP | EPA | Fixed group size (max.) | [FN96] |
| Forward, backward, total interference | | | | | | |
| Max. orthogonality factor | | No/Yes | GEP | EPA | | [STSK98], [STKL01], [KTSK01] |
| Min. channel gain ratio | | | | | | |
| SINR margin | | | | | | |
| MaxMin SINR | | Yes | Semidefinite Programming | | Random, HPF | [SBO06] |
| Assignment preference factor | | Yes | GEP | Adaptive bit loading | Fixed group size (max.) | [KT02] |
| Sum of channel gains with SPs | | No | ZF with DPC | WFA | | [CS03], [TUBN06] |
| Min. normalized scalar product | | No | ZF | WFA, EPA, MaxMin SINR | | [Cal04] |
| Sum of channel gains with SPs | | No | ZF | WFA | | [YG05] |
| Sum of channel gains with SPs | | No | BD | WFA | | [TJ05] |
| Group capacity | | Yes | ZF | WFA | | [DS05] |
| Projection-based group capacity | | No/Yes | BD | WFA, EPA | | [FDH07] |
| SINR margin | First Fit Algorithm (FFA) | Yes | GEP | EPA | Fixed group size (max.) | [STSK98], [STKL01], [KTSK01] |
| MaxMin SINR | | Yes | Semidefinite Programming | | Random, HPF | [SBO06] |
| Weighted sum of channel gains | Random Grp. Alg. (RGA) | No | UL SIC | EPA | Fixed group size (max.) | [Lau02] |
| Max. sum of eigenvalues | | | BD | WFA | | [TJ05] |
| Frobenius norm-based total spatial correlation | COA | No | - | - | | [SS04a] |
| Normalized scalar product | MS partitioning | No | ZF | Adaptive bit loading | | [ZL05] |
| Projection-based group capacity | Tree-based grouping | No | BD | EPA, WFA | - | [FDH05], [FDH07] |
| Average group SNR | | | | | | |
| Max. sum rate | Three-step algorithm | Yes | ZF | WFA | | [TC04a], [TC04b] |
| Max. sum rate | Greedy and convex optimization | Yes | ZF | WFA | | [CC05], [CC07] |
| Max. (weighted) sum rate | Exhaustive Search | Yes | ZF | (Weighted) WFA | | [Lau04], [Lau05] |
| | Genetic Alg. | | | | | |
| MaxMin SINR | Simulated Annealing | No | ZF | MaxMin SINR | Fixed group size (max.) | [CPIPN03] |
| Max. sum of scores | MS insertion | No | GEP | EPA | | [Wil06] |
| Min. power under rate constraints | Integer optimization | Yes | ZF | Integer optimization | | [ZTC06] |
| MaxMin SINR | Admit all | Yes | Semidefinite Programming | | Random, HPF | [SBO06] |
| Max. sum rate | Convex optimization | Yes | Semidefinite Programming | | Max. SINR gap | [MSLT07] |

In [FN96], a measure of the distance between the subspaces spanned by the channels of the MSs and different measures of the interference generated by introducing new MSs in an SDMA group have been used as grouping metric and combined with the BFA. Therein, the total interference generated by the admission of each candidate MS has provided the best results as grouping metric.

The BFA has been introduced in [STSK98, STKL01, KTSK01], where it has been combined with different grouping metrics. Therein, the admission of a new MS to the SDMA group is subject to an SINR margin constraint, so that an MS is only accepted in the group if the SINR of all MSs in the group becomes not lower than a given target

SINR. Indeed, in [STSK98, STKL01, KTSK01] different SDMA algorithms have been proposed, which involved different grouping metrics and grouping algorithms. For example, the maximum orthogonality factor, which is a function of the cosine of the angle between the vector channels of any two MSs, has been used as grouping metric, so that MSs whose spatial channels are close to orthogonal are put in the same group, similarly to [FN96] when considering the subspace distance metric. A minimum channel gain ratio has also been considered, which puts in the same group MSs with similar channel gains. In both cases, however, the admission of the MSs has been constrained by the SINR margin. In [STSK98, STKL01, KTSK01], the First Fit Algorithm (FFA) has also been proposed as grouping algorithm. The FFA is also considered in this thesis and described in more details in Section 3.2.2. Finally, the SINR margin has been directly used with the BFA in [STSK98, STKL01, KTSK01] in order to build the SDMA groups, so that the MS leading to the maximum minimum margin between the perceived and target SINRs of the MSs is admitted to the group. This SDMA algorithm provided the best results, but at the expense of increased complexity. In [FN96, STSK98, STKL01, KTSK01], a Generalized Eigen-Precoding (GEP) [Zet95, Zet99] and Equal Power Allocation (EPA) are used as precoding and power allocation algorithms, respectively.

In [SBO06], SDMA algorithms employing different grouping algorithms are considered. Therein, the RA strategy aims at maximizing the minimum SINR among the MSs in the SDMA group while respecting fixed minimum target SINR values for the MSs. Differently from [FN96, STSK98, STKL01, KTSK01], a joint optimization of precoding and power allocation based on Semidefinite Programming (SDP) relaxations is employed in [SBO06], which performs a weighted balancing of the SINRs perceived by the MSs in the group. Weighted SINR balancing with fixed target SINRs is usually the objective considered by most joint precoding and power allocation algorithms, as e.g. [SB04, SBO06]. In [SBO06], RA strategies using a random and a Highest Power First (HPF) removal criterion for the SRA are investigated in combination with the BFA and FFA. Moreover, Eigen-Precoding (EP) [PNG03] and GEP [Zet95, Zet99] are also considered as precoding and power allocation algorithm besides the optimal joint solution based on SDP. Again, it is shown in [SBO06] that the RA strategy employing the BFA in the SDMA algorithm and considering the optimal joint precoding and power allocation algorithm based on SDP provides the best results, but at the expense of increased complexity.

In [KT02, CS03, DS04, Cal04, YG05, TJ05, DS05, TUBN05, B$^+$06, TUBN06, Wil06, FDH07], SDMA algorithms employing the BFA, different grouping metrics, and aiming at the maximization of the sum rate have been proposed.

In [KT02], an assignment preference factor has been proposed as grouping metric. The assignment preference factor captures the interference and the rate increment induced by the admission of an MS into an SDMA group. The less interference and the higher the rate increment an MS generates, the most suited for admission in the SDMA group this MS is.

In [CS03], it has been shown that the maximum sum rate of the system can be achieved by using DPC. Therein, it is proposed to admit the MSs to the SDMA group in order of channel gain after a projection onto the null space of the channels of the already admitted MSs, so that previously admitted MSs do not see any interference from MSs posteriorly admitted to the group. After building the group, the interference from previously admitted MSs to the posteriorly admitted MSs is suppressed using DPC, and after that the WFA is used for power allocation. Thus, the RA strategy in [CS03] has the sum of the channel gains after null space Successive Projections (SPs) as grouping metric, the BFA as grouping algorithm, Zero-Forcing (ZF) and DPC as precoding algorithm, and the WFA as power allocation algorithm. A more comprehensive description of the RA strategy in [CS03] can be found in [DS05]. RA strategies very similar to the one in [CS03] can be found in [YG05, TJ05, TUBN05, B+06, TUBN06].

In [YG05], it is shown that for a relatively large number of MSs in the system, the maximum sum rate of the system can be achieved using linear ZF precoding instead of non-linear DPC techniques, as in [CS03]. In [TJ05, TUBN05, B+06, TUBN06], multiple antennas are considered at the receivers. The algorithms in [TJ05, TUBN05, B+06, TUBN06] take into account the receive beamformers, i.e., the vectors containing the beamforming weights used at receive AAs, at the transmitter side when performing the SDMA grouping. Similarly to [YG05], linear precoding is considered in [TJ05], where Block Diagonalization (BD) with transmit-receive cooperation is applied [SSH04]. In [TUBN05, B+06, TUBN06], adaptive beamforming based on an SVD is performed at the receiver side and is taken into account at the transmitter side by the null space SPs, after which DPC techniques are applied, as in [CS03]. In [DS04, DS05], it is shown that using the group capacity as grouping metric and the BFA as grouping algorithm, the achieved sum rate of the system is only 13% lower than the value achieved by the DPC-based proposal of [CS03]. Similar results have been independently achieved in [MK06] and in [FDH07]. However, in [FDH07] the complexity of the SDMA algorithm is reduced by employing a projection-based group capacity, which has lower complexity than the conventional group capacity considered in [DS05, MK06]. The RA strategy in [FDH07] considers BD and EPA as precoding and power allocation algorithms, respectively. In order to further reduce complexity, the groups built on previous runs of the SDMA algorithm in [FDH07] are just updated by removing one MS, adding a new MS, or potentially replacing one MS while preserving the current group size. For this update step, the group capacity considering BD and the WFA is adopted as grouping metric, since it has to be computed for at most three groups.

The SDMA algorithm in [Cal04] uses the sum of the maximum normalized scalar product between the vector channel of the candidate MS and the spatial channels of the MSs in the SDMA group as grouping metric. The BFA is used as grouping algorithm and the MS with the minimum metric value is admitted to the group. Therein, ZF precoding combined with the WFA and with an SINR balancing are applied as precoding and power allocation algorithm. In [Cal04], different SRA criteria have also be considered in order to implement spatial bit loading. Additionally, in [CPIPN03] simulated annealing and MS partitioning have also been considered as grouping algorithms.

The normalized scalar product, or a simple function of it, has also been considered as grouping metric in [OOI97, STSK98, STKL01, KTSK01, DH04, Wil06], as well as by the author in [MK06, MK07a, MK07b, MK07c]. Differently from [Cal04], the maximum normalized scalar product is compared in [Wil06] to an orthogonality threshold as part of an MS insertion algorithm that first builds a set of candidate MSs for admission to the SDMA group and then admits to the group the MS from the candidate set having the highest traffic priority. The main advantages of grouping metrics based simply on the maximum normalized scalar product are their ability to capture the spatial correlation among the channels of the MSs and their low complexity compared to metrics like the group capacity used e.g. in [DS05], the projection-based metrics used e.g. in [CS03, YG05, FDH05, TJ05, TUBN05, $B^{+}06$, TUBN06, FDH07], or the rate and SINR-based metrics used in [KT02] and [STSK98, STKL01, KTSK01], respectively.

Many proposals to solve the RA problem with reduced complexity can be found in the literature. Therein, optimal and suboptimal solutions with high and low complexity, respectively, are usually proposed and compared. For example, in [STSK98, STKL01, KTSK01] the BFA (with high complexity) and the FFA (with low complexity) are proposed and investigated. Similarly, in [SBO06] RA strategies involving the BFA, the FFA, and a third grouping algorithm, as well as different precoding and power allocation algorithms, are compared. In [DS05], a RA strategy involving linear ZF precoding (with low complexity) is compared to the case using DPC (with high complexity). Similarly, in [YG05] it is shown that linear ZF precoding (with low complexity) may achieve the same performance of DPC (with high complexity).

In [SS04a], a scaled Frobenius-norm [GL96, Mey01] is used to measure the spatial correlation between the MIMO channels of each pair of MSs. The metric values are organized in a matrix, which is then given to a Compatibility Optimization Algorithm (COA) that divides the MSs in as many groups as resources while minimizing the total spatial correlation among the MSs in each group. In [FDH05, FDH07], a tree structure is used to limit the number of candidate SDMA groups to a low value compared to an Exhaustive Search Algorithm (ESA). Moreover, a projection-based group capacity is used to reduce the complexity of the SDMA algorithm. In [ZL05], the maximum normalized scalar product is compared to a threshold as to divide the MSs into sets of highly spatially correlated MSs and an SDMA group is built by picking one MS from each set.

In [Lau02, TJ05], examples considering the Random Grouping Algorithm (RGA) as grouping algorithm can be found, which of course have a complexity considerably lower than the BFA, FFA, and than some other grouping algorithms. However, because they neglect the spatial compatibility among the MSs, their performance, e.g., in terms of average sum rate, is hardly compromised.

More complex RA strategies to maximize the sum rate of the system can be found, e.g., in [TC04a, TC04b, CC05, ZTC06, CC07, MSLT07]. In [TC04a, TC04b], a three-step algorithm is proposed which first reduces the search space by determining the optimal SDMA groups for each acceptable SDMA group size on each resource. Then, it employs convex relaxations on an iterative step to find an optimal assignment of groups of this reduced search space to the resources. In [TC04a, TC04b], ZF precoding and the WFA are considered for precoding and power allocation, respectively. Because the non-relaxed problem is non-convex and the obtained solution might correspond to a local optimum, a perturbation step is proposed in [TC04a, TC04b] to try to escape from local optima and achieve the global optimum. Approaches very similar to this strategy have been presented later in [CC05, CC07].

In [Lau04, Lau05], an ESA is used as grouping algorithm in order to find the SDMA group that maximizes a given utility function. Because the search space inspected by the ESA may become huge when the number of MSs increases, an alternative grouping algorithm based on a genetic algorithm is also considered to reduce the complexity of the RA strategy.

In [ZTC06], the SDMA grouping and resource assignment problems are formulated as an integer optimization problem, which is solved using integer optimization tools [NW99]. Therein, the minimization of the total transmit power subject to achieving certain rates is considered as optimization objective and ZF precoding is employed to spatially separate the signals intended for the MSs.

In [MSLT07], the maximization of the sum rate subject to target SINR constraints is formulated as an SDP problem, which is solved using convex optimization. Because predicting the feasibility of the problem is as hard as solving the problem itself, the problem is solved in [MSLT07] by considering all MSs in the SDMA group and sequentially removing the MS whose gap between target and attained SINR is maximum until obtaining a feasible solution, i.e., an SRA is employed together with the SDMA algorithm.

Many of the proposals listed in Table 1.1 consider only a single resource and focused on the SDMA grouping problem and a few precoding and power allocation algorithms. However, some of them also consider multiple resources and embed the resource assignment problem in the proposed RA strategies. These works are listed in Table 1.2, in which it is indicated if the GA algorithm solves the SDMA grouping problem and the resource assignment problem jointly or separately. Moreover, some remarks about the RA strategies are also provided in Table 1.2.

In the following, the previous works listed in Table 1.2 are shortly discussed in the order in which they are referenced in the last column of Table 1.2.

It can be seen in Table 1.2 that some RA strategies simplify the RA problem by separately solving the SDMA grouping and the resource assignment problem, as in [STSK98, STKL01, KTSK01, TUBN05, B$^+$06, TUBN06]. While in [STSK98, STKL01, KTSK01] an SDMA-TDMA system disposing of the same amount of power for each

Table 1.2. RA strategies – State-of-the-art – Multiple resources.

| SDMA grouping & resource assignment | Multiple Access | Remarks | Ref. |
|---|---|---|---|
| Separated Grouping and Assignment (SGA) | SDMA-TDMA | SDMA, precoding, and power allocation algorithms are applied on a resource-by-resource basis. | [STSK98], [STKL01], [KTSK01] |
| | SDMA-FDMA | SDMA algorithm and precoding algorithm are applied on a resource-by-resource basis. Power allocation is applied over the resources altogether. | [TUBN05], [B$^+$06], [TUBN06] |
| Joint Grouping and Assignment (JGA) | - | MSs admitted according to the BFA applied over the multiple resources. Precoding and power allocation applied on a resource-by-resource basis. | [FN96] |
| | - | MSs admitted according to the BFA applied over the multiple resources using the assignment preference factor. Power allocation involves an adaptive bit loading algorithm. | [KT02] |
| | SDMA-FDMA | Single MSs are allocated to the subcarriers and the groups are extended by building SDMA groups on each resource. ZF precoding and adaptive bit and power loading are applied. | [Cal04] |
| | - | SDMA groups simultaneously built on the different resources by partitioning the MSs into groups using a Compatibility Optimization Algorithm (COA). | [SS04a] |
| | SDMA-FDMA | Highly spatially correlated MSs partitioned into sets allocated on orthogonal bands. MSs from different sets share resources through SDMA. Single-User adaptive bit and power loading are applied. | [ZL05] |
| | SDMA-FDMA | The SDMA grouping, resource assignment and power allocation problems are jointly solved as a convex optimization problem. ZF precoding is employed. | [TC04a], [TC04b], [CC05], [CC07] |
| | SDMA-FDMA | SDMA groups, resource assignment and power allocation are solved altogether as an integer optimization problem. | [ZTC06] |

TS is assumed, in [TUBN05, B$^+$06, TUBN06] a grouping metric based on null space SPs and totally independent of the precoding and power allocation is adopted in an SDMA-FDMA system. Both designs lead to a predefined power distribution among resources for the SDMA and assignment algorithms, thus making the SDMA grouping independent from resource to resource and simplifying the RA problem. In particular in [TUBN05, B$^+$06, TUBN06], after solving the SDMA grouping, the resource assignment, and the precoding problems, a power allocation based on the WFA is applied over the resources altogether, i.e., in space and frequency dimensions.

Some other strategies jointly solve the SDMA grouping and resource assignment problems, as in [FN96, KT02, Cal04, SS04a, ZL05], and sometimes the power allocation problem too, as in [TC04a, TC04b, CC05, ZTC06, CC07]. In this case, one decides not only the SDMA group in which an MS can be adequately accommodated, but also which resource is assigned to the considered MS.

Additional complexity results whenever the SDMA grouping and the resource assignment problems are jointly solved. This additional complexity may be compensated by the use of simpler SDMA and assignment algorithms. In [FN96], an initial MS is considered on each resource and the additional MSs are allocated into the SDMA group of a given resource based on a BFA. In [Cal04], the initial MS on each resource is assigned based on its channel gain and the SDMA groups are extended afterwards based on an adaptive bit and power loading algorithm. In [SS04a], the MSs are partitioned in SDMA groups to be allocated on different TSs based on a relatively simple grouping metric, which is a function of the spatial correlation, and using a relatively simple COA. In [ZL05], MSs are partitioned into sets containing only highly spatially correlated MSs, orthogonal resources are allocated to each set, and MSs from different sets are selected to share the resources through SDMA. In [ZL05], after the SDMA grouping and the resource assignment problems are solved, SU bit and power loading are applied to the scheduled MSs.

In [TC04a, TC04b, CC05, ZTC06, CC07], either integer or convex optimization tools are used to jointly solve the SDMA grouping, resource assignment, and power allocation problems while considering ZF precoding. In [ZTC06], integer optimization methods are directly applied while in [TC04a, TC04b, CC05, CC07], an integer relaxation is first applied to convert the intrinsically integer problem into a convex problem, whose solution might be either rounded at the end or considered as a resource sharing over a long-term time perspective. In spite of achieving a high fraction of the maximum sum rate of the system, such solutions incur a considerably high computational effort.

## 1.4 Open Problems

In this section, some open problems and important aspects for investigation are discussed.

From the review of the RA strategies in Table 1.1 and Table 1.2, it can be noted that RA in MIMO-OFDMA systems has become a very active research field in the last few years and a considerable number of investigations has already been conducted. However, there are still open problems which could be characterized as follows:

1. There is a large number of suboptimal RA strategies, which follow the most varied approaches to solve the RA problem. This complicates their comparison and a model is required in order to identify the main elements of the RA strategies and to help classifying and comparing them.
2. In order to reduce complexity, many suboptimal SDMA algorithms simplify their grouping metric and grouping algorithm and might present different performance-complexity trade-offs. In this context, two additional problems can be characterized as follows:

(a) The grouping metric has a large impact on the performance-complexity trade-off of the SDMA algorithms, especially if it has to be computed for a large number of candidate groups and if it depends on the actual precoding and power allocation, cf. Section 1.2 and Section 1.3. Thus, an important issue in this context is to determine whether grouping metrics that do not depend on the actual precoding and power allocation can perform as good as metrics that depend on the actual precoding and power allocation, i.e., if there are low-complexity rather efficient grouping metrics.

(b) The grouping algorithm has also influence on the complexity of the SDMA algorithms. Indeed, there are grouping algorithms based on simple heuristics, such as the BFA, and based on complex methods, such as ES or SDP, cf. Section 1.3. Analogously to the grouping metric, it is also important to determine whether simple grouping algorithms can perform as good as the complex ones.

These are important issues since they may provide important insight into the performance-complexity trade-off of the different RA strategies.

3. Because suboptimal SDMA algorithms are usually employed by the RA strategies, cf. Section 1.2 and Section 1.3, an important issue is to determine whether such strategies can attain a considerably high fraction of the maximum achievable sum rate of the system.

4. Most of the analyses performed in the works listed in Table 1.1 and Table 1.2 consider perfect Channel State Information (CSI) at the transmitter side, which is required in order to perform RA. Because the quality and availability of CSI might be constrained by several practical aspects, such as limited signaling bandwidth or imperfect channel estimation, it is important to study how robust (or sensitive) the RA strategies are to the available CSI and how their performance degrades with imperfect CSI. It is also important to investigate whether more complex SDMA algorithms offer any particular advantage over the simpler ones in this case.

5. Most joint precoding and power allocation algorithms consider fixed target SINRs values and perform a weighted SINR balancing, cf. Section 1.3. For the maximization of the sum rate, fixed target SINRs values are hard to determine and joint precoding and power allocation algorithm with target SINRs varying within a certain range of values might be more adequate.

6. Precoding and power allocation can be performed either separately or jointly. One importante issue is to compare the performance and complexity of joint and separate precoding and power allocation algorithms in RA strategies aiming at the maximization of the sum rate of the system.

7. Many of the RA strategies in Table 1.1 consider a fixed SDMA group size. However, the selection of the size of the groups built by SDMA algorithms unaware of the precoding and power allocation may affect the performance of the RA strategies, cf. Section 1.2, and is also a relevant aspect for investigation.

8. Because the complexity of Joint Grouping and Assignment (JGA) approaches is usually higher than that of Separated Grouping and Assignment (SGA), cf. Section 1.3, it is important to investigate whether there are considerable advantages of

JGA over SGA.

9. The maximization of the sum rate of the system is an important problem in MU MIMO-OFDMA systems. However, QoS and fairness aspects are also relevant and have been taken into account by some RA strategies discussed in Section 1.3. In many RA strategies, parameters introduced to simplify the RA problem may be used to provide throughput fairness among the MSs. For RA strategies aiming at the maximization of the sum rate of the system, a relevant aspect for investigation is to determine whether considerable improvements on the throughput fairness among the MSs can be obtained at the expense of only small reductions of the sum rate.

## 1.5 Contents and Contributions of the Thesis

This thesis consists of five chapters, whose contents and contributions are briefly described in this section.

In this chapter, a framework for suboptimal RA strategies intended for MU MIMO-OFDMA systems has been proposed in order to provide some insight into the overall problem of maximizing the sum rate of the system, the problem complexity, the component subproblems and their interdependencies. This framework is depicted in Fig. 1.3 and helps to address problem 1. Indeed, with help of this framework several suboptimal RA strategies have been put together in Table 1.1 and Table 1.2 and have been discussed in Section 1.3.

Chapter 2 introduces the models adopted in the investigations conducted in this thesis. It describes the model for the DL of the MU MIMO-OFDMA system considered in this thesis, which assumes a single cell and multiple MSs being multiplexed in frequency using block OFDMA, in time using TDMA, and in space using SDMA. The modeling of the DL wireless channel is also provided in Chapter 2, as well as the models for the CSI required in order to perform RA. In Chapter 2, the standard mathematical formulation of the problem of maximizing the sum rate of the system is provided, as well as a new alternative formulation based on mixed integer optimization. This new formulation allows to further discuss some of the particular characteristics of the framework proposed in Section 1.2.

Chapter 3 concentrates on the RA problem considering a single resource, cf. Section 1.2, and which is indicated by the dashed line in Fig. 1.3. In Chapter 3, RA strategies employing several SDMA algorithms are studied. These SDMA algorithms consider different grouping metrics and grouping algorithms. In particular, two new grouping metrics are proposed, which do not depend on the actual precoding and power allocation but only on the spatial correlation and channel gain of the MSs and which have low complexity. Thus, these metrics offer a solution to problem 2a. A new grouping algorithm is also proposed, namely the Convex Grouping Algorithm (CGA), which is based on convex optimization and also considers only the spatial correlation and channel gain of the MSs. This grouping algorithm is relatively general and can be combined with different grouping metrics, as it will be discussed in Chapter 3. The CGA offers a suitable solution to problem 2b.

The new grouping metrics are combined with existing grouping algorithms, namely the Best Fit Algorithm (BFA) and First Fit Algorithm (FFA) [STSK98, STKL01, KTSK01]. Additionally, one of the proposed metrics is also combined with the CGA. These combinations define new SDMA algorithms, which are compared to SDMA algorithms employing more complex grouping metrics and grouping algorithms. Such comparisons address problem 2 of Section 2.3.

The performance of RA strategies employing the new algorithms is evaluated through simulations and compared in terms of the average sum rate to the performance of RA strategies employing existing algorithms. These investigations will determine the efficiency of the proposed algorithms and are related to problem 3. The performance of the RA strategies in terms of the average sum rate of the system considering both perfect and imperfect CSI is also investigated in Chapter 3 in order to determine whether any of the RA strategies is particularly more efficient or more robust than the others, cf. problem 4 in Section 2.3. In particular, it is interesting to see whether RA strategies employing simple SDMA algorithms are as robust as RA strategies using more complex SDMA algorithms when imperfect CSI is considered.

Regarding the precoding and power allocation, two existing precoding algorithms are considered, namely the linear Zero-Forcing (ZF) precoding and the Generalized Eigen-Precoding (GEP). These precoding algorithms are combined with the Water Filling Algorithm (WFA) and a new Soft Dropping Algorithm (SDA), respectively, which perform power allocation. In particular, an iterative joint precoding and power allocation algorithm combining the GEP and the SDA is proposed, which dynamically adapts the target SINR of the MSs and can pursue either a maximization of the sum rate or a weighted SINR balancing among the MSs, thus offering a solution to problem 5. RA strategies considering both separated and joint precoding and power allocation are considered in Chapter 3 and their performances in terms of average sum rate are compared considering both perfect and imperfect CSI at the transmitter. These investigations address problem 6 of Section 2.3.

Regarding the potentially required removal of MSs from the SDMA groups in order to obtain a suitable group size, a new SRA algorithm is proposed, which considers two different criteria to determine which MS to remove from the SDMA group. The SRA offers a solution to problem 7 and its impact on the performance of the RA strategies is also investigated in Chapter 3.

The overall complexity of the RA strategies is also assessed in Chapter 3, in which formulas for the complexity of several RA strategies are provided, as well as formulas for their rough complexity order. Thus, from the results presented in Chapter 3, it is possible to compare the performance-complexity trade-off of the different RA strategies, cf. problem 2 and problem 3 in Section 2.3.

Throughput fairness aspects are also studied as special cases in Chapter 3, where particular parameters of the SDMA algorithms are used by the time-domain scheduler in order to prioritize some MSs in the system. At this point, it is interesting to observe whether a considerably higher improvement on the throughput fairness among MSs can be achieved at the expense of only a small reduction of the average sum rate of the system, cf. problem 9.

Chapter 4 studies the performance of a selected subset of the RA strategies of Chapter 3 in a scenario considering the existence of multiple resources. For the scenario considered in Chapter 4, the resource assignment problem recovers its role in the RA problem and different assignment algorithms are proposed and introduced into the Grouping & Assignment (GA) algorithm in combination with the selected SDMA algorithms. In particular, RA strategies are considered in which SDMA grouping and resource assignment are performed either separately or jointly and their performances are compared. These investigations are related to problem 8. In Chapter 4, priority schemes are also considered and, in particular, assignment algorithms that assign resources either to initial MSs or to whole SDMA groups, that work on a resource basis or over the multiple resources altogether, and that consider an a-priori power allocation among resources or a joint power allocation among the resources are proposed and their performances are compared, thus addressing problem 8 and problem 9.

Chapter 5 summarizes the main results obtained in this thesis and concludes this work.

# Chapter 2

# Multi-User MIMO-OFDMA System Modeling and Resource Allocation Problem Formulation

## 2.1 Introduction

This chapter describes the system model considered in this thesis, which concerns adaptive RA performed by the transmitter in the DL of a MU MIMO-OFDMA system. This chapter also presents the optimization problem that will be the focus in this work, as well as it discusses additional aspects of the proposed framework for suboptimal RA strategies for MIMO-OFDMA systems, which has been introduced in Chapter 1.

This chapter is organized as follows. Section 2.2 presents the system model adopted in this work. It describes the general scenario and the main assumptions made in this thesis, which are discussed in Section 2.2.1. In Section 2.2.2, the definition of the considered frame structure and of the radio resource units are made. In Section 2.2.3, the models adopted to characterize the radio channel are provided. Section 2.2.4 introduces the models employed to represent the CSI available at the transmitter side, which are used later by the strategies proposed in Chapter 3 and Chapter 4 to perform adaptive RA. Section 2.3 introduces the optimization objective pursued in this work. The general optimization problem of maximizing the sum rate of the system in the DL is formulated in this section. Then, a new alternative formulation of the problem as a mixed integer optimization problem is given, which is used to discuss some particular aspects of the proposed framework for suboptimal RA strategies.

## 2.2 System Model

### 2.2.1 Overall Scenario and Assumptions

In this section, the overall scenario and the main assumptions made in this thesis are presented.

The DL of a single BS located on the corner of a hexagonal cell sector is considered. Considering only a single BS is a reasonable assumption, since performing RA jointly over multiple cells would require a rather large signaling overhead to share control information among them [Liu04]. Therefore, RA is usually performed by each BS individually [ZKAQ01, HRM02, HT02].

It is assumed that the BS disposes of a maximum transmit power $P$, which has to be adequately distributed among the MSs being served by the BS. Interference from other BSs is assumed to be Gaussian-distributed and is directly incorporated into the Additive White Gaussian Noise (AWGN) perceived in the system. In other words, the interference plus noise are considered as the perceived AWGN in the system. Despite the fact that only a limited number of interferers exist in realistic scenarios, this is a widely adopted assumption, which becomes more and more accurate the larger the number of interferers becomes [WIN05a].

The BS is equipped with an $M$-element AA and associated with the BS there is a total number $K$ of MSs. Let $k, k = 1, \ldots, K$, indicate the $k^{\text{th}}$ MS in the system. Each MS $k$ has an $N_k$-element AA, so that there is a total number $N = \sum_{k=1}^{K} N_k$ of receive antennas in the sector. Fig. 2.1 shows the considered scenario. The AAs at the BS and MSs can be seen as spatial filters responsible for performing spatial processing of the signals transmitted from and received by each AA element [LR99, Van02, Hay02]. Indeed, using the AAs, the system can simultaneously transmit on the same frequency-time resource up to $\min\{N, M\}$ data streams [Tel95, FG98, PNG03].

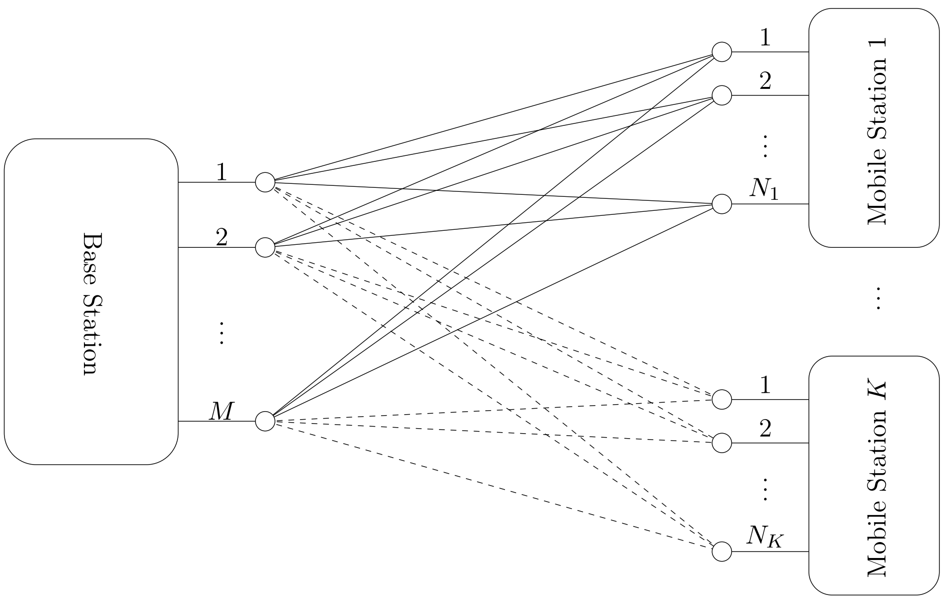



Figure 2.1. DL of a single BS and $K$ MSs equipped with AAs.

In this thesis, the focus is on transmitter side processing and a receiver-oriented design [MBQ04, Qiu04] is pursued, so that processing is mostly concentrated at the BS. It is assumed that the BS knows the spatial processing techniques used by each MS. This permits the BS to adequately perform spatial processing while taking into account spatial processing applied by the MSs [MBQ04, Qiu04]. It is assumed that the BS has CSI about the DL channels to the MSs, which is a requirement in order to perform adaptive RA. The modeling of the CSI available at the BS will be presented later in

Section 2.2.4. Additionally, because it is well-known that adaptive RA is best suited to low mobility environments [LL05], only low mobility is considered in this work.

Gaussian signaling is considered and the data symbols transmitted by the BS to the MSs are assumed to be uncorrelated with unit average power. Moreover, it is assumed that the BS always has data to transmit to the MSs, i.e., a full-buffer traffic model is assumed at the BS for all MSs.

Further, more specific assumptions directly related to the particular topics discussed in the next sections will be introduced in the corresponding section.

### 2.2.2 Frame Structure and Resource Definition

In this section, the frame structure adopted in the system is described and the resource units in frequency, time, and space are defined.

The system employs a combination of OFDMA, TDMA, and SDMA as multiple access scheme. Fig. 2.2 shows the frame structure in time and frequency.

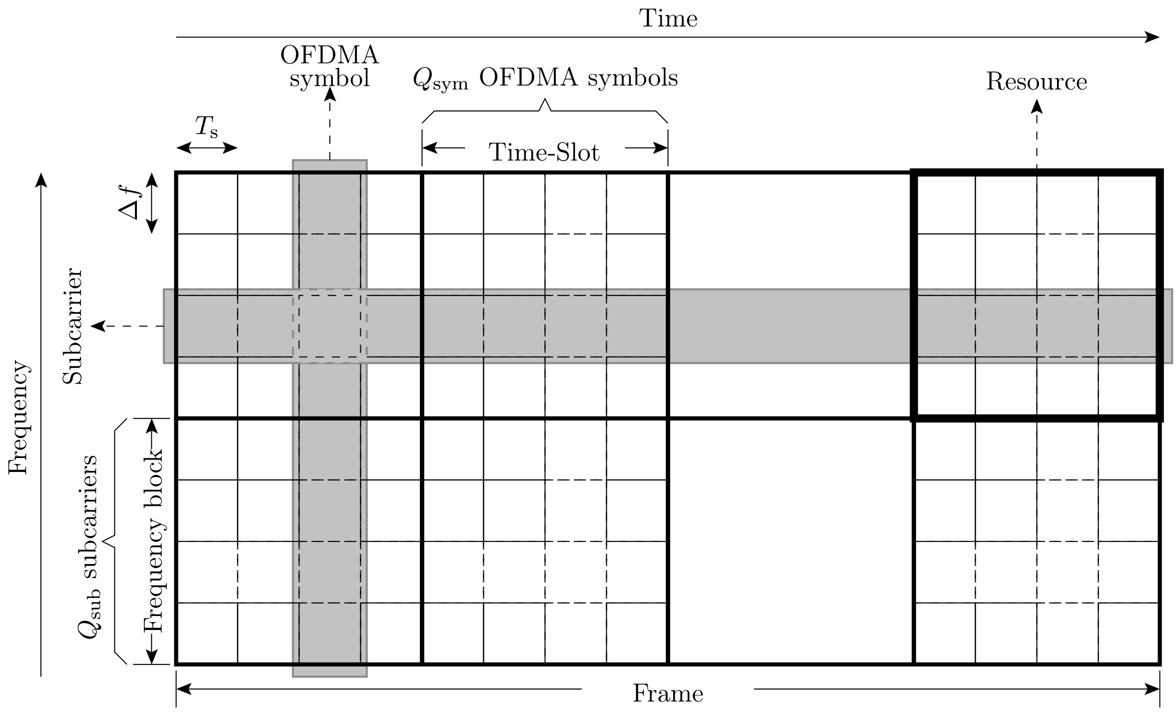



Figure 2.2. Frame structure.

Regarding the OFDMA component of the multiple access scheme, the BS bandwidth $B_{\text{sys}}$ is divided into $S$ subcarriers, which are assumed to be perfectly orthogonal, i.e., there is no Inter-Carrier Interference (ICI) in the system. The channel coherence bandwidth $B_c$ is the bandwidth over which the channel transfer function remains almost flat [Pro95, Skl97]. By selecting a subcarrier bandwidth $\Delta f \ll B_c$, an almost flat channel transfer function over a frequency block composed of $Q_{\text{sub}}$ adjacent subcarriers can be obtained [VVE+00, AHSB+06]. One frequency block is indicated in Fig. 2.2. Frequency blocks are the minimum allocable resource unit in frequency. Thus, a block OFDMA multiple access in frequency domain is used by the system [LL05, FKC06]. Denoting by $\lfloor\cdot\rfloor$ the nearest integer smaller than or equal to the argument, there is a total number $B = \left\lfloor \frac{S}{Q_{\text{sub}}} \right\rfloor$ of frequency blocks in the system, which will be indicated by $b, b = 1, \ldots, B$.

Regarding the TDMA component of the multiple access scheme, frames of duration $T_{\text{FRM}}$ are considered. Each frame is divided into a number $T$ of TSs of duration $T_{\text{TS}} = \frac{T_{\text{FRM}}}{T}$. One TS and one frame are indicated in Fig. 2.2. TSs are the minimum allocable resource unit in time and each TS transports a number $Q_{\text{sym}}$ of subsequent OFDMA symbols. Each OFDMA symbol transports $S$ data symbols. Perfect time synchronization is assumed and a Guard Interval (GI) in form of a Cyclic Prefix (CP) of adequate length is employed. Thus, there is no Inter-Symbol Interference (ISI) in the system. The channel coherence time $T_{\text{c}}$ is the time over which the channel transfer function remains almost constant [Pro95, Skl97]. By selecting an OFDMA symbol duration $T_{\text{s}} \ll T_c$, an almost constant channel transfer function over several OFDMA symbols can be assumed. Considering $T_{\text{s}} \ll T_c$, low MS mobility, and a short frame duration $T_{\text{FRM}}$, the channel transfer function over a whole frame is not expected to vary considerably [LL05, WIN05b].

The multiple access scheme in frequency and time domains defined above is very similar to that considered in the adaptive transmission modes of several candidates for future mobile radio systems [IEE04, WIN05a, DSK+06, 3GP06a]. In this thesis, a radio resource is defined as a frequency-time resource unit described by one frequency block and one TS, as shown in Fig. 2.2. In the literature, such radio resources are also called slots [IEE04], chunks [WIN05a, DSK+06], or Physical Resource Blocks (PRBs) [3GP06a, 3GP06b]. Further on, radio resources will be termed resources and are the minimum allocable radio resource unit in the system. Consequently, the amount of signaling needed to inform MSs about resources allocated to them might be substantially reduced compared to a case in which each subcarrier and OFDMA symbol could be allocated to a different MS [LL05, WIN05b, 3GP06a, 3GP06b]. Because the channel transfer function within a resource is not expected to vary substantially, signaling requirements and channel estimation efforts can also be reduced.

The SDMA component of the multiple access scheme is implemented by multiplexing on each resource up to $M$ data streams, which are separated in space through spatial processing techniques, such as linear precoding [PNG03, JUN05]. In this way, virtual resources are created for each frequency-time resource. These virtual resources are usually termed spatial layers [WIN05a, DSK+06] and are the space resource units in the system. Because in general the BS can transmit $M$ interference-free data streams [SSH04], it is assumed that there are $M$ spatial layers for each resource, as illustrated in Fig. 2.3. On each spatial layer of a resource, the BS can transmit a data stream to a different MS.

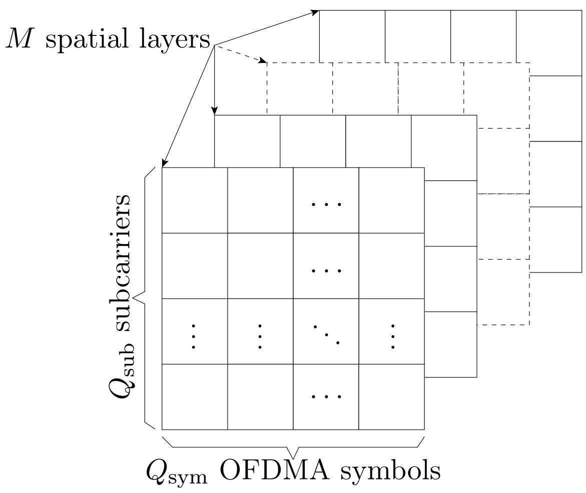



Figure 2.3. Spatial layers of one frequency-time resource.

It is assumed that MSs have not more antennas than the BS has, i.e., $N_k \leq M$, for $k = 1, \ldots, K$. This is a reasonable assumption, due to space limitations at the MSs to install larger AAs, which are much more restrictive than for the BS. Thus, each MS $k$ can occupy at most $N_k$ spatial layers of each resource.

It is assumed that the BS power $P$ can be arbitrarily shared among resources. Let $N_0$ denote the one-sided power spectral density of the AWGN in the system. Then, the average Signal-to-Noise Ratio (SNR) $\gamma$ can be defined as

$$\gamma = \frac{P}{B_{\text{sys}} N_0} = \frac{P}{S \Delta f N_0} = \frac{P}{S} \frac{1}{\sigma^2}, \tag{2.1}$$

where $\sigma^2$ is the average AWGN power per subcarrier. $\gamma$ is used to model the average SNR in the system.

### 2.2.3 Channel Model

This section describes how the radio channel is modeled in this work. In the investigations that will be performed in this thesis, only fast fading will be considered, which is a common assumption when studying RA in MIMO-OFDMA systems. Indeed, most of the RA strategies in Table 1.1 and Table 1.2 have considered only fast fading.

The capacity gains that MIMO techniques can provide to a radio system strongly depend on the spatial properties of the radio channel [PNG03] and, therefore, the adopted spatial channel model is very important. However, because the distributions of Angle of Arrivals (AoAs) and Angle of Departures (AoDs) depend on the geometric properties of the propagation environment, a unique general model for the spatial characteristics of the radio channel cannot be developed. Instead of this, several spatial channel models have been proposed for different particular scenarios, e.g., in [FLFV00, DH03, YO02, PNG03, WIN05c, 3GP07]. For a survey of spatial channel models, the works [YO02], [PNG03, chapters 2 and 3], [WIN05c], and the references therein can be consulted.

In this thesis, the WINNER Phase I Channel Model (WIM) described in [WIN05c] has been selected to model MIMO channels for the following reasons. It is a geometric-based stochastic MIMO channel model whose parameters have been determined based on a large measurement campaign conducted within the Wireless World Initiative New Radio (WINNER) project. Geometric-based stochastic MIMO channel models have also been widely adopted, e.g., within the 3rd Generation Partnership Project (3GPP) [3GP07]. Since the WIM parameters are based on measurements, the WIM models realistic propagation scenarios. The WIM models both the space and time characteristics of the radio channel. Besides that, a free implementation of the model has also been made available in [SDS+05] and can be used to reproduce the results presented later in this thesis.

In the following, a brief description of the main characteristics of the WIM is presented. In the WIM, scenarios representing different propagation environments are modeled. The spatial characteristics of the propagation environment in each scenario are parametrically modeled by groups of scatterers, which are termed clusters. According to the scenario, the WIM models up to $L = 20$ clusters, which will be indexed further on by $l, l = 1, \ldots, L$. Fig. 2.4 shows one cluster as modeled by the WIM and considering Uniform Linear Arrays (ULAs) at the BS and MS.

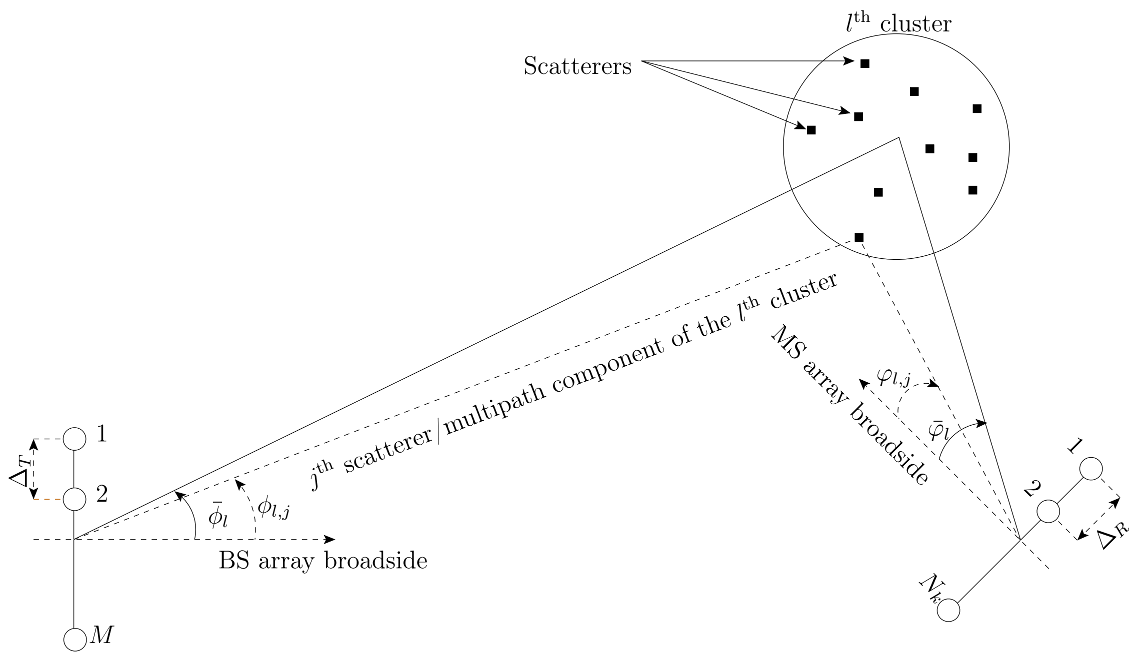



Figure 2.4. Cluster modeling in the WIM.

In Fig. 2.4, $\Delta_T$ and $\Delta_R$ are the distance between adjacent elements of the AAs at the BS and the MS, respectively. According to the positions of BS and MS, each cluster $l$ has associated with it an average AoD $\bar{\phi}_l$, an average AoA $\bar{\varphi}_l$, and an excess delay $\tau_l$, thus building a single bounce between transmitter and receiver [LR99], as illustrated in Fig. 2.4. $\bar{\phi}_l$ and $\bar{\varphi}_l$ are measured with respect to the transmit and receive AA broadsides, respectively [WIN05c, Bal05]. $\bar{\phi}_l$, $\bar{\varphi}_l$ and $\tau_l$ are all statistically modeled by Random Variables (RVs) in the WIM. Each cluster $l$ is composed of 10 scatterers, indexed by $j, j = 1, \ldots, 10$, which have associated AoDs $\phi_{l,j}$ and AoAs $\varphi_{l,j}$ randomly distributed around $\bar{\phi}_l$ and $\bar{\varphi}_l$, respectively. The AoDs $\phi_{l,j}$ and AoAs $\varphi_{l,j}$ are statistically modeled by RVs whose distributions have been determined based on the existing literature and on the performed measurement campaigns [WIN05c]. The multipath components associated with each scatterer $j$ of a cluster $l$ are assumed to have the same gain $\alpha_l^2$ and the same excess delay $\tau_l$. However, for each scatterer $j$ of cluster $l$ a different random phase $\theta_{l,j}$ for its associated multipath component is assumed.

In this thesis, the channel modeling is considered in the frequency domain. For an MS $k, k = 1, \ldots, K$, and a subcarrier $s, s = 1, \ldots, S$, let the channel coefficient $h_{n,m}$ denote the sampled frequency response of the channel between the $m^{\text{th}}$ antenna of the BS and the $n^{\text{th}}$ antenna of the MS. Thus, the MIMO channel between the BS and an MS $k$ on subcarrier $s$ can be illustrated as shown in Fig. 2.5.

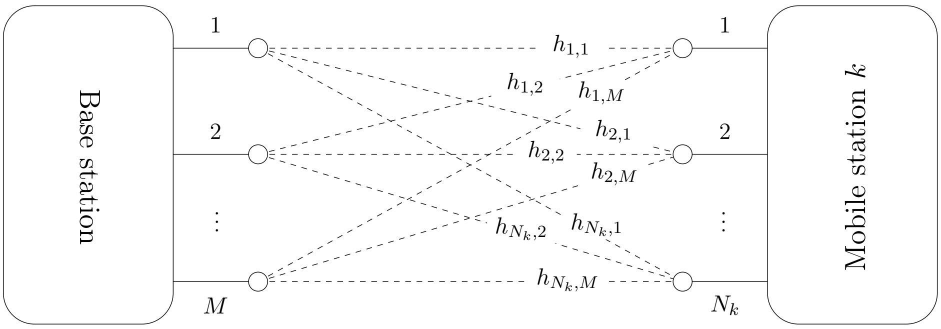



Figure 2.5. MIMO channel between the BS and MS $k$ on subcarrier $s$.

Then, the channel coefficients $h_{n,m}$ of the MIMO channel between the BS and an MS $k, k = 1, \ldots, K$, on a subcarrier $s, s = 1, \ldots, S$, can be organized in a channel matrix $\mathbf{H}_{k,s} \in \mathbb{C}^{N_k \times M}$ as

$$\mathbf{H}_{k,s} = \begin{bmatrix} h_{1,1} & h_{1,1} & \ldots & h_{1,M} \\ h_{2,1} & h_{2,1} & \ldots & h_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_k,1} & h_{N_k,1} & \ldots & h_{N_k,M} \end{bmatrix} \tag{2.2}$$

which is obtained using the implementation of the WIM provided in [SDS+05]. The channel matrix $\mathbf{H}_{k,s}$ of (2.2) encompasses all the relevant propagation characteristics of the channel.

On subcarrier $s$ each link between the BS and an MS $k$ has an associated channel matrix $\mathbf{H}_{k,s}$ given by (2.2). Let $(\cdot)^{\text{T}}$ denote vector/matrix transposition. Then, using (2.2), the channel matrix $\mathbf{H}_s \in \mathbb{C}^{N \times M}$ of all MSs on subcarrier $s$ can be written by stacking the channel matrices $\mathbf{H}_{k,s}$ as

$$\mathbf{H}_s = \left[\mathbf{H}_{1,s}^{\text{T}} \ \mathbf{H}_{2,s}^{\text{T}} \ \ldots \ \mathbf{H}_{K,s}^{\text{T}}\right]^{\text{T}}. \tag{2.3}$$

Analogously to (2.3), an estimated channel matrix $\hat{\mathbf{H}}_s$ of all MSs on subcarrier $s$ can be built by stacking the estimated channel matrices $\hat{\mathbf{H}}_{k,s}$ of the MSs. Both $\mathbf{H}_s$ and $\hat{\mathbf{H}}_s$ are used later in the description of the Channel State Information at the Transmitter (CSIT) models considered in this work.

It must be mentioned that, in general, the performance of RA strategies strongly depends on the considered scenario and channel model [PNG03]. However, the relative performance among the RA strategies studied in this work is expected to be relatively well preserved when considering other scenarios and channel models. The study of the RA strategies of this thesis considering other scenarios and channel models is left for future investigation.

### 2.2.4 Channel State Information Models

In this thesis, Perfect Channel State Information at the Receiver (P-CSIR) is assumed, which means that the MSs perfectly know the channel transfer function of the link between the BS and themselves. In this section, the models adopted to represent the CSI required at the transmitter side in order to perform adaptive RA are discussed. For the CSIT, three models are considered:

- Perfect Channel State Information at the Transmitter (P-CSIT), in which the BS has perfect knowledge about the instantaneous channel transfer function on each subcarrier $s$ for all the MSs.
- Block Channel State Information at the Transmitter (B-CSIT), in which the BS perfectly knows, for each link to an MS, the instantaneous channel transfer function of the middle subcarrier of each frequency block, cf. Section 2.2.2.
- Second-order Channel State Information at the Transmitter (S-CSIT), in which the BS has only knowledge about the second-order statistics of the links to the MSs due to signaling or complexity constraints.

First, the P-CSIT is discussed, which is employed to study the RA strategies proposed in this thesis under idealized conditions. P-CSIT corresponds here to the case in which the BS perfectly knows the matrices $\mathbf{H}_s$, i.e., $\hat{\mathbf{H}}_s = \mathbf{H}_s$, for $s = 1, \ldots, S$. Assuming perfect knowledge about $\mathbf{H}_s$ is to some extent realistic for Time Division Duplexing (TDD), because Uplink (UL) and DL channels are reciprocal whenever the duplexing time $T_\mathrm{d}$, i.e., the interval in which the system alternates the usage of the channel between UL and DL, is much shorter than the channel coherence time $T_\mathrm{c}$ [PNG03]. In this case, channel transfer functions estimated by the BS in the UL are equivalent to the channel transfer functions in the DL and could be used as CSI to perform adaptive RA for each subcarrier individually. However, adaptation on a subcarrier basis implies large signaling overheads to notify the MSs about the subcarriers assigned to them and, therefore, does not represent a realistic system design [LL05].

In this thesis, CSIT is mostly assumed to be obtained on a frame basis through channel estimation and to be available at the first TS of each frame. This CSIT is then considered for all the TSs of a frame. Such an assumption is also considered by the proposals for future mobile radio systems [IEE04, DSK+06, 3GP06a] and results in lower amounts of signaling compared to obtaining CSIT at higher rates, e.g., on a TS basis. Considering that CSIT is obtained on a frame basis and that the channel transfer function is expected to be almost constant within a resource, the CSIT models derived in the sequence are based on the channel transfer function of the frequency blocks during the first TS of each frame.

Secondly, the B-CSIT model is introduced, which is a more realistic CSIT model. It is the main model considered in this thesis. According to Section 2.2.2, an almost flat channel transfer function is assumed for the frequency blocks. Thus, the channel transfer function of a frequency block can be efficiently represented by the channel transfer function of a single subcarrier of the frequency block. Let $s_b$ denote the first subcarrier of the $b^{\text{th}}$ frequency block. Then, the middle subcarrier $\overline{s}_b$ of the frequency block is given by

$$\overline{s}_b = s_b + \lfloor Q_{\text{sub}}/2 \rfloor \,. \tag{2.4}$$

The channel transfer function of the subcarrier $\overline{s}_b$ is assumed to be known by the BS and is used as CSI to represent the channel transfer function of all the $Q_{\text{sub}}$ subcarriers of the frequency block. The reason for choosing the middle subcarrier is that the average distance between it and any other subcarrier within the frequency block is minimum. Because there is a clear and unique mapping between $\overline{s}_b$ and $b$, let $\hat{\mathbf{H}}_{k,b}$ denote the estimated channel matrix of MS $k$ on frequency block $b$, and let $\hat{\mathbf{H}}_b$ denote the estimated channel matrix of all MSs on frequency block $b$, where the index $b$ is used instead of $\overline{s}_b$ for simplicity of notation. $\hat{\mathbf{H}}_{k,b}$ and $\hat{\mathbf{H}}_b$ are given by

$$\hat{\mathbf{H}}_{k,b} = \mathbf{H}_{k,b} = \mathbf{H}_{k,\overline{s}_b}, \tag{2.5a}$$

and

$$\hat{\mathbf{H}}_b = \left[\hat{\mathbf{H}}_{1,b}^{\text{T}} \;\; \hat{\mathbf{H}}_{2,b}^{\text{T}} \;\; \dots \;\; \hat{\mathbf{H}}_{K,b}^{\text{T}}\right]^{\text{T}} = \left[\hat{\mathbf{H}}_{1,\overline{s}_b}^{\text{T}} \;\; \hat{\mathbf{H}}_{2,\overline{s}_b}^{\text{T}} \;\; \dots \;\; \hat{\mathbf{H}}_{K,\overline{s}_b}^{\text{T}}\right]^{\text{T}}, \tag{2.5b}$$

respectively.

Thirdly, the S-CSIT model is discussed, which is a realistic model when signaling constraints are present in the system [BO02]. This model might help to get insight into the performance of adaptive RA when only a very limited amount of CSIT is available, for example when considering Frequency Division Duplexing (FDD). While for TDD assuming P-CSIT or B-CSIT might be realistic, for FDD it might be not. For FDD, channel reciprocity cannot be exploited, since UL and DL frequencies are located in different and usually far apart portions of the spectrum. Therefore, the DL channel matrices have to be estimated by the MSs and transmitted back to the BS. A prohibitively large signaling amount through a feedback channel would be required to report the complex channel transfer factors of each MS on each subcarrier or frequency block back to the BS.

In the following, a representation for the S-CSIT is provided. Let $\mathcal{E}\{\cdot\}$ denote the expectation operator, here over time, and $(\cdot)^{\mathrm{H}}$ denote the conjugate transpose of a vector/matrix. Assuming that MSs' channels are wide-sense stationary [PP02, PNG03], let $\mathbf{R}_{k,b}$ denote the spatial covariance matrix for MS $k$ on frequency block $b$. Applying an EVD to $\mathbf{R}_{k,b}$, one obtains the eigenvector matrix $\mathbf{V}_{k,b}$ and eigenvalue matrix $\mathbf{\Lambda}_{k,b}$. $\mathbf{R}_{k,b}$ can be expressed as

$$\mathbf{R}_{k,b} = \mathcal{E}\left\{\mathbf{H}_{k,b}^{\mathrm{H}}\mathbf{H}_{k,b}\right\} = \mathbf{V}_{k,b}\mathbf{\Lambda}_{k,b}\mathbf{V}_{k,b}^{\mathrm{H}} \tag{2.6}$$

[BO02, SB04, RFH08]. For S-CSIT, the matrix $\mathbf{R}_{k,b}$ is assumed to be known by the BS for each MS $k$ and frequency block $b$. Note that the channel matrix $\mathbf{H}_{k,b}$ is obtained for the middle subcarrier $\bar{s}_b$ of the frequency block $b$, cf. (2.4).

According to [VM01, BO02, SB04], the spatial covariance matrix $\mathbf{R}_{k,b}$ preserves the spatial structure of the channel, i.e., characteristics of the slowly changing geometry of the environment and, according to [SB04, BO02], it can be assumed to be roughly the same for both UL and DL. Because spatial covariance matrices are long-term statistics, they do not need to be reported very often. Therefore, their estimation imposes neither prohibitive computational effort nor signaling overhead. However, the radio channel is not necessarily wide-sense stationary and, indeed, the channel mean may vary due to MS mobility and large-scale fading. Moreover, the spatial covariance matrix $\mathbf{R}_{k,b}$, cf. (2.6), only exists in a statistical sense while for practical purposes it has to be approximated. Let $\tilde{f}$ indicate the $\tilde{f}^{\text{th}}$ frame, which should not be confused with the frequency $f$. Consider a given number $W$ of past frames taken into account to compute a spatial covariance matrix $\mathbf{R}_{k,b}$ to be used as CSIT by the RA strategies during the frames $\tilde{f}, \ldots, \tilde{f}+W-1$, i.e., during the $W$ subsequent frames. In the following, $W$ will be termed the window size. Let $\hat{\mathbf{H}}_{k,b}^{(j)}$ denote the estimated channel matrix of MS $k$ on frequency block $b$ for the frame $j$. Then, an approximation for the spatial covariance matrix $\mathbf{R}_{k,b}$ can be written as

$$\mathbf{R}_{k,b} \approx \frac{1}{W} \sum_{j=\tilde{f}-W}^{\tilde{f}-1} \hat{\mathbf{H}}_{k,b}^{(j),\mathrm{H}} \hat{\mathbf{H}}_{k,b}^{(j)}. \tag{2.7}$$

If wide-sense stationarity can be assumed and $W$ is sufficiently large, $\mathbf{R}_{k,b}$ in (2.7) converges to $\mathbf{R}_{k,b}$ in (2.6). Otherwise, $\mathbf{R}_{k,b}$ in (2.7) is only a statistic taken over a window containing $W$ past samples of the channel matrix $\hat{\mathbf{H}}_{k,b}^{(j)}$. Anyway, for the scenarios investigated in this work the approximation in (2.7) is considered valid.

Using (2.7), the channel matrices $\hat{\mathbf{H}}_{k,b}$ considered in the case of S-CSIT are derived in the sequence. According to [NLTW98, VM01, GJJV03, PNG03, RFH08, KT08], when only second-order channel statistics are available it is recommended to transmit along the $N_k \leq M$ dominant eigenmodes of $\mathbf{R}_{k,b}$, since they represent the most relevant components in the spatial structure of the considered channel. The dominant eigenmodes are characterized by the dominant eigenvalues $\lambda_1(\mathbf{R}_{k,b}) \geq \lambda_2(\mathbf{R}_{k,b}) \geq \ldots \geq \lambda_{N_k}(\mathbf{R}_{k,b})$ of $\mathbf{R}_{k,b}$ and their associated eigenvectors $\mathbf{v}_1(\mathbf{R}_{k,b}), \mathbf{v}_2(\mathbf{R}_{k,b}), \ldots, \mathbf{v}_{N_k}(\mathbf{R}_{k,b})$. Then, for S-CSIT the channel estimate $\hat{\mathbf{H}}_{k,b}$ for MS $k$ on frequency block $b$ is defined as

$$\hat{\mathbf{H}}_{k,b} = \begin{bmatrix} \lambda_1^{1/2}(\mathbf{R}_{k,b}) & 0 & \ldots & 0 \\ 0 & \lambda_2^{1/2}(\mathbf{R}_{k,b}) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_{N_k}^{1/2}(\mathbf{R}_{k,b}) \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^{\mathrm{H}}(\mathbf{R}_{k,b}) \\ \mathbf{v}_2^{\mathrm{H}}(\mathbf{R}_{k,b}) \\ \vdots \\ \mathbf{v}_{N_k}^{\mathrm{H}}(\mathbf{R}_{k,b}) \end{bmatrix}, \tag{2.8}$$

which can be used in the same way as, e.g., $\hat{\mathbf{H}}_{k,b}$ in (2.5a). For single-antenna MSs, the rank-one approximation that results from using (2.8) effectively represents the MSs' channels only if their spatial covariance matrices in (2.7) have low rank too [NLTW98, VM01].

Up to now, the CSIT models considered in this thesis have been introduced and discussed. Because adaptive RA depends on the available CSIT, the accuracy of the channel estimates is also relevant. In the following, a model for erroneous CSIT is introduced, which applies in particular to the B-CSIT model and enables to evaluate the impact of erronenous CSIT on the performance of the RA strategies.

Channel estimation is not perfect and involves channel estimation errors which might originate from AWGN and interference in the system, from suboptimal channel estimation algorithms, from inherent processing or feedback delays, and from too short or too correlated training sequences, among others. Usually, estimation errors are simply neglected by the system so that estimated channel matrices are assumed to be the actual ones. However, this simplifying assumption leads to suboptimal performance [HH03].

Both channel estimation errors and imperfections due to processing or feedback delays can be modeled by an additive Zero Mean Circularly Symmetric Complex Gaussian (ZMCSCG) error term [HH03, YG06, KK07b, KK07a], which is denoted here by $\mathbf{E}_{k,b} \in \mathbb{C}^{N_k \times M}$ for MS $k$ on frequency block $b$. The entries of $\mathbf{E}_{k,b}$ have a variance $\sigma_{\mathbf{E}}^2$. Let $\sigma_{\mathbf{H}}^2$ denote the variance of the entries of the channel matrix $\mathbf{H}_{k,b}$ described in Section 2.2.3. Moreover, consider that $\sigma_{\mathbf{H}}^2$ is known, which is a common assumption, e.g., when ZMCSCG fading is considered and $\sigma_{\mathbf{H}}^2 = 1$ [HH03, YG06, KK07b, KK07a]. Let $\sigma_{\hat{\mathbf{H}}}^2$ denote the variance of the entries in the estimated channel matrices $\hat{\mathbf{H}}_{k,b}$ given by (2.5a). Let $\sigma_{\mathbf{E}}^2$ be defined to be equal to $\sigma_{\mathbf{H}}^2$ and let $0 \leq \nu \leq 1$ be a model parameter controlling the amounts of the actual channel matrix $\mathbf{H}_{k,b}$ and of the additive error term $\mathbf{E}_{k,b}$ in the estimated channel matrix $\hat{\mathbf{H}}_{k,b}$. Thus, different amounts of imperfection in the CSIT can be obtained by varying $\nu$. The estimated channel matrices $\hat{\mathbf{H}}_{k,b}$ of each MS $k$, on frequency block $b$, in the presence of errors can be modeled as

$$\hat{\mathbf{H}}_{k,b} = \sqrt{1-\nu}\mathbf{H}_{k,b} + \sqrt{\nu}\mathbf{E}_{k,b}, \tag{2.9}$$

where $\sqrt{1-\nu}\mathbf{H}_{k,b}$ and $\sqrt{\nu}\mathbf{E}_{k,b}$ model the contributions present in $\hat{\mathbf{H}}_{k,b}$ due to $\mathbf{H}_{k,b}$ and $\mathbf{E}_{k,b}$, respectively. The model in (2.9) is structurally equivalent to the models in [HH03, YG06, KK07b, KK07a]. The main difference relies on the fact that (2.9) ensures $\sigma^2_{\hat{\mathbf{H}}} = \sigma^2_{\mathbf{H}}$, while in [HH03, YG06, KK07b, KK07a] $\sigma^2_{\hat{\mathbf{H}}}$ changes according to the variance of the entries in $\mathbf{E}_{k,b}$. Anyway, defining $\sigma^2_{\mathbf{E}} = \sigma^2_{\mathbf{H}}$ and $\hat{\mathbf{H}}_{k,b}$ according to (2.9) is reasonable because a normalization of $\hat{\mathbf{H}}_{k,b}$ is usually included in the channel estimation.

Similarly to [SH05], the model in (2.9) allows to describe the amount of imperfection in the CSIT as follows. Let $[\mathbf{H}_{k,b}]_{i,j}$ and $[\mathbf{E}_{k,b}]_{i,j}$ denote an entry of the actual channel matrix and of the error term associated with the link of MS $k$ on frequency block $b$, respectively. Considering (2.9), on can define a measure

$$\gamma_{\text{CSI}} = \frac{\mathcal{E}\left\{\left[\sqrt{1-\nu}\,[\mathbf{H}_{k,b}]_{i,j}\right]^2\right\}}{\mathcal{E}\left\{\left[\sqrt{\nu}\,[\mathbf{E}_{k,b}]_{i,j}\right]^2\right\}} = \frac{(1-\nu)}{\nu}\frac{\mathcal{E}\left\{[\mathbf{H}_{k,b}]^2_{i,j}\right\}}{\mathcal{E}\left\{[\mathbf{E}_{k,b}]^2_{i,j}\right\}} = \frac{(1-\nu)}{\nu}\frac{\sigma^2_{\mathbf{H}}}{\sigma^2_{\mathbf{H}}} = \frac{1-\nu}{\nu}, \quad (2.10)$$

which expresses the relationship between the magnitude of the term due to the actual channel matrix $\mathbf{H}_{k,b}$ and the magnitude of the additive estimation error matrix $\mathbf{E}_{k,b}$ present in the estimated channel matrix $\hat{\mathbf{H}}_{k,b}$ under the assumption that all entries in $\mathbf{H}_{k,b}$ are identically distributed [SH05]. Thus, $\gamma_{\text{CSI}}$ describes the quality of the CSI available at the BS. The values that $\gamma_{\text{CSI}}$ can assume in a realistic situation depend on the amount of training symbols and on the transmit power dedicated to them, as well as on the particular properties of the employed channel estimation algorithm [HH03, YG06, KK07b, KK07a].

## 2.3 Optimization Problem

In this section, the problem of maximizing the sum rate of the system is mathematically formulated. Maximizing the sum rate of the system is the main problem studied in this thesis and the problem for which new adaptive RA strategies are developed. The referred formulation will be used later in this section to derive a new mixed integer formulation of the same problem, which is the basis for the framework for suboptimal RA strategies introduced in Section 1.2 and further discussed in this section.

For a scenario with MSs having multiple antennas, it would be necessary to jointly optimize the spatial filters at both the transmitter and receiver sides in order to maximize the sum rate of the system. However, such a joint optimization would lead to increased complexity of MSs and would be in opposition to the receiver-oriented design [MBQ04, Qiu04] pursued in this work. Moreover, if both spatial transmit and receive filters are jointly computed at the BS and informed to the MSs, a too large amount of control information would have to be exchanged between BS and MSs, which is also not desired in practice. Therefore, the problem of maximizing the sum rate of the system considered in this thesis assumes some simplifications, which are discussed in the sequence.

In this work, only linear algorithms are considered for spatial filtering because they are less complex than non-linear algorithms, such as those based on DPC techniques, and allow for a more tractable and simplified mathematical formulation of the problem of maximizing the sum rate of the system. Indeed, linear spatial filtering techniques [PNG03, MBQ04, JUN05] are the ones most often considered in the proposals for future mobile radio systems [IEE04, DSK$^+$06, 3GP06a].

Considering linear spatial filtering, the receiver-oriented design from [MBQ04, Qiu04] can be used to take into account the receive beamformers of the MSs in the problem of maximizing the sum rate of the system without incurring additional complexity of MSs or signaling overheads. Let $\check{\mathbf{w}}_{k,b,n} \in \mathbb{C}^{1\times N_k}$ denote the receive beamformer employed by MS $k$ on frequency block $b$ for the spatial layer $n$, which is computed at the MS. Because it is assumed that the BS knows the spatial processing technique employed by the MSs, it can compute an approximate receive beamformer $\tilde{\mathbf{w}}_{k,b,n}$ for $\check{\mathbf{w}}_{k,b,n}$ based on the available CSI. Let $\tilde{\mathbf{W}}_{k,b} = \left[\tilde{\mathbf{w}}_{k,b,1}^{\mathrm{T}} \;\; \tilde{\mathbf{w}}_{k,b,2}^{\mathrm{T}} \;\; \ldots \;\; \tilde{\mathbf{w}}_{k,b,N_k}^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{C}^{N_k\times N_k}$ be a receive beamforming matrix containing the $N_k$ approximated receive beamformers computed by the BS considering the MS $k$ and frequency block $b$. Then, according to [MBQ04, Qiu04], the BS can adequately take the receive beamformers of MS $k$ into account by using an equivalent channel matrix $\hat{\mathbf{H}}'_{k,b} = \tilde{\mathbf{W}}_{k,b}\hat{\mathbf{H}}_{k,b}$ as estimated channel matrix instead of the original $\hat{\mathbf{H}}_{k,b}$. Considering this approach, each receive antenna $n$ of an MS $k$ can be assumed as a sort of virtual single-antenna MS [SS04b] and the MU MIMO system can be represented as an equivalent MU Multiple Input Single Output (MISO) system. Such an approach has been often employed in different contexts [SSH04, SS04b, CC05, TUBN06, CC07]. However, it copes only partially with the cooperation among antennas of the same MS and allows to obtain approximate solutions to the referred problem of maximizing the sum rate of the system. Anyway, the referred approach is followed in the remainder of this work and only single-antenna MSs are assumed, which might represent either real MSs or virtual single-antenna MSs.

In the following, the problem of maximizing the sum rate of the system is formulated. Because CSIT is made available on a frame basis, cf. Section 2.2.4, it is possible to formulate the problem for the resources described by the frequency blocks $b = 1, \ldots, B$, and the first TS of each frame only and a solution obtained for these resources applies to the resources associated with the remaining TSs of the frame. Then, according to Section 2.2.4 an estimated channel matrix $\hat{\mathbf{H}}_b \in \mathbb{C}^{N\times M}$ of the actual channel matrix $\mathbf{H}_b$ of all MSs is available at the BS for each frequency block $b$ of the system. The estimated channel matrix $\hat{\mathbf{H}}_b$ represents the CSI available at the BS and is given by one of the models introduced in Section 2.2.4. $\hat{\mathbf{H}}_b$ is assumed as the actual channel matrix $\mathbf{H}_b$ without loss of generality, since it reduces to it in the case of perfect channel estimation, thus not affecting the problem formulation [HH03, YG06]. Let $p_{k,b}$ denote the allocated power and $\mathbf{w}_{k,b}$ denote the precoding vector associated with MS $k$ on frequency block $b$. $p_{j,b}$ and $\mathbf{w}_{j,b}$, with $j \neq k$, are the allocated power and the precoding vector of MS $j$ interfering with MS $k$ on frequency block $b$. Let $\|\cdot\|_2$ denote the 2-norm of a vector. Then, the DL SINR $\gamma_{k,b}$ of MS $k$ on frequency block $b$ is given by

$$\gamma_{k,b} = \frac{p_{k,b}\|\hat{\mathbf{h}}_{k,b}\mathbf{w}_{k,b}\|_2^2}{\sigma^2 + \sum\limits_{j=1, j\neq k}^{K} p_{j,b}\|\hat{\mathbf{h}}_{k,b}\mathbf{w}_{j,b}\|_2^2}, \tag{2.11}$$

where the second term in the denominator corresponds to the interference caused by MSs $j \neq k$ sharing in space the frequency block $b$ with MS $k$. Then, using (2.11) the problem of maximizing the sum rate of the system can be formulated as

$$\{p^\star_{k,b}, \mathbf{w}^\star_{k,b}\} = \underset{\{p_{k,b}, \mathbf{w}_{k,b}\}}{\arg\max} \left\{ \sum_{b=1}^{B}\sum_{k=1}^{K} \log_2 \left( 1 + \frac{p_{k,b}\|\hat{\mathbf{h}}_{k,b}\mathbf{w}_{k,b}\|_2^2}{\sigma^2 + \sum\limits_{j=1, j\neq k}^{K} p_{j,b}\|\hat{\mathbf{h}}_{k,b}\mathbf{w}_{j,b}\|_2^2} \right) \right\} \tag{2.12a}$$

$$\text{subject to: } p_{k,b} \geq 0, \forall k, b, \tag{2.12b}$$

$$\sum_{b=1}^{B}\sum_{k=1}^{K} p_{k,b} = \frac{PB}{S}, \tag{2.12c}$$

$$\|\mathbf{w}_{k,b}\|_2 = 1, \forall k, b, \tag{2.12d}$$

where the constraint (2.12b) ensures non-negative power for all MSs and frequency blocks, the constraint (2.12c) limits the total transmit power that the BS can spend, and (2.12d) is the unit-norm constraint for the precoding vector [PNG03, Cal04, TC04a, TC04b, FDH05, DS05, CC05, CC07, FDH07]. The objective function in problem (2.12) is neither a convex nor a concave function of the optimization variables $p_{k,b}$ and $\mathbf{w}_{k,b}$ [BV04], so that convex optimization methods cannot be directly employed to solve (2.12) efficiently. Convex and concave optimization problems are easily convertable into each other [BV04]. Since a set of up to $M$ MSs has to be selected for each frequency block $b$, problem (2.12) is a complex non-convex combinatorial problem [CC05, LZ06, CC07].

In the following, a new formulation for problem (2.12) is developed, which leads to a mixed integer optimization problem [NW99] and that allows to explicitly characterize the four subproblems implicitly embedded in (2.12), i.e., the SDMA grouping problem, the precoding problem, the power allocation problem, and the resource assignment problem discussed in Section 1.1.3. Through the division of the problem into four subproblems, the proposed formulation allows to obtain useful insight information that can be used in the design and comparison of efficient suboptimal RA strategies.

The combinatorial nature of (2.12) can be better seen by introducing some auxiliary binary variables. Let $u_{k,b} \in \{0, 1\}$ be a binary variable indicating whether MS $k$ is allocated on frequency block $b$. Of course, whenever $u_{k,b} = 1$, one has $p_{k,b} > 0$, so that

$$u_{k,b} = \begin{cases} 1, & \text{for } p_{k,b} > 0, \\ 0, & \text{for } p_{k,b} = 0. \end{cases} \tag{2.13}$$

Note that there exists a maximum number

$$G = \sum_{j=1}^{M} \binom{K}{j} = \sum_{j=1}^{M} \frac{K!}{j!(K-j)!} \tag{2.14}$$

of sets of MSs that can be formed using $u_{k,b}$, with each set determining an SDMA group. On frequency block $b$, one of these sets of MSs shares the resource in space through SDMA.

Let $g, g = 1, \ldots, G$, indicate the SDMA groups, $\mathcal{G}_g$ denote the $g^{\text{th}}$ group, and $v_{g,b} \in \{0,1\}$ be a binary variable indicating whether the $g^{\text{th}}$ SDMA group is allocated on frequency block $b$. Note that according to (2.13), if the frequency block $b$ is allocated to an SDMA group $\mathcal{G}'_b$ and $\exists k \notin \mathcal{G}'_b, p_{k,b} > 0$, then there exists another SDMA group $\mathcal{G}_b \neq \mathcal{G}'_b$ for which $p_{k,b} > 0 \Leftrightarrow k \in \mathcal{G}_b$, and the SDMA group $\mathcal{G}_b$ can be seen as the group effectively allocated on frequency block $b$. Considering this, $v_{g,b}$ can defined as

$$v_{g,b} = \begin{cases} 1, & \text{if } p_{k,b} > 0 \Leftrightarrow k \in \mathcal{G}_g, \\ 0, & \text{otherwise.} \end{cases} \tag{2.15}$$

Now, using $u_{k,b}$ and $v_{g,b}$, problem (2.12) can be rewritten as

$$\{p^\star_{k,b}, \mathbf{w}^\star_{k,b}, u^\star_{k,b}, v^\star_{g,b}\} = \underset{\{p_{k,b}, \mathbf{w}_{k,b}, u_{k,b}, v_{g,b}\}}{\arg\max} \left\{ \sum_{b=1}^{B} \sum_{g=1}^{G} v_{g,b} \sum_{k=1}^{K} u_{k,b} \log_2 \left( 1 + \frac{p_{k,b} \|\hat{\mathbf{h}}_{k,b} \mathbf{w}_{k,b}\|_2^2}{\sigma^2 + \sum_{j=1, j\neq k}^{K} p_{j,b} \|\hat{\mathbf{h}}_{k,b} \mathbf{w}_{j,b}\|_2^2} \right) \right\} \tag{2.16a}$$

$$\text{subject to: } p_{k,b} \geq 0, \forall k, b, \tag{2.16b}$$

$$\sum_{b=1}^{B} \sum_{k=1}^{K} p_{k,b} = \frac{PB}{S}, \tag{2.16c}$$

$$\|\mathbf{w}_{k,b}\|_2 = 1, \forall k, b, \tag{2.16d}$$

$$u_{k,b} \in \{0,1\}, \forall k, b, \tag{2.16e}$$

$$\sum_{g=1}^{G} v_{g,b} \leq 1, \forall b, \tag{2.16f}$$

$$v_{g,b} \in \{0,1\}, \forall g, b. \tag{2.16g}$$

The new formulation in (2.16) for maximizing the sum rate of the system clearly characterizes the problem as a mixed integer optimization problem [NW99]. Another formulation for the maximization of the sum rate of the system as a mixed integer optimization problem can be found, e.g., in [CC05, CC07]. However, in [CC05, CC07] it is assumed that MSs sharing the same resource in space do not interfere with each other, i.e., the interference term shown in (2.11) is forced to zero for all MSs on each resource by means of precoding.

Differently from [CC05, CC07], the introduction of $u_{k,b}$ and $v_{g,b}$ in (2.16) together with $\mathbf{w}_{k,b}$ and $p_{k,b}$ allow to explicitly characterize the four subproblems in (2.12) as follows.

- The variables $u_{k,b}$ are related to the **SDMA grouping problem**, in which on each frequency block $b$ up to $M$ of the $K$ MSs have to be optimally selected. Let $\mathcal{G}^{\star}_{g,b}$ denote the optimal SDMA group on frequency block $b$. Then, $u_{k,b} = 1, \forall k \in \mathcal{G}^{\star}_{g,b}$, i.e., for the up to $M$ selected MSs that build the optimal group $\mathcal{G}^{\star}_{g,b}$, and $u_{k,b} = 0, \forall k \notin \mathcal{G}^{\star}_{g,b}$, i.e., for all the other MSs.
- In order to adequately multiplex the MSs in $\mathcal{G}^{\star}_{g,b}$ in space, precoding vectors $\mathbf{w}_{k,b}$ are required, now with $k \in \mathcal{G}^{\star}_{g,b}$. Thus, the variables $\mathbf{w}_{k,b}$ are directly related to the **precoding problem**.
- Similarly, powers $p_{k,b}$ have to be determined and allocated from the BS power $P$ to the MSs in $\mathcal{G}^{\star}_{g,b}$ on frequency block $b$, so that $p_{k,b}$ are directly associated with the **power allocation problem**.
- Finally, $v_{g,b}$ relates to the **resource assignment problem** which corresponds to allocating an SDMA group on each frequency block $b$ while ensuring through the constraint (2.16f) that not more than one SDMA group is assigned on each resource.

This new formulation in (2.16) provides interesting insight into the elements of the original optimization problem (2.12), as discussed in the sequel.

The first interesting aspect made clearer by the proposed formulation of (2.16) is that the distribution of the transmit power is the element that keeps the four subproblems interdependent. It can be seen from (2.16) that if the power is distributed a priori among resources, e.g., using EPA, the resource assignment can be performed on a resource-by-resource basis. Additionally, the precoding vector $\mathbf{w}_{k,b}$ of MS $k$ on frequency block $b$ plays no role if the power $p_{k,b}$ allocated to MS $k$ is zero. The same applies to the precoding vector $\mathbf{w}_{j,b}$ of an interfering MS $i$ on frequency block $b$ whenever the power $p_{j,b}$ of the interfering MS is zero. In the same way, the binary selection variables $u_{k,b}$, which determine SDMA groups, and $v_{g,b}$, which determine which SDMA group is allocated on frequency block $b$, depend only on $p_{k,b}$, as shown in (2.13) and (2.15), respectively.

A second interesting aspect is that the SDMA grouping problem is responsible for yielding (2.16) combinatorial. It can be seen from (2.14) that the combinatorial increase of $G$ affecting $v_{g,b}$ is due to $u_{k,b}$. If SDMA groups are already defined on each resource, the problem is no longer NP-C.

A third aspect of interest in the proposed framework resides on the fact that dividing problem (2.16) into subproblems allows to adapt the solutions given to each subproblem individually, as discussed in Section 1.2. By efficiently solving the subproblems, efficient solutions for the complete problem can be obtained. This is the main approach in this work, in which solutions for each of the four subproblems are proposed, carefully combined, and shown to perform a suboptimal but efficient RA. Moreover, the individual adaptation of the solution given to each subproblem allows to derive RA strategies providing a trade-off between the maximization of the sum rate of the system and the degree of throughput fairness among the MSs. According to problems (2.12) and (2.16), after performing the RA, a large amount of resources might be assigned to some MSs while other MSs might get no resources, which is in accordance with the objectives in (2.12) and (2.16). However, this fact might generate insatisfaction among the MSs, which is an undesired effect in a realistic scenario. Therefore, it might be desired to have some degree of fairness on the distribution of resources in the system without significantly compromising the performance of the system in terms of its sum rate, which might be obtained by considering the proposed framework given by problem (2.16).

# Chapter 3

# Resource Allocation in MIMO-OFDMA Systems: Single Resource

## 3.1 Introduction

In the chapter, the problem of maximizing the sum rate of the system is investigated for the particular case in which a single resource is considered at a time. The single-resource case is indicated by the dashed line in Fig. 1.3. Considering a single resource, problem (2.16) can be considerably simplified. Nevertheless, RA strategies designed for this case can be employed in a MIMO-OFDMA systems with multiple resources by considering resources one-by-one and shall present lower complexity than RA strategies that take multiple resources into account at a time.

Because the transmit power is the element coupling the subproblems of (2.16), as discussed in Section 2.3, each resource can be considered individually if the power allocated to each of them is determined a priori, e.g., through an EPA. Considering a single frequency block $b$ and the first TS of a frame, the resource assignment problem is no longer part of the optimization problem in (2.16). Therefore, the summation in $b$ can be disregarded in (2.16), as well as the allocation variables $v_{g,b}$ and the associated constraints (2.16f) and (2.16g). Removing the index $b$ from the optimization variables and denoting by $P_b$ the available power for the resource under consideration, problem (2.16) can be simplified to

$$\{p_k^\star, \mathbf{w}_k^\star, u_k^\star\} = \underset{\{p_k, \mathbf{w}_k, u_k\}}{\arg\max} \left\{ \sum_{k=1}^{K} u_k \log_2 \left( 1 + \frac{p_k \|\hat{\mathbf{h}}_k \mathbf{w}_k\|_2^2}{\sigma^2 + \sum\limits_{j=1, j\neq k}^{K} p_j \|\hat{\mathbf{h}}_k \mathbf{w}_j\|_2^2} \right) \right\} \tag{3.1a}$$

$$\text{subject to: } p_k \geq 0, \forall k, \tag{3.1b}$$

$$\sum_{k=1}^{K} p_k = P_b, \tag{3.1c}$$

$$\|\mathbf{w}_k\|_2 = 1, \forall k, \tag{3.1d}$$

$$u_k \in \{0, 1\}, \forall k. \tag{3.1e}$$

In spite of being more simple than problem (2.16), problem (3.1) is still a combinatorial problem since the SDMA grouping problem remains included into it. Thus, problem (3.1) cannot be solved in an optimal way unless through an ES for the optimal SDMA group, precoding vectors, and power allocation. Similarly to problem (2.16), problem (3.1) also asks for a suboptimal rather efficient solution.

For large numbers of MSs, the number $G$ of SDMA groups, given by (2.14), becomes too large and an ES requiring to compute precoding vectors and power allocation for each of the $G$ groups becomes too complex. However, if the SDMA grouping problem is suboptimally solved and a suboptimal but adequate SDMA group is determined, the precoding vectors and power allocation must be computed only for this single group. In this chapter, most of the proposed RA strategies follow the approach of solving the SDMA grouping problem independently of the precoding and power allocation algorithm. Nevertheless, some RA strategies involving SDMA algorithms that depend on the precoding and power allocation are also considered for comparison.

This chapter is organized as follows. Section 3.2 describes the proposed SDMA algorithms, which combine a grouping metric and a grouping algorithm, cf. Section 1.2. Section 3.2.1 describes the grouping metrics. In this section, two new grouping metrics are proposed, namely the weighted norm of the total spatial correlation and the convex combination of the total spatial correlation and channel gains. Section 3.2.2 describes the grouping algorithms. In this section, a new grouping algorithm is proposed, namely the Convex Grouping Algorithm (CGA), which is formulated as a convex optimization problem and, therefore, is not NP-C. In Section 3.2.3, the SDMA algorithms are defined by adequately combining a grouping metric of Section 3.2.1 and a grouping algorithm of Section 3.2.2. In this section, combinations of the new grouping metrics and grouping algorithms are discussed, as well as some combinations of existing grouping metrics and grouping algorithms not yet considered in previous works.

In Section 3.3, the precoding and power allocation algorithms are discussed. Section 3.3.1 and Section 3.3.2 describe the precoding algorithms and the power allocation algorithms considered in this work, respectively. In these two sections, precoding and power allocation are considered separately. In Section 3.3.2, a new Soft Dropping Algorithm (SDA) is proposed to perform power allocation. The SDA can be seen as an extension of the Soft Dropping Power Control (SDPC) of [YGRS97] to MU MIMO systems. The SDA includes an initial power allocation and takes into account the total power constraint of (3.1c). In Section 3.3.3, the SRA algorithm is proposed, which is responsible for adjusting the size of groups built by SDMA algorithms unaware of the actual precoding and power allocation. The SRA is aware of the actual precoding and power allocation and is able to improve the sum rate of the system by eventually removing MSs from SDMA groups. In Section 3.3.4, the precoding and power allocation algorithms are defined by combining a precoding algorithm, a power allocation algorithm, and, whenever necessary, the SRA with an adequate criterion for MS removal, cf. Section 1.2. In Section 3.3.4, also the alternative of performing joint precoding and power allocation is discussed, and a new iterative algorithm for joint precoding and power allocation is proposed, which combines Generalized Eigen-Precoding and the SDA. Moreover, a convergence proof for the SDA applied to MU MIMO scenarios is newly provided in the same section.

In Section 3.4, the RA strategies considered in this chapter are defined by combining an SDMA algorithm of Section 3.2 and a precoding and power allocation algorithm of Section 3.3.

In Section 3.5, the performance and complexity of the RA strategies considered in this chapter are investigated. In Section 3.5.1, the values of the system parameters and the values of the parameters of the RA strategies are defined. In Section 3.5.2, Section 3.5.3, and Section 3.5.4 the performance of the RA strategies in terms of the achieved average sum rate is investigated considering the P-CSIT, B-CSIT, and S-CSIT models of Section 2.2.4, respectively. Similarly, in Section 3.5.5, the performance of the RA strategies is investigated considering erroneous B-CSIT, cf. Section 2.2.4. In Section 3.5.6, the impact of the SRA on the average sum rate achieved by the RA strategies is discussed. In Section 3.5.7, the impact of a parameter $\beta$, which will be introduced in Section 3.2.1, on the performance of two of the newly proposed SDMA algorithms is investigated. In Section 3.5.8, expressions for the complexity of the studied RA strategies are newly derived and, using them, the RA strategies are compared in terms of complexity. In Section 3.5.9, the fairness of the RA strategies is investigated considering different criteria for the MS priority. Finally, in Section 3.5.10 some results of this chapter are shortly summarized.

## 3.2 SDMA Algorithm

### 3.2.1 Grouping Metric

This section deals with the grouping metrics, which are employed by SDMA algorithms to measure the spatial compatibility among MSs in an SDMA group, cf. Section 1.2. Several grouping metrics have been considered in previous works, as it has been shown in Table 1.1. In general, all the grouping metrics make use of the channel matrices $\hat{\mathbf{H}}_b$ representing the CSI available at the BS, cf. Section 2.2.4. Therefore, grouping metrics are usually functions of the CSIT that try to map the characteristics of the spatial channels of the MSs to a scalar value quantifying how efficiently these MSs can be separated in space.

In this section, four grouping metrics are considered:

1. The **group capacity**, which considers the capacity of an SDMA group as spatial compatibility metric and takes into account the actual precoding and power allocation algorithm of the system.
2. The **sum of channel gains with null space Successive Projections (SPs)**, which successively performs projections of the estimated channel matrices of one MS onto the null space of the channel matrices of other MSs.
3. The new **weighted norm of the total spatial correlation**, which has been proposed by the author in [MK06]. It is a heuristic grouping metric with low complexity which is based only on the spatial correlation and gain of the channels of the MSs.

4. The new **convex combination of the total spatial correlation and channel gains**, which has been proposed by the author in [MK07a]. It is a convex combination of the total spatial correlation and gain of the channels of the MSs in a SDMA group and also has low complexity.

In the following, the group capacity is described. Let $\mathcal{G}$ denote a candidate SDMA group on the considered resource and let $K_{\mathcal{G}}$ denote the number of MSs in $\mathcal{G}$. The channel matrix $\hat{\mathbf{G}}$ for the SDMA group $\mathcal{G}$ can be obtained from the estimated channel matrix $\hat{\mathbf{H}}$ in (2.5b) by considering only the rows of $\hat{\mathbf{H}}$ corresponding to the channels of the MSs that belong to $\mathcal{G}$. For example, if the MSs 1, 2, and $K$ belong to the SDMA group $\mathcal{G}$, the matrix $\hat{\mathbf{G}} \in \mathbb{C}^{K_{\mathcal{G}} \times M}$ contains the $1^{\text{st}}$, the $2^{\text{nd}}$, and the $K^{\text{th}}$ rows of $\hat{\mathbf{H}}$. Let $i, i = 1, \ldots, K_{\mathcal{G}}$, indicate the MSs in $\mathcal{G}$, while the index $k, k = 1, \ldots, K$, still indicates the $k^{\text{th}}$ MS in the system. The vector channel of the $i^{\text{th}}$ MS in $\mathcal{G}$ is given by the $i^{\text{th}}$ row $\hat{\mathbf{g}}_i \in \mathbb{C}^{1 \times N}$ of $\hat{\mathbf{G}}$. Further on the term vector will be omitted when referring to the vector channels of the MSs. $\mathbf{w}_i \in \mathbb{C}^{M \times 1}$ and $p_i \in \mathbb{R}^+$ are the precoding vector and the allocated power of the $i^{\text{th}}$ MS in $\mathcal{G}$, respectively. Then, the group capacity $f_{\text{CAP}}(\mathcal{G})$ of the SDMA group $\mathcal{G}$ is written as

$$f_{\text{CAP}}(\mathcal{G}) = \sum_{i=1}^{K_{\mathcal{G}}} \log_2 \left( 1 + \frac{p_i \left\| \hat{\mathbf{g}}_i \mathbf{w}_i \right\|_2^2}{\sigma^2 + \sum\limits_{j=1, j \neq i}^{K_{\mathcal{G}}} p_j \left\| \hat{\mathbf{g}}_i \mathbf{w}_j \right\|_2^2} \right), \quad \text{with} \quad i, j \in \mathcal{G}, \tag{3.2}$$

where CAP stands for capacity.

Assuming adequate precoding and power allocation, the higher the group capacity $f_{\text{CAP}}(\mathcal{G})$ is, the more spatially compatible the MSs in an SDMA group $\mathcal{G}$ are. Since the group capacity $f_{\text{CAP}}(\mathcal{G})$ reflects the effective capacity of the SDMA group considering precoding and power allocation, it is a reliable grouping metric [FDH05, MK06, FDH07, MK07a, MK07b, MK08].

A drawback of $f_{\text{CAP}}(\mathcal{G})$ is that new precoding vectors and a new power allocation must be computed for all MSs in $\mathcal{G}$ whenever the composition of $\mathcal{G}$ changes. Because precoding and power allocation might involve complex operations, such as matrix inversions or decompositions, the complexity of an SDMA algorithm employing $f_{\text{CAP}}(\mathcal{G})$ as grouping metric might become high if $f_{\text{CAP}}(\mathcal{G})$ must be computed for a large number of candidate SDMA groups.

Considering multi-antenna MSs, the group capacity has been considered, e.g., in [FDH05, MK06, FDH07]. In this case, the main changes concern precoding and power allocation which must take into account the multiple antennas at the MSs.

In the following, the sum of channel gains with null space SPs is described. The gain of the channel $\hat{\mathbf{h}}_k$ of MS $k$ can be simply calculated as its squared 2-norm, i.e, $\|\hat{\mathbf{h}}_k\|_2^2$. In general, the higher the channel gain $\|\hat{\mathbf{h}}_k\|_2^2$ of MS $k$, the higher its achievable capacity.

However, considering null space projections the effective gains of the channels of the MSs in an SDMA group are conditioned to the degree of spatial correlation among the channels [DH04, FDH05, FDH07]. This fact is illustrated in Fig. 3.1 for two MSs. Let $\mathcal{N}_1$ and $\mathcal{N}_2$ denote the null spaces of the channels $\hat{\mathbf{h}}_1$ and $\hat{\mathbf{h}}_2$, respectively, and consider that $\hat{\mathbf{h}}_1$ is projected onto $\mathcal{N}_2$ and $\hat{\mathbf{h}}_2$ is projected onto $\mathcal{N}_2$, as shown in Fig. 3.1. Then, if the channels $\hat{\mathbf{h}}_1$ and $\hat{\mathbf{h}}_2$ are highly spatially uncorrelated, as in Fig. 3.1 left, after the projections much of the gain of the original channels is preserved. However, if the channels $\hat{\mathbf{h}}_1$ and $\hat{\mathbf{h}}_2$ are highly spatially correlated, as in Fig. 3.1 right, a considerable part of the channel gains get lost after the projection. Therefore, the sum of the channel gains considering null space projections becomes high if the channels of the MSs in the SDMA group are highly spatially uncorrelated and have high gain. This principle is also valid for SDMA groups containing several MSs. However, for an SDMA group $\mathcal{G}$ containing several MSs the channel $\hat{\mathbf{g}}_i$ of each MS $i \in \mathcal{G}$ would have to be projected onto the joint null space of the MSs $i' \in \mathcal{G}, i' \neq i$, which can be determined, e.g., using SVDs [SSH04, FDH05, FDH07].

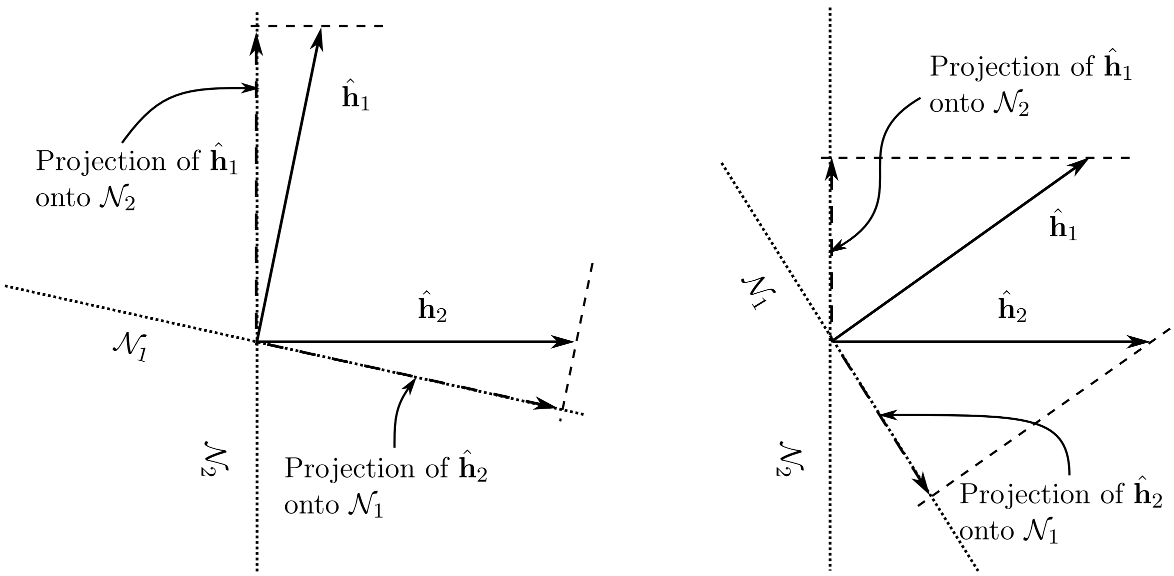



Figure 3.1. Null space projections.

Using SPs, only the channel $\hat{\mathbf{g}}_i$ of MS $i$ is projected onto the null space of the channels $\hat{\mathbf{g}}_{i'}$ of all MSs $i', i' = 1, 2, \ldots, i-1$. Let $\mathbf{I}_M$ denote an $M \times M$ identity matrix and let $\mathbf{T}_i \in \mathbb{C}^{M \times M}$ denote the matrix that projects the channel of MS $i$ onto the null space of the channels of MSs $i'$. Then, $\mathbf{T}_i$ is written as

$$\mathbf{T}_i = \begin{cases} \mathbf{I}_M, & \text{for } i = 1, \\ \mathbf{T}_{i-1} - \dfrac{\mathbf{T}_{i-1}^{\mathrm{H}} \hat{\mathbf{g}}_{i-1}^{\mathrm{H}} \hat{\mathbf{g}}_{i-1} \mathbf{T}_{i-1}}{\|\hat{\mathbf{g}}_{i-1} \mathbf{T}_{i-1}\|_2^2}, & \text{for } i = 2, \ldots, K_{\mathcal{G}}, \end{cases} \tag{3.3}$$

where for $i = 1$ no projections are needed and $\mathbf{T}_1 = \mathbf{I}_M$ [DS05, FDH05, YG05, TJ05, TUBN06, FDH07, MK07b, MK08]. Then, using (3.3), the sum of channel gains with null space SPs $f_{\mathrm{SP}}(\mathcal{G})$ is written as

$$f_{\mathrm{SP}}(\mathcal{G}) = \sum_{i=1}^{K_{\mathcal{G}}} \|\hat{\mathbf{g}}_i \mathbf{T}_i\|_2^2 \tag{3.4}$$

[DS05, FDH05, YG05, TJ05, TUBN06, FDH07, MK07b, MK08].

The channel gain of an MS $i \notin \mathcal{G}$ after a SP onto the null space of the channels of the MSs $i' \in \mathcal{G}$ is $\|\hat{\mathbf{g}}_i \mathbf{T}_i\|_2^2$. The higher the channel gain $\|\hat{\mathbf{g}}_i\|_2^2$ of MS $i$ is and the more spatially uncorrelated the channel of MS $i$ with respect to the channels of the MSs $i' \in \mathcal{G}$ is, the higher $\|\hat{\mathbf{g}}_i \mathbf{T}_i\|_2^2$ might become and the more spatially compatible MS $i$ and the MSs $i'$ are considered to be. Therefore, the higher the gains of the channels of the MSs in $\mathcal{G}$ and the more orthogonal the channels of the MSs in $\mathcal{G}$ are, the higher the values that $f_{\mathrm{SP}}(\mathcal{G})$ assumes and the more spatially compatible the MSs in $\mathcal{G}$ are considered to be. Since $f_{\mathrm{SP}}(\mathcal{G})$ favors SDMA groups whose MSs have high channel gain and are highly spatially uncorrelated, it can be efficiently used as grouping metric [DS05, FDH05, YG05, TJ05, TUBN06, FDH07, MK07b, MK08].

For single-antenna MSs, $f_{\mathrm{SP}}(\mathcal{G})$ involves no SVDs or matrix inversions, depends neither on precoding vector nor on the power allocation, and requires a reduced number of null space projections using (3.3) since the channels of an MS $i'$ in $\mathcal{G}$ is not projected onto the null space of the channels of the MSs $i > i'$. Thus, $f_{\mathrm{SP}}(\mathcal{G})$ has a relatively low complexity, e.g., compared to $f_{\mathrm{CAP}}(\mathcal{G})$ [MK08]. A potential drawback of $f_{\mathrm{SP}}(\mathcal{G})$ is that it explicitly depends on the admission order of the MSs in the SDMA group and, therefore, yields a total number

$$G' = \sum_{j=1}^{M} j! \binom{K}{j} = \sum_{j=1}^{M} \frac{K!}{(K-j)!}, \tag{3.5}$$

of possible candidate groups, which might be much larger than $G$ in (2.14). Nevertheless, by taking care of the admission order in the design of the SDMA algorithm, $f_{\mathrm{SP}}(\mathcal{G})$ can be efficiently used as grouping metric [DS05, FDH05, YG05, TJ05, TUBN06, FDH07, MK07b, MK08].

Considering multi-antenna MSs, null-space successive projections have been considered, e.g., in [FDH05, TUBN06, FDH07]. In particular, a low-complexity group capacity metric based on null-space projections is considered in [FDH05, FDH07] while the sum of channel gains with null-space SPs is used in [TUBN06].

In the following, the weighted norm of the total spatial correlation is discussed. Let $|\cdot|$ denote the absolute value of a complex number, $\mathrm{Re}\{\cdot\}$ denote the real part of a complex number, $\mathcal{D}\{\cdot\}$ denote a diagonal matrix whose diagonal elements are given in the vector argument, and $\|\cdot\|_{\mathrm{F}}$ denote the Frobenius norm of a matrix/vector. Given the channels $\hat{\mathbf{h}}_j$ and $\hat{\mathbf{h}}_k$ of MSs $j$ and $k$, respectively, the spatial correlation among these two MSs is given by the maximum normalized scalar product

$$\rho_{j,k} = \frac{|\hat{\mathbf{h}}_j \hat{\mathbf{h}}_k^{\mathrm{H}}|}{\|\hat{\mathbf{h}}_j\|_2 \|\hat{\mathbf{h}}_k\|_2} \tag{3.6}$$

[Cal04, MK06, MK07a, MK07b, MK07c, MK08], which is an upper bound on the standard normalized scalar product $\tilde{\rho}_{j,k}$ [Mey01]. The standard normalized scalar product $\tilde{\rho}_{j,k}$ corresponds to the cosine of the angle between the vector channels of the two MSs and is given by $\tilde{\rho}_{j,k} = \frac{\mathrm{Re}\{\hat{\mathbf{h}}_j \hat{\mathbf{h}}_k^{\mathrm{H}}\}}{\|\hat{\mathbf{h}}_j\|_2 \|\hat{\mathbf{h}}_k\|_2}$ [Mey01]. Because $\tilde{\rho}_{j,k}$ suppress the imaginary part of $\hat{\mathbf{h}}_j \hat{\mathbf{h}}_k^{\mathrm{H}}$, the normalized scalar product does not capture the spatial correlation among the two complex vector channels as efficiently as the maximum normalized scalar product $\rho_{j,k}$ [SS04a].

Let the attenuation vector $\mathbf{a}_{\mathcal{G}} \in \mathbb{R}_+^{K_{\mathcal{G}} \times 1}$ be defined as a vector containing the inverse of the channel gains of all the $K_{\mathcal{G}}$ MSs in $\mathcal{G}$, i.e.,

$$\mathbf{a}_{\mathcal{G}} = \left[ \|\hat{\mathbf{g}}_1\|_2^{-2} \;\; \|\hat{\mathbf{g}}_2\|_2^{-2} \;\; \ldots \;\; \|\hat{\mathbf{g}}_{K_{\mathcal{G}}}\|_2^{-2} \right]^{\mathrm{T}}. \tag{3.7}$$

Then, using (3.6), (3.7), and $\hat{\mathbf{G}}$, one can write the spatial correlation matrix $\mathbf{C}_{\mathcal{G}} \in \mathbb{R}_+^{K_{\mathcal{G}} \times K_{\mathcal{G}}}$ as

$$\mathbf{C}_{\mathcal{G}} = \left| \sqrt{\mathcal{D}\{\mathbf{a}_{\mathcal{G}}\}} \hat{\mathbf{G}} \hat{\mathbf{G}}^{\mathrm{H}} \sqrt{\mathcal{D}\{\mathbf{a}_{\mathcal{G}}\}} \right|, \tag{3.8}$$

which contains the spatial correlation $\rho_{i,j}, \forall i,j \in \mathcal{G}$, with the operator $|\cdot|$ being applied elementwise to $\mathbf{C}_{\mathcal{G}}$.

Let $\mu_{\mathcal{G}}$ denote the geometric mean of the square root of the channel gains of the MSs in $\mathcal{G}$, i.e.,

$$\mu_{\mathcal{G}} = \left( \prod_{i=1}^{K_{\mathcal{G}}} \|\hat{\mathbf{g}}_i\|_2 \right)^{1/K_{\mathcal{G}}}, \quad \text{with} \quad i \in \mathcal{G}. \tag{3.9}$$

Then, using (3.8) and (3.9), the weighted norm of the total spatial correlation is defined as

$$f_{\mathrm{WN}}(\mathcal{G}) = \frac{\mu_{\mathcal{G}} K_{\mathcal{G}}}{\|\mathbf{C}_{\mathcal{G}}\|_{\mathrm{F}}} \tag{3.10}$$

[MK06], where WN stands for weighted norm. This heuristic grouping metric has been proposed by the author in [MK06] and it has been designed to present the following properties:

1. It captures the total spatial correlation among the channels of all MSs in the SDMA group.
2. It favors large SDMA groups.
3. It favors SDMA groups containing MSs whose channel gains are high and uniformly distributed.

In the following, the motivation for the choice of the above properties is provided. The lower $\rho_{j,k}$ for the channels of the MSs $j$ and $k$ is, the less spatially correlated the channels of the MSs $j$ and $k$ are. Therefore, $\rho_{j,k}$ efficiently measures the spatial compatibility between the MSs $j$ and $k$ and it has often been used as grouping metric [STSK98, KTSK01, Cal04, MK06, MK07a, MK07b, MK07c, MK08]. However, $\rho_{j,k}$ can only be calculated pairwise. Because an SDMA group can contain more than two MSs, a measure of the total spatial correlation among the channels of all MSs in $\mathcal{G}$ is required. In (3.10), the total spatial correlation among the MSs in $\mathcal{G}$ is captured by the Frobenius norm $\|\mathbf{C}_{\mathcal{G}}\|_{\mathrm{F}}$, which involves the spatial correlation of all MSs in $\mathcal{G}$. A similar approach employing the Frobenius norm to capture the total spatial correlation among several vector channels has been employed in [SS04a]. The lower the values of the elements $\rho_{i,j}$ of $\mathbf{C}_{\mathcal{G}}$ are, the lower its Frobenius norm $\|\mathbf{C}_{\mathcal{G}}\|_{\mathrm{F}}$ is, and the less spatially correlated the MSs in $\mathcal{G}$ are.

The larger the SDMA group size $K_{\mathcal{G}}$ is, the higher the spatial multiplexing gains that can be achieved [FDH05]. Therefore, larger SDMA groups should be preferred, which is accomplished by the factor $K_{\mathcal{G}}$ placed in the numerator of (3.10).

The higher the gain of the channel of an MS, the higher its achievable rate. Therefore, MSs with high channel gain should be preferred. It is also well-known that in order to maximize capacity an unequal power distribution among MSs must be preferred [PNG03, Cal04, CT06]. In an SDMA group, allocating power to the MS with the worst channel gain might be less efficient with respect to the group capacity than giving the same power to an MS with higher channel gain. Thus, if the channel gains of the MSs in a group strongly differ, it becomes more probable that the MS with the worst channel gain gets no power allocated to it. This is often the case when applying the Water Filling Algorithm (WFA) for power allocation [PNG03, BV04, PF05, JRV$^{+}$05]. Nevertheless, if MSs getting non-zero and zero power belong to the same group, an unnecessary separation in space of the ones with respect to the others might still be performed. With a more uniform distribution of the channel gains of the MSs in an SDMA group, such a situation becomes more unlikely. Therefore, a uniform distribution of the channel gains of the MSs in $\mathcal{G}$ is also desirable. The channel gains have been considered as grouping metric, e.g., in [TJ05], while putting MSs with similar channel gains in the same group has been considered in [STSK98, STKL01, KTSK01]. In (3.10), the geometric mean $\mu_{\mathcal{G}}$ favors SDMA groups containing MSs with high and more uniformly distributed channel gains.

Since $f_{\mathrm{WN}}(\mathcal{G})$ is inversely proportional to $\|\mathbf{C}_{\mathcal{G}}\|_{\mathrm{F}}$ and directly proportional to $K_{\mathcal{G}}$ and $\mu_{\mathcal{G}}$, the less spatially correlated the MSs in $\mathcal{G}$ are, the larger the SDMA group size $K_{\mathcal{G}}$ is, and the higher and the more uniformly distributed the gains of the channels of the MSs in $\mathcal{G}$ are, the higher values $f_{\mathrm{WN}}(\mathcal{G})$ assumes and the more spatially compatible the MSs in $\mathcal{G}$ are considered to be. $f_{\mathrm{WN}}(\mathcal{G})$ does not depend on precoding or power allocation and does not involve matrix inversions, decompositions, or projections, and it has lower complexity than $f_{\mathrm{CAP}}(\mathcal{G})$ and $f_{\mathrm{SP}}(\mathcal{G})$.

Considering multi-antenna MSs, this metric has been investigated by the author in [MK06] and has been shown to perform as good as $f_{\mathrm{CAP}}(\mathcal{G})$.

In the following, the convex combination of the total spatial correlation and channel gains is discussed. For this metric, the motivation for combining spatial correlation and channel gains is the same as described for $f_{\mathrm{WN}}(\mathcal{G})$, i.e., to favor SDMA groups containing MSs whose channels have high gain and are highly spatially uncorrelated.

Extending the definitions of (3.7) and (3.8) to all the $K$ MSs, the attenuation vector $\mathbf{a} \in \mathbb{R}_{+}^{K\times 1}$ and the spatial correlation matrix $\mathbf{C} \in \mathbb{R}_{+}^{K\times K}$ can be expressed as

$$\mathbf{a} = \left[\|\hat{\mathbf{h}}_1\|_2^{-2} \ \|\hat{\mathbf{h}}_2\|_2^{-2} \ \ldots \ \|\hat{\mathbf{h}}_K\|_2^{-2}\right]^{\mathrm{T}}, \tag{3.11a}$$

and

$$\mathbf{C} = \left|\sqrt{\mathcal{D}\left\{\mathbf{a}\right\}}\hat{\mathbf{H}}\hat{\mathbf{H}}^{\mathrm{H}}\sqrt{\mathcal{D}\left\{\mathbf{a}\right\}}\right|, \tag{3.11b}$$

respectively.

Let the binary selection vector $\mathbf{u}$ be defined as

$$\mathbf{u} = \left[u_1 \ u_2 \ \ldots \ u_K\right]^{\mathrm{T}}, \tag{3.12}$$

where the index $b$ is omitted. For any group $\mathcal{G}$, it holds that $u_k = 1, \forall k \in \mathcal{G}$, otherwise $u_k = 0$. Thus, each value that $\mathbf{u}$ can assume is uniquely associated with a given SDMA group $\mathcal{G}$. Then, combining (3.11) and (3.12) the convex combination of the total spatial correlation and channel gains $f_{\mathrm{CC}}(\mathcal{G})$ is defined as

$$f_{\mathrm{CC}}(\mathcal{G}) = \frac{(1-\beta)}{\|\mathbf{C}\|_{\mathrm{F}}}\mathbf{u}^{\mathrm{T}}\mathbf{C}\mathbf{u} + \frac{\beta}{\|\mathbf{a}\|_{\mathrm{F}}}\mathbf{a}^{\mathrm{T}}\mathbf{u}, \tag{3.13}$$

where $0 \leq \beta \leq 1$ is a control parameter establishing the trade-off between spatial correlation and channel gain. The factors $\frac{1}{\|\mathbf{C}\|_{\mathrm{F}}}$ and $\frac{1}{\|\mathbf{a}\|_{\mathrm{F}}}$ are normalization factors introduced to balance $\mathbf{C}$ and $\mathbf{a}$, i.e., to compensate for their absolute difference and have an unbiased $\beta$ [MK07a, MK07b, MK08]. In this thesis, the scaling factors $\frac{1}{\|\mathbf{C}\|_{\mathrm{F}}}$ and $\frac{1}{\|\mathbf{a}\|_{\mathrm{F}}}$ have been experimentally chosen as to obtain the highest average sum rate figures for values of $\beta$ around 0.5. In (3.13), CC stands for convex combination.

In contrast to $f_{\mathrm{CAP}}(\mathcal{G})$, $f_{\mathrm{SP}}(\mathcal{G})$, and $f_{\mathrm{WN}}(\mathcal{G})$, which increase when the spatial compatibility among the MSs in an SDMA increases, the convex combination of the total spatial correlation and channel gains $f_{\mathrm{CC}}(\mathcal{G})$ decreases. Thus, the lower the value that $f_{\mathrm{CC}}(\mathcal{G})$ assumes, the more spatially compatible the MSs in $\mathcal{G}$ are.

Similarly to $f_{\mathrm{WN}}(\mathcal{G})$, $f_{\mathrm{CC}}(\mathcal{G})$ is less complex than $f_{\mathrm{CAP}}(\mathcal{G})$ and $f_{\mathrm{SP}}(\mathcal{G})$ and depends neither on precoding matrices, as do the metrics in [STKL01, FDH05], nor on a computationally complex operation, such as the SVD in [TJ05, SS04a]. Differently from the grouping metric in [TJ05] that considers only the channel gains, the spatial correlation among the MSs is suitably taken into account by $f_{\mathrm{CC}}(\mathcal{G})$. In contrast to $f_{\mathrm{WN}}(\mathcal{G})$,

$f_{CC}(\mathcal{G})$ allows to control the importance given to spatial correlation and to channel gains through the parameter $\beta$.

Among the four grouping metrics described in this section, the group capacity $f_{CAP}(\mathcal{G})$ is the only one that takes into account the precoding and power allocation algorithm employed by the system. Grouping metrics that take into account the precoding and the power allocation algorithms, such as the group capacity [FDH05] or the SINR of the MSs in an SDMA group [STKL01, FDH05], efficiently measure spatial compatibility and help to achieve good performance in terms of average sum rate. However, because precoding and power allocation usually depend on relatively complex vector/matrix operations, the good performance achieved by such metrics comes at the expense of increased complexity [MK08]. The sum of channel gains with SPs $f_{SP}(\mathcal{G})$ does not involve the actual precoding and power allocation, but requires relatively complex vector/matrix operations to quantify the spatial compatibility among MSs in an SDMA group. Grouping metrics that involve operations like projections [FDH05, YG05, TUBN06] or SVDs of the channel matrix of the SDMA group [SS04a, TJ05] have also been shown to provide good performance in terms of average sum rate. Anyway, both $f_{CAP}(\mathcal{G})$ and $f_{SP}(\mathcal{G})$ can be considered as relatively complex grouping metrics, especially if they have to be computed for a large number of candidate SDMA groups.

Alternatively, grouping metrics that are based only on the spatial correlation, e.g., measured using (3.6), and on the channel gains of the MSs involve much simpler vector/matrix operations. Indeed, spatial correlation and channel gains have been successfully used to design low-complexity rather efficient grouping metrics [FN96, STSK98, STKL01, KTSK01, DH04, Cal04, SS04a, MK06, MK07a, MK07c]. The proposed weighted norm of the total spatial correlation $f_{WN}(\mathcal{G})$ and the convex combination of the total spatial correlation and channel gains $f_{CC}(\mathcal{G})$ belong to this group of metrics and offer an efficient and less complex way of measuring the spatial compatibility among MSs.

The performance and complexity of SDMA algorithms employing one of the four grouping metrics discussed in this section will be investigated later in this chapter.

### 3.2.2 Grouping Algorithm

In this section, the grouping algorithms are discussed, which combined with the grouping metrics presented in Section 3.2.1 will compose the SDMA algorithms investigated in this work. The task of the grouping algorithm is to determine an efficient SDMA group on a given resource with acceptable complexity compared to an ES. Five grouping algorithms are considered here:

1. The Exhaustive Search Algorithm (ESA), which performs an ES for the best SDMA group, i.e., which looks for the group maximizing a certain grouping metric $f_{(\cdot)}(\mathcal{G}_g)$ by comparing all possible candidate groups.

2. The Convex Grouping Algorithm (CGA), which is a new grouping algorithm proposed by the author in [MK07a] and formulated as a quadratic convex optimization problem. This grouping algorithm finds an SDMA group composed of spatially compatible MSs using convex optimization methods.
3. The Best Fit Algorithm (BFA), which is a greedy algorithm that, starting from a group containing an initial MS, extends the group by sequentially admitting the most spatially compatible MS with respect to the MSs already admitted to the SDMA group [STKL01]. This grouping algorithm is based on a simple heuristic rule and can build SDMA groups containing spatially compatible MSs with less computational effort than the previous two grouping algorithms.
4. The First Fit Algorithm (FFA), which is a simplification of the BFA that does not admit the most spatially compatible MS to the SDMA group, but the first spatially compatible MS that it finds [STKL01].
5. The Random Grouping Algorithm (RGA), which just randomly builds an SDMA group of specific size while ignoring any grouping metric. This is the simplest grouping algorithm.

In the following, each of the above grouping algorithms will be discussed. The ESA is the optimal grouping algorithm, since it finds a globally optimal SDMA group $\mathcal{G}^\star$ that maximizes the considered grouping metric, i.e.,

$$\mathcal{G}^\star = \arg\max_{\mathcal{G}_g} \left\{ f_{(\cdot)}(\mathcal{G}_g) \right\}, \text{ with } g = 1, \ldots, G. \tag{3.14}$$

It can be observed in (3.14) that the ESA uses the grouping metric to compare the different candidate SDMA groups $\mathcal{G}_g$. For a given grouping metric $f_{(\cdot)}(\mathcal{G}_g)$, the ESA can be used to upper bound the performance of other grouping algorithms. In spite of being the optimal grouping algorithm, the ESA might be prohibitively complex, since it computes the grouping metric $f_{(\cdot)}(\mathcal{G}_g)$ for each SDMA group $\mathcal{G}_g$ and, according to (2.14), the number $G$ of SDMA groups combinatorially increases with $K$.

In the following, the proposed CGA is described, which has been designed by the author in [MK07a] together with $f_{\mathrm{CC}}(\mathcal{G})$ of (3.13). Herein, a more detailed derivation of the CGA compared to that in [MK07a] is provided. According to Section 3.2.1, there is a unique mapping between $\mathbf{u}$ in (3.12) and an SDMA group $\mathcal{G}$ and, consequently, $f_{\mathrm{CC}}(\mathcal{G})$ can be expressed as a function of $\mathbf{u}$. Let $\|\cdot\|_1$ denote the 1-norm of a vector. Then, using (3.13) one can write the following quadratic integer optimization problem

$$\mathbf{u}^\star = \arg\min_{\mathbf{u}} \left\{ \frac{(1-\beta)}{\|\mathbf{C}\|_{\mathrm{F}}} \mathbf{u}^{\mathrm{T}} \mathbf{C} \mathbf{u} + \frac{\beta}{\|\mathbf{a}\|_{\mathrm{F}}} \mathbf{a}^{\mathrm{T}} \mathbf{u} \right\}, \tag{3.15a}$$

$$\text{subject to: } \|\mathbf{u}\|_1 = K_{\mathcal{G}}^\star, \tag{3.15b}$$

$$u_k \in \{0,1\}, \forall k, \tag{3.15c}$$

whose solution $\mathbf{u}^\star$ can be directly mapped to the best SDMA group $\mathcal{G}^\star$ of size $K_{\mathcal{G}}^\star$ containing MSs that have low total spatial correlation and low total channel attenuation, which are weighted by $\frac{(1-\beta)}{\|\mathbf{C}\|_{\mathrm{F}}}$ and $\frac{\beta}{\|\mathbf{a}\|_{\mathrm{F}}}$ in (3.15a), respectively. The constraint (3.15b) introduces an additional parameter, namely the target SDMA group size $K_{\mathcal{G}}^{\star}$, which is the number of MSs that $\mathcal{G}^{\star}$ must contain and which can assume values between 1 and $M$. Several SDMA algorithms make use of a target SDMA group size in order to simplify the determination of the best group $\mathcal{G}^{\star}$.

In the following, the CGA is derived as a quadratic convex approximation of problem (3.15). A convex optimization problem is desired because it has a unique global optimal solution, is not NP-C, and can be efficiently solved using convex optimization methods. In the following, the changes required to obtain a quadratic convex optimization problem from (3.15) are described.

Let the components $u_k$ of binary selection vector $\mathbf{u}$ of (3.15) be relaxed to belong to the interval $[0, 1]$ and let $\tilde{u}_k \in [0, 1]$ and $\tilde{\mathbf{u}}$ denote the relaxed versions of $u_k$ and $\mathbf{u}$, respectively. Then, a relaxed quadratic optimization problem derived from (3.15) can be written as

$$\tilde{\mathbf{u}}^{\star} = \arg\min_{\tilde{\mathbf{u}}} \left\{ \frac{(1-\beta)}{\|\mathbf{C}\|_{\mathrm{F}}} \tilde{\mathbf{u}}^{\mathrm{T}} \mathbf{C} \tilde{\mathbf{u}} + \frac{\beta}{\|\mathbf{a}\|_{\mathrm{F}}} \mathbf{a}^{\mathrm{T}} \tilde{\mathbf{u}} \right\}, \tag{3.16a}$$

$$\text{subject to: } \|\tilde{\mathbf{u}}\|_1 = K_{\mathcal{G}}^{\star}, \tag{3.16b}$$

$$0 \leq \tilde{u_k} \leq 1, \forall k. \tag{3.16c}$$

Because quadratic integer optimization problems might still be NP-C and have consequently a prohibitively high complexity [GJ79], this relaxation is necessary to obtain a quadratic convex optimization problem from (3.15). However, this is not sufficient. Indeed, if $\mathbf{C}$ in (3.16a) is not positive semidefinite, problem (3.16) is still non-convex and NP-C [GJ79, BV04], and the application of convex optimization methods might find only a local minimum for (3.16). Because $\mathbf{C}$ in (3.11b) is not necessarily positive semidefinite, a quadratic convex optimization problem for the CGA can be derived from (3.16) by replacing $\mathbf{C}$ in (3.16a) by a positive semidefinite matrix $\tilde{\mathbf{C}}$, which can be obtained through diagonal loading, i.e., by defining $\tilde{\mathbf{C}} \triangleq \mathbf{C} + \epsilon \mathbf{I}_K$, where $\epsilon \geq 0$ is the diagonal loading factor. Then, by replacing $\mathbf{C}$ by $\tilde{\mathbf{C}}$ in problem (3.16), the quadratic convex optimization problem considered by the CGA is obtained as

$$\tilde{\mathbf{u}}^{\star} = \arg\min_{\tilde{\mathbf{u}}} \left\{ \frac{(1-\beta)}{\|\mathbf{C}\|_{\mathrm{F}}} \tilde{\mathbf{u}}^{\mathrm{T}} \tilde{\mathbf{C}} \tilde{\mathbf{u}} + \frac{\beta}{\|\mathbf{a}\|_{\mathrm{F}}} \mathbf{a}^{\mathrm{T}} \tilde{\mathbf{u}} \right\}, \tag{3.17a}$$

$$\text{subject to: } \|\tilde{\mathbf{u}}\|_1 = K_{\mathcal{G}}^{\star}, \tag{3.17b}$$

$$0 \leq \tilde{u_k} \leq 1, \forall k. \tag{3.17c}$$

The CGA is presented in Table 3.1.

Table 3.1. Convex Grouping Algorithm.

1. Solve problem (3.17).
2. Determine $\mathcal{G}^{\star}$ by rounding to one the $K_{\mathcal{G}}^{\star}$ largest components and to zero the other $K - K_{\mathcal{G}}^{\star}$ components of $\tilde{\mathbf{u}}^{\star}$.

In the following, the impact of the changes made to problem (3.15) to formulate the CGA are discussed. Integer relaxation and diagonal loading are techniques often used to simplify and solve optimization problems [NW99, Set04, BV04]. In problem (3.17), the components $\tilde{u}_k^\star$ of the obtained continuous solution $\tilde{\mathbf{u}}^\star$ can be interpreted as the probability of the corresponding MS $k$ belonging to the best SDMA group $\mathcal{G}^\star$. Thus, in order to determine $\mathcal{G}^\star$, the solution $\tilde{\mathbf{u}}^\star$ of (3.17) has to be converted into an integer solution $\mathbf{u}^\star$. This is done in step 2 of Table 3.1 by rounding to one the $K_\mathcal{G}^\star$ largest components $\tilde{u}_k^\star$ and to zero the other $K - K_\mathcal{G}^\star$ components of $\tilde{\mathbf{u}}^\star$, thus determining which MSs belong to $\mathcal{G}^\star$ with highest probability.

Consider the cost function in (3.15a) and the constraints (3.15b) and (3.15c). Then, replacing $\mathbf{u}^\mathrm{T}\mathbf{C}\mathbf{u}$ in (3.15a) by $\mathbf{u}^\mathrm{T}\tilde{\mathbf{C}}\mathbf{u}$ results in

$$\begin{aligned}\frac{(1-\beta)}{\|\mathbf{C}\|_\mathrm{F}}\mathbf{u}^\mathrm{T}\tilde{\mathbf{C}}\mathbf{u} + \frac{\beta}{\|\mathbf{a}\|_\mathrm{F}}\mathbf{a}^\mathrm{T}\mathbf{u} &= \frac{(1-\beta)}{\|\mathbf{C}\|_\mathrm{F}}\mathbf{u}^\mathrm{T}\mathbf{C}\mathbf{u} + \frac{\beta}{\|\mathbf{a}\|_\mathrm{F}}\mathbf{a}^\mathrm{T}\mathbf{u} + \frac{(1-\beta)\epsilon}{\|\mathbf{C}\|_\mathrm{F}}\|\mathbf{u}\|_2^2 \\ &= \frac{(1-\beta)}{\|\mathbf{C}\|_\mathrm{F}}\mathbf{u}^\mathrm{T}\mathbf{C}\mathbf{u} + \frac{\beta}{\|\mathbf{a}\|_\mathrm{F}}\mathbf{a}^\mathrm{T}\mathbf{u} + \frac{(1-\beta)\epsilon K_\mathcal{G}^\star}{\|\mathbf{C}\|_\mathrm{F}},\end{aligned} \tag{3.18}$$

which is equivalent to (3.15a), since the term $\frac{(1-\beta)\epsilon K_\mathcal{G}^\star}{\|\mathbf{C}\|_\mathrm{F}}$ is a constant for all $\mathbf{u}$ and, consequently, does not change problem (3.15). However, replacing $\tilde{\mathbf{u}}^\mathrm{T}\mathbf{C}\tilde{\mathbf{u}}$ by $\tilde{\mathbf{u}}^\mathrm{T}\tilde{\mathbf{C}}\tilde{\mathbf{u}}$ in (3.17a) affects the optimization problem. Indeed, expanding (3.17a) analogously to (3.18), the obtained last term $\frac{(1-\beta)\epsilon}{\|\mathbf{C}\|_\mathrm{F}}\|\tilde{\mathbf{u}}\|_2^2$ is no longer constant for all $\tilde{\mathbf{u}}$. Consequently, problem (3.17) is not equivalent but only an approximation of problem (3.15). The smaller the term $\frac{(1-\beta)\epsilon}{\|\mathbf{C}\|_\mathrm{F}}\|\tilde{\mathbf{u}}\|_2^2$ is, the closer the approximation of problem (3.17) to problem (3.15) becomes. Due to (3.17b) and (3.17c), the $\|\tilde{\mathbf{u}}\|_2^2$ is constrained and limited to $\frac{(K_\mathcal{G}^\star)^2}{K} \leq \|\tilde{\mathbf{u}}\|_2^2 \leq K_\mathcal{G}^\star$, so that the approximation of problem (3.17) to problem (3.15) is improved by selecting a loading factor $\epsilon$ as small as possible. Let $\lambda^-(\mathbf{C})$ denote the minimum eigenvalue of $\mathbf{C}$. Therefore, the smallest $\epsilon$ value that makes $\tilde{\mathbf{C}}$ positive semidefinite is given by

$$\epsilon = -\min\left\{0, \lambda^-(\mathbf{C})\right\}, \tag{3.19}$$

which turns problem (3.17) into a convex problem and well approximates problem (3.17) to problem (3.15). However, in order to determine $\epsilon$, cf. (3.19), an EVD of $\mathbf{C}$ is required, which is an undesired operation due to its complexity. This EVD can be avoided by upper bounding $\epsilon$. Unfortunately, the author does not know the existence of a tight upper bound for the absolute value of the minimum eigenvalue of a real non-negative symmetric matrix like $\mathbf{C}$.

In the following, alternatives that do not involve complex operations are proposed to upper bound $\epsilon$ with a value $\varepsilon$. Let $\mathbf{e}_k, k = 1, \ldots, K$, denote the $k^\mathrm{th}$ column of the $K \times K$ identity matrix $\mathbf{I}_K$ and let $\mathbf{1}_K$ denote a $K \times 1$ vector of ones. According to the Perron-Frobenius theorem, the dominant eigenvalue $\lambda^+(\mathbf{C})$ of $\mathbf{C}$ is always non-negative and is lower bounded by $\min\left\{\mathbf{1}_K^\mathrm{T}\mathbf{C}\mathbf{e}_1, \ldots, \mathbf{1}_K^\mathrm{T}\mathbf{C}\mathbf{e}_K\right\}$ [Mey01]. Because a loading factor $\epsilon$ is only needed if $\lambda^-(\mathbf{C}) < 0$, the lower bound of $\lambda^+(\mathbf{C})$ can be used to upper bound

$\epsilon$, i.e., a reliable upper bound $\varepsilon$ for $\epsilon$ can be defined as

$$0 \leq \epsilon \leq \varepsilon = \min\left\{\mathbf{1}_K^{\mathrm{T}}\mathbf{C}\mathbf{e}_1, \mathbf{1}_K^{\mathrm{T}}\mathbf{C}\mathbf{e}_2, \ldots, \mathbf{1}_K^{\mathrm{T}}\mathbf{C}\mathbf{e}_K\right\} \leq \lambda^+(\mathbf{C}), \tag{3.20}$$

and $\tilde{\mathbf{C}}$ can be redefined as $\tilde{\mathbf{C}} \triangleq \mathbf{C} + \varepsilon\mathbf{I}_K$. Because the upper bound $\varepsilon$ in (3.20) is often not a tight bound to $\epsilon$, a non-negative value for $\varepsilon$ can be alternatively determined based on measurements. In the investigations conducted in this thesis, no value larger than or equal to 1 has been observed for $\epsilon$ so that it was decided to arbitrarily set $\varepsilon$ to one, which means $\tilde{\mathbf{C}} \triangleq \mathbf{C} + \mathbf{I}_K$.

In the following, the potential combination of CGA with other grouping metrics is discussed and the conditions that grouping metrics must fulfill in order to be used with the CGA are stated. Until now, it might seem that CGA can only be used with the $f_{\mathrm{CC}}(\mathcal{G})$. However, the CGA is in essence the convex relaxation of the quadratic integer optimization problem (3.15) and it can be used with any grouping metric that allows stating a real, symmetric, and positive semidefinite matrix $\mathbf{C}$. Indeed, even a hermitian matrix could also be used. However, due to the quadratic form of (3.15), only its real part would be taken into account in the optimization. Moreover, no restrictions on the definition of $\mathbf{a}$ are imposed. Let $\varrho_{j,k}$ be a spatial compatibility metric among the channels of MSs $j$ and $k$ used to compose $\mathbf{C}$ analogously to $\rho_{j,k}$ of (3.6). Thus, $\varrho_{j,k}$ must fulfill

$$\varrho_{k,j} = \varrho_{j,k}, \tag{3.21a}$$

$$\varrho_{k,j} \geq 0, \tag{3.21b}$$

in order to be used with CGA. Condition (3.21a) ensures symmetry for $\mathbf{C}$, but not necessarily positive semidefiniteness. In this case, diagonal loading of $\mathbf{C}$ with $\varepsilon$ can be applied again to build a positive semidefinite matrix $\tilde{\mathbf{C}}$ to be used in the approximation problem. Condition (3.21b) ensures that the entries in $\mathbf{C}$ are non-negative so that the upper bound $\varepsilon$ in (3.20) can be applied. As an example, the squared cosine of the angle between two vector channels $\hat{\mathbf{h}}_j$ and $\hat{\mathbf{h}}_k$ given by $\frac{\mathrm{Re}^2\{\hat{\mathbf{h}}_j\hat{\mathbf{h}}_k^{\mathrm{H}}\}}{\|\hat{\mathbf{h}}_j\|_2^2\|\hat{\mathbf{h}}_k\|_2^2}$, the grouping metrics proposed in [SS04a], as well as the equal norm grouping metric in [STKL01], respect (3.21) and could be used with the CGA. Anyway, in the investigations conducted in this thesis only $f_{\mathrm{CC}}(\mathcal{G})$ will be combined with the CGA.

In spite of solving only an approximation problem for (3.15), the CGA builds efficient SDMA groups, as it will be seen later in this chapter. Moreover, the complexity of the CGA, which relies on numerical convex optimization methods, will also be investigated later in this chapter.

In the following, the BFA is described. The BFA has been proposed in [STSK98, STKL01]. Starting with an SDMA group containing a single MS, the BFA sequentially extends the group by admitting to it the MS that most increases the grouping metric. Let $\mathcal{G} = \{k'\}$ be the initial SDMA group containing only the MS $k'$ and let $K_{\mathcal{G}}$ be the size of $\mathcal{G}$. Then, the BFA temporarily admits one MS $k \notin \mathcal{G}$ to the SDMA group

$\mathcal{G}$ and computes the grouping metric $f_{(\cdot)}(\mathcal{G})$ for this extended group. This MS $k$ is removed from $\mathcal{G}$ and the process is repeated with the next MS. After the grouping metrics for the groups built by temporarily admitting each of the $K - K_{\mathcal{G}}$ MSs have been computed, the MS which resulted in the highest value for the grouping metric when admitted to $\mathcal{G}$ is permanently inserted into the group. Then, the same procedure is repeated with the remaining MSs for the extended group until the group size $K_{\mathcal{G}}$ reaches the target SDMA group size $K^\star_{\mathcal{G}}$ or until no more MSs able to increase the grouping metric exist. The BFA is presented in Table 3.2.

Table 3.2. Best Fit Algorithm.

1. Select an initial MS $k'$ and build $\mathcal{G} = \{k'\}$.
2. While $K_{\mathcal{G}} < K^\star_{\mathcal{G}}$
   (a) Set $\mathcal{G}' = \mathcal{G} \cup \left\{ \arg\max_{k \notin \mathcal{G}} \left\{ f_{(\cdot)}(\mathcal{G} \cup \{k\}) \right\} \right\}$.
   (b) If $f_{(\cdot)}(\mathcal{G}') > f_{(\cdot)}(\mathcal{G})$, set $\mathcal{G} = \mathcal{G}'$, otherwise stop.
3. Define the best group as $\mathcal{G}^\star = \mathcal{G}$.

An example of the BFA is shown in Fig. 3.2 for a total number $K = 9$ of single-antenna MSs and $M = 4$ antennas at the BS. A target SDMA group size $K^\star_{\mathcal{G}} = M = 4$ is assumed. Therein, MS $k' = 1$ is selected as the MS for the initial SDMA group $\mathcal{G}$. Then, MSs 2 to 9 are temporarily admitted to $\mathcal{G}$. The grouping metric is computed and one finds out that MS 3 is the MS that at most increases the grouping metric when admitted to $\mathcal{G}$. MS 3 is then permanently added to $\mathcal{G}$, and the process is repeated now for MSs 2, and 4 to 9, with MS 9 being added in the sequence, followed by MS 4, when the size of $\mathcal{G}$ size reaches the target SDMA group size $K^\star_{\mathcal{G}}$ and the BFA finishes.

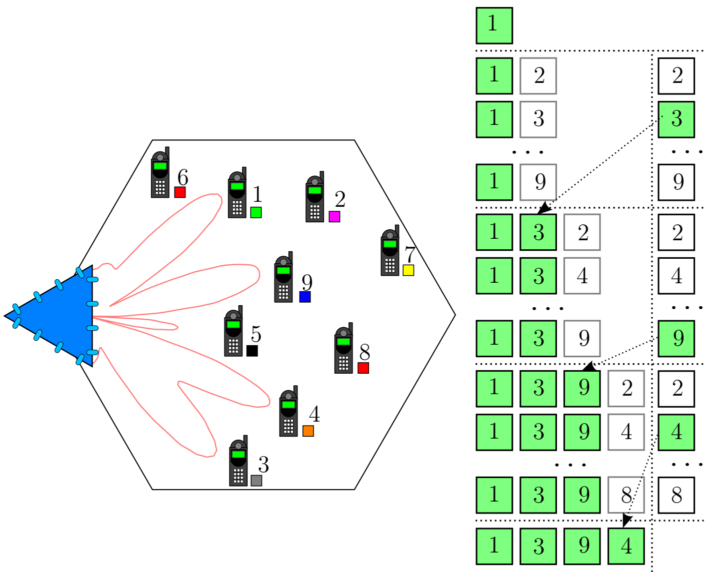


Figure 3.2. Example of BFA.

Because the BFA tests only a small number of candidate SDMA groups and relies on a simple heuristic, it is less complex than the ESA and CGA.

In the following, the FFA is described. The FFA has also been proposed in [STKL01]. Starting with an initial SDMA group $\mathcal{G}$ containing a single MS $k'$, the FFA sequentially extends $\mathcal{G}$ by immediately admitting an MS $k \notin \mathcal{G}$ whenever the admission of this MS improves the grouping metric. Then, this procedure is repeated considering the next MS and the extended group until the target SDMA group size $K_{\mathcal{G}}^{\star}$ is reached or until no more MSs able to increase the grouping metric exist. The FFA is presented in Table 3.3.

Table 3.3. First Fit Algorithm.

1. Select an initial MS $k'$ and build $\mathcal{G} = \{k'\}$.
2. While $K_{\mathcal{G}} < K_{\mathcal{G}}^{\star}$
   (a) Set $\mathcal{G}' = \mathcal{G} \cup \{k\}$, with $k \notin \mathcal{G}$.
   (b) If $f_{(\cdot)}(\mathcal{G}') > f_{(\cdot)}(\mathcal{G})$, set $\mathcal{G} = \mathcal{G}'$.
3. Define the best group as $\mathcal{G}^{\star} = \mathcal{G}$.

Differently from the BFA, which test all MSs that do not belong to $\mathcal{G}$ and admits the one that at most increases the grouping metric, the FFA admits permanently an MS $k$ as soon as the grouping metric computed for $\mathcal{G} \cup \{k\}$ is larger than the grouping metric computed for the group $\mathcal{G}$ without the new MS $k$. Thus, in order to build $\mathcal{G}^{\star}$ the FFA involves less computations than the BFA.

The RGA needs not many comments. As stated before, it just randomly selects a group of $K_{\mathcal{G}}^{\star}$ MSs to build the SDMA group $\mathcal{G}^{\star}$. Consequently, it is less complex than the other grouping algorithms in this section.

### 3.2.3 SDMA Algorithm Definition

In this section, the SDMA algorithms that will be investigated in this work are defined by combining the grouping metrics from Section 3.2.1 and the grouping algorithms from Section 3.2.2.

The SDMA algorithms will be named after the grouping metric and grouping algorithm. For example, the SDMA algorithm employing the group capacity $f_{\mathrm{CAP}}(\mathcal{G})$ as grouping metric and ESA as grouping algorithm will be named CAP-ESA. The SDMA algorithms and their names are listed in Table 3.4.

Table 3.4 does not contain all the possible SDMA algorithms that could be defined by combining the grouping metrics from Section 3.2.1 and the grouping algorithms from Section 3.2.2, but only those SDMA algorithms in which grouping metric and grouping algorithm can be suitably combined. In the following, the reasons for not combining certain grouping metrics and grouping algorithms are discussed.

The CAP-ESA exhaustively searches for the group that maximizes the sum rate on the considered resource. In this chapter, the main function of the CAP-ESA is to upper

Table 3.4. SDMA algorithms.

| SDMA algorithm | Grouping metric | Grouping algorithm |
|---|---|---|
| **CAP-ESA** | $f_{\mathrm{CAP}}(\mathcal{G})$ | ESA |
| **CC-CGA** | $f_{\mathrm{CC}}(\mathcal{G})$ | CGA |
| **CAP-BFA** | $f_{\mathrm{CAP}}(\mathcal{G})$ | BFA |
| **SP-BFA** | $f_{\mathrm{SP}}(\mathcal{G})$ | |
| **WN-BFA** | $f_{\mathrm{WN}}(\mathcal{G})$ | |
| **CC-BFA** | $f_{\mathrm{CC}}(\mathcal{G})$ | |
| **WN-FFA** | $f_{\mathrm{WN}}(\mathcal{G})$ | FFA |
| **RGA** | - | RGA |

bound the performance of the other SDMA algorithms in terms of the sum rate and, although possible, combinations of other grouping metrics with the ESA would not fulfill this requirement and are consequently unnecessary.

In general, $f_{\mathrm{CAP}}(\mathcal{G})$ and $f_{\mathrm{SP}}(\mathcal{G})$ do not fulfill the constraints in (3.21) and, therefore, cannot be efficiently combined with CGA. In spite of being possible, a combination of $f_{\mathrm{WN}}(\mathcal{G})$ with the CGA would not be in accordance with properties described in Section 3.2.1 and which underlie the design of $f_{\mathrm{WN}}(\mathcal{G})$.

Grouping metrics that necessarily increase when a new MS is added to the SDMA group $\mathcal{G}$, such as $f_{\mathrm{SP}}(\mathcal{G})$ and $f_{\mathrm{CC}}(\mathcal{G})$, cannot be efficiently combined with the FFA, since the first $K_{\mathcal{G}}^{\star}$ MSs would always be selected to build the best SDMA group. Combining $f_{\mathrm{CAP}}(\mathcal{G})$ with FFA is also not recommendable. In this case, FFA would admit an MS to the SDMA group whenever the group capacity increases. Because it is often the case that each MS increases the group capacity at least very slightly, this would lead very often to wrong admissions and correspondingly degrade the performance of the SDMA algorithm.

The RGA does not make use of any grouping metric and its main function is to lower bound the performance of the other SDMA algorithms in terms of the sum rate.

In Table 3.4, the CAP-ESA and the CAP-BFA are the only SDMA algorithms that are aware of the actual precoding and power allocation because the group capacity $f_{\mathrm{CAP}}(\mathcal{G})$ of (3.2) used by these SDMA algorithms depends on the precoding vectors and on the allocated powers. The remaining algorithms of Table 3.4 are consequently unaware of the actual precoding and power allocation since the grouping metrics that they employ do not depend on the precoding vectors or on the allocated powers.

It is also worth noting that the convex combination of the total spatial correlation and channel gains $f_{\mathrm{CC}}(\mathcal{G})$ of (3.13) is to be minimized for the MSs in an SDMA group. Therefore, when combined with the BFA, the best MS to be admitted to the SDMA group in step 2a of Table 3.2 is the one that increases the grouping metric minimally. Anyway, this fact has no impact on the SDMA algorithms, since changing from maximization to minimization in the algorithms is straightforward.

In the following, some remarks are made regarding the novelty of the SDMA algorithms employing the CGA, BFA, and FFA as grouping algorithms. The CC-CGA is a new SDMA algorithm proposed here. As approximation for the SDMA grouping problem, the CC-CGA formulates a quadratic convex optimization problem which depends only on the spatial correlation and channel gains of the MSs and which can be efficiently solved using convex optimization.

The BFA has been proposed in [STKL01], but its combination with the metrics $f_{\mathrm{WN}}(\mathcal{G})$ and $f_{\mathrm{CC}}(\mathcal{G})$ is newly introduced here. The CAP-BFA has been investigated by the author in [MK06] while there were parallel studies in [DS04, DS05, FDH07]. The SP-BFA has been considered, e.g., in [YG05, TUBN06, MSLT07]. The SDMA algorithms employing BFA are considered here as low-complexity heuristic alternatives to solve the SDMA grouping problem and whose computational effort is mainly concentrated on the computation of the grouping metric.

The SDMA algorithm WN-FFA is a suboptimal SDMA algorithm with even lower complexity than those employing ESA, CGA, or BFA as grouping algorithm. The combination of the FFA with $f_{\mathrm{WN}}(\mathcal{G})$ is newly introduced in this thesis.

In the following, some aspects related to the computational efficiency to be taken into account when defining the SDMA algorithm are discussed. For grouping algorithms that admit one MS at a time to the SDMA group, such as the BFA and the FFA, a grouping metric that can be written in a recursive form is preferred, since the decision about the new MS to be admitted can be taken by reusing the metric values previously computed and by calculating only a complementary term associated with the respective new MS. In this way, considerable computational effort can be saved. In general, $f_{\mathrm{CAP}}(\mathcal{G})$ cannot be put in a recursive form due to its dependency on the precoding and power allocation. $f_{\mathrm{SP}}(\mathcal{G})$ is essentially recursive. As it will be shown in the sequel, for $f_{\mathrm{WN}}(\mathcal{G})$ a simple recursive version can be derived, which can be combined with the BFA and the FFA. Similarly, a simple recursive version of $f_{\mathrm{CC}}(\mathcal{G})$ can also be derived and combined with the BFA.

In the following, a recursive form for $f_{\mathrm{WN}}(\mathcal{G})$ is derived. Let $\mathcal{G}$ be the SDMA group containing the $K_{\mathcal{G}}$ MSs already admitted by the BFA or FFA. Let $j \notin \mathcal{G}$ denote the index of the MS that might be admitted to $\mathcal{G}$. Let $\mathcal{G}' = \mathcal{G} \cup \{j\}$ denote the extended SDMA group of size $K'_{\mathcal{G}} = K_{\mathcal{G}} + 1$ obtained by admitting the MS $j$ to $\mathcal{G}$. Using (3.9) and the geometric mean $\mu_{\mathcal{G}}$ associated with the SDMA group $\mathcal{G}$, the geometric mean $\mu_{\mathcal{G}'}$ associated with the extended SDMA group $\mathcal{G}'$ can be written as

$$\mu_{\mathcal{G}'} = \left( \prod_{j \in \mathcal{G}'} \|\hat{\mathbf{g}}_j\|_2 \right)^{1/K_{\mathcal{G}'}} = \left( \|\hat{\mathbf{g}}_j\|_2 \prod_{i \in \mathcal{G}} \|\hat{\mathbf{g}}_i\|_2 \right)^{1/K_{\mathcal{G}'}} = \Big( \underbrace{\|\hat{\mathbf{g}}_j\|_2}_{\ell_1} \mu_{\mathcal{G}}^{K_{\mathcal{G}}} \Big)^{1/K_{\mathcal{G}'}}. \tag{3.22}$$

Let $[\cdot]_{r,c}$ denote the element in the $r^{\mathrm{th}}$ row and $c^{\mathrm{th}}$ column of a matrix. Remembering that the MS $j$ is the unique MSs that belongs to $\mathcal{G}'$ and does not belong to $\mathcal{G}$, the

Frobenius norm in the denominator of (3.10) can be written for the extended SDMA group $\mathcal{G}'$ as

$$\begin{aligned}\|\mathbf{C}_{\mathcal{G}'}\|_{\mathrm{F}} &= \left(\sum_{\tilde{\imath}\in\mathcal{G}'}\sum_{\tilde{\jmath}\in\mathcal{G}'}[\mathbf{C}_{\mathcal{G}'}]^2_{\tilde{\imath},\tilde{\jmath}}\right)^{1/2} = \left(\sum_{\tilde{\imath}\in\mathcal{G}}\sum_{\tilde{\jmath}\in\mathcal{G}}[\mathbf{C}_{\mathcal{G}}]^2_{\tilde{\imath},\tilde{\jmath}} + 1 + 2\sum_{i\in\mathcal{G}}[\mathbf{C}_{\mathcal{G}}]^2_{i,j}\right)^{1/2} \\ &= \left(\|\mathbf{C}_{\mathcal{G}}\|^2_{\mathrm{F}} + 1 + 2\sum_{i\in\mathcal{G}}[\mathbf{C}_{\mathcal{G}}]^2_{i,j}\right)^{1/2} = \Bigg(\|\mathbf{C}_{\mathcal{G}}\|^2_{\mathrm{F}} + \underbrace{1 + 2\sum_{i\in\mathcal{G}}\rho^2_{i,j}}_{\ell_2}\Bigg)^{1/2},\end{aligned} \tag{3.23}$$

and, if $\mu_{\mathcal{G}}$ and $\|\mathbf{C}_{\mathcal{G}}\|_{\mathrm{F}}$ associated with the SDMA group $\mathcal{G}$ are known, only the terms $\ell_1$ in (3.22) and $\ell_2$ in (3.23) need to be computed to determine the value of $f_{\mathrm{WN}}(\mathcal{G}')$.

In the following, a recursive form for $f_{\mathrm{CC}}(\mathcal{G})$ is derived. Let $\mathbf{u}$ of (3.12) denote the binary selection vector associated with the SDMA group $\mathcal{G}$, so that $u_k|_{k=j} = 0$. Using (3.13), $f_{\mathrm{CC}}(\mathcal{G}')$ can be written for the extended SDMA group $\mathcal{G}'$ as

$$\begin{aligned}f_{\mathrm{CC}}(\mathcal{G}') &= \frac{(1-\beta)}{\|\mathbf{C}\|_{\mathrm{F}}}\left(\mathbf{u}^{\mathrm{T}}\mathbf{C}\mathbf{u} + 2\mathbf{u}_K^{\mathrm{T}}\mathbf{C}\mathbf{e}_j + 1\right) + \frac{\beta}{\|\mathbf{a}\|_{\mathrm{F}}}\left(\mathbf{a}^{\mathrm{T}}\mathbf{u} + \mathbf{a}^{\mathrm{T}}\mathbf{e}_j\right) \\ &= f_{\mathrm{CC}}(\mathcal{G}) + \frac{(1-\beta)}{\|\mathbf{C}\|_{\mathrm{F}}}\left(2\mathbf{u}_K^{\mathrm{T}}\mathbf{C}\mathbf{e}_j + 1\right) + \frac{\beta}{\|\mathbf{a}\|_{\mathrm{F}}}\|\hat{\mathbf{h}}_j\|_2^{-2} \\ &= f_{\mathrm{CC}}(\mathcal{G}) + \underbrace{\frac{(1-\beta)}{\|\mathbf{C}\|_{\mathrm{F}}}\left(1 + 2\sum_{i\in\mathcal{G}}\rho_{i,j}\right) + \frac{\beta}{\|\mathbf{a}\|_{\mathrm{F}}}\|\hat{\mathbf{h}}_j\|_2^{-2}}_{\ell_3},\end{aligned} \tag{3.24}$$

and, if $f_{\mathrm{CC}}(\mathcal{G})$ is known, only the term $\ell_3$ in (3.24) needs to be computed to determine $f_{\mathrm{CC}}(\mathcal{G}')$.

In the following, some remarks related to the target SDMA group size $K^\star_{\mathcal{G}}$ and the initial MS $k'$ are done. Both $K^\star_{\mathcal{G}}$ and $k'$ are parameters that restrict the number of candidate SDMA groups to be considered when looking for the best SDMA group $\mathcal{G}^\star$ and they correspond to simplifications introduced into the design of the grouping algorithms, e.g., in order to allow an adequate mathematical formulation or to limit their complexity. Indeed, there exist only $\binom{K}{K^\star_{\mathcal{G}}}$ candidate groups of size $K^\star_{\mathcal{G}}$, which can be a much smaller number than the total number $G$ of SDMA groups given by (2.14).

The group capacity does not monotonically increase with the SDMA group size $K_{\mathcal{G}}$ [MK06, MK07a, MK07c]. Actually, the group capacity depends on both size and composition of the SDMA group, as well as on the precoding and power allocation algorithm employed by the system [Cal04, FDH05, LZ06, MK06, MK07a, MK07c]. The ideal value for the target SDMA group size $K^\star_{\mathcal{G}}$ can not be determined a priori [MK06, FDH07, MK07a]. However, because the larger the SDMA group, the higher the potential spatial multiplexing gains [FDH05], it is a common practice to set $K^\star_{\mathcal{G}}$ to the maximum admissible size, i.e., to set $K^\star_{\mathcal{G}} = M$ [YG05, MK06, Wil06, TUBN06, MK07b, MK08].

For the BFA and FFA of Table 3.2 and Table 3.3, respectively, the target SDMA group size $K_{\mathcal{G}}^{\star}$ corresponds to the largest group size that the grouping algorithms will build, so that groups of smaller size can be obtained, e.g., if there are no more MSs that increase the grouping metric and the condition in step 2b of Table 3.2 or the condition in step 2b of Table 3.3 is not fulfilled. Differently from BFA and FFA, for the CGA and RGA the target SDMA group size $K_{\mathcal{G}}^{\star}$ is definitely the size of the best group $\mathcal{G}^{\star}$.

As discussed in Section 1.2, a posterior adjustment of the SDMA group size by means of an SRA might improve the sum rate of the system. The SRA is indicated in the lowest block of Fig. 1.3 and will be discussed in Section 3.3.3.

The selection of the initial MS $k'$ leaves only $K-1$ MSs to be potentially admitted to an SDMA group and, consequently, also reduces the number of candidate groups considered by a grouping algorithm.

The BFA and FFA require that an initial MS $k'$ be defined. Anyway, for the CGA an initial MS $k'$ can be selected by adding a constraint $\tilde{u}_{k'} = 1$ to problem (3.17), while for the ESA and RGA there is no need to define an initial MS $k'$.

## 3.3 Precoding and Power Allocation Algorithms

### 3.3.1 Precoding Algorithm

In this section, the precoding algorithms shown in the lower half of Fig. 1.3 are discussed. They have the task of efficiently separating in space the signals transmitted to MSs of an SDMA group.

There exist different spatial processing techniques that allow to separate signals sent over the same resource [PNG03]. Nevertheless, in this thesis only linear precoding techniques are going to be considered [SSH04, JUN05], cf. Section 2.3. They are usually simpler to implement and to mathematically deal with, but might not perform as well as non-linear techniques [Qiu04, Joh04, SH05].

In this thesis, two precoding algorithms are considered: the Zero-Forcing (ZF) precoding algorithm [PNG03, SSH04, JUN05] and a Generalized Eigen-Precoding (GEP) algorithm [SB04, BS05, Ger05].

In the following, ZF precoding is described. Using the channel matrix $\hat{\mathbf{G}}$ of the SDMA group $\mathcal{G}$, which has been defined in Section 3.2.1, the DL SINR $\gamma_i$ of an MS $i, i = 1, \ldots, K_{\mathcal{G}}$, in $\mathcal{G}$ can be written, similarly to (2.11), as

$$\gamma_i = \frac{p_i \left\| \hat{\mathbf{g}}_i \mathbf{w}_i \right\|_2^2}{\sigma^2 + \sum\limits_{j=1, j \neq i}^{K_{\mathcal{G}}} p_j \left\| \hat{\mathbf{g}}_i \mathbf{w}_j \right\|_2^2}. \tag{3.25}$$

ZF precoding imposes a ZF constraint on the interference among MSs in the SDMA group $\mathcal{G}$, so that in (3.25) one must have

$$\sum_{j=1,j\neq i}^{K_{\mathcal{G}}} p_j \left\|\hat{\mathbf{g}}_i \mathbf{w}_j\right\|_2^2 = 0, \forall i \in \mathcal{G}. \tag{3.26}$$

The condition in (3.26) implies that the precoding vector $\mathbf{w}_i$ of MS $i$ must belong to the joint null space of the channels of all the other MSs in $\mathcal{G}$ [SSH04]. Let $\tilde{\mathbf{G}}_i \in \mathbb{C}^{(K_{\mathcal{G}}-1)\times M}$ denote a matrix containing the channels of all the MSs $j \neq i, j \in \mathcal{G}$. Let $\tilde{\mathbf{U}}_i \in \mathbb{C}^{(K_{\mathcal{G}}-1)\times(K_{\mathcal{G}}-1)}$, $\tilde{\mathbf{\Lambda}}_i \in \mathbb{R}_+^{(K_{\mathcal{G}}-1)\times M}$ and $\tilde{\mathbf{V}}_i \in \mathbb{C}^{M\times M}$ denote the unitary matrix of the left singular vectors, the matrix of the singular values, and the unitary matrix of the right singular vectors of $\tilde{\mathbf{G}}_i$, respectively, which result from the SVD of $\tilde{\mathbf{G}}_i$. Let $\tilde{r}_i \leq (K_{\mathcal{G}}^{\star}-1)$ denote the rank of $\tilde{\mathbf{G}}_i$. Then, the first $\tilde{r}_i$ columns of $\tilde{\mathbf{V}}_i$ are organized in the matrix $\tilde{\mathbf{V}}_i^{(1)}$ and draw an orthonormal basis to the range space of $\tilde{\mathbf{G}}_i$ [GL96, SSH04]. The last $K_{\mathcal{G}} - \tilde{r}_i$ columns of $\tilde{\mathbf{V}}_i$ are organized in the matrix $\tilde{\mathbf{V}}_i^{(0)}$ and draw an orthonormal basis to the null space of $\tilde{\mathbf{G}}_i$ [GL96, SSH04]. Then, one has

$$\tilde{\mathbf{G}}_i = \left[\hat{\mathbf{g}}_1^{\mathrm{T}}\ \hat{\mathbf{g}}_2^{\mathrm{T}}\ \ldots\ \hat{\mathbf{g}}_{i-1}^{\mathrm{T}}\ \hat{\mathbf{g}}_{i+1}^{\mathrm{T}}\ \ldots\ \hat{\mathbf{g}}_{K_{\mathcal{G}}^{\star}-1}^{\mathrm{T}}\ \hat{\mathbf{g}}_{K_{\mathcal{G}}^{\star}}^{\mathrm{T}}\right]^{\mathrm{T}} = \tilde{\mathbf{U}}_i \tilde{\mathbf{\Lambda}}_i \tilde{\mathbf{V}}_i^{\mathrm{H}} = \tilde{\mathbf{U}}_i \tilde{\mathbf{\Lambda}}_i \left[\tilde{\mathbf{V}}_i^{(1)}\ \tilde{\mathbf{V}}_i^{(0)}\right]^{\mathrm{H}}, \tag{3.27}$$

and in order to fulfill (3.26), $\mathbf{w}_i$ must belong to the space spanned by the columns of $\tilde{\mathbf{V}}_i^{(0)}$ [SSH04]. Note that, if $\hat{\mathbf{G}}$ is full-rank, the null space of $\tilde{\mathbf{G}}_i$ has dimension one, i.e., $(K_{\mathcal{G}} - \tilde{r}_i) = 1, \forall i$, and $\mathbf{w}_i$ is uniquely determined and equal to the last column of $\tilde{\mathbf{V}}_i$. With the pseudo-inverse $\hat{\mathbf{G}}^{\dagger}$ of $\hat{\mathbf{G}}$ given by

$$\hat{\mathbf{G}}^{\dagger} = \hat{\mathbf{G}}^{\mathrm{H}}\left(\hat{\mathbf{G}}\hat{\mathbf{G}}^{\mathrm{H}}\right)^{-1} \tag{3.28}$$

[GL96, PNG03] and denoting by $\hat{\mathbf{g}}_i^{\dagger}$ the $i^{\text{th}}$ column of $\hat{\mathbf{G}}^{\dagger}$, the precoding vector $\mathbf{w}_i$ can be more simply obtained as

$$\mathbf{w}_i = \frac{\hat{\mathbf{g}}_i^{\dagger}}{\|\hat{\mathbf{g}}_i^{\dagger}\|_2} \tag{3.29}$$

[PNG03]. The precoding vectors $\mathbf{w}_i$ given by (3.29) are considered in this work to perform ZF precoding. Note that, by defining the precoding matrix $\mathbf{W}$ as

$$\mathbf{W} = \left[\mathbf{w}_1\ \mathbf{w}_2\ \ldots\ \mathbf{w}_{K_{\mathcal{G}}^{\star}}\right], \tag{3.30}$$

the equivalent channel given by the product $\hat{\mathbf{G}}\mathbf{W}$ is a diagonal matrix, i.e.,

$$\hat{\mathbf{G}}\mathbf{W} = \begin{bmatrix} \hat{\mathbf{g}}_1 \\ \hat{\mathbf{g}}_2 \\ \vdots \\ \hat{\mathbf{g}}_{K_{\mathcal{G}}^{\star}} \end{bmatrix} \left[\mathbf{w}_1\ \mathbf{w}_2\ \ldots\ \mathbf{w}_{K_{\mathcal{G}}^{\star}}\right] = \begin{bmatrix} \hat{\mathbf{g}}_1\mathbf{w}_1 & 0 & \ldots & 0 \\ 0 & \hat{\mathbf{g}}_2\mathbf{w}_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \hat{\mathbf{g}}_{K_{\mathcal{G}}^{\star}}\mathbf{w}_{K_{\mathcal{G}}^{\star}} \end{bmatrix}, \tag{3.31}$$

and fulfills the ZF condition in (3.26). An interesting characteristic of ZF precoding is that it fully decouples the MU MIMO channel matrix $\hat{\mathbf{G}}$ into $K_{\mathcal{G}}$ effective SISO channels. Because the effective channels of the MSs become independent, the power allocated to one MS in the SDMA group $\mathcal{G}$ has no counter effect in terms of interference to the other MSs in $\mathcal{G}$, and the precoding algorithm and the power allocation algorithm are fully separated of each other. Nevertheless, precoding algorithms that do not diagonalize $\hat{\mathbf{G}}$ are also relevant and may outperform ZF precoding in terms of sum rate [PNG03, SPSH04, MBQ04, JUN05].

In the following, the GEP is discussed [Zet95, Zet99, SBO06]. The GEP is an algorithm that maximizes the SINR of the MSs in an SDMA group $\mathcal{G}$ for a fixed power allocation by solving a set of generalized eigenproblems. The GEP formulation considered here is based on the duality between the UL MAC and DL BC MIMO channels [VT03, VJG03, SB04, BS05, Ger05], which is more general. Anyway, the formulation considered here is structurally equivalent to that in [Zet95, Zet99, SBO06] by means of a simple change of parameters.

According to the duality theory between MAC and BC MIMO channels, precoding vectors optimal for the dual UL problem remain optimal for the DL, so that the same sum rate can be achieved in both the dual UL and the actual DL channels using the same precoding vectors and an adequate power allocation [VT03, VJG03, SB04, BS05]. The dual UL channel corresponds to the actual UL channel, if the actual UL and DL channels are reciprocal. Otherwise, the dual UL channel is a sort of virtual UL channel associated with the actual DL channel. More details about the duality theory between MAC and BC MIMO channels can be found, e.g., in [VT03, VJG03, SB04, BS05]. Further on, the term dual will be omitted.

Let $q_i$ denote the UL power allocated to MSs $i, i = 1, \ldots, K_{\mathcal{G}}$, in SDMA group $\mathcal{G}$. Then, the UL SINR $\gamma_i'$ of MS $i$ in $\mathcal{G}$ can be written as

$$\gamma_i' = \frac{q_i \mathbf{w}_i^{\mathrm{H}} \hat{\mathbf{g}}_i^{\mathrm{H}} \hat{\mathbf{g}}_i \mathbf{w}_i}{\mathbf{w}_i^{\mathrm{H}} \left( \sigma^2 \mathbf{I}_{K_{\mathcal{G}}^\star} + \sum\limits_{j=1, j\neq i}^{K_{\mathcal{G}}^\star} q_j \hat{\mathbf{g}}_j^{\mathrm{H}} \hat{\mathbf{g}}_j \right) \mathbf{w}_i}. \tag{3.32}$$

[SB04, BS05]. Comparing (3.25) and (3.32), it can be noted that the precoding vector $\mathbf{w}_i$ of an MS $i$ that achieves a given DL SINR depends on the precoding vectors $\mathbf{w}_j$ of MS $j \neq i$, while in the UL this dependency does not exist, which in many cases simplifies the determination of optimal precoding vectors for DL problems by solving equivalent UL problems [VT03, VJG03, SB04, BS05].

In the following, (3.32) is used to determine the precoding vectors $\mathbf{w}_i$ of the GEP, i.e., the precoding vectors $\mathbf{w}_i^\star$ that maximize (3.32) for each MS $i$ given a fixed UL power allocation $q_i$, with $i = 1, \ldots, K_{\mathcal{G}}$. The optimal precoding vectors $\mathbf{w}_i^\star$ for the GEP are the solution of the optimization problem

$$\mathbf{w}_i^\star = \arg\max_{\mathbf{w}} \left\{ \frac{q_i \mathbf{w}^{\mathrm{H}} \tilde{\mathbf{R}}_i \mathbf{w}_i}{\mathbf{w}_i^{\mathrm{H}} \tilde{\mathbf{Q}}_i \mathbf{w}_i} \right\}, \tag{3.33a}$$

with

$$\tilde{\mathbf{R}}_i = \hat{\mathbf{g}}_i^{\mathrm{H}} \hat{\mathbf{g}}_i, \tag{3.33b}$$

and

$$\tilde{\mathbf{Q}}_i = \sigma^2 \mathbf{I}_{K_{\mathcal{G}}^\star} + \sum_{j=1, j\neq i}^{K_{\mathcal{G}}^\star} q_j \hat{\mathbf{g}}_j^{\mathrm{H}} \hat{\mathbf{g}}_j = \sigma^2 \mathbf{I}_{K_{\mathcal{G}}^\star} + \sum_{j=1, j\neq i}^{K_{\mathcal{G}}^\star} q_j \tilde{\mathbf{R}}_j. \tag{3.33c}$$

In [Ger05], it is shown that solving a problem of the form (3.33) is equivalent to maintaining constant the distortionless response to the desired signal in the numerator of (3.33a) while minimizing the interference plus noise in the denominator of (3.33a), i.e., to solve

$$\mathbf{w}_i^\star = \arg\min_{\mathbf{w}} \left\{ \mathbf{w}_i^{\mathrm{H}} \tilde{\mathbf{Q}}_i \mathbf{w}_i \right\}, \tag{3.34a}$$

$$\text{subject to: } \mathbf{w}^{\mathrm{H}} \tilde{\mathbf{R}}_i \mathbf{w}_i = 1. \tag{3.34b}$$

Taking the first Karush-Kuhn-Tucker (KKT) condition [BV04] for the problem (3.34), i.e., by setting the derivative $\frac{\partial \mathcal{L}(\mathbf{w}_i, \lambda_i)}{\partial \mathbf{w}_i}$ of the Lagrangian function $\mathcal{L}(\mathbf{w}_i, \lambda_i)$ of (3.34) equal to zero, where $\lambda_i$ is the Lagrange multiplier, the generalized eigenvalue problem

$$\frac{\partial \mathcal{L}(\mathbf{w}_i, \lambda_i)}{\partial \mathbf{w}_i} = 0 \quad \Leftrightarrow \quad \tilde{\mathbf{Q}}_i \mathbf{w}_i - \lambda_i \tilde{\mathbf{R}}_i \mathbf{w}_i = 0 \quad \Leftrightarrow \quad \boxed{\tilde{\mathbf{Q}}_i \mathbf{w}_i = \lambda_i \tilde{\mathbf{R}}_i \mathbf{w}_i} \tag{3.35}$$

is obtained. According to [Ger05, SB04, BS05], the optimal precoding vector $\mathbf{w}_i^\star$ that solves (3.33) corresponds to the dominant eigenvector of the generalized eigenvalue problem (3.35). Consequently, the precoding vectors of the GEP algorithm are found by determining the dominant eigenvector of (3.35) for each MS individually.

In the following, two remarks regarding the GEP presented in this section are provided. The problem formulation in (3.33) is different of the GEP originally formulated in [Zet95, Zet99, SBO06]. The differences lie on the fact that in [Zet95, Zet99, SBO06] $q_i = 1, \forall i$, $\mathbf{R}_i$ of (2.6) is employed instead of $\tilde{\mathbf{R}}_i$ of (3.33b), and a scaling factor $\tilde{\sigma}$ is considered instead of $\sigma^2$. Since in (3.33), $\mathbf{w}_i$ is the optimization variable and $\tilde{\mathbf{R}}_i$, $q_i$, and $\sigma^2$ are only the problem data [BV04], problem (3.33) matches the formulation in [Zet95, Zet99, SBO06] by means of a simple change of parameters. Thus, the formulation in [Zet95, Zet99, SBO06] can be seen as particular instance of problem (3.33).

The GEP of (3.33) is based on the framework proposed in [SB04, BS05]. Differently from the ZF precoding, the GEP does not diagonalize the group channel matrix $\hat{\mathbf{G}}$ and, in spite of being independent of each other, the optimal precoding vectors $\mathbf{w}_i^\star$ of MS $i$ still depend on the allocated power $q_j$ of the MSs $j \neq i$. According to [SB04, BS05], this dependency suggests an iterative alternate optimization of precoding vectors and power allocation and, consequently, a joint optimization of precoding and power allocation. A joint precoding and power allocation algorithm for a weighted SINR balancing problem has been investigated in [SB04, BS05]. Later in Section 3.3.4, a joint precoding and power allocation algorithm employing the GEP will be discussed.

Considering multi-antenna terminals, a comprehensive set of linear and non-linear precoding algorithms can be found in [Sta06]. In particular, generalizations of ZF and Minimum Mean Square Error (MMSE) precoding can be found, e.g., in [SSH04, SH08]. The precoding algorithms in [Sta06] fully cope with receive antenna cooperation at the MSs and can be efficiently employed in RA strategies to maximize the sum rate [FDH05, FDH07].

### 3.3.2 Power Allocation Algorithm

In this section, the two power allocation algorithms considered in this thesis are described, namely, the Water Filling Algorithm (WFA) and the Soft Dropping Algorithm (SDA).

The WFA is a well-known iterative solution to the convex problem of maximizing the sum rate of a set of independent channels [PNG03, BV04, PF05, CT06]. The WFA finds application in many multi-channel systems, e.g., for subcarrier power allocation in SISO OFDMA systems [Liu04] or for power allocation to MSs on the spatial layers of one or more resources [SSH04, Cal04, DS05, MK06, TUBN06, FDH07, MK07a]. This latter is the case of interest in this chapter, which considers a single resource.

Let the power allocation vector $\mathbf{p}$ containing the power $p_i$ allocated to each MS $i, i = 1, \ldots, K_{\mathcal{G}}$, in the SDMA group $\mathcal{G}$ be defined as

$$\mathbf{p} = \begin{bmatrix} p_1 & p_2 & \ldots & p_{K_{\mathcal{G}}} \end{bmatrix}^{\mathrm{H}}. \tag{3.36}$$

The WFA corresponds to solving the convex optimization problem

$$\mathbf{p}^{\star} = \arg\max_{\{\mathbf{p}\}} \left\{ \sum_{i=1}^{K_{\mathcal{G}}^{\star}} \log_2 \left( 1 + p_i \frac{\|\hat{\mathbf{g}}_i \mathbf{w}_i\|_2^2}{\sigma^2} \right) \right\} \tag{3.37a}$$

$$\text{subject to: } p_i \geq 0, \forall i \in \mathcal{G} \tag{3.37b}$$

$$\sum_{i=1}^{K_{\mathcal{G}}} p_i = P_b \tag{3.37c}$$

[PNG03, BV04, PF05, CT06], which is accomplished by employing the iterative algorithm of Table 3.5 [PF05].

In the following, a new SDA is described. The SDA is an iterative power allocation algorithm derived from the SDPC algorithm of [YGRS97], which performs distributed Power Control (PC) among co-channel links in a multi-cell SISO system. The SDA can

Table 3.5. Water Filling Algorithm.

1. Sort the MSs in $\mathcal{G}$ so that $\|\hat{\mathbf{g}}_1\mathbf{w}_1\|_2^2 \geq \|\hat{\mathbf{g}}_2\mathbf{w}_2\|_2^2 \geq \ldots \geq \left\|\hat{\mathbf{g}}_{K_\mathcal{G}}\mathbf{w}_{K_\mathcal{G}}\right\|_2^2$.
2. For $\tilde{K} = K_\mathcal{G}$ to 1,

   If $\left(\frac{\sigma^2}{\left\|\hat{\mathbf{g}}_{\tilde{K}}\mathbf{w}_{\tilde{K}}\right\|_2^2}\right)\tilde{K} - \left(P_b + \sum_{i=1}^{\tilde{K}} \frac{\sigma^2}{\|\hat{\mathbf{g}}_i\mathbf{w}_i\|_2^2}\right) < 0$, go to step 3.
3. Compute the water level $\mu = \frac{1}{\tilde{K}}\left(P_b + \sum_{i=1}^{\tilde{K}} \frac{\sigma^2}{\|\hat{\mathbf{g}}_i\mathbf{w}_i\|_2^2}\right)$.
4. For $i = 1$ to $K_\mathcal{G}$

   If $i \leq \tilde{K}$, set $p_i = \mu - \frac{\sigma^2}{\|\hat{\mathbf{g}}_i\mathbf{w}_i\|_2^2}$, else set $p_i = 0$.
5. Sort the MSs back to their original order.

be seen as an extension of the SDPC to a MU MIMO system and includes an initial power allocation and takes into account the total power constraint of (3.1c). In this case, differently from [YGRS97], the interferers are no longer located in other cells, but correspond to the MSs which belong to the same SDMA group.

In order to allocate power to the MSs in an SDMA group, the SDA considers a power-dependent target SINR which decreases with the amount of power allocated to an MS, so that an MS with bad channel condition and which requires more power will target at lower SINR values than another MS whose channel condition is good. Consequently, a self-regulation of the target SINR of the MS takes place and a margin to a more efficient power usage among several MSs is created.

In the following, the power-dependent target SINR $\bar{\gamma}_i$ and the power allocation of SDA are described. Let $\check{\gamma}$ and $\hat{\gamma}$ denote the minimum and maximum target SINR that an MS can aim at, respectively, and let $\check{\Gamma}$ and $\hat{\Gamma}$ be their respective values in dB. Similarly, let $\check{p}$ and $\hat{p}$ denote the minimum and maximum power that can be allocated to an MS, respectively, and let $\check{P}$ and $\hat{P}$ be their corresponding values in dBW. The parameters $\check{\gamma}$, $\hat{\gamma}$, $\check{p}$, and $\hat{p}$ might be adjusted for each MS individually. However, the analysis in this work is limited to the case in which $\check{\gamma}$, $\hat{\gamma}$, $\check{p}$, and $\hat{p}$ are the same for all MSs. For the SDA, the power-dependent target SINR $\bar{\Gamma}_i$ is defined as illustrated in Fig. 3.3 [YGRS97].

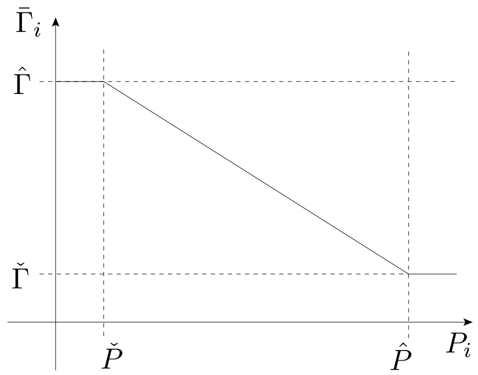



Figure 3.3. Power-dependent target SINR $\bar{\Gamma}_i$ for SDA.

Let $\tilde{t}$ denote the iteration of the SDA, which should not be confused with the time $t$. At iteration $\tilde{t}$, let $\bar{\gamma}_i^{(\tilde{t})}$ and $\bar{\Gamma}_i^{(\tilde{t})}$ denote the target SINR of MS $i$ in linear and in dB scale, respectively, and let $p_i^{(\tilde{t})}$ and $P_i^{(\tilde{t})}$ denote the power allocated to MS $i$ in W and in dBW, respectively. Then, according to Fig. 3.3, the target SINRs $\bar{\Gamma}_i^{(\tilde{t})}$ and $\bar{\gamma}_i^{(\tilde{t})}$ for the SDA can be written as

$$\begin{aligned} \bar{\Gamma}_i^{(\tilde{t}+1)} &= \min\left\{\max\left\{\left(\frac{\hat{\Gamma}-\check{\Gamma}}{\check{P}-\hat{P}}\right)P_i^{(\tilde{t})}+\check{\Gamma},\check{\Gamma}\right\},\hat{\Gamma}\right\} \quad \Leftrightarrow \\ \bar{\gamma}_i^{(\tilde{t}+1)} &= \min\left\{\max\left\{\check{\gamma}\left(\frac{p_i^{(\tilde{t})}}{\hat{p}}\right)^{\zeta},\check{\gamma}\right\},\hat{\gamma}\right\}, \text{ with } \zeta=\frac{\log_{10}(\hat{\gamma}/\check{\gamma})}{\log_{10}(\check{p}/\hat{p})} \end{aligned} \tag{3.38}$$

[MK07c], while the power allocated to MS $i$ at iteration $\tilde{t}$ [MK07a] is given similarly by

$$P_i^{(\tilde{t}+1)} = P_i^{(\tilde{t})} + \xi(\bar{\Gamma}_i^{(\tilde{t})} - \Gamma_i^{(\tilde{t})}) \quad \Leftrightarrow \quad p_i^{(\tilde{t}+1)} = p_i^{(\tilde{t})}\left(\frac{\bar{\gamma}_i^{(\tilde{t})}}{\gamma_i^{(\tilde{t})}}\right)^{\xi}, \tag{3.39}$$

where $\xi$ is a feedback parameter controlling the fraction of the difference between the target and the current SINRs that should be compensated at each iteration of the SDA [AEW94, ZKAQ01, Gun00, PED04, MK07c].

In the following, the SDA is presented. Using (3.36), let $\mathbf{p}^{(\tilde{t})}$ denote the power allocation vector at iteration $\tilde{t}$. Starting from an initial power allocation vector $\mathbf{p}^{(0)}$, the current and the target SINRs of each MS can be calculated using (3.25) and (3.38), respectively. For $\mathbf{p}^{(0)}$, an EPA among the MSs in $\mathcal{G}$ is a good choice since it is fair and does not bias the power allocation to any MS in particular. Then, each component $p_i^{(\tilde{t})}$ of the power vector $\mathbf{p}^{(\tilde{t})}$ can be iteratively updated according to (3.39) until the power allocation vector $\mathbf{p}^{(\tilde{t})}$ converges to an optimal power allocation vector $\mathbf{p}^\star$ to a specified precision $\epsilon_\mathrm{p}$. Using these definitions, the SDA is presented in Table 3.6.

Table 3.6. Soft Dropping Algorithm.

1. Set $\tilde{t}=0$ and $\mathbf{p}^{(\tilde{t})}=\mathbf{p}^{(0)}=(P_b/K_\mathcal{G})\mathbf{1}_{K_\mathcal{G}}$.
2. Set $\tilde{t}=\tilde{t}+1$.
3. For $i=1$ to $K_\mathcal{G}$, compute $\mathbf{w}_i^{(\tilde{t}-1)}$ using (3.33).
4. For $i=1$ to $K_\mathcal{G}$, compute $\gamma_i^{(\tilde{t}-1)}$ using (3.25).
5. For $i=1$ to $K_\mathcal{G}$, compute $\bar{\gamma}_i^{(\tilde{t})}$ using (3.38).
6. For $i=1$ to $K_\mathcal{G}$, update the power $p_i^{(\tilde{t})}$ using (3.39).
7. Set $\mathbf{p}^{(\tilde{t})}=\frac{\mathbf{p}^{(\tilde{t})}}{\|\mathbf{p}^{(\tilde{t})}\|_1}P_b$ in order to fulfill the constraint (3.1c).
8. If $\min\left\{\left|p_i^{(\tilde{t})}-p_i^{(\tilde{t}-1)}\right|,\ldots,\left|p_{K_\mathcal{G}}^{(\tilde{t})}-p_{K_\mathcal{G}}^{(\tilde{t}-1)}\right|\right\}>\epsilon_\mathrm{p}$, go to step 2, otherwise stop.

Note that the normalization of $\mathbf{p}^{(t)}$ in step 7 of Table 3.6 is required in order to fulfill the constraint (3.1c). Furthermore, note that by adequately setting the parameters $\check{\gamma}$, $\hat{\gamma}$, $\check{p}$ $\hat{p}$, the target SINRs aimed at the MSs can be varied and different sum rate values can be achieved in the system. In fact, as it has been discussed by the author in [MK07c], an SINR balancing as well as the maximization of the sum rate of the system might be pursued by the SDA through an adequate setting of the parameters $\check{\gamma}$, $\hat{\gamma}$, $\check{p}$, and $\hat{p}$. Anyway, the parameter settings considered in this work will aim only at maximizing the sum rate of the system.

### 3.3.3 Sequential Removal Algorithm

In this section, the Sequential Removal Algorithm (SRA) is described, which is responsible for adjusting the size of the SDMA groups as to enhance the sum rate of the system. Some of the SDMA algorithms introduced in Section 3.2.3 are unaware of the actual precoding and power allocation and, consequently, these SDMA algorithms might build groups containing MSs that do not contribute to enhance the sum rate of the system, as discussed in Section 1.2. Consider for example the case in which ZF precoding and the WFA are used. Whenever the WFA allocates null power to an MS, this MS does not contribute to enhance the group capacity anymore. On the contrary, since the channels of the others MSs in the SDMA group are projected onto the null space of the channel of this one MS, its removal from the group can only improve the group capacity [Cal04, MK06].

In order to decide which MSs should be removed from the SDMA group by the SRA, the two following criteria are considered:

1. The Lowest Gain First (LGF), in which the MS in the group with the lowest effective channel gain is removed, i.e., the MS $i^\star$ to be removed is defined as

$$i^\star = \arg\min_{i \in \mathcal{G}} \left\{ \|\hat{\mathbf{g}}_i \mathbf{w}_i\|_2^2 \right\}. \tag{3.40}$$

2. The Highest Power First (HPF), in which the MS demanding the highest amount of power is removed, i.e., the MS $i^\star$ to be removed is defined as

$$i^\star = \arg\max_{i \in \mathcal{G}} \left\{ p_i \right\}. \tag{3.41}$$

The LGF criterion is reasonable, since the lower the effective channel gain of an MS in an SDMA group is, the lower its achievable capacity is. Thus, removing the MS with the lowest channel gain might improve the performance of the other MSs in the group. This criterion has been adopted, e.g., in [Cal04, MK06].

The HPF criterion is also a reasonable criterion, since the worse the channel condition of an MS in an SDMA group is, e.g., in terms of channel gain, the more power the MS requires to attain a certain SINR. Thus, removing the MS requiring the largest amount of power makes available a considerable amount of power that might be used to improve the performance of the other MSs in the SDMA group. This criterion has been adopted, e.g., in [SBO06, MK07c].

In the following, the SRA is presented. It removes one MS from the SDMA group $\mathcal{G}$ according to the LGF or the HPF criterion. Then, using (3.2), the SRA computes and stores the capacity for the resulting SDMA group taking into account the actual precoding and power allocation algorithms. Then, the process is repeated and another MS is removed, and so on. At the end, the SDMA group with the highest capacity is kept as the best SDMA group $\mathcal{G}^\star$. The proposed SRA is presented in Table 3.7.

Table 3.7. Sequential Removal Algorithm.

1. Set $\mathcal{G} = \mathcal{G}^\star$ and compute the group capacity $f_{\mathrm{CAP}}(\mathcal{G})$ using (3.2).
2. While the size $K_\mathcal{G}$ of $\mathcal{G}$ is greater than or equal to one
   (a) Determine the MS $i^\star$ using (3.40) or (3.41) and remove it from $\mathcal{G}$, i.e., set $\mathcal{G} = \mathcal{G} \setminus \{i^\star\}$.
   (b) Compute the group capacity $f_{\mathrm{CAP}}(\mathcal{G})$ of $\mathcal{G}$ using (3.2).
   (c) If $f_{\mathrm{CAP}}(\mathcal{G}) > f_{\mathrm{CAP}}(\mathcal{G}^\star)$, define $\mathcal{G}^\star = \mathcal{G}$.

Note that, for an SDMA group $\mathcal{G}$ of size $K_\mathcal{G}$, the SRA needs to compute $K_\mathcal{G}$ group capacities using (3.2). However, because $K_\mathcal{G}$ is relatively small and because the size of the SDMA group $\mathcal{G}$ is sequentially reduced, these computations add only slightly to the complexity of the RA strategies that consider the SRA. Nevertheless, the SRA can provide considerable gains to the system in terms of sum rate, as it has been shown by the author in [MK06, MK07a].

### 3.3.4 Precoding and Power Allocation Algorithm Definition

In this section, the precoding and power allocation algorithms are defined, which combine a precoding algorithm from Section 3.3.1, a power allocation algorithm from Section 3.3.2 and, if adequate, the SRA from Section 3.3.3 with a suitable MS removal criterion. Furthermore, in this section iterative joint precoding and power allocation is also approached for the algorithms in which precoding vectors and the allocated powers present an interdependency. In this chapter, the precoding and power allocation algorithms listed in Table 3.8 are considered.

Table 3.8. Precoding and power allocation algorithms.

| **Algorithm** | **Precoding** | **Power allocation** | **SRA criterion** |
|---|---|---|---|
| **ZF** | ZF | WFA | LGF |
| **GEP** | GEP | SDA | HPF |

As it can be seen in Table 3.8, the precoding and power allocation algorithms are named only after the precoding algorithm since a single power allocation algorithm and a single removal criterion for the SRA are associated with each precoding algorithm. Therefore, the term ZF will be used to indicate either the ZF precoding algorithm defined in Section 3.3.1, or the combination of ZF precoding, WFA, and SRA using the LGF criterion, cf. Table 3.8. Analogously, the term GEP might also mean GEP algorithm from Section 3.3.1 or the its combination with the SDA and the SRA using the HPF criterion, cf. Table 3.8. Nevertheless, whenever required by the context, a clear distinction about these cases will be explicitly made.

In the following, the combinations of the precoding algorithm, power allocation algorithm, and the SRA criteria are briefly justified. The combination of ZF precoding and WFA is justified by the fact that the WFA is designed for parallel channels, which are obtained with ZF precoding as effective channels, cf. (3.31). If the employed precoding algorithm does not decouple the MIMO channel matrix into a set of independent SISO channels, the WFA will no longer maximize the sum rate in problem (3.37) since interference among MSs sharing a same resource in space takes place and the assumption of having parallel independent channels does not hold anymore. Because the WFA is employed, the LGF criterion for the SRA is a suitable choice, since the MS with the lowest effective channel gains are the ones that will most probably get no power [Cal04, MK06].

The combination of GEP and SDA is justified by the fact that there is a dependency between precoding vectors and allocated powers by the GEP, cf. Section 3.3.1, which can be iteratively handled by the SDA in order to attain a suitable trade-off between the achieved sum rate, the intra-cell interference, and the efficient usage of the available power. Nevertheless, the SDA might be combined with other precoding algorithms. Moreover, because the SDA is employed, the HPF criterion for the SRA is a suitable choice, since the MS demanding the highest amount of power is probably the one with the worst SINR and its removal will make available the highest amount of power, which might be more efficiently used by the other MS in the SDMA group [MK07c].

In fact, the combination of GEP and SDA results into a joint optimization of precoding and power allocation which follows the iterative alternate optimization framework of [SB04, BS05]. In the following, the joint optimization of precoding and power allocation considering the GEP and the SDA is described. Similarly to [SB04, BS05], the precoding and power allocation problems are solved in the UL and a power allocation for the DL is obtained afterwards. The new joint precoding and power allocation algorithm considers target SINRs that dynamically vary according to (3.38) and, consequently, is more flexible than the algorithms in [SBO06, SB04, BS05].

Initially a convergence proof for the SDA in UL is provided. Then, the power allocation in the DL is determined from the power allocation in the UL and, finally, the joint precoding and power allocation algorithm combining the GEP and SDA is presented.

In the following, the required conditions for the convergence of the SDA in the UL are newly provided. Because the SDA is iterative, its convergence acquires an important role. Analogously to (3.36), let the power allocation vector $\mathbf{q}$ containing the power $q_i$ allocated in the UL to each MS $i, i = 1, \ldots, K_{\mathcal{G}}$, in the SDMA group $\mathcal{G}$ be defined as

$$\mathbf{q} = \begin{bmatrix} q_1 & q_2 & \ldots & q_{K_{\mathcal{G}}} \end{bmatrix}^{\mathrm{H}} \tag{3.42}$$

In [Yat95], a framework for UL PC has been proposed, which shows that a power iteration of the form

$$\mathbf{q}^{(\tilde{t}+1)} = \mathcal{I}(\mathbf{q}^{(\tilde{t})}) \tag{3.43}$$

always converges to the optimal power vector $\mathbf{q}^{\star}$ whenever $\mathcal{I}(\mathbf{q})$ is a standard interference function, i.e., whenever $\mathcal{I}(\mathbf{q})$ satisfies the following properties for any $\mathbf{q}$ with $q_i \geq 0$ and $i = 1, \ldots, K_{\mathcal{G}}$:

1. Positivity: $\mathcal{I}(\mathbf{q}) \geq 0$.
2. Monotonicity: $\mathcal{I}(\mathbf{q}) \geq \mathcal{I}(\mathbf{q}')$, for $\mathbf{q} \geq \mathbf{q}'$, i.e., $q_i \geq q_i', \forall i$.
3. Scalability: $a\mathcal{I}(\mathbf{q}) \geq \mathcal{I}(a\mathbf{q})$, for $a \geq 1$.

In [PED04], the SDPC of [YGRS97] is extended to the UL of a SU MIMO system and is shown to be a standard interference function in this scenario. Because the SDA is derived from the SDPC, the convergence proof provided in [PED04] can be adapted to the SDA, thus extending the results in [PED04] to the MU MIMO case. In fact, this extension has been done by the author in [MK07c], where the joint optimization of precoding and power allocation of [SB04, BS05] has been adapted to employ the SDA. Moreover, in [MK07c] it has also been shown that the proposed algorithm combining the GEP and SDA converges after a few iterations, just like the weighted SINR balancing algorithm of [SB04, BS05].

The proof of the three properties of standard interference functions for SDA are provided in the sequel. These proofs are very similar to that provided in [PED04] for the UL of a SU MIMO channel, since no large structural differences between the SU MIMO and the MU MIMO channels exist. Let $\mathbf{W}^{(\tilde{t})}$ be the precoding matrix at iteration $\tilde{t}$ of the SDA. $\mathbf{W}^{(\tilde{t})}$ has the same form of (3.30), however each column $\mathbf{w}_i^{(\tilde{t})}$ of $\mathbf{W}^{(\tilde{t})}$ is obtained according to the GEP by solving (3.33) for the $i^{\text{th}}$ MS considering the power allocation vector $\mathbf{q}^{(\tilde{t})}$. Therefore, using (3.32), (3.33), (3.38), and (3.39), one has

$$q_i^{(\tilde{t}+1)} = \mathcal{I}(q_i^{(\tilde{t})}) = q_i^{(\tilde{t})} \left[ \min\left\{ \max\left\{ \check{\gamma} \left( \frac{q_i^{(\tilde{t})}}{\hat{p}} \right)^{\zeta}, \check{\gamma} \right\}, \hat{\gamma} \right\} \left( \frac{\mathbf{w}_i^{(\tilde{t}),\mathrm{H}} \tilde{\mathbf{Q}}_i^{(\tilde{t})} \mathbf{w}_i^{(\tilde{t})}}{q_i^{(\tilde{t})} \mathbf{w}_i^{(\tilde{t}),\mathrm{H}} \tilde{\mathbf{R}}_i^{(\tilde{t})} \mathbf{w}_i^{(\tilde{t})}} \right) \right]^{\xi}. \tag{3.44}$$

In spite of being possible to put (3.44) in matrix form, it cannot be solved as a linear system, as done for the weighted SINR balancing problem in [SB04, BS05], since (3.44) is non-linear in $q_i$. Anyway, the standard interference properties can be demonstrated component-wise as follows.

Positivity for $\mathcal{I}(q_i^{(\tilde{t})})$ in (3.45) follows directly, since all involved terms are positive for all $\bar{\gamma}_i$.

Using (3.44), for $\check{\gamma} \leq \bar{\gamma}_i \leq \hat{\gamma}$ one has

$$\mathcal{I}(q_i^{(\tilde{t})}) = q_i^{(\tilde{t})} \left[ \check{\gamma} \left( \frac{q_i^{(\tilde{t})}}{\hat{p}} \right)^{\zeta} \left( \frac{\mathbf{w}_i^{(\tilde{t}),\mathrm{H}} \tilde{\mathbf{Q}}_i^{(\tilde{t})} \mathbf{w}_i^{(\tilde{t})}}{q_i^{(\tilde{t})} \mathbf{w}_i^{(\tilde{t}),\mathrm{H}} \tilde{\mathbf{R}}_i^{(\tilde{t})} \mathbf{w}_i^{(\tilde{t})}} \right) \right]^{\xi} = \left( q^{(\tilde{t})} \right)^{1+\zeta\xi-\xi} \left[ \frac{\check{\gamma} \mathbf{w}_i^{(\tilde{t}),\mathrm{H}} \tilde{\mathbf{Q}}_i^{(\tilde{t})} \mathbf{w}_i^{(\tilde{t})}}{\hat{p}^{\zeta} \mathbf{w}_i^{(\tilde{t}),\mathrm{H}} \tilde{\mathbf{R}}_i^{(\tilde{t})} \mathbf{w}_i^{(\tilde{t})}} \right]^{\xi}$$
$$= \left( q^{(\tilde{t})} \right)^{1+\zeta\xi-\xi} I_i(\mathbf{W}^{(\tilde{t})}, \mathbf{q}^{(\tilde{t})}). \tag{3.45}$$

For each iteration $\tilde{t}$ of the power allocation in (3.45), the precoding matrix $\mathbf{W}^{(\tilde{t})}$ is constant, and vice-versa. Omitting the iteration index $\tilde{t}$ and considering that the powers $q_j$ of all the MSs $j \neq i$ are constant, monotonicity is obtained for $q_i \geq q_i'$ if

$$\begin{aligned} &\mathcal{I}(q_i) \geq \mathcal{I}(q_i') \Leftrightarrow q_i^{1+\zeta\xi-\xi} I_i(\mathbf{W}, \mathbf{q}) \geq q_i'^{1+\zeta\xi-\xi} I_i'(\mathbf{W}, \mathbf{q}) \Leftrightarrow q_i^{1+\zeta\xi-\xi} \geq q_i'^{1+\zeta\xi-\xi} \Leftrightarrow \\ &\Leftrightarrow 1 + \zeta\xi - \xi \geq 0 \Leftrightarrow \boxed{\xi \leq (1-\zeta)^{-1}}, \end{aligned} \tag{3.46}$$

while for scalability one needs

$$\begin{aligned} &a\mathcal{I}(q_i) \geq \mathcal{I}(aq_i) \Leftrightarrow aq_i^{1+\zeta\xi-\xi} I_i(\mathbf{W}, \mathbf{q}) \geq (aq_i)^{(1+\zeta\xi-\xi)} I_i(\mathbf{W}, \mathbf{q}) \Leftrightarrow \\ &\Leftrightarrow 1 + \zeta\xi - \xi \leq 1 \Leftrightarrow \xi \geq \zeta\xi \Leftrightarrow \boxed{\zeta \leq 1}. \end{aligned} \tag{3.47}$$

The conditions on $\xi$ and $\zeta$ in (3.46) and (3.47) will be used later to ensure the convergence of (3.44). Considering for each MS that the powers allocated to the other MSs are constant does not compromise the proofs. In fact, this assumption corresponds to the same type of relaxation introduced in [SB04, BS05] when an extended power vector $[\mathbf{q}\ \ 1]^{\mathrm{T}}$ is used to solve the weighted SINR balancing problem. Moreover, the normalization step 7 of Table 3.6 ensures that this assumption will not lead to a violation of the sum power constraint of (3.1c).

For $\bar{\gamma}_i \leq \check{\gamma}$ or $\bar{\gamma}_i \geq \hat{\gamma}$, one has

$$\begin{aligned} \mathcal{I}(q_i^{(\tilde{t})}) &= q_i^{(\tilde{t})} \left( \frac{\bar{\gamma}_i \mathbf{w}_i^{(\tilde{t}),\mathrm{H}} \tilde{\mathbf{Q}}_i^{(\tilde{t})} \mathbf{w}_i^{(\tilde{t})}}{q_i^{(\tilde{t})} \mathbf{w}_i^{(\tilde{t}),\mathrm{H}} \tilde{\mathbf{R}}_i^{(\tilde{t})} \mathbf{w}_i^{(\tilde{t})}} \right)^{\xi} = \left( q^{(\tilde{t})} \right)^{1-\xi} \left( \frac{\bar{\gamma}_i \mathbf{w}_i^{(\tilde{t}),\mathrm{H}} \tilde{\mathbf{Q}}_i^{(\tilde{t})} \mathbf{w}_i^{(\tilde{t})}}{\mathbf{w}_i^{(\tilde{t}),\mathrm{H}} \tilde{\mathbf{R}}_i^{(\tilde{t})} \mathbf{w}_i^{(\tilde{t})}} \right)^{\xi} \\ &= \left( q^{(\tilde{t})} \right)^{1-\xi} I_i(\mathbf{W}^{(\tilde{t})}, \mathbf{q}^{(\tilde{t})}), \end{aligned} \tag{3.48}$$

with monotonicity being ensured whenever $\xi \leq 1$ and scalability being obtained whenever $\xi \geq 0$, i.e.,

$$0 \leq \xi \leq 1. \tag{3.49}$$

Finally, combining (3.46), (3.47), and (3.49) results in

$$-\infty \leq \zeta \leq 0, \tag{3.50a}$$

and

$$0 < \xi \leq (1-\zeta)^{-1}, \tag{3.50b}$$

where (3.50a) is fulfilled by adequately setting the parameters $\check{\gamma}$, $\hat{\gamma}$, $\check{p}$, and $\hat{p}$, which consequently fixes the valid range of values for $\xi$ in (3.50b). The parameters $\zeta$ and $\xi$ control the SDA and by obeying the conditions in (3.50), the convergence of SDA is ensured. Providing the same proof for the DL of a MU MIMO system is straightforward because $I_i(\mathbf{W}, \mathbf{q})$ in (3.45) and in (3.48) is the only term that changes to $I_i(\mathbf{W}, \mathbf{p})$ when considering the DL.

In the following, the power allocation for the DL is derived from the power allocation in the UL. Let $\bar{\gamma}_i^\star$ denote the optimal target SINRs and let $\mathbf{w}_i^\star, i = 1, \ldots, K_\mathcal{G}$, denote the optimal precoding vectors, which are the same for both UL and DL and are known after $\mathbf{q}$ converged to $\mathbf{q}^\star$. Let the gain matrix $\boldsymbol{\Upsilon}$ and noise vector $\boldsymbol{\eta}$ be defined as

$$[\boldsymbol{\Upsilon}]_{i,j} = \begin{cases} \frac{1}{\bar{\gamma}_i^\star}, & \text{for } i = j, \\ \frac{\mathbf{w}_j^{\star,\mathrm{H}} \tilde{\mathbf{R}}_i \mathbf{w}_j^\star}{\mathbf{w}_i^{\star,\mathrm{H}} \tilde{\mathbf{R}}_i \mathbf{w}_i^\star}, & \text{for } i \neq j, \end{cases} \tag{3.51a}$$

and

$$\boldsymbol{\eta} = \begin{bmatrix} \frac{\sigma^2}{\mathbf{w}_1^{\star,\mathrm{H}} \tilde{\mathbf{R}}_1 \mathbf{w}_1^\star} & \frac{\sigma^2}{\mathbf{w}_2^{\star,\mathrm{H}} \tilde{\mathbf{R}}_2 \mathbf{w}_2^\star} & \cdots & \frac{\sigma^2}{\mathbf{w}_{K_\mathcal{G}}^{\star,\mathrm{H}} \tilde{\mathbf{R}}_{K_\mathcal{G}} \mathbf{w}_{K_\mathcal{G}}^\star} \end{bmatrix}^\mathrm{T}, \tag{3.51b}$$

respectively [MK07c]. Then, writting (3.25) in matrix form [ZKAQ01], using (3.51), and using the fact that $\gamma_i = \bar{\gamma}_i^\star, \forall i$, when $\mathbf{q} = \mathbf{q}^\star$ [VT03, VJG03, SB04, BS05], one can obtain the optimal power vector $\mathbf{p}^\star$ in the DL as

$$\mathbf{p}^\star = \left( \begin{bmatrix} \boldsymbol{\Upsilon} \\ \mathbf{1}_{K_\mathcal{G}^\star}^\mathrm{T} \end{bmatrix}^\mathrm{T} \begin{bmatrix} \boldsymbol{\Upsilon} \\ \mathbf{1}_{K_\mathcal{G}^\star}^\mathrm{T} \end{bmatrix} \right)^{-1} \begin{bmatrix} \boldsymbol{\Upsilon} \\ \mathbf{1}_{K_\mathcal{G}^\star}^\mathrm{T} \end{bmatrix}^\mathrm{T} \begin{bmatrix} \boldsymbol{\eta} \\ P_b \end{bmatrix} \tag{3.52}$$

[MK07c]. Finally, the procedure for the joint optimization of precoding and power allocation combining the GEP and the SDA is presented Table 3.9.

Table 3.9. Joint precoding and power allocation using GEP and SDA.

1. Set $\tilde{t} = 0$ and $\mathbf{q}^{(\tilde{t})} = \mathbf{q}^{(0)} = \frac{P_b}{K_\mathcal{G}} \mathbf{1}_{K_\mathcal{G}}$.
2. Set $\tilde{t} = \tilde{t} + 1$.
3. For $i = 1$ to $K_\mathcal{G}$, compute $\mathbf{w}_i^{(\tilde{t}-1)}$ using (3.33).
4. For $i = 1$ to $K_\mathcal{G}$, compute $\gamma_i'^{(\tilde{t}-1)}$ using (3.32).
5. For $i = 1$ to $K_\mathcal{G}$, compute $\bar{\gamma}_i^{(\tilde{t})}$ using (3.38).
6. For $i = 1$ to $K_\mathcal{G}$, update the powers $q_i^{(\tilde{t})}$ using (3.44).
7. Set $\mathbf{q}^{(\tilde{t})} = \frac{\mathbf{q}^{(\tilde{t})}}{\left\|\mathbf{q}^{(\tilde{t})}\right\|_1} P_b$ in order to fulfill the constraint (3.1c).
8. If $\min\left\{\left|q_1^{(\tilde{t})} - q_1^{(\tilde{t}-1)}\right|, \ldots, \left|q_{K_\mathcal{G}}^{(\tilde{t})} - q_{K_\mathcal{G}}^{(\tilde{t}-1)}\right|\right\} > \epsilon_\mathrm{p}$, go to step 2.
9. Calculate $\mathbf{p}^\star$ using (3.52).

Note that the algorithms in Table 3.6 and Table 3.9 are quite similar. Similarly to step 7 of Table 3.6, the normalization step 7 of Table 3.9 ensures that (3.1c) is fulfilled. Differently from Table 3.6, the SDA in Table 3.9 is performed in the UL with the power allocation in the DL being determined in step 9 of Table 3.9.

## 3.4 Resource Allocation Strategy Definition

In this section, the RA strategies considered in this chapter are defined. They consist of combining one of the SDMA algorithms listed in Table 3.4 with one of the precoding and power allocation algorithms listed in Table 3.8.

The RA strategies will be named after the SDMA algorithm and the precoding and power allocation algorithm that they employ. For example, the RA strategy employing the CAP-ESA of Table 3.4 and the ZF algorithm of Table 3.8 will be named CAP-ESA-ZF. The RA strategies considered in this chapter and their names are listed in Table 3.10.

Table 3.10. RA strategies: single resource.

| **SDMA algorithm** | **Precoding and power allocation algorithm** | |
|---|---|---|
| | **ZF** | **GEP** |
| **CAP-ESA**[a] | CAP-ESA-ZF | (CAP-ESA-GEP)[b] |
| **CC-CGA** | CC-CGA-ZF | CC-CGA-GEP |
| **CAP-BFA**[a] | CAP-BFA-ZF | CAP-BFA-ZF |
| **SP-BFA** | SP-BFA-ZF | SP-BFA-GEP |
| **WN-BFA** | WN-BFA-ZF | WN-BFA-GEP |
| **CC-BFA** | CC-BFA-ZF | CC-BFA-GEP |
| **WN-FFA** | WN-FFA-ZF | WN-CGA-GEP |
| **RGA** | RGA-ZF | RGA-GEP |

[a] For CAP-ESA and CAP-BFA, the SRA is not needed and no longer employed.
[b] This RA strategy is too complex and is not considered further.

The RA strategies of Table 3.10 cover all the combinations between the SDMA algorithms of Table 3.4 and the precoding and power allocation algorithms of Table 3.8. In particular, the CAP-ESA-GEP strategy is not considered further because it is too complex due to the large number of SDMA groups considered by the CAP-ESA and the joint iterative optimization of precoding and power allocation of the GEP of Table 3.8.

## 3.5 Performance and Complexity of the Resource Allocation Strategies

### 3.5.1 System and Resource Allocation Strategy Parameters

In this section, the values of the system parameters and the values of the parameters of the RA strategies are defined. These parameter values are used in the simulations conducted in this work to assess the performance of the RA strategies.

In the following, the values of the system parameters are defined. A center frequency $f_0 = 5$ GHz and a system bandwidth $B_{\text{sys}} = 468.75$ MHz are considered. A total number $S = 48$ of subcarriers with bandwidth $\Delta_f \approx 9.766$ kHz is considered for data transmission. For the channel realizations considered in the simulations, the WIM considering the Non Line Of Sight (NLOS) propagation scenario C2 is employed, which presents a root mean squared excess delay $\tau_{\text{rms}}$ of approximately 0.8 $\mu s$ [WIN05c, Skl97]. Considering the coherence bandwidth $B_{\text{c}} \triangleq \frac{1}{5\tau_{\text{rms}}}$, cf. [Skl97], one has for the adopted scenario that $B_{\text{c}} \approx 250$ kHz. The subcarriers of the system are organized into $B = 8$ frequency blocks of $Q_{\text{sub}} = 6$ adjacent subcarriers. However, it must be remembered that for the analyses conducted in this chapter, a single resource is considered. The values of $Q_{\text{sub}}$ and $\Delta f$ result in a frequency block of bandwidth $Q_{\text{sub}}\Delta f < 0.25B_{\text{c}}$. An average MS speed $v_{\text{MS}} \approx 2.78$ m/s is assumed. Considering the coherence time $T_{\text{c}} \triangleq \frac{\lambda}{2v_{\text{MS}}}$, cf. [Skl97], one has for the referred scenario that $T_{\text{c}} \approx 11$ ms. Frames of duration $T_{\text{FRM}} = 1$ ms are considered, which corresponds to have $T_{\text{FRM}} < 0.1T_{\text{c}}$. Each frame is composed of $T = 4$ TSs. A single BS located at the corner of a hexagonal cell sector and equipped with a 4-element ULA is considered. A total number $K = 16$ of single-antenna MSs are associated with the BS. The values of the system parameters are listed in Table 3.11.

In the following, the values of the parameters of the RA strategies are defined. These parameters are related to the SDMA algorithms and to the precoding and power allocation algorithms discussed in Section 3.2 and Section 3.3, respectively. The values of the parameters are presented in Table 3.12 and are discussed in the sequel.

In the following, the values given to the parameters related to the SDMA algorithms are discussed. For the RA strategies whose SDMA algorithm employs the CGA, BFA, FFA, or the RGA grouping algorithms, a target SDMA group size $K_{\mathcal{G}}^{\star} = M = 4$ is considered. This value is adopted for the same reasons as discussed in Section 3.2.3. Moreover, if needed the SRA of Section 3.3.3 can reduce the SDMA group size afterwards.

For the RA strategies whose SDMA algorithm employs the CGA, BFA or the FFA as grouping algorithm, the MS with the highest channel gain is chosen as the initial MS $k'$. This is a reasonable choice because the composition of the SDMA group is unknown at the time of the selection of $k'$ and for a SU case the MS with the highest channel

Table 3.11. System parameters.

| Parameter | Symbol | Value | Unit | Remark |
|---|---|---|---|---|
| Center frequency | $f_0$ | 5.0 | GHz | - |
| System bandwidth | $B_{\text{sys}}$ | 468.75 | kHz | - |
| # of subcarriers | $S$ | 48 | - | - |
| Subcarrier spacing | $\Delta f$ | 9.766 | kHz | - |
| Channel model | - | WIM | - | Scenario C2, NLOS |
| # of resources | $B$ | 8 | - | In this chapter, a single resource is considered |
| # of subcarriers per block | $Q_{\text{sub}}$ | 6 | - | < 25% of the coherence bandwidth $B_{\text{c}}$ |
| Average speed of the MSs | $v_{\text{MS}}$ | ≈2.78 | m/s | 10 km/h |
| Frame duration | $T_{\text{FRM}}$ | 1 | ms | < 10% of the coherence time $T_{\text{c}}$ |
| # of TSs per frame | $T$ | 4 | - | - |
| TS duration | $T_{\text{TS}}$ | 0.25 | ms | - |
| BS array size | $M$ | 4 | - | ULA with omnidirectional elements separated by half wavelength |
| # of MSs | $K$ | 16 | - | Single-antenna |

Table 3.12. Parameters for the RA strategies.

| Parameter | Symbol | Value | Unit | Remark |
|---|---|---|---|---|
| **SDMA algorithms** | | | | |
| Target SDMA group size | $K_{\mathcal{G}}^{\star}$ | 4 | MSs | $K_{\mathcal{G}}^{\star} = 4 = M$, cf. Table 3.11 |
| Initial MS | $k'$ | $\arg\max_k\{\Vert\hat{\mathbf{h}}_k\Vert_2^2\}$ | - | MS with the highest channel gain $\Vert\hat{\mathbf{h}}_k\Vert_2^2$ |
| Parameter for $f_{\text{CC}}(\mathcal{G})$ | $\beta$ | 0.5 | - | - |
| Loading factor for the CGA | $\varepsilon$ | 1 | - | - |
| **Precoding and power allocation (for the SDA)** | | | | |
| Minimum target SINR | $\check{\gamma}$ | 0.26 | - | $\check{\Gamma} \approx -5.85$ dB, $\log_2(1+\check{\gamma}) = \frac{1}{3}$ |
| Maximum target SINR | $\hat{\gamma}$ | $M\gamma$ | - | $\hat{\Gamma} = 10\log_{10}(\gamma) + 10\log_{10}(M)$, i.e., $10\log_{10}(\gamma)$ increased by the average array gain |
| Minimum allocable power | $\check{p}$ | $10^{-3}$ | W | $\check{P} = 0$ dBm |
| Maximum allocable power | $\hat{p}$ | 1 | W | $\hat{P} = 30$ dBm |
| Feedback parameter | $\xi$ | $(1-\zeta)^{-1}$ | - | Leads to fastest convergence |
| Power precision | $\epsilon_{\text{p}}$ | $10^{-4}$ | W | - |

gain is the one maximizing the rate of the system. Consequently, this choice is in accordance with the objective of maximizing the sum rate. It might also be considered as an MS priority, which is part of the GA algorithm illustrated in Fig. 1.3 and will only be considered as such in Section 3.5.9.

In particular, for the CC-CGA-ZF and CC-CGA-GEP strategies, the parameter $\beta$ of (3.13) is set to 0.5 in order to give almost the same weight to the spatial correlation and channel gains in the $f_{\text{CC}}(\mathcal{G})$ of (3.13). The impact of the choice of $\beta$ on the performance of the RA strategies whose SDMA algorithm employs $f_{\mathrm{CC}}(\mathcal{G})$ as grouping metric will be investigated in Section 3.5.7.

For the CC-CGA-ZF and CC-CGA-GEP strategies, the parameter $\varepsilon$ is set to 1. This value has been determined by inspecting the histogram of the minimum eigenvalue $\lambda^{-}(\mathbf{C})$ considered in (3.19), which has been obtained by computing $\lambda^{-}(\mathbf{C})$ for a large number of channel realizations while considering the parameters in Table 3.11. In Fig. 3.4, the histogram of the minimum eigenvalue of the diagonally loaded matrix $\tilde{\mathbf{C}} = \mathbf{C} + \mathbf{I}_K$ is shown and it can seen that setting $\varepsilon = 1$ efficiently upper bounds $\epsilon$ of (3.19) so that $\tilde{\mathbf{C}}$ is always positive semidefinite and the need of performing an EVD of $\mathbf{C}$ is eliminated, as discussed in Section 3.2.2.

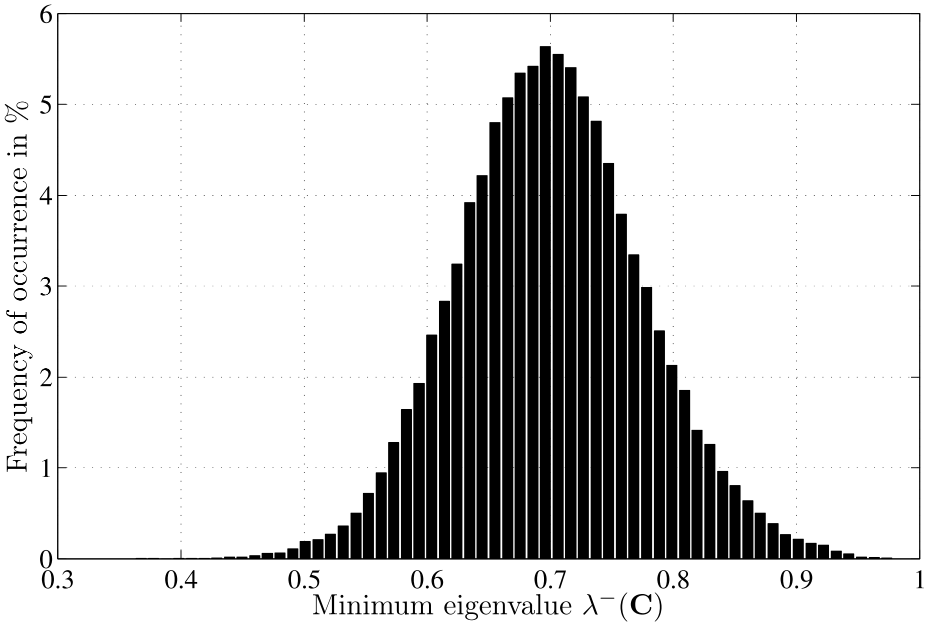



Figure 3.4. Histogram of the minimum eigenvalue $\lambda^{-}(\mathbf{C} + \mathbf{I}_K)$.

In the following, the values given to parameters related to the precoding and power allocation algorithms are discussed. The RA strategies employing the GEP as precoding and power allocation need to adequately set the parameters $\hat{\gamma}$, $\check{\gamma}$, $\hat{p}$, $\check{p}$, and $\xi$. The highest allowed target SINR $\hat{\gamma}$ has been set to the value of the average SNR $\gamma$ of (2.1) scaled by the average array gain that an MS would perceive if it were alone in an SDMA group [PNG03]. The lowest allowed target SINR $\check{\gamma}$ has been just set to a very low value, which corresponds to a rate of $\frac{1}{3}$ bit per channel access. In this way, the dynamic SNR range that might be expected in each scenario is relatively well covered. The maximum allocable power $\hat{p}$ has been set to the total available power $P_b = 1$ W. The minimum allocable power $\check{p}$ has been set as to provide a dynamic power range of 30 dBm, which is a common dynamic range for PC in wireless communication systems [HRM02, HT02]. The value of the parameter $\zeta$ is derived directly from $\check{\gamma}$, $\hat{\gamma}$, $\check{p}$, and $\hat{p}$. Because the higher the value of $\xi$, the faster the convergence of the SDA, $\xi$ has been set to its maximum allowed value, cf. (3.50b) [YGRS97, PED04, MK07c]. The convergence of the GEP algorithm of Table 3.9 is considered when the maximum difference between components of $\mathbf{q}$ and $\mathbf{q}^{\star}$ is lower than or equal to the power precision

$\epsilon_p = 10^{-4}$, i.e., one power of ten lower than the minimum allocable power $\check{p}$. Usually, the GEP algorithm of Table 3.9 requires less than ten iterations to converge [MK07c].

### 3.5.2 Performance with Different Amounts of Channel State Information at the Transmitter

In this section, the impact of the amount of CSIT on the performance of the RA strategies of Table 3.10 is investigated considering the P-CSIT and B-CSIT models introduced in Section 2.2.4. In this section, it is of interest to observe how the performance of the RA strategies degrades in terms of the average sum rate when moving from the case in which P-CSIT is available to the case in which B-CSIT is available on a frame basis. This allows to study the trade-off existing between the average sum rate and the amount of signaling required to achieve that sum rate. For this sake, the three cases below are considered:

1. P-CSIT with adaptive RA per subcarrier and TS: SDMA, precoding, and power allocation algorithms are applied for each subcarrier and TS of the frame. This case assumes that the channel matrix $\mathbf{H}_s$ of (2.3) is known for each subcarrier $s$ and for each TS of a frame.
2. B-CSIT with adaptive RA per TS: SDMA, precoding, and power allocation algorithms are applied for each TS, but reused over all the subcarriers of the resource. This case assumes that the channel matrix $\hat{\mathbf{H}}_b$ of (2.5b) is known for the middle subcarrier $\bar{s}_b$ of each frequency block and for each TS of a frame.
3. B-CSIT with adaptive RA per frame: SDMA, precoding, and power allocation algorithms are applied once per frame and reused over the subcarriers of the resource and over all TSs of the frame. This case assumes that the channel matrix $\hat{\mathbf{H}}_b$ of (2.5b) is known for the middle subcarrier $\bar{s}_b$ of the frequency block $b$ and for the first TS of a frame.

For these three cases, Fig. 3.5 shows the average sum rate of the system in bits/channel use achieved by the CAP-ESA-ZF strategy as a function of the average SNR $\gamma$ given by (2.1). A channel use corresponds to one subcarrier being used for the duration $T_s$ of one OFDMA symbol. The average sum rate of the system in bits/channel use will be the metric mostly adopted for comparing the RA strategies in this section.

In Fig. 3.5, the average sum rate of the system obtained using the algorithm in [JRV+05] is additionally included as an absolute upper bound for the average sum rate that can be achieved using DPC techniques, cf. Section 1.3, and considering P-CSIT with adaptive RA per subcarrier and TS. In Fig. 3.5, it can be observed that the CAP-ESA-ZF with P-CSIT achieves about 90% of the maximum achievable average sum rate. The lower average sum rate achieved by the CAP-ESA-ZF strategy compared to the algorithm in [JRV+05] results because the CAP-ESA-ZF strategy does not jointly optimize precoding vectors and power allocation as the algorithm in [JRV+05] does.

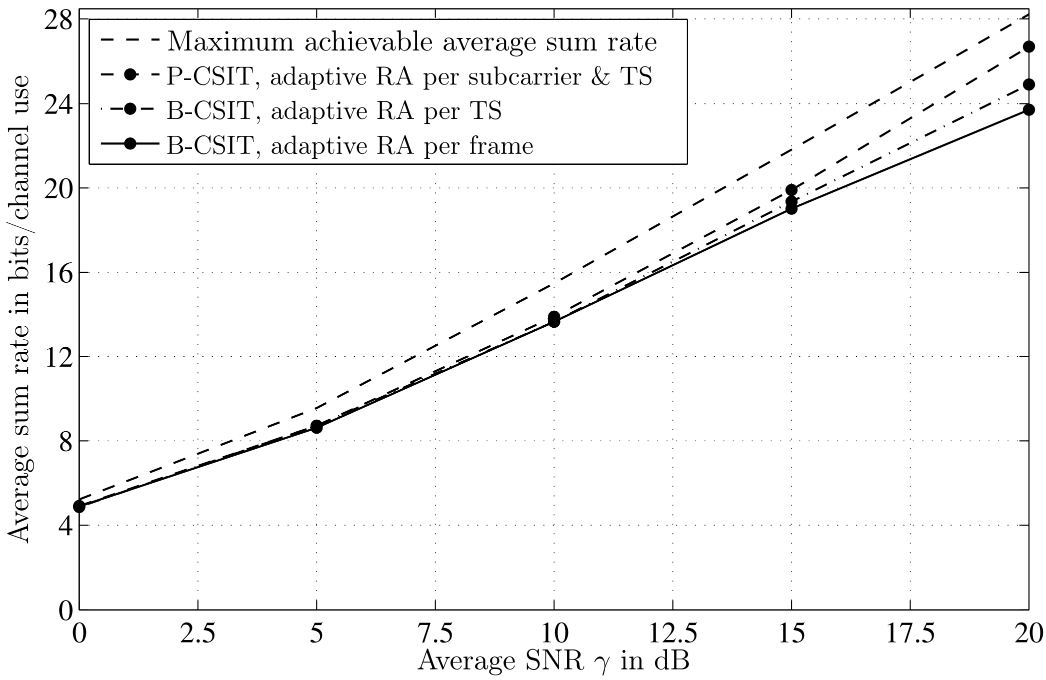


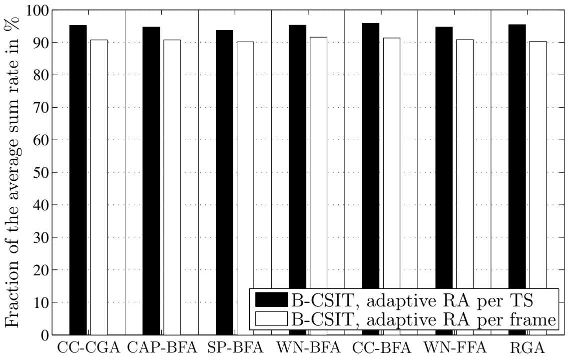

Figure 3.5. Average sum rate of the system considering CAP-ESA-ZF.

Considering now the different cases of the CAP-ESA-ZF strategy, it can be seen that they achieve almost the same average sum rate for low to moderate values of $\gamma$. For higher average SNR values the average sum rates achieved considering B-CSIT with adaptive RA per TS and per frame are only 7% and 11% lower than the average sum rate achieved with P-CSIT and adaptive RA per subcarrier and TS, respectively.

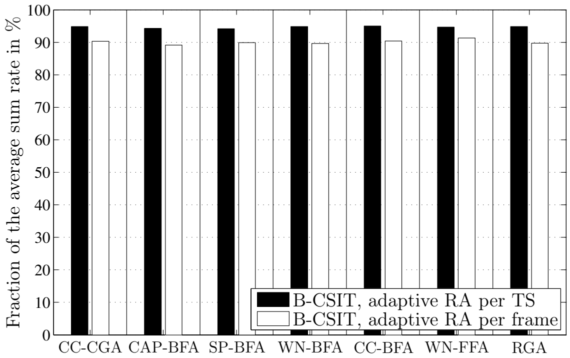
The same comparison conducted for the CAP-ESA-ZF strategy in Fig. 3.5 has been conducted for all RA strategies of Table 3.10. The same trend shown for the CAP-ESA-ZF in Fig. 3.5 has been observed for all the RA strategies, which achieve almost the same average sum rate with P-CSIT and B-CSIT for low to moderate average SNR values and slightly lower average sum rate values for high SNR values. Fig. 3.6 shows the fraction of the average sum rate considering P-CSIT with adaptive RA per subcarrier and TS that is achieved by the RA strategies of Table 3.10 when considering B-CSIT with adaptive RA per TS and per frame for an average SNR $\gamma$ of 20 dB.

The reduction of about 5% to 10% of the average sum rate observed in Fig. 3.5 and Fig. 3.6 result from the reduced adaptivity of the RA strategies due to reduced amount of CSIT compared to the case with P-CSIT.

As it can be noted in Fig. 3.5 and Fig. 3.6, all the RA strategies are affected almost in the same way. On the one hand, performing adaptive RA per frame using B-CSIT leads to a reduction of the average sum rates. On the other hand, a considerable reduction of signaling requirements is obtained. Indeed, for the adaptive RA per subcarrier and TS the MSs need to be informed about the individual subcarriers allocated to them on each TS, which requires roughly a factor of $T \cdot Q_{\text{sub}}$ more signaling than in the case with B-CSIT and adaptive RA on a frame basis.



(a) ZF-WFA-LGF.



(b) GEP-SDA-HPF.

Figure 3.6. Fraction of the average sum rate considering P-CSIT with adaptive RA per subcarrier and TS that is achieved by the RA strategies when considering B-CSIT with adaptive RA per TS and per frame. Average SNR $\gamma = 20$ dB.

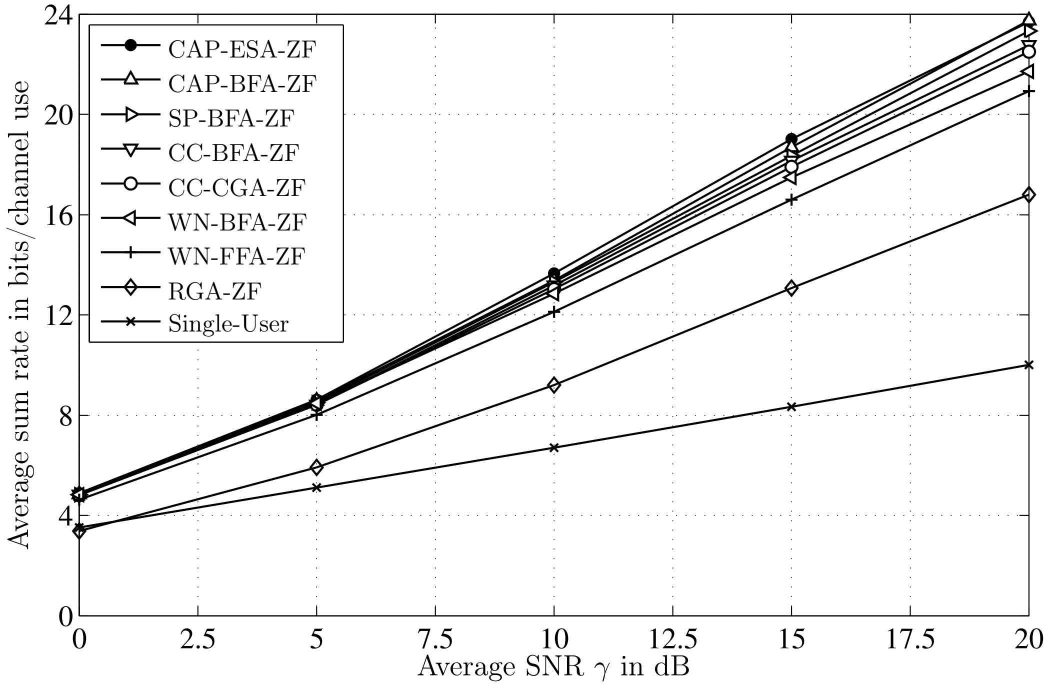
Further on, P-CSIT with adaptive RA per subcarrier and TS will not be considered, adaptive RA per TS will be locally considered in Section 3.5.9, and adaptive RA per frame will be considered in most of the cases. Moreover, the performance of the CAP-ESA-ZF considering adaptive RA per frame will be considered to upper bound the performance of the other RA strategies in terms of average sum rate.

### 3.5.3 Performance with Block Channel State Information at the Transmitter

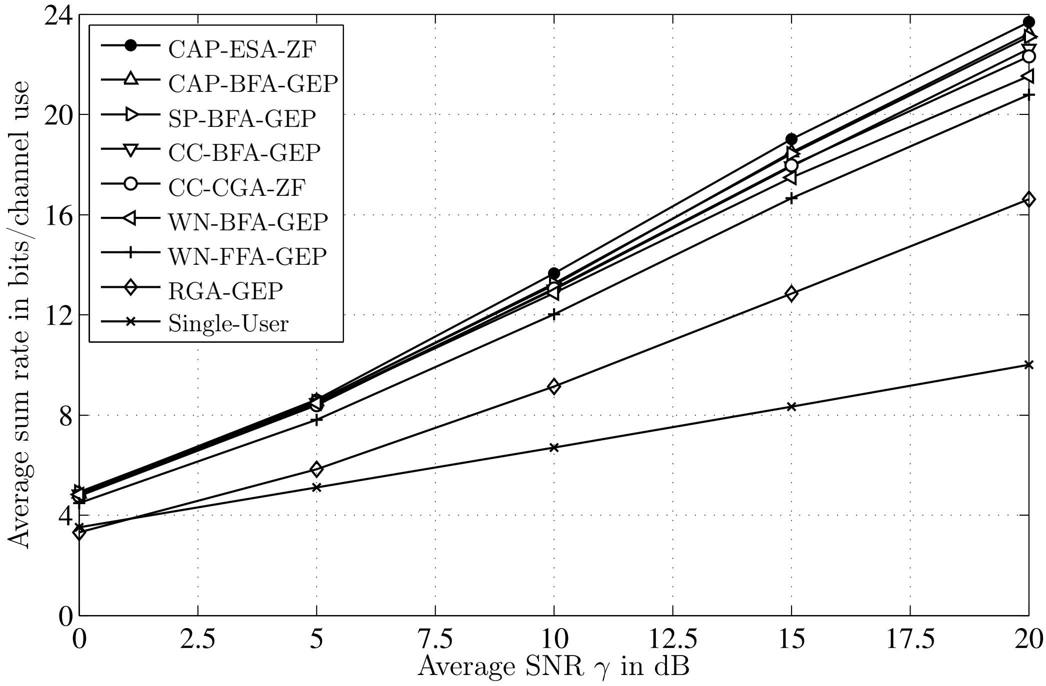
In this section, the performance of the RA strategies is compared considering B-CSIT with adaptive RA per frame. Fig. 3.7 shows the average sum rate achieved by the RA strategies of Table 3.10 as a function of the average SNR $\gamma$ in dB, with the results achieved considering the ZF and the GEP defined in Section 3.3 being shown in Fig. 3.7(a) and Fig. 3.7(b), respectively. Additionally, the average sum rate achieved considering in the Single-User (SU) case, i.e., when transmitting to the MS with the highest channel gain, is also presented for comparison.

In the following, the average sum rate performance shown in Fig. 3.7(a) is discussed from the best- to the worst-performing RA strategy. Because all the RA strategies in Fig. 3.7(a) employ the ZF algorithm of Table 3.8, the performance differences in terms of the average sum rate can be attributed to the different SDMA algorithms employed by each RA strategy.

In Fig. 3.7(a), CAP-BFA-ZF is the strategy that best approximates the average sum rate of the CAP-ESA-ZF strategy. Because the number of candidate SDMA groups considered by the BFA might be much smaller than the number of candidate groups which are considered by the ESA, cf. (2.14), this result shows that a high fraction of the average sum rate of the system can be achieved by employing simple suboptimal RA strategies. The reason for the good performance of the CAP-BFA-ZF strategy is




(a) ZF-WFA-LGF.




(b) GEP-SDA-HPF.

Figure 3.7. Average sum rate of the system for the RA strategies with B-CSIT.

that it is aware of the actual precoding and power allocation, which allows to accurately estimate the capacity of the candidate groups and the MSs admitted to the group in step 2a of Table 3.2 are the ones that at most increase the capacity of the group.

The average sum rate achieved by the CAP-BFA-ZF strategy is followed very closely by that achieved by the SP-BFA-ZF strategy. Differently from the CAP-BFA-ZF strategy, the SP-BFA-ZF strategy is unaware of the precoding and power allocation. The grouping metric considered in the SP-BFA is based on the sum of the channel gains with null space SPs and, consequently, is quite similar to the null space projections performed by the ZF precoding algorithm discussed in Section 3.3.1. Combined with the BFA, this metric effectively captures the spatial correlation among the MSs, which leads to its good performance in terms of average capacity.

In Fig. 3.7(a), the next best performing RA strategies are CC-BFA-ZF and CC-CGA-ZF, which have very similar performances. These two strategies are also unaware of the precoding and power allocation, but differently from the SP-BFA-ZF strategy, the grouping metric employed by their SDMA algorithms involves only the pairwise spatial correlation among the MSs measured using (3.6). Consequently, the CC-BFA-ZF and CC-CGA-ZF strategies are slightly less accurate in estimating the spatial compatibility among the MSs, which leads to their slightly lower performance in terms of average sum rate compared to the CAP-BFA-ZF and SP-BFA-ZF strategies.

For the WN-BFA-ZF and the WN-FFA-ZF strategies, the employed grouping metric also captures only the pairwise spatial correlation among the MSs. Moreover, the total spatial correlation is not added up, but averaged by the Frobenius norm in (3.10) [SS04a, MK06], which is responsible for the slightly lower average sum rate achieved by the WN-BFA-ZF strategy. The lower average sum rate achieved by the WN-FFA-ZF strategy compared to some of the RA strategies is due to the FFA, which simplifies the search for the best SDMA group and consequently leads to worse performance. Nevertheless, the WN-FFA-ZF obtains over 85% of the average sum rate achieved by the CAP-ESA-ZF strategy.

Comparing now the average sum rate figures achieved by the SP-BFA-ZF, CC-BFA-ZF, and CC-CGA-ZF strategies with that of the CAP-ESA-ZF strategy, it can be concluded that RA strategies employing SDMA algorithms unaware of the actual precoding and power allocation are able to efficiently approximate the average sum rate achieved through an ES for the SDMA group of highest capacity. In fact, these RA strategies reach over 90% of the average sum rate achieved by the CAP-ESA-ZF strategy.

The RGA-ZF performs the worst, which is an expected result. Anyway, it can be observed in Fig. 3.7(a) that the RGA-ZF strategy achieves about 70% of the capacity of the CAP-ESA-ZF strategy, which is a relatively good performance for such a simple RA strategy. However, as it will be seen in Section 3.5.6, a considerable fraction of the average sum rate achieved by the RA strategies, and in particular by the RGA-ZF strategy, is due to the SRA, which is aware of the precoding and power allocation.

In general, the higher the average SNR, the more the CAP-ESA-ZF strategy can take advantage of its complete knowledge about all possible SDMA groups, so that the gap between the performance of this and the other RA strategies slightly increases for higher average SNR values. Moreover, the sum rate increases more than linearly with the average SNR $\gamma$, which also contributes to increase the performance gap between the

CAP-ESA-ZF strategy and the other RA strategies in terms of the achieved average sum rates for high average SNR values.

Compared to average sum rate achieved in the SU case, it can be seen in Fig. 3.7 that the RA strategies provide gains ranging from 30% for low SNR to more than 100% for high SNR. In particular, the RGA-ZF strategy presents lower gains compared to the SU case and performs even slightly worse than it for low values of $\gamma$. For low SNR $\gamma$, SDMA groups containing a single MS are often built by the RGA-ZF strategy and because the MS with highest channel gain does not necessarily belongs to the group built by this strategy, its achieved average sum rate might be lower than the sum rate achieved in the SU case.

The same remarks made for the RA strategies considering ZF-WFA-LGF in Fig. 3.7(a) can be made for the corresponding RA strategies considering GEP-SDA-HPF shown in Fig. 3.7(b). Indeed, the RA strategies present quite similar performance independent of the precoding and power allocation algorithm being considered.

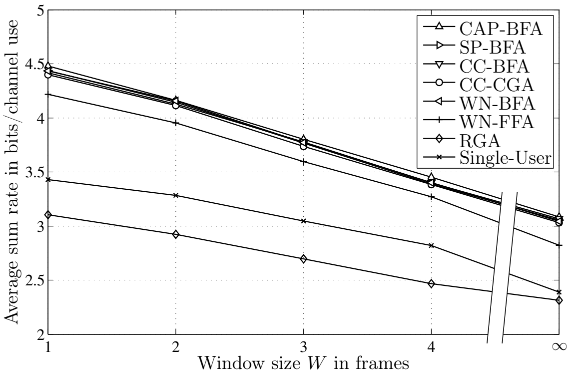
### 3.5.4 Performance with Second-order Channel State Information at the Transmitter

In this section, the performance of the RA strategies is investigated considering S-CSIT. The performance of the RA strategies strongly depends on the available CSIT. Obtaining up to date CSIT is particularly challenging for FDD systems, in which channel reciprocity cannot be exploited and CSI has to be acquired by the MSs and fed back to the BS. Consequently, the signaling overhead to obtain CSIT in these systems may become prohibitively large. As discussed in Section 2.2.4, this signaling overhead can be considerably reduced by using S-CSIT, however at the expense of reduced performance of the RA strategies in terms of average sum rate, as it will be shown in the sequel.

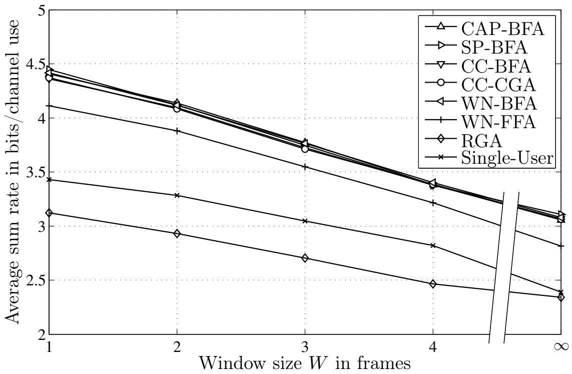
Fig. 3.8 shows the average sum rate in bits/channel use achieved by the RA strategies for a varying window size $W$, cf. (2.7), and for average SNRs $\gamma = 0$ dB, $\gamma = 10$ dB, and $\gamma = 20$ dB. The performances of the RA strategies considering ZF-WFA-LGF and GEP-SDA-HPF are shown on the left and right sides of Fig. 3.8, respectively.

In Fig. 3.8, it can be seen that the average sum rate achieved by the RA strategies rapidly decays when the window size $W$ increases. In the following, the reasons for this reduction of the average sum rates are discussed.

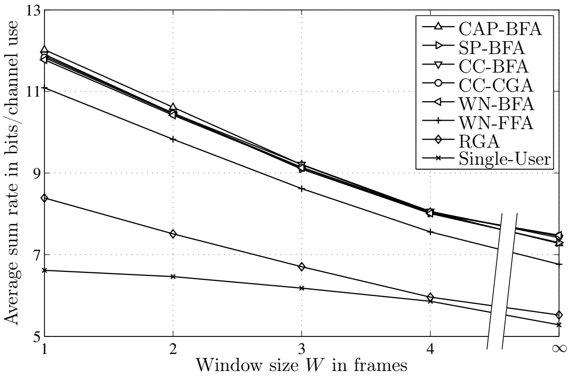
Firstly, the S-CSIT model considers that the CSI obtained during the $W$ past frames is used during the $W$ subsequent frames and, consequently, it inherently involves a delay that affects the quality of the CSIT. For example, for $W = 1$, the S-CSIT corresponds indeed to the B-CSIT of the previous frame, since the dominant eigenmode of $\mathbf{R}_k$ given by (2.7) is equivalent to $\hat{\mathbf{h}}_k^{(\tilde{f}-1)}$. Because a frame takes less than 10% of the coherence



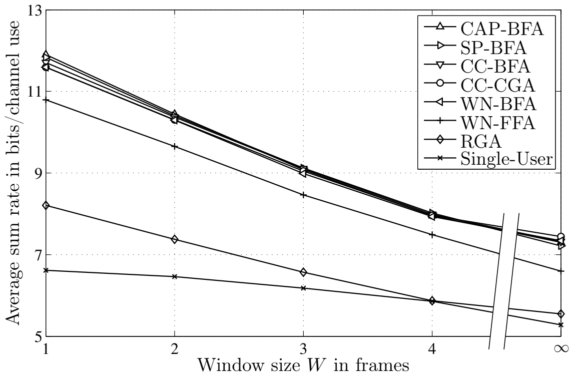

(a) ZF-WFA-LGF, $\gamma = 0$ dB.



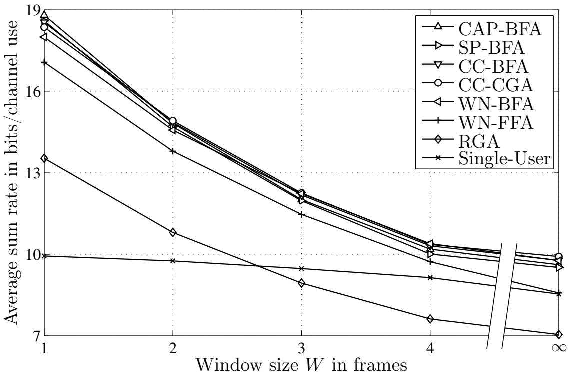

(b) GEP-SDA-HPF, $\gamma = 0$ dB.



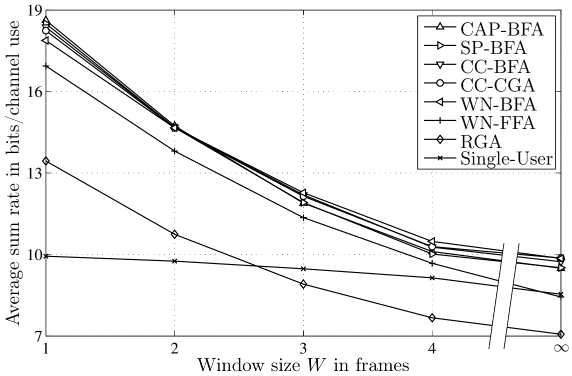

(c) ZF-WFA-LGF, $\gamma = 10$ dB.




(d) GEP-SDA-HPF, $\gamma = 10$ dB.




(e) ZF-WFA-LGF, $\gamma = 20$ dB.




(f) GEP-SDA-HPF, $\gamma = 20$ dB.

Figure 3.8. Average sum rate achieved by the RA strategies considering S-CSIT with different window sizes $W$.

time $T_c$, the B-CSIT of the past frame is highly correlated with the B-CSIT of the current frame, which explains the correspondingly better results obtained with $W = 1$ compared to $W \rightarrow \infty$. As the window size increases, this correlation decreases and the achieved average sum rates also decrease. In fact, for $W = 4$ frames, it can be seen in Fig. 3.8 that the achieved average sum rates already approximate the values obtained with a window size $W \rightarrow \infty$.

Secondly, the considered channel model and scenario involve highly uncorrelated scattering [WIN05c], so that the covariance matrices of (2.7) are full-rank and have a non-concentrated energy distribution among their eigenmodes. Consequently, the rank-one assumption made on (2.8) leads only to a coarse representation of the channel of each

MSs and the transmission along the dominant spatial propagation mode of the channels leads only to poor performance in terms of average sum rate. Indeed, according to [VM01], the optimal transmission scheme when the spatial covariance matrix is full-rank corresponds to transmit along all their eigenmodes, which would on the other hand require a rank-one approximation for the SDMA algorithm [FN96, DH04].

Similarly to the results in Fig. 3.7, the reduction of the average sum rate of the RA strategies in Fig. 3.8 is larger for higher average SNR values than for lower average SNR values for the same reasons as discussed in Section 3.5.3. The higher the average SNR in Fig. 3.8, the more advantage the RA strategies can take from its better CSIT, as discussed in Section 3.5.3. Moreover, the sum rate increases more than linearly with the average SNR.

The results presented in Fig. 3.8(a), Fig. 3.8(c), and Fig. 3.8(e) considering the ZF algorithm of Table 3.8 are very similar to the results presented in Fig. 3.8(b), Fig. 3.8(d), and Fig. 3.8(f), which consider the GEP algorithm of Table 3.8, so that ZF and GEP do not offer any particular advantage over each other regarding the utilization of S-CSIT. This result has also been observed in Fig. 3.7, in which B-CSIT is considered. Moreover, none of the RA strategies is considerably more robust against the imperfections present in the S-CSIT and the same performance trend among RA strategies seen in Fig. 3.7 holds here. For $W = 1$, the RA strategies are capable of obtaining about 80% to 90% of the average sum rate values shown in Fig. 3.7 considering B-CSIT, and for $W = 4$ they achieve only 45% to 70% of the average sum rates shown in Fig. 3.7. In order to ensure about 60% of the achievable average sum rate considering B-CSIT, a window size of at most two frames has to be considered with S-CSIT, which according to the parameters in Table 3.11 approximately corresponds to 40% of the coherence time $T_c$ considering the inherent delay. This window size also permits the proposed RA strategies to perform better than the SU case in all the considered configurations.

In Fig. 3.8, it can be noted that the reductions in the achieved average sum rates due to the use of S-CSIT are less pronounced in the SU case than in the cases considering the studied RA strategies. Due to the imperfect CSIT, spatial compatibility cannot be measured very efficiently by the SDMA algorithms. Similarly, the performance of precoding and power allocation algorithm is also compromised to some extent. Especially for the large SDMA groups built by the RA strategies when considering high SNR values, elevated levels of spatial interference arise and lead to the more pronounced reductions of the average sum rates observed in Fig. 3.8 when comparing the RA strategies and the SU case. In the SU case, a single MS is served at a time and, consequently, there is no spatial interference, which explains the less pronounced reductions of the average sum rates in this case.

From the discussions presented in this section, it can be concluded that the potential of SDMA of strongly increasing the sum rate of the system might be considerably compromised if stringent constraints on the signaling amount for CSI are imposed to the system, e.g., when considering FDD systems. However, the S-CSIT model involves only $\frac{1}{W}$ of the signaling amount required by the B-CSIT model and, consequently, could compensate for the reduction of the sum rate of the system. Nevertheless, better results in terms of average sum rate using S-CSIT than those presented in this section could be expected in scenarios involving, e.g., a strong Line Of Sight (LOS), low angular spread, and a small number of propagation paths, which would better support the rank-one assumption made in (2.8).

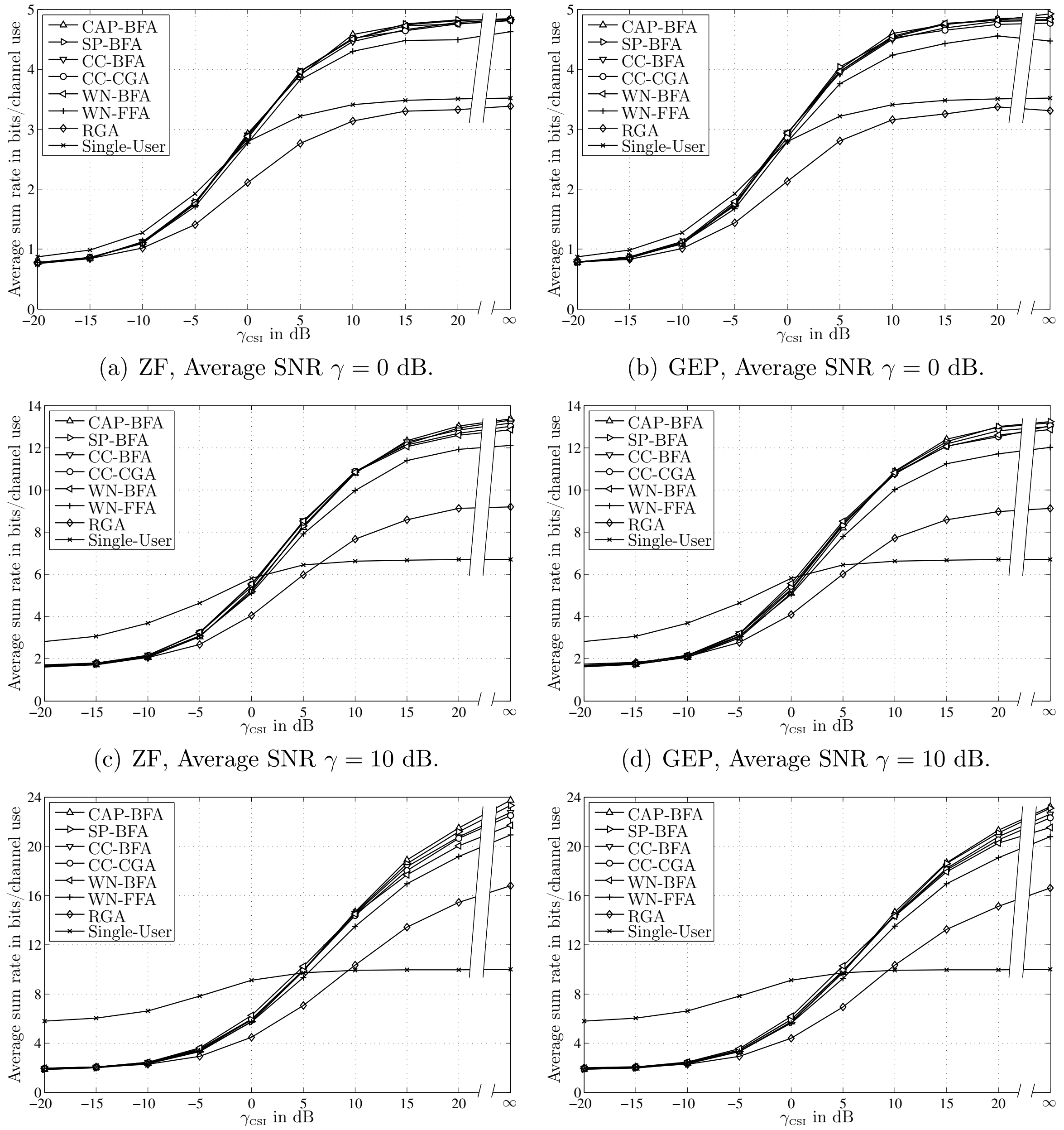
### 3.5.5 Performance with Erroneous Block Channel State Information at the Transmitter

In this section, the performance of the RA strategies of Table 3.10 considering erroneous B-CSIT, cf. (2.9), is evaluated. In Section 3.5.3, B-CSIT is assumed to be instantaneously known for the first TS of each frame and to be error-free. However, in practice neither instantaneous nor error-free CSI is available. In fact, processing and feedback delays, as well as imperfect channel estimation corrupt the CSIT and ultimately affect the performance of the RA strategies. The adopted error model may represent both CSIT imperfections due to channel estimation errors and delays, as discussed in Section 2.2.4.

Fig. 3.9 shows the average sum rate achieved by the RA strategies considering erroneous B-CSIT as a function of $\gamma_{\text{CSI}}$, cf. (2.10), which characterizes the amount of error present in the CSIT. Average SNR values of $\gamma = 0$ dB, $\gamma = 10$ dB, and $\gamma = 20$ dB are considered and the average sum rate results for the RA strategies employing the ZF and GEP algorithms of Table 3.8 are shown on the left and right sides of Fig. 3.9, respectively.

Analogously to the relative performance observed in Section 3.5.4 with S-CSIT, none of the RA strategies presents itself particularly more efficient than the others when considering erroneous B-CSIT. As the amount of error in the CSIT increases, i.e., as $\gamma_{\text{CSI}}$ of (2.10) decreases, the performance of the RA strategies in terms of the average sum rate degrades rapidly. In order to obtain at least 60% of the average sum rate achieved by the RA strategies with error-free B-CSIT, channel estimation errors and processing/feedback delays should lead to $\gamma_{\text{CSI}}$ values not lower than 10 dB, so that the RA strategies of Table 3.10 are considerably sensitive to imperfections in the CSIT. This value of $\gamma_{\text{CSI}}$ also permits the proposed RA strategies to perform better than the SU case in all the considered configurations.

Similarly to the results shown in Fig. 3.8, it can be seen in Fig. 3.9 that the reductions in the achieved average sum rates due to erroneous B-CSIT are less pronounced in the SU case than in the cases considering the studied RA strategies. The reasons for the higher sensitivity of the studied RA strategies to imperfect CSI here are the same as discussed in Section 3.5.4, i.e., less accurate measurement of spatial compatibility and less accurate determination of precoding vectors and power allocation, which lead to increased spatial interference. Differently from the results shown in Section 3.5.4,




(a) ZF, Average SNR $\gamma = 0$ dB.

(b) GEP, Average SNR $\gamma = 0$ dB.

(c) ZF, Average SNR $\gamma = 10$ dB.

(d) GEP, Average SNR $\gamma = 10$ dB.

(e) ZF, Average SNR $\gamma = 20$ dB.

(f) GEP, Average SNR $\gamma = 20$ dB.

Figure 3.9. Impact of erroneous CSIT on the average sum rate achieved the RA strategies.

where the proposed RA strategies always outperformed the SU case, the results in Fig. 3.9 show that the contrary may occur if the amount of errors in the CSIT reaches very high levels. In this case, the performance of the RA strategies converges to that of the RGA-ZF and RGA-GEP, which perform worse than the SU case due to the higher levels of spatial interference perceived by the MSs in the SDMA groups.

Results similar to that shown in Fig. 3.9 can be found in [YG05], where imperfections in the CSIT are due to the feedback delay $\tilde{\tau}$ only. If errors in the CSIT are only due to feedback delays, $\gamma_{\text{CSI}}$ of (2.10) is related to the feedback delay $\tilde{\tau}$ as $\sqrt{\frac{\gamma_{\text{CSI}}}{1+\gamma_{\text{CSI}}}} = J_0\left(\frac{2\pi v_{\text{MS}}\tilde{\tau}}{\lambda}\right)$ [HH03, YG05, YG06, KK07b, KK07a], where $J_0(\cdot)$ is the zero-order Bessel function of the first kind and corresponds to the spaced-time correlation of the channel [Skl97]. Table 3.13 shows the values of $J_0\left(\frac{2\pi v_{\text{MS}}\tilde{\tau}}{\lambda}\right)$ for the different values of $\gamma_{\text{CSI}}$. An average MS speed $v_{\mathbf{MS}} = 10$ km/h and an wavelength $\lambda = \frac{c}{f_0} = 6$ cm are considered, cf. Table 3.11.

Table 3.13. Spaced-time correlation values for the different values of $\gamma_{\text{CSI}}$.

| $\gamma_{\mathbf{CSI}}$ **in dB** | **-20** | **-15** | **-10** | **-5** | **0** | **5** | **10** | **15** | **20** |
|---|---|---|---|---|---|---|---|---|---|
| $\tilde{\tau}$ **in ms** | 7.6 | 7.2 | 6.4 | 5.3 | 3.9 | 2.5 | 1.5 | 0.9 | 0.5 |
| $\frac{2\pi v_{\text{MS}}\tilde{\tau}}{\lambda}$ | 0.35 | 0.33 | 0.30 | 0.24 | 0.18 | 0.12 | 0.07 | 0.04 | 0.02 |
| $J_0\left(\frac{2\pi v_{\text{MS}}\tilde{\tau}}{\lambda}\right)$ | 0.100 | 0.175 | 0.302 | 0.490 | 0.707 | 0.872 | 0.953 | 0.985 | 0.995 |

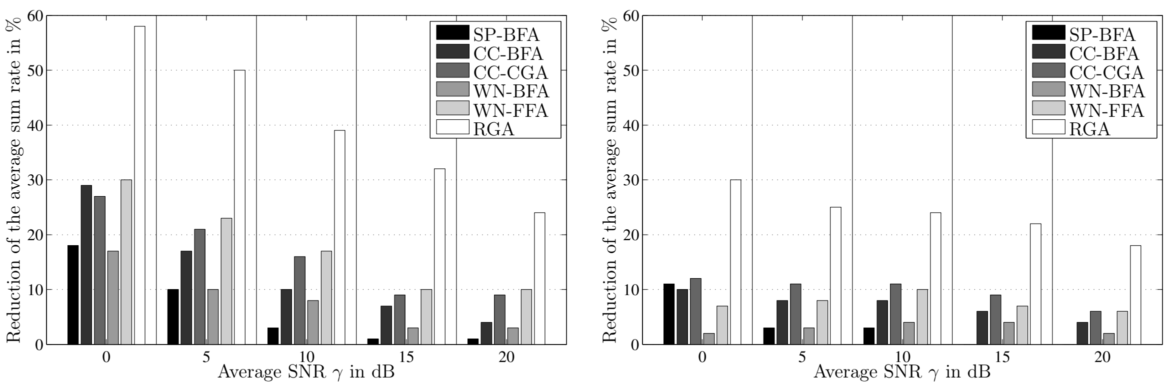
In Table 3.13, note that for $\gamma_{\mathbf{CSI}} \leq -5$ dB, the erroneous B-CSIT, i.e., the delayed B-CSIT, is already highly uncorrelated with the actual B-CSIT. In particular, considering the values shown Table 3.13, the results shown in Fig. 3.9(c) and Fig. 3.9(d) can be easily compared to the results in [YG05].

Similarly to the results presented in Fig. 3.7 and Fig. 3.8, also in Fig. 3.9 the RA strategies employing the ZF algorithm achieve almost the same performance in terms of average sum rate as the corresponding RA strategies employing the GEP algorithm. Larger reductions in the average sum rate for higher average SNR values than for lower average SNR values are also observed and are due to the same reasons as discussed in Section 3.5.3. Considering the results presented in Fig. 3.7, Fig. 3.8, and in Fig. 3.9, it can be concluded that the ZF and GEP algorithms do not offer any particular advantage over each other regarding the different CSIT models. Consequently, RA strategies employing a simple precoding and power allocation algorithm, such as the ZF algorithm of Table 3.8, can perform as good as RA strategies employing an iterative joint precoding and power allocation algorithm, such as the GEP algorithm of Table 3.9 and reach almost the same average sum rate achieved by the CAP-ESA-ZF strategy, as shown in Fig. 3.7.

### 3.5.6 Impact of the Sequential Removal Algorithm on the Performance

In this section, the impact of the SRA on the performance of the RA strategies is discussed. In Section 3.5.3, Section 3.5.4, and Section 3.5.5, it has been observed that the RA strategies employing rather simple SDMA algorithms, such as the RGA, reach over 70% of the average sum rate achieved by the CAP-ESA-ZF strategy, which was a considerably good performance. In Section 3.5.3, Section 3.5.4, and Section 3.5.5, it has also been observed that the GEP algorithm of Table 3.9 provided no advantages compared to the more simple ZF algorithm. However, the similar performances observed in Section 3.5.3, Section 3.5.4, and Section 3.5.5 are intrinsically related to the SRA, as it is discussed in the sequel.

The SRA is responsible for a considerable fraction of the average sum rate achieved by the RA strategies that use the SP-BFA, CC-BFA, CC-CGA, WN-BFA, WN-FFA, or RGA, which are SDMA algorithms unaware of the actual precoding and power allocation. For these RA strategies, Fig. 3.10 shows the percentual reduction of the average sum rate values presented in Fig. 3.7 that result if the SRA is disabled and a target SDMA group size $K_{\mathcal{G}}^{\star} = 4$ MSs is considered.




(a) ZF-WFA. (b) GEP-SDA.

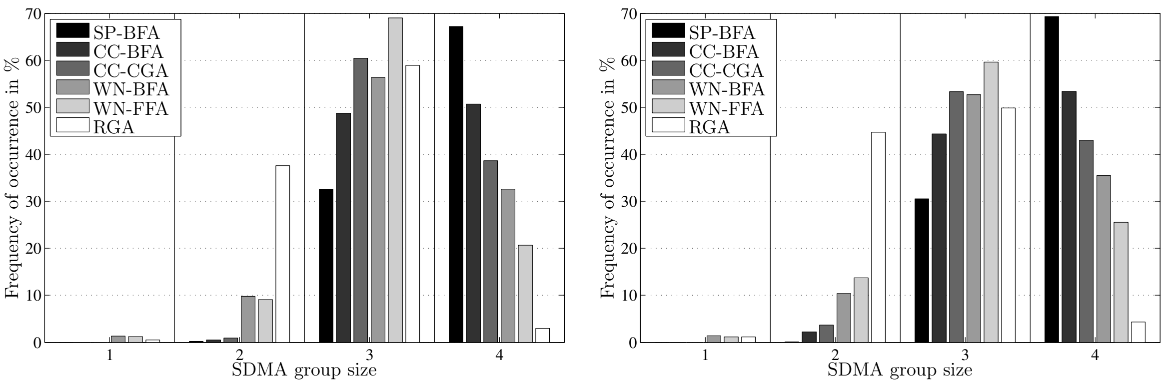
Figure 3.10. Percentual reduction of the average sum rate of the RA strategies when the SRA is disabled. Target SDMA group size $K_{\mathcal{G}}^{\star} = 4$.

For low average SNR values, it can be seen in Fig. 3.10(a) that disabling the SRA results in reductions of about 10% to 30% of the average sum rates for most of the RA strategies. In particular, for the RGA-ZF-WFA strategy this reduction is above 50% for low average SNR values and, consequently, the RGA-ZF-WFA strategy would achieve only 35% of the average sum rate obtained by CAP-ESA-ZF strategy in this case. Thus, the good performance of the RGA-ZF-WFA in Fig. 3.7 is in great part due to the SRA. Due to their design, the SDMA algorithms considered in Fig. 3.10 cannot estimate the group capacity and, consequently, the SDMA groups that they build have suboptimal size and are suboptimal in terms of group capacity. The SRA improves the average sum rate achieved by these RA strategies by removing MSs from the SDMA groups and keeping the group with the best capacity, as discussed in Section 3.3.3.

In the following, the impact of the SRA on the RA strategies employing the GEP algorithm is discussed. Comparing Fig. 3.10(b) and Fig. 3.10(a), it can be noted that the reductions of the average sum rate of the RA strategies considering the GEP-SDA without the SRA are about the half of that observed when using the ZF-WFA without the SRA too. Except for the RGA-GEP-SDA without the SRA, the RA strategies in Fig. 3.10(b) are able to achieve about 85% or more of the average sum rate that they achieve using the SRA. Without the SRA, the better performance of the GEP-SDA in Fig. 3.10(b) compared to the ZF-WFA in Fig. 3.10(a) results from the fact that GEP and SDA iteratively establish a trade-off between spatial separation and interference among MSs in the SDMA group. Thus, the prejudicial effect of having too large SDMA groups is effectively compensated by a more efficient adjustment of precoding vectors and power allocation. In fact, it can be said that the GEP-SDA is a kind of MMSE precoding approach [BS05] and, as such, shall outperform ZF precoding for low SNR values and have almost the same performance as ZF for high SNR values [PNG03, MBQ04]. Anyway, differently from the standard MMSE linear precoding [JUN05], the GEP-SDA is relatively more complex since it relies on an iterative procedure that requires to solve $K_{\mathcal{G}}^{\star}$ generalized eigenvalue problems.

In the following, the impact of an adequate selection of the target SDMA group size is discussed for the case in which the SRA is disabled. The performance of the SDMA algorithms considered in Fig. 3.10 is strongly dependent on the selection of the target SDMA group size. This dependency can be recognized by observing the decrease of the average sum rate losses in Fig. 3.10 when the average SNR increases. For higher average SNR values, larger group are supported and the removals performed by SRA have a lower impact on the overall performance of the RA strategies. Thus, in Fig. 3.10(a), the reason for reductions of the average sum rate are the same as discussed in Section 3.3.3, i.e., the unnecessary null space projections with respect to MSs that get no power allocated. In Fig. 3.10(b), the impact of an inadequate group size is compensated by the iterative joint precoding and power allocation.

If the optimum target SDMA group size is known a priori and is employed by the SDMA algorithms, the impact of the SRA on the average sum rate achieved by the RA strategies becomes less relevant. Considering an average SNR $\gamma = 10$ dB, Fig. 3.11 shows the distribution of the group size obtained after applying the SRA for each of the RA strategies in Fig. 3.10.

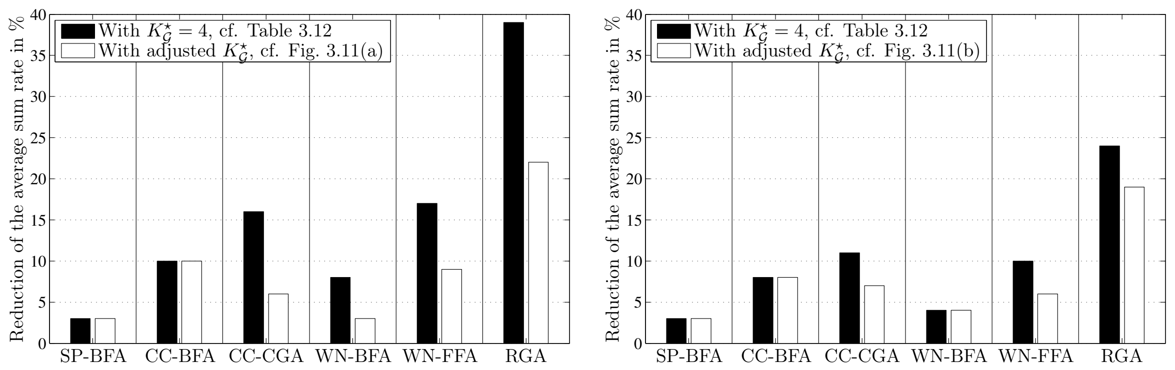



(a) ZF-WFA-LGF (b) GEP-SDA-HPF

Figure 3.11. SDMA group size distribution for RA strategies unaware of the actual precoding and power allocation algorithm. Average SNR $\gamma = 10$ dB.

In Fig. 3.11, it can be noted that often an SDMA group size $K_{\mathcal{G}}^{\star} \leq 4$ results for most of the RA strategies. In order to verify the impact of an adequate selection of $K_{\mathcal{G}}^{\star}$, the reductions of the average sum rate achieved by each RA strategy considering $K_{\mathcal{G}}^{\star} = 4$, cf. Table 3.12, and considering the most frequent group size $K_{\mathcal{G}}^{\star}$, cf. Fig. 3.11, are compared. In both cases, the SRA is disabled and an average SNR $\gamma = 10$ dB is considered. Fig. 3.12 shows the reduction of the average sum rate achieved by the RA strategies for $K_{\mathcal{G}}^{\star} = 4$, cf. Table 3.12, and with $K_{\mathcal{G}}^{\star}$ set to the most frequent value according to the results shown in Fig. 3.11.

In Fig. 3.12, it can be seen that in the cases in which $K_{\mathcal{G}}^{\star} = 4$ is also the most frequent group size, cf. Fig. 3.11, the reductions of the average sum rate remain naturally unchanged, while in the other cases the reductions of the average sum rate of the RA




(a) ZF-WFA. (b) GEP-SDA.

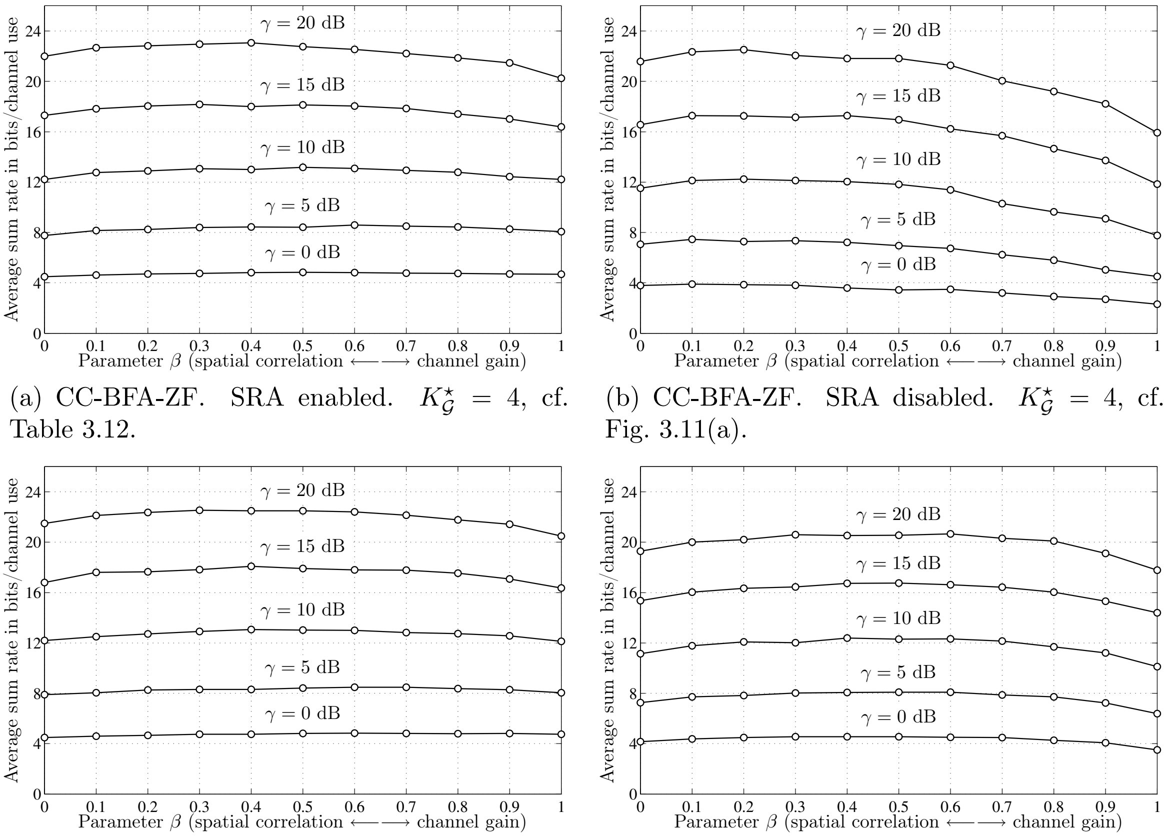
Figure 3.12. Fraction of the average sum rate achieved by the RA strategies without the SRA for $K_{\mathcal{G}}^{\star} = 4$, cf. Table 3.12, and for adjusted $K_{\mathcal{G}}^{\star}$, cf. Fig. 3.11. Average SNR $\gamma = 10$ dB.

strategies with $K_{\mathcal{G}}^{\star} = 4$ are often the double of that when $K_{\mathcal{G}}^{\star}$ is selected as the most frequent group size, cf. Fig. 3.11. Except for the RGA-ZF-WFA strategy in Fig. 3.12(a) and the RGA-GEP-SDA strategy in Fig. 3.12(b), the RA strategies in Fig. 3.12 lose 10% or less of the average sum rates that they achieve with the SRA enabled if the target SDMA group size is adequately selected. However, as discussed in Section 3.3.3, the ideal value of $K_{\mathcal{G}}^{\star}$ cannot be determined a priori and an SRA is required. Anyway, adequate values for $K_{\mathcal{G}}^{\star}$ could be learnt by the system, as suggested e.g. in [FDH05, FDH07] and by the author in [MK07a, MK07c], and considerable computational effort could be saved by considering a fixed target SDMA group size or a small set of target SDMA group size values to be tested by the SRA [FDH05, MK06, FDH07, MK07a, MK07c]. Nevertheless, considering the results presented in Fig. 3.10(b) and Fig. 3.12(b), the application of an MMSE-like algorithm for precoding is recommended whenever a fixed target SDMA group size $K_{\mathcal{G}}^{\star}$ is used, since the precoding algorithm partially compensates the usage of suboptimal values of $K_{\mathcal{G}}^{\star}$, as it was discussed in this section for the GEP-SDA of Table 3.9.

### 3.5.7 Impact of the Parameter $\beta$ on the Performance

In this section, the impact of the parameter $\beta$ on the performance of the RA strategies that use the CC-BFA and CC-CGA is investigated. As discussed in Section 3.2.1, the parameter $\beta$ of (3.13) controls the relevance given to the spatial correlation and channel gains of the MSs in an SDMA group.

Fig. 3.13 shows the average sum rate achieved by the CC-BFA-ZF and CC-CGA-ZF strategies for varying $\beta$ and different values of the average SNR $\gamma$. Fig. 3.14 shows the equivalent results for the CC-BFA-GEP and CC-CGA-GEP strategies. The results on the left and right sides of Fig. 3.13 and Fig. 3.14 consider the SRA enabled and disabled, respectively. For the cases in which the SRA is disabled the target SDMA group size $K_{\mathcal{G}}^{\star}$ has been set to 4 in Fig. 3.13(b) and Fig. 3.14(b), and to 3 in Fig. 3.13(d) and




(a) CC-BFA-ZF. SRA enabled. $K_{\mathcal{G}}^{\star} = 4$, cf. Table 3.12.

(b) CC-BFA-ZF. SRA disabled. $K_{\mathcal{G}}^{\star} = 4$, cf. Fig. 3.11(a).

(c) CC-CGA-ZF. SRA enabled. $K_{\mathcal{G}}^{\star} = 4$, cf. Table 3.12.

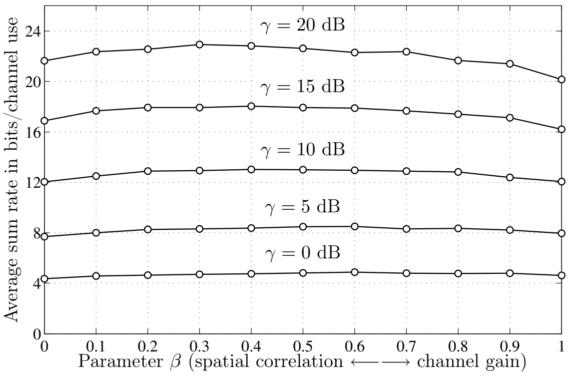
(d) CC-CGA-ZF. SRA disabled. $K_{\mathcal{G}}^{\star} = 3$, cf. Fig. 3.11(a).

Figure 3.13. Impact of the parameter $\beta$ on the average sum rate of the CC-CGA-ZF and CC-BFA-ZF strategies. Left: SRA is enabled. Right: SRA is disabled.

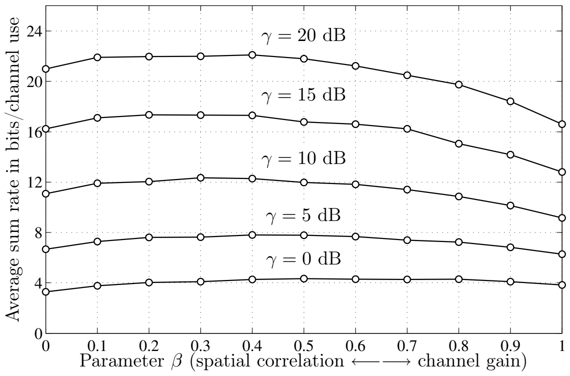
Fig. 3.14(d), for the RA strategies employing the CC-BFA and CC-CGA, respectively, in a similar way as it has been done for the results shown in Fig. 3.11.

In both Fig. 3.13 and Fig. 3.14, it can be seen that the average sum rates achieved by the RA strategies are relatively independent of the parameter $\beta$. In Fig. 3.13 and Fig. 3.14, it can be seen that good results in terms of average sum rate can be obtained by the RA strategies by employing the scaling factors $\frac{1}{\|\mathbf{C}\|_{\mathrm{F}}}$ and $\frac{1}{\|\mathbf{a}\|_{\mathrm{F}}}$ adopted in this work and a value of $\beta$ around 0.5. Anyway, compared to the extreme cases in which $\beta = 0$ or $\beta = 1$, it can be seen that some gain can still be obtained by varying $\beta$, especially if the SRA is not used, as in Fig. 3.13(b), Fig. 3.13(d), Fig. 3.14(b), and Fig. 3.14(d). Thus, by adjusting $\beta$ and finding an adequate trade-off between spatial correlation and channel gain, the CC-CGA and the CC-BFA proposed herein can outperform SDMA algorithms that take into account only the spatial correlation among the MSs, as those in [STKL01, Cal04], or only the channel gains, as that in [TJ05], in terms of the achieved average sum rate.

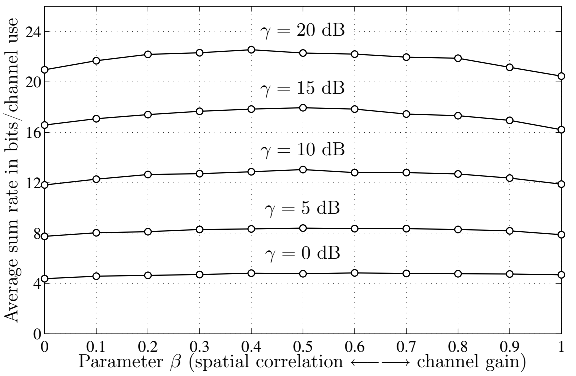
Analogously to the analyses in Section 3.3.3 regarding the SRA, there is also a dependency between the target SDMA group size $K_{\mathcal{G}}^{\star}$ and the parameter $\beta$, which is discussed in the sequel. Comparing the average sum rates obtained by the RA strategies with and without the SRA in Fig. 3.13 and Fig. 3.14, it can be noted that when



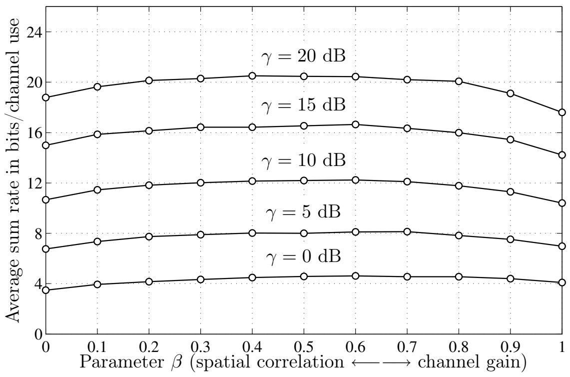

(a) CC-BFA-GEP. SRA enabled. $K_{\mathcal{G}}^{\star} = 4$, cf. Table 3.12.




(b) CC-BFA-GEP. SRA disabled. $K_{\mathcal{G}}^{\star} = 4$, cf. Fig. 3.11(b).




(c) CC-CGA-GEP. SRA enabled. $K_{\mathcal{G}}^{\star} = 4$, cf. Table 3.12.




(d) CC-CGA-GEP. SRA disabled. $K_{\mathcal{G}}^{\star} = 3$, cf. Fig. 3.11(b).

Figure 3.14. Impact of the parameter $\beta$ on the average sum rate of the CC-CGA-GEP and CC-BFA-GEP strategies. Left: SRA is enabled. Right: SRA is disabled.

the SRA is enabled, the parameter $\beta$ loses relevance. This occurs because the SDMA group size is adjusted by the SRA afterwards. For high average SNR values, larger SDMA groups result after the application of the SRA and the relevance of the spatial correlation among MSs in such large groups is higher than that of the channel gains. Consequently, the average sum rate of the RA strategies for values of $\beta$ around 0.5 are slightly better than for values of $\beta$ near 1.0, as it can be seen in Fig. 3.13(a), Fig. 3.13(c), Fig. 3.14(a), and Fig. 3.14(c). For low SNR values, smaller SDMA groups containing only one or two MSs result after the application of the SRA and relevance of the spatial correlation among MSs in such small groups is lower than that of the channel gains. Consequently, almost the same performance in terms of average sum rate is obtained independent of the value of $\beta$.

For the results in Fig. 3.13(b), Fig. 3.13(d), Fig. 3.14(b), and Fig. 3.14(d), the SRA is disabled and the target SDMA group size $K_{\mathcal{G}}^{\star}$ is fixed. Consequently, relatively large SDMA groups are always built and spatial correlation among MSs becomes more relevant than the channel gains. In particular, in Fig. 3.13(b) and Fig. 3.14(b), it can be clearly noted that values of $\beta \leq 0.5$ lead to better performance in terms of average sum rate than higher values of $\beta$. In Fig. 3.13(d) and Fig. 3.14(d), the same trend is less pronounced because the considered target SDMA group size is smaller. Moreover, the CGA builds a whole SDMA group at once using convex optimization while the BFA considered in Fig. 3.13(b) and Fig. 3.14(b) builds the group iteratively, as discussed in Section 3.2.2. Consequently, the groups built by each SDMA algorithm may differ, as well as the best values of $\beta$ for RA strategies using CC-CGA and CC-BFA.

### 3.5.8 Complexity Analysis

In this section, the complexity of some of the RA strategies of Table 3.10 is discussed. The results presented in Section 3.5.3, Section 3.5.4, and Section 3.5.5 have shown that the RA strategies of Table 3.10 have quite similar performances in terms of average sum rate. However, these RA strategies have quite different computational costs and, consequently, it is possible to select RA strategies with low complexity that are able to attain a high fraction of the maximum average sum rate of the system.

The number of complex multiplications required by an RA strategy or algorithm is used here to measure its complexity. Further on, the term complex will be omitted when referring to complex multiplications. Because square roots and divisions have the same complexity than a multiplication if they are efficiently implemented using Newton's method [BV04], they are counted as such. Additions, subtractions, and logical operations are not taken into account. The algorithms are assumed to be implemented as efficiently as possible. Repeated operations within an algorithm do not increase its complexity, since the results of the operations can be stored and reused within the algorithm whenever necessary. For a given problem, it is often of interest to know the complexity order $\mathcal{O}(\cdot)$ of an algorithm used to solve the problem as a function of the problem dimensions. The complexity order $\mathcal{O}(\cdot)$ of an algorithm considers the big O [GL96], can be approximated by the dominant term in the expression describing the complexity of the algorithm [Leu04], and is often more useful than having a complicated complexity expression. The complexities of some mathematical operations and functions are listed in Table A.1 in Appendix A and are used in the determination of the complexity of the RA strategies presented in this section.

In order to estimate the complexity of the RA strategies, the complexity in number of multiplications required by certain operations and algorithms involved in the RA strategies has been expressed in Table 3.14. The complexity of these operations and algorithms is a function of the number $K$ of MSs, the number $M$ of transmit antennas, and the size $K_{\mathcal{G}}$ of an SDMA group $\mathcal{G}$.

The complexity analyses considered in this section take into account only a subset of the RA strategies of Table 3.10. Indeed, only the complexity of the RA strategies employing the ZF algorithm of Table 3.8 is considered in this section for the following reasons. The GEP algorithm of Table 3.9 is an iterative procedure and, to converge, it takes a certain number of iterations (usually $< 10$), denoted here by $I_{\text{SDA}}$. Considering the expressions in Table 3.14 and the fact that $K_{\mathcal{G}}$ is upper limited by $M$, the complexity order $\mathcal{O}\left(\frac{I_{\text{SDA}} 7M^4}{6}\right)$ of the GEP of Section 3.3.1 is much higher than the complexity order $\mathcal{O}\left(\frac{3M^3}{2}\right)$ of the ZF precoding of the same section. Analogously, the complexity

Table 3.14. Complexity of certain operations and algorithms involved in the RA strategies.

| Operation/Algorithm | Number of multiplications | $\mathcal{O}(\cdot)$ | Remark |
|---|---|---|---|
| **Operations related to the grouping metrics** | | | |
| $\mathbf{T}_i$ of (3.3) | $\frac{5M^2+5M+2}{2}$ | $\mathcal{O}\left(\frac{5M^2}{2}\right)$ | $i>1$ |
| $\mathbf{a}$ of (3.11a) | $KM+K$ | $\mathcal{O}(KM)$ | - |
| $\mathbf{C}$ of (3.11b) | $\frac{K^2M+4K^2-KM-2K}{2}$ | $\mathcal{O}\left(\frac{K^2M}{2}\right)$ | Known $\mathbf{a}$ |
| $\frac{(1-\beta)\mathbf{C}}{\lVert\mathbf{C}\rVert_{\mathrm{F}}}$ and $\frac{\beta\mathbf{a}}{\lVert\mathbf{a}\rVert_{\mathrm{F}}}$ in (3.13) | $2K^2+K+3$ | $\mathcal{O}(2K^2)$ | Known $\mathbf{C}$ and $\mathbf{a}$ |
| **Grouping metrics** | | | |
| $f_{\mathrm{CAP}}(\mathcal{G})$ of (3.2) | $2K_{\mathcal{G}}^2(M+1)+K_{\mathcal{G}}$ | $\mathcal{O}(2K_{\mathcal{G}}^2M)$ | Known $\mathbf{w}_i$ and $p_i$ |
| $f_{\mathrm{WN}}(\mathcal{G})$ of (3.10) | $K_{\mathcal{G}}^2+K_{\mathcal{G}}+3$ | $\mathcal{O}(K_{\mathcal{G}}^2)$ | Known $\mathbf{C}$ and $\mathbf{a}$ |
| $f_{\mathrm{WN}}(\mathcal{G}')$, cf. Section 3.2.3 | $K_{\mathcal{G}'}+8$ | $\mathcal{O}(K_{\mathcal{G}'})$ | Known $\mu_{\mathcal{G}}$, $\lVert\mathbf{C}_{\mathcal{G}}\rVert_{\mathrm{F}}$, $\mathbf{C}$ and $\mathbf{a}$ |
| $f_{\mathrm{CC}}(\mathcal{G})$ of (3.13) | $K^2+2K$ | $\mathcal{O}(K^2)$ | Known $\frac{(1-\beta)}{\lVert\mathbf{C}\rVert_{\mathrm{F}}}\mathbf{C}$ and $\frac{\beta}{\lVert\mathbf{a}\rVert_{\mathrm{F}}}\mathbf{a}$ |
| $\frac{\partial f_{\mathrm{CC}}(\mathcal{G})}{\partial\mathbf{u}}$ | $K^2$ | $\mathcal{O}(K^2)$ | Known $\mathbf{C}$ and $\mathbf{a}$ |
| **Precoding and power algorithm** | | | |
| ZF, cf. Section 3.3.1 | $\frac{K_{\mathcal{G}}^3+3K_{\mathcal{G}}^2(M+1)+5K_{\mathcal{G}}M+4K_{\mathcal{G}}}{2}$ | $\mathcal{O}\left(\frac{K_{\mathcal{G}}^3+3K_{\mathcal{G}}^2M}{2}\right)$ | - |
| GEP, cf. Section 3.3.1 | $\frac{7K_{\mathcal{G}}M^3+12K_{\mathcal{G}}M^2+9K_{\mathcal{G}}M+6K_{\mathcal{G}}}{6}+\mathcal{O}(K_{\mathcal{G}}M^2)$ | $\mathcal{O}\left(\frac{7K_{\mathcal{G}}M^3}{6}\right)$ | Per iteration |
| WFA of Table 3.5 | $2K_{\mathcal{G}}M+6K_{\mathcal{G}}$ | $\mathcal{O}(2K_{\mathcal{G}}M)$ | - |
| SDA of Table 3.6 | $2K_{\mathcal{G}}M^2+2K_{\mathcal{G}}M+9K_{\mathcal{G}}+10$ | $\mathcal{O}(2K_{\mathcal{G}}M^2)$ | Per iteration |
| SRA of Table 3.7 | $KM+6M+10+\sum_{K_{\mathcal{G}}=2}^{K_{\mathcal{G}}^{\star}}\frac{K_{\mathcal{G}}^3+7K_{\mathcal{G}}^2(M+1)+9K_{\mathcal{G}}M+18K_{\mathcal{G}}}{2}$ | $\mathcal{O}\left(\frac{K_{\mathcal{G}}^{\star 3}+7K_{\mathcal{G}}^{\star 2}M}{2}\right)$ | - |

order $\mathcal{O}(I_{\mathrm{SDA}}2M^3)$ of the SDA is also larger than then complexity order $\mathcal{O}(2M^2)$ of the WFA. Even for a single iteration of the GEP algorithm of Table 3.9, it would be more complex than the ZF algorithm of the same table. Additionally, the GEP algorithm of Table 3.9 still needs to determine the optimal DL powers, cf. (3.51). Finally, the results previously presented have shown that the performance of the RA strategies is as good with the ZF algorithm as with the GEP algorithm.

The complexity of the SRA shown in Table 3.14 is not take into account in the complexity of the RA strategies discussed in this section due to the following reasons. Firstly, the complexity of the SRA might be eventually saved, or at least reduced, by employing adequate target SDMA group sizes, as discussed in Section 3.3.3. Secondly, except for the CAP-ESA-ZF and the CAP-BFA-ZF strategies, all the other RA strategies employ the SRA so that its complexity represents mainly a common offset in the complexity of these strategies. It can be noted that the SRA is relatively similar to the CAP-BFA-ZF strategy, but sequentially removes MSs from the group instead of sequentially admitting MSs to the group as the CAP-BFA-ZF strategy does. However, because the $K_{\mathcal{G}}^{\star}$ is usually smaller than $K$, the number of group capacity calculations done by the SRA is considerably smaller than that performed by the CAP-BFA-ZF strategy. Consequently, omitting the complexity of the SRA will not lead to a considerable underestimation of the complexity of the other RA strategies compared to the CAP-BFA-ZF strategy.

The following assumption is made regarding the complexity of the CC-CGA of Table 3.4. According to [BV04], the complexity order of a quadratic optimization problem like the problem (3.17) of the CC-CGA is roughly proportional to the cost of evaluating its cost function $f_{\mathrm{CC}}(\mathcal{G})$ and the first derivative of $f_{\mathrm{CC}}(\mathcal{G})$, yielding $2K^2+2K$ multiplications, cf. Table 3.14. Furthermore, the complexity of the CC-CGA depends on the number $I_{\mathrm{CGA}}$ of iterations required by the CGA to converge. Here, one evaluation of the cost function and of its first derivative are assumed per iteration of the CC-CGA.

Because all RA strategies considered in this section employ the ZF algorithm and because the complexity of the SRA is not considered, the complexity of the RA strategies will mainly differ due to the different SDMA algorithms that they employ. Moreover, the RA strategies using the SP-BFA, CC-BFA, CC-CGA, WN-BFA, WN-FFA, or RGA need to compute the precoding vectors and the allocated powers only for the final group while the RA strategies using the CAP-ESA or CAP-BFA need to compute precoding vectors and the allocated powers several times since they employ the group capacity $f_{\mathrm{CAP}}(\mathcal{G})$ as grouping metric.

Considering the previous remarks, the complexity of the RA strategies studied in this section is presented in Table 3.14. The RA strategies are listed in Table 3.14 in decreasing order of complexity.

Table 3.15. Complexity of the RA strategies.

| **RA strategy** | **Number of multiplications** | $\mathcal{O}(\cdot)$ |
|---|---|---|
| CAP-ESA-ZF | $\sum_{K_{\mathcal{G}}=2}^{K_{\mathcal{G}}^{\star}} \frac{K!}{K_{\mathcal{G}}!(K-K_{\mathcal{G}})!}\left(\frac{K_{\mathcal{G}}(K_{\mathcal{G}}^2+7K_{\mathcal{G}}(M+1)+9M+18)}{2}\right)+(KM+6M+10)$ | Non-Polynomial |
| CAP-BFA-ZF | $\sum_{K_{\mathcal{G}}=2}^{K_{\mathcal{G}}^{\star}}(K-K_{\mathcal{G}}+1)\left(\frac{K_{\mathcal{G}}(K_{\mathcal{G}}^2+7K_{\mathcal{G}}(M+1)+9M+18}{2}\right)+(KM+6M+10)$ | $\mathcal{O}\left(\frac{K{K_{\mathcal{G}}^{\star}}^2(K_{\mathcal{G}}^{\star}+7M)}{2}\right)$ |
| CC-CGA-ZF | $I_{\mathrm{CGA}}(2K^2+2K)+\frac{K^2(M+8)+K(M+2)+6}{2}+\left(KM+\frac{K_{\mathcal{G}}^{\star}({K_{\mathcal{G}}^{\star}}^2+3K_{\mathcal{G}}^{\star}(M+1)+9M+16)}{2}\right)$ | $\mathcal{O}\left(\frac{K^2(M+4I_{\mathrm{CGA}})}{2}\right)$ |
| SP-BFA-ZF | $\sum_{K_{\mathcal{G}}=2}^{K_{\mathcal{G}}^{\star}}(K-K_{\mathcal{G}}+1)\left(\frac{5M^2+5M+2}{2}\right)+\left(KM+\frac{K_{\mathcal{G}}^{\star}({K_{\mathcal{G}}^{\star}}^2+3K_{\mathcal{G}}^{\star}(M+1)+9M+16)}{2}\right)$ | $\mathcal{O}\left(\frac{5KM^2}{2}\right)$ |
| CC-BFA-ZF | $\frac{K_{\mathcal{G}}^{\star}}{K}\,\frac{K^2(M+8)+K(M+2)+6}{2}+\left(KM+\frac{K_{\mathcal{G}}^{\star}({K_{\mathcal{G}}^{\star}}^2+3K_{\mathcal{G}}^{\star}(M+1)+9M+16)}{2}\right)$ | $\mathcal{O}\left(\frac{KK_{\mathcal{G}}^{\star}(M+8)}{2}\right)$ |
| WN-BFA-ZF | $\sum_{K_{\mathcal{G}}=2}^{K_{\mathcal{G}}^{\star}}(K-K_{\mathcal{G}}+1)(K_{\mathcal{G}}+8)+\left(K+\frac{K_{\mathcal{G}}^{\star}}{K}\,\frac{K^2(M+4)+KM}{2}+KM+\frac{K_{\mathcal{G}}^{\star}({K_{\mathcal{G}}^{\star}}^2+3K_{\mathcal{G}}^{\star}(M+1)+9M+16)}{2}\right)$ | $\mathcal{O}\left(\frac{KK_{\mathcal{G}}^{\star}(M+4)}{2}\right)$ |
| WN-FFA-ZF | $\sum_{K_{\mathcal{G}}=2}^{K_{\mathcal{G}}^{\star}}(K_{\mathcal{G}}+8)+\left(K+\frac{K_{\mathcal{G}}^{\star}}{K}\,\frac{K^2(M+4)+KM}{2}+KM+\frac{K_{\mathcal{G}}^{\star}({K_{\mathcal{G}}^{\star}}^2+3K_{\mathcal{G}}^{\star}(M+1)+9M+16)}{2}\right)$ | $\mathcal{O}\left(\frac{KK_{\mathcal{G}}^{\star}(M+4)}{2}\right)$ |
| RGA-ZF | $\frac{K_{\mathcal{G}}^{\star}({K_{\mathcal{G}}^{\star}}^2+3K_{\mathcal{G}}^{\star}(M+1)+9M+16)}{2}$ | $\mathcal{O}\left(\frac{{K_{\mathcal{G}}^{\star}}^3+3{K_{\mathcal{G}}^{\star}}^2M}{2}\right)$ |

In the following, the complexity order of the RA strategies is discussed considering $K_{\mathcal{G}}^{\star}=M$. Because of the number $G$ of SDMA groups given by (2.14) that are considered by the ESA, the complexity of the CAP-ESA-ZF strategy increases combinatorially in $K$. Moreover, the group capacity and, consequently, precoding vectors and allocated powers, have to be computed for each of these $G$ SDMA groups. Both these facts cause the complexity of the CAP-ESA-ZF strategy to be non-polynomial and extremely high.

The CAP-BFA-ZF strategy has also a relatively high complexity, since for each candidate group, precoding vectors and power allocation have to be computed. Considering

$K_{\mathcal{G}}^{\star} = M$, its complexity order is $\mathcal{O}\left(\frac{7KM^3}{2}\right)$, which causes its complexity to become particularly high when considering large AAs.

Differently from CAP-BFA-ZF strategy, the complexity order $\mathcal{O}\left(\frac{K^2(M+4I_{\text{CGA}})}{2}\right)$ of the CC-CGA-ZF strategy increases only linearly with $M$, but increases quadratically with $K$. Thus, for large numbers of MSs, the complexity of this strategy increases rapidly. Nevertheless, as it will be seen later in this section, for large AA and low to moderate numbers of MSs, the CC-CGA-ZF strategy can be less complex than the CAP-BFA-ZF and SP-BFA-ZF strategies.

The SP-BFA-ZF strategy has a complexity order $\mathcal{O}\left(\frac{5KM^2}{2}\right)$ which increases only linearly with $K$, but quadratically with $M$. Similarly to the CAP-BFA-ZF strategy, the complexity of the SP-BFA-ZF strategy becomes particularly high when considering large AAs.

The CC-BFA-ZF strategy makes use of the same input data as the CC-CGA-ZF strategy, i.e., the scaled versions of $\mathbf{C}$ and $\mathbf{a}$ used in (3.13) and (3.24). However, the CC-BFA-ZF strategy employs the BFA which requires only a fraction $\frac{K_{\mathcal{G}}^{\star}}{K}$ of the scaled versions of $\mathbf{C}$ and $\mathbf{a}$ and the used grouping metric given by (3.24) does not involve any further multiplication whenever the scaled versions of $\mathbf{C}$ and $\mathbf{a}$ are known. Consequently, the CC-BFA-ZF strategy is less complex than the CC-CGA-ZF strategy. The CC-BFA-ZF strategy has a complexity order $\mathcal{O}\left(\frac{KM^2}{2}\right)$ that increases quadratically with $M$, so that it is also less complex than the CAP-BFA-ZF strategy.

The WN-BFA-ZF and WN-FFA-ZF strategies also require only a fraction $\frac{K_{\mathcal{G}}^{\star}}{K}$ of $\mathbf{C}$ and $\mathbf{a}$. However, they consider smaller and non-scaled spatial correlation matrices and attenuation vectors, cf. Section 3.2.1, and, consequently, have lower complexity than the CC-BFA-ZF strategy. Both WN-BFA-ZF and WN-FFA-ZF strategies have complexity order $\mathcal{O}\left(\frac{KM^2}{2}\right)$. However, because the BFA is more complex than the FFA, the WN-BFA-ZF strategy is also more complex than the WN-FFA-ZF strategy.

The RGA-ZF is of course the least complex among the RA strategies. It involves only the precoding and power allocation algorithm applied to the final group and, therefore, is independent of the number $K$ of MSs and has a complexity order $\mathcal{O}\left(\frac{3M^3}{2}\right)$.

In the following, the complexity of the CAP-BFA-ZF, SP-BFA-ZF, CC-CGA-ZF, and CC-BFA-ZF strategies are compared. They have been selected since they have very similar performances which surpass 90% of the average sum rate obtained by the CAP-ESA-ZF strategy and are also superior to that of the remaining RA strategies in Table 3.15. Therefore, the CAP-BFA-ZF, SP-BFA-ZF, CC-CGA-ZF, and CC-BFA-ZF strategies offer a good trade-off between capacity and complexity. For simplicity, the target SDMA group size $K_{\mathcal{G}}^{\star} = M$ has been assumed.

The complexity of the CC-CGA-ZF strategy depends on the number $I_{\text{CGA}}$ of iterations required by the CC-CGA to converge. By counting $I_{\text{CGA}}$ for a large number of runs of the

CC-CGA-ZF strategy considering the parameter values in Table 3.12, the distribution of $I_{\text{CGA}}$ has been obtained and is presented in Fig. 3.15.

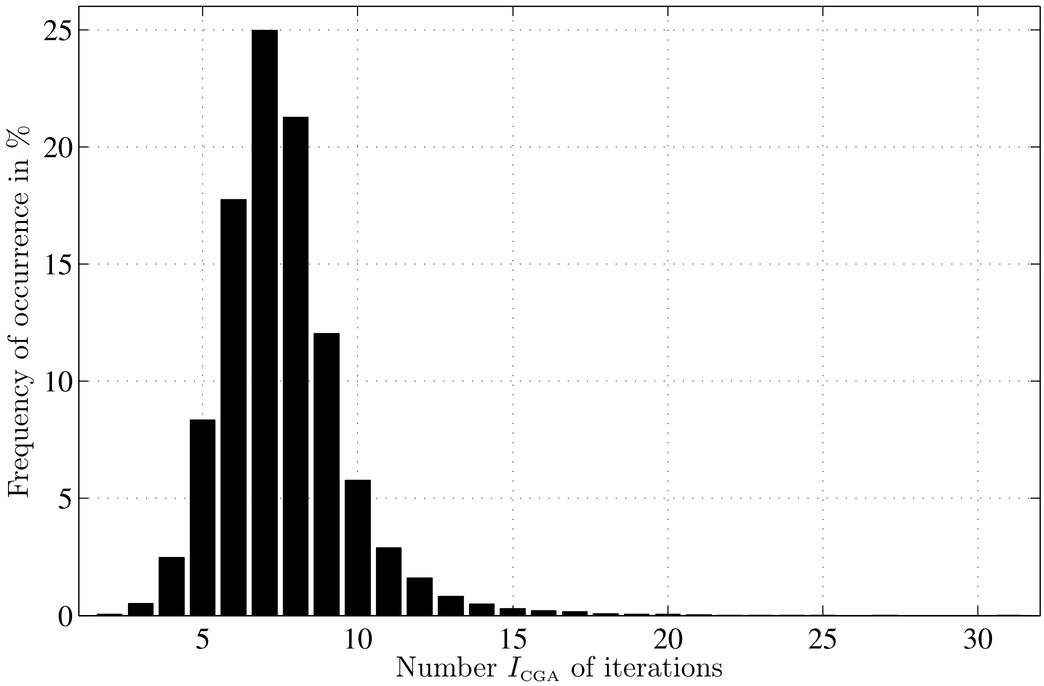



Figure 3.15. Distribution of the number $I_{\text{CGA}}$ of iterations required by the CC-CGA.

In Fig. 3.15, it can be seen that the CC-CGA requires in average a number $\bar{I}_{\text{CGA}} \approx 8 = \frac{K}{2}$ of iterations to converge when considering the parameter values in Table 3.12. Based on this result and in order to simplify the complexity analyses, it is assumed that a number $\bar{I}_{\text{CGA}} \triangleq \frac{K}{2}$ is required by the CC-CGA to converge.

For the complexity comparisons, the number $M$ of transmit antennas has been fixed and the number $K$ of MSs has been varied. Moreover, it is assumed that there are at least $K = 1.5M$ and at most $K = 32$ MSs in the cell. Considering these assumptions, Fig. 3.16 shows the complexity of the RA strategies for different values of $M$ and varying number $K$ of MSs.

In Fig. 3.16, it can be noted that for small values of $M$ the complexity of the CAP-BFA-ZF and SP-BFA-ZF strategies is usually smaller than that of the CC-CGA-ZF. However, the complexity of the CAP-BFA-ZF and SP-BFA-ZF strategies increases faster for larger values of $M$ and larger number $K$ of MSs and surpass the complexity of the CC-CGA-ZF strategy. In particular, the CC-BFA-ZF strategy presents the lowest complexity in all the cases. Considering the average sum rates shown in Fig. 3.7, Fig. 3.8, and Fig. 3.9, and the complexity results shown in Fig. 3.16, it can be noted that the CC-BFA-ZF strategy offers the best performance-complexity trade-off among the RA strategies considered in this section.

In particular, the complexity of the CC-CGA-ZF and CC-BFA-ZF strategies can still be reduced. For the CC-CGA-ZF strategy, the tolerances for the optimal solution of the quadratic optimization algorithm can be adjusted as to find a solution in a shorter

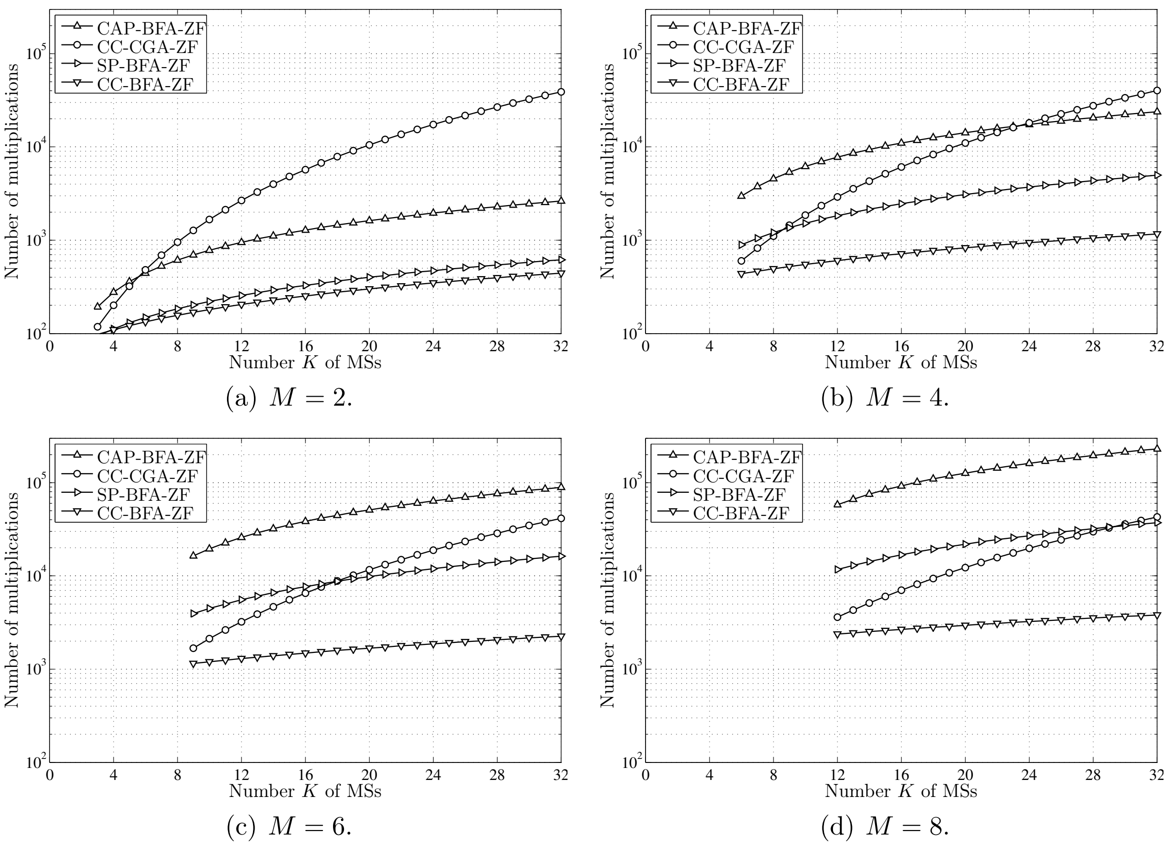



Figure 3.16. Complexity of the CAP-BFA-ZF, CC-CGA-ZF, SP-BFA-ZF, and CC-BFA-ZF strategies for different number $M$ of transmit antennas as a function of the number $K$ of MSs. $1.5M \leq K \leq 32$.

number $I_{\text{CGA}}$ of iterations. A maximum number of iterations can also be imposed, so that a suboptimal group can be found in a shorter time. For both CC-CGA-ZF and CC-BFA-ZF strategies, the scaling factors $\frac{1}{\|\mathbf{C}\|_{\text{F}}}$ and $\frac{1}{\|\mathbf{a}\|_{\text{F}}}$ can be arbitrarily defined, thus avoiding the computation that they involve.

Considering the results presented in this section and in Section 3.5.3, Section 3.5.4, and Section 3.5.5, it can noted that the proposed CC-CGA-ZF and CC-BFA-ZF strategies offer a good trade-off between the achieved average sum rate and computational complexity.

### 3.5.9 Fairness Analysis Considering Different MS Priority Criteria

In this section, the fairness of the RA strategies is investigated considering different criteria for the definition of MS priorities affecting the selection of the initial MS $k'$. QoS aspects are taken into account in this section in order to complement the analysis of the RA strategies considered in this work, which are mainly intended to maximize the sum rate of the system. In this section, it is not an objective to design QoS-oriented RA strategies, but to investigate whether considerable enhancements in terms of throughput fairness can be achieved in the system at the expense of rather small reductions of the average sum rate of the system.

In the following, the consideration of QoS aspects by the RA strategies is shortly motivated. In this chapter, adaptive RA on a frame basis has been considered most of the time. Because B-CSIT has been assumed to be available only for the first TS of each frame, cf. Section 2.2.4, the MSs to which the first TS of a frame is allocated are also the MSs to which all the subsequent TSs of the frame are allocated, which is in accordance with the objective of maximizing the sum rate of the system. Thus, considering B-CSIT and the fact that a single resource is considered in this chapter, the frequency and time components do not play a decisive role in the RA to maximize the sum rate of the system during one allocation period, i.e., one frame. Furthermore, according to the assumptions in Chapter 2, all the MSs will attain the same throughput on a long-term perspective. This is satisfactory if only best-effort services are considered. However, data services having QoS requirements, e.g., in terms of average throughput or maximum delay, are also expected in future mobile radio systems.

The selection of the initial MS $k'$ can be performed according to a different criterion than the one introduced in Table 3.12 in order to follow some QoS-oriented objective, as mentioned in Section 3.5.1. Thus, the consideration of QoS aspects may be incorporated in the RA strategies by adjusting some of their parameters. However, an enhancement of the QoS perceived by the MSs is expected to be obtained at the expense of a reduction of the average sum rate of the system.

In particular, the utilization of a priority criterion for the assignment of MSs or SDMA groups is foreseen in the framework discussed in Section 1.2 and Section 2.3. Indeed, MS priorities are part of the GA algorithm in Fig. 1.3 and have been employed in this chapter. However, they have not been referred to as such because only the maximization of the sum rate has been pursued.

The concept of MS priorities has often been used in time-scheduling algorithms to manage the QoS of the MSs in a system [Zha95, FL02] and it can be directly applied in order to select the initial MS $k'$. In this case, the fairness among the MSs in terms, e.g., of the throughput, can be enhanced. Moreover, it is straightforward to modify the SDMA algorithms of Section 3.2 and the SRA of Section 3.3.3 to ensure that a given MS $k'$ always belongs to the SDMA group assigned on a given resource.

In the following, MS priorities are explicitly defined according to different criteria. Let $\mathfrak{u}_k$ denote the priority of MS $k$. Considering the selection of the initial MS $k'$ in Table 3.12, it can be noted that the MS priority $\mathfrak{u}_k$ being adopted up to now corresponded to

$$\mathfrak{u}_k = \log_2\left(1 + \frac{p_k \|\hat{\mathbf{h}}_i \mathbf{w}_k\|_2}{\sigma^2}\right), \tag{3.53}$$

which is equivalent to have

$$\mathfrak{u}_k = \|\hat{\mathbf{h}}_k\|_2^2. \tag{3.54}$$

Consequently, the initial MS $k'$, which is now defined as the MS with the highest priority, corresponds to

$$k' = \arg\max_k \{\mathfrak{u}_k\}, \tag{3.55}$$

which meets the definition in Table 3.12 when $\mathfrak{u}_k$ is given by (3.54).

Note that according to (3.54), the same MS $k'$ is selected for all the TSs of a frame. In spite of considering the same B-CSIT for the whole frame, performing RA on a TS basis becomes important in this case because different criteria can be adopted to define MS priorities and, consequently, a potentially different MS $k'$ might be selected for each TS in a frame. Further on, the criterion associated with $\mathfrak{u}_k$ in (3.54) will be termed the Capacity Maximization (CM) criterion and two additional criteria for MS priorities are defined in the sequel, namely the Weighted Round Robin (WRR) and the Weighted Proportional Fair (WPF) criteria. Moreover, one allocation period will correspond to one TS further on. However, the same CSIT is considered for all the TSs of a frame.

For the WRR criterion, the initial MS $k'$ is selected according to a Round Robin (RR) policy with static priorities. Let $\tilde{p}$ denote the current allocation period, which should not be confused with the power $p$. Let $\mathcal{G}^\star$ denote the SDMA group to which the considered resource has been assigned in the current allocation period. Let $\mathfrak{u}_k^{(0)}$ and $\mathfrak{u}_k^{(\tilde{p})}$ denote the static priority of MS $k$ and the priority of MS $k$ at the allocation period $\tilde{p}$, respectively. For the WRR criterion, the priority $\mathfrak{u}_k^{(\tilde{p})}$ is written as

$$\mathfrak{u}_k^{(\tilde{p}+1)} = \begin{cases} \mathfrak{u}_k^{(0)}, & \text{for } k \in \mathcal{G}^\star, \\ \mathfrak{u}_k^{(\tilde{p})} + \mathfrak{u}_k^{(0)}, & \text{for } k \notin \mathcal{G}^\star. \end{cases} \tag{3.56}$$

The static priorities in (3.56) are constants determining the fraction of resources that each MS obtains on the long term. If SDMA is not considered, (3.56) ensures that on the long term each MS is allocated a number of resources directly proportional to its static priority and that MSs are served in a WRR way.

For the WPF criterion, let $R_k^{(0)}$ and $\bar{R}_k^{(\tilde{p})}$ denote the contracted and the perceived average throughputs of MS $k$, respectively. Thus, the throughput ratio $\frac{R_k^{(0)}}{\bar{R}_k^{(\tilde{p})}}$ measures how well the MS has met its QoS requirements [SWJO07]. For the WPF criterion, the priority of MS $k$ is written as

$$\mathfrak{u}_k^{(\tilde{p}+1)} = \frac{R_k^{(0)}}{\bar{R}_k^{(\tilde{p})}} \log_2\left(1 + \frac{p_k \|\hat{\mathbf{h}}_k \mathbf{w}_k\|_2}{\sigma^2}\right), \tag{3.57}$$

which is equivalent to set

$$\mathfrak{u}_k^{(\tilde{p}+1)} = \frac{R_k^{(0)}}{\bar{R}_k^{(\tilde{p})}} \|\hat{\mathbf{h}}_k\|_2^2, \tag{3.58}$$

so that the initial MS $k'$ is selected as the one providing the best trade-off between its QoS, represented by the throughput ratio, and its currently achievable rate, similarly as in [SWJO07]. For the same reasons presented for the CM criterion, the rate of the MS can be replaced here by its channel gain, which leads to the selection of the same initial MS $k'$ when using (3.58) instead of (3.57).

In this section, only the initial MS $k'$ is selected according to (3.55), with $\mathfrak{u}_k$ given by (3.54), (3.56) and (3.58) for the CM, WRR, and WPF criteria, respectively. Therefore, MS priority and grouping metric are kept relatively separated from each other. This is a reasonable choice since the SDMA group is not known a priori and the SDMA algorithm should not be penalized by the choice of $\mathfrak{u}_k$. However, because a group of MSs is allocated on the resource instead of a single MS, the priority of all the MSs in $\mathcal{G}^\star$ needs to be updated after each allocation period. In particularly for the WRR criterion, the average fraction of resources that an MS $k$ gets is not $\frac{\mathfrak{u}_k^{(0)}}{\sum_{j=1}^{K}\mathfrak{u}_j^{(0)}}$, as in the case without SDMA, but this average fraction of resources now depends on the spatial compatibility among the MSs, which finally determines the composition of the SDMA groups.

In the following, the relative performance of the SP-BFA-ZF, CC-BFA-ZF, CC-CGA-ZF strategies is investigated considering the different criteria for MS priorities. These three strategies have been selected because they presented the best trade-off between performance in terms of the average sum rate and computational complexity. In order to have different QoS requirements in the system, the MSs have been divided into two categories: the low-priority and the high-priority category. One half of the MSs belongs to each category. For the WRR criterion, $\mathfrak{u}_i^{(0)} = 1$ and $\mathfrak{u}_j^{(0)} = 2\mathfrak{u}_i^{(0)} = 2$ have been set for all MSs $i$ and $j$ belonging to the low- and high-priority categories, respectively. For the WPF criterion, $R_i^{(0)} = \mathfrak{u}_i^{(0)} R^{(0)} = R^{(0)}$ and $R_j^{(0)} = \mathfrak{u}_j^{(0)} R^{(0)} = 2R^{(0)}$ have been adopted for all MSs $i$ and $j$ belonging to the low- and high-priority categories, respectively. $R^{(0)}$ is an arbitrary contracted throughput, whose absolute value is not relevant for a relative performance analysis.

Fig. 3.17 shows the average sum rate achieved by the SP-BFA-ZF, CC-BFA-ZF, and CC-CGA-ZF strategies considering the CM, WPF, and WRR criteria and an average SNR $\gamma$ of 10 dB.

Because only the selection of the initial MS $k'$ is subject to the priority criterion, and not the composition of the whole SDMA group, it can be seen in Fig. 3.17 that reductions of only 4% and 12% of the average sum rate result when comparing the WPF and WRR criteria to the CM criterion, respectively. The small reduction of the average sum rate is an expected result since it has been assumed that the MS priorities should not interfere with the group composition determined by the SDMA algorithm.

The small reduction of the average sum rate changes the throughput distribution among the MSs in the system. In order to determine how fair the throughput distribution in the system is, Jain's Index of Fairness (JIF) has been applied to the average throughput

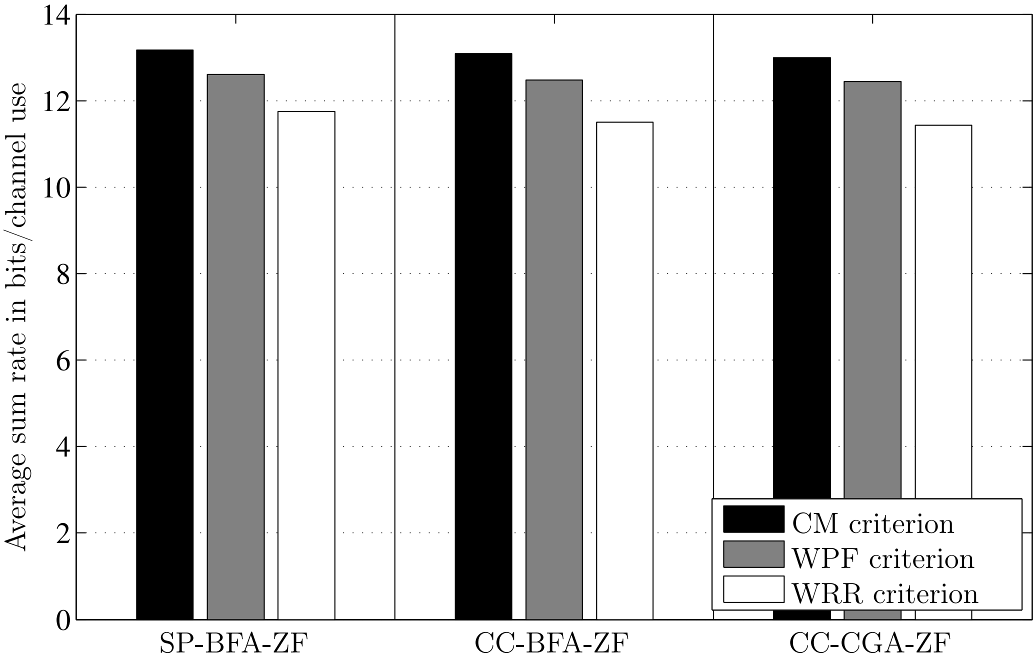



Figure 3.17. Average sum rate achieved by the SP-BFA-ZF, CC-BFA-ZF, and CC-CGA-ZF strategies considering different criteria for MS priorities. Average SNR $\gamma = 10$ dB.

$\bar{R}_k^{(\tilde{p})}$ perceived by the MSs [Jai91, Cal04]. JIF will be denoted by $\mathcal{J}(\bar{R}_k^{(\tilde{p})})$ and is defined as

$$\mathcal{J}(\bar{R}_k^{(\tilde{p})}) = \frac{\left(\sum\limits_{k=1}^{K} \bar{R}_k^{(\tilde{p})}/\mathfrak{u}_k^{(0)}\right)^2}{K \sum\limits_{k=1}^{K} \left(\bar{R}_k^{(\tilde{p})}/\mathfrak{u}_k^{(0)}\right)^2}, \tag{3.59}$$

where the throughput perceived by the MSs has been scaled in order to reflect their static priorities. For JIF, values close to zero mean an unfair throughput distribution among the MSs and values close to one a fair throughput distribution. A JIF value of, e.g., $0.7 \cdot 100$ can also be interpreted as having 70% of the MSs fairly served with the remaining 30% of the MSs getting absolutely no throughput [Jai91].

For the SP-BFA-ZF, CC-BFA-ZF, and CC-CGA-ZF strategies, Fig. 3.18 shows JIF observed in the system after a varying number of frames and considering an average SNR $\gamma$ of 10 dB.

In Fig. 3.18, it can be seen that the RA strategies present similar throughput fairness for the same MS priority criterion. Moreover, it can be noted that the WPF and WRR criteria lead to a more fair throughput distribution compared to the CM criterion. By comparing Fig. 3.18 and Fig. 3.17, it can be observed that the more fair throughput distribution among the MSs in Fig. 3.18 when considering the WPF and WRR comes at the expense of only small reductions of the average sum rate of the system in Fig. 3.17.

Since low- and high-priority categories of MSs are considered, it is important to verify whether the high-priority MSs are getting higher throughput than the low-priority MSs. Fig. 3.19 shows the percentual difference between the throughput of high- and

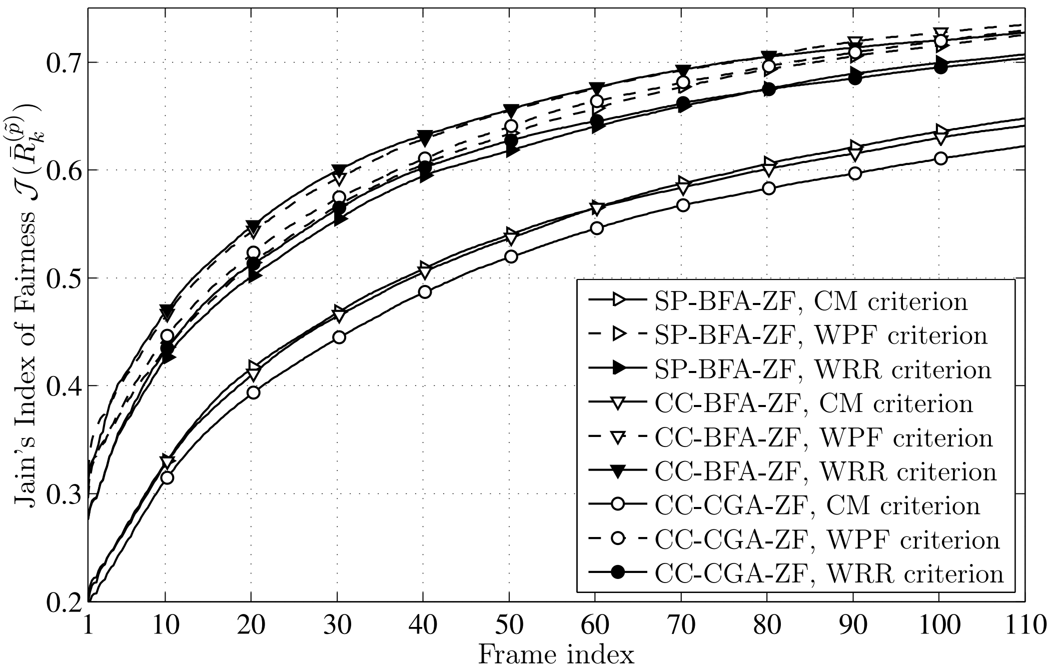



Figure 3.18. JIF obtained by the SP-BFA-ZF, CC-BFA-ZF, and CC-CGA-ZF strategies after a varying number of frames. Average SNR $\gamma = 10$ dB.

low-priority MSs considering the CM, WPF, and WRR criteria and an average SNR $\gamma$ of 10 dB. In other words, Fig. 3.19 shows how much higher the throughput of the high-priority MSs is on average than the throughput of the low-priority MSs.

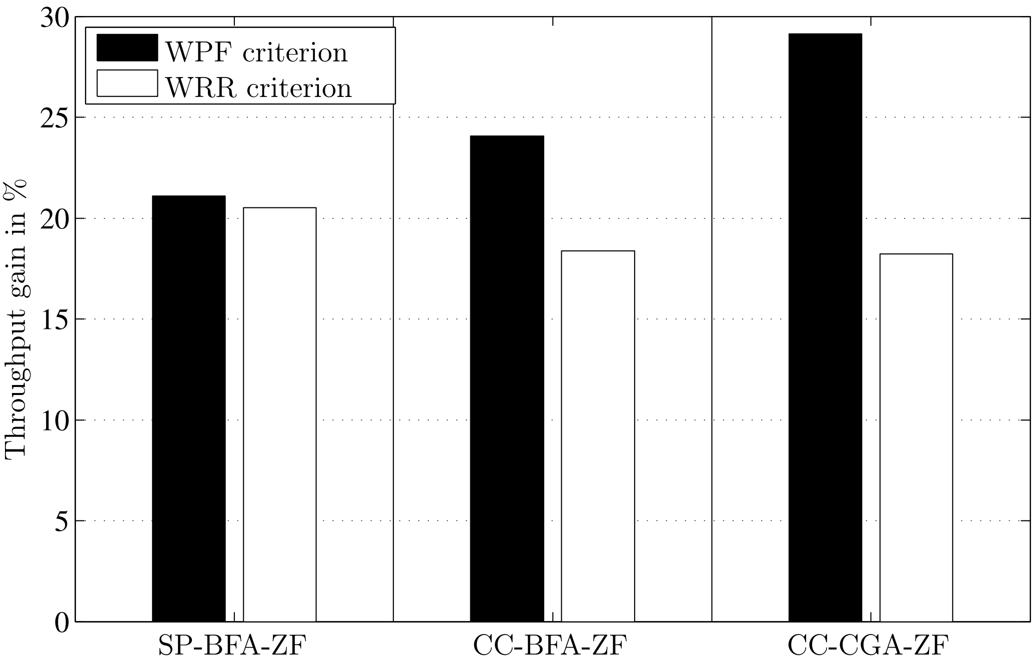



Figure 3.19. Percentual throughput difference for high-priority MSs compared to low-priority MSs. Average SNR $\gamma = 10$ dB.

In Fig. 3.19, it can be seen that the throughput of high-priority MSs is only about 20% to 30% higher than that of MSs with low priority. Because the SP-BFA-ZF, CC-BFA-ZF, and CC-CGA-ZF strategies defined in this chapter are oriented towards the maximization of the sum rate of the system, providing QoS in terms of a contracted average throughput to the MSs is not possible. Consequently, only a slightly improved throughput for the high-priority MSs can be expected, as shown in Fig. 3.19. Moreover, the criteria for the MS priorities adopted here only determine the selection of one MS in the SDMA group, i.e., the initial MS $k'$, and because it is assumed that the BS always has data to send to the MSs, cf. Section 2.2.1, it is possible that MSs that have already reached their average throughput requirements get resources anyway. Consequently, the low-priority MSs are expected to perceive average throughputs higher than the required values.

Comparing Fig. 3.17 and Fig. 3.19, it can be seen that the WPF criterion provides the best throughput gains to the high-priority MSs compared to the low-priority MSs if compared to the WRR criterion. Moreover, these throughput gains are obtained with the WPF criterion at the expense of only a small reduction of the average sum rate, which is also smaller than that observed when considering the WRR criterion. This result is expected because the WPF criterion takes into account an estimate of the perceived average throughput of the MSs when selecting the initial MS $k'$.

The further study and design of QoS-oriented RA strategies would ask for different precoding and power allocation algorithms and would impose additional constraints on the SDMA algorithms. This topic is out of the scope of this thesis and is left open for future investigation. Some results involving QoS aspects have been investigated by the author in [MK07b, MK07c, MK08]. Some additional results involving MS priorities, group priorities, and throughput fairness are provided in the next chapter.

### 3.5.10 Summary

In this section, different suboptimal RA strategies aiming at the maximization of the sum rate of the system have been studied in a scenario in which a single resource is considered.

It has been seen that the use of B-CSIT and adaptive RA on a frame basis leads only to a small degradation of the average sum rate of the system but to a manyfold reduction of the signaling demands compared to the use of P-CSIT and adaptive RA on a subcarrier and TS basis.

The suboptimal RA strategies have been shown to achieve over 90% of the average sum rate obtained by an ES for the SDMA group maximizing the sum rate of the system. The complexity of such an ES increases combinatorially with the problem dimensions and therefore the studied suboptimal RA strategies offer a rather efficient and low-complexity solution to the considered RA problem.

While employing B-CSIT leads to smaller signaling demands, further reductions of the signaling overhead can be obtained by employing S-CSIT. However, considering

S-CSIT and the parameters introduced in Section 3.5.1, the studied RA strategies have been shown to achieve only about 45% and 70% of the average sum rate that they achieve with B-CSIT for high and low average SNRs values, respectively. In particular, estimating and reporting S-CSIT at each half channel coherence time allows the strategies to obtain at least 60% of their average sum rate with B-CSIT.

The impact of erroneous CSIT has also been investigated. The RA strategies have been shown to be substantially dependent on the quality of the CSIT described by $\gamma_{\text{CSI}}$, cf. (2.10). It has been shown that in order to attain at least 60% of the average sum rate achieved with error-free CSIT, the errors due to processing and feedback delays, and due to imperfect channel estimation should lead to $\gamma_{\text{CSI}}$ values not lower than 10 dB. The amount and quality of the CSIT is therefore fundamental to realize the potential gains of adaptive RA in frequency, time, and space.

In particular, the RA strategies employing the CAP-BFA and SP-BFA and the proposed CC-CGA and CC-BFA have presented almost the same performance and achieved the best results in terms of average sum rate, which surpassed 90% of the average sum rate achieved by the CAP-ESA-ZF. None of the referred RA strategies has been shown to be particularly more robust against imperfect CSIT. Oppositely, they have been shown to be almost equally sensitive to the amount and quality of the CSIT.

The dependencies of the RA strategies on the target SDMA group size have been discussed considering the SRA. It has been shown that the SRA can provide gains about 10% to 30% in terms of the average sum rate to the RA strategies, but that its complexity can be saved or at least reduced by considering adequate target SDMA group sizes.

The impact of the paratemer $\beta$ of the CC-BFA and CC-CGA on the average sum rate of the RA strategies has also been investigated. For the RA strategies employing these SDMA algorithms, it has been shown that the performances in terms of the average sum rate are almost independent of the value of $\beta$.

Expressions for the complexity of the RA strategies considering the ZF algorithm of Table 3.8 have been derived and it has been shown that the proposed RA strategies considering the CC-CGA and especially the CC-BFA have in many cases a considerably lower computational complexity than those employing CAP-BFA and SP-BFA as SDMA algorithm. Nevertheless, the SDMA algorithms WN-BFA e WN-FFA proposed here offer a suboptimal solution to the SDMA grouping problem with even lower complexity than the CC-CGA and CC-BFA, but with slightly worse performance in terms of average sum rate.

Additionally, the SP-BFA-ZF, CC-BFA-ZF, and CC-CGA-ZF strategies have been analyzed considering different MS priority criteria. It has also been shown that an enhancement of the throughput fairness among the MSs can be obtained by suitably choosing the initial MS $k'$ and that the considered RA strategies can be modified to take QoS aspects into account and can provide high-priority MSs with a higher throughput compared to low-priority MSs.

The RA strategies studied in this chapter could be extended to fully cope with receive antenna cooperation of multi-antenna MSs. Indeed, only minor changes are required to extend the SDMA algorithms and the SRA to the case with multi-antenna MSs. However, major changes are required for the precoding and power allocation algorithms, which become more complex. A comprehensive set of precoding and power allocation algorithms for multi-antenna MSs can be found in [Sta06]. Multi-antenna MSs have been considered by the author in [MK06], where BD [SSH04] is employed. An extensive performance and complexity analysis of RA strategies considering precoding and power allocation algorithms especially designed for multi-antenna MSs is left for future investigation.

# Chapter 4

# Resource Allocation in MIMO-OFDMA Systems: Multiple Resources

## 4.1 Introduction

In this chapter, RA strategies are investigated considering the existence of multiple resources in the system. The multiple-resource case is indicated by the dot-dashed line in Fig. 1.3. Differently from Chapter 3, the resource assignment problem discussed in Section 1.2 and Section 2.3 has also to be taken into account in this chapter. Consequently, the complete problem in (2.16) is considered. Nevertheless, the division of the RA problem of (2.16) into subproblems is still considered, so that specific algorithms are applied to each of the subproblems, with the combination of these algorithms defining a suboptimal RA strategy.

Suboptimal SDMA algorithms are still needed here and for the investigations conducted in this chapter, a subset composed by the CAP-BFA, SP-BFA, CC-BFA and CC-CGA introduced in Section 3.2 is considered. These SDMA algorithms have been selected because the RA strategies employing them have been shown in Chapter 3 to well approximate the performance of the CAP-ESA-ZF strategy. Moreover, because the GEP algorithm of Table 3.8 has been shown to be considerably more complex and to perform only as good as the ZF algorithm of Table 3.8, only this latter will be considered in this chapter.

Because the resource assignment problem is considered in this chapter, different RA strategies can be designed depending on how the SDMA grouping and the resource assignment problems are solved. As discussed in Section 2.3 and Section 3.1, whenever the power $P_b$ allocated to the resource $b$, with $b = 1, \ldots, B$, are known a priori, the SDMA grouping problem can be solved on a resource-by-resource basis. In this case, the assignment of resources to SDMA groups is performed sequentially and the SDMA grouping and the resource assignment problems are solved separately from each other. In this chapter, the system resources correspond to the frequency blocks associated with each TS of a frame and the allocation period $\tilde{p}$ corresponds to one TS, as it has been the case in Section 3.5.9.

There are different alternatives to separately solve the SDMA grouping problem and resource assignment problem. Considering RA strategies whose SDMA algorithms are unaware of the actual precoding and power allocation, one alternative corresponds to first solving the resource assignment problem and allocating each resource to an initial MS, e.g., using an assignment algorithm and the MS priorities discussed in

Section 3.5.9, and afterwards build the SDMA groups on each resource considering the initial MSs to which resources have been assigned. Another alternative corresponds to solving first the SDMA grouping problem and building a set of SDMA groups on each resource and, afterwards, assigning each resource to one of the SDMA groups in the corresponding set using an assignment algorithm. These two alternatives correspond to two new algorithms for performing Separated Grouping and Assignment (SGA) which will be considered in this chapter. Considering RA strategies whose SDMA algorithms are aware of the actual precoding and power allocation, the SDMA grouping and resource assignment problems can be solved either separately, e.g., according to one of the above alternatives, or jointly. An algorithm performing Joint Grouping and Assignment (JGA) is also considered in this chapter.

Regarding the precoding and power allocation, it can be noted that ZF precoding described in Section 3.3.1 works on a resource basis and remains unchanged when multiple resources are considered. On the other hand, the WFA needs to be adapted in order to take multiple resources into account whenever an adaptive distribution of the power among all the resources is considered. Both the cases in which the power is divided among resources a priori or adaptively are considered in this chapter.

This chapter is organized as follows. In Section 4.2, the Grouping & Assignment (GA) algorithms are described, which involve the MS or group priority, the SDMA algorithm, and the assignment algorithm, as illustrated in Fig. 1.3.

In Section 4.2.1, the MS priorities of Section 3.5.9 are revisited and their concepts are extended to group priorities considering the CM, WRR, and WPF criteria. Because MS priorities have already been discussed in Section 3.5.9, they receive less attention in this section.

In Section 4.2.2, the SDMA and assignment algorithms are discussed. However, because the SDMA algorithms considered in this chapter, i.e., the CAP-BFA, SP-BFA, CC-BFA and CC-CGA, have already been detailed in Section 3.2, they will receive less focus in this section. The assignment algorithms are new and play a major role in this section and, consequently, will receive more attention. The proposed algorithms for performing SDMA grouping and resource assignment separately or jointly are described in Section 4.2.2.1 and Section 4.2.2.2, respectively.

In Section 4.2.2.1, three new algorithms for performing SDMA grouping and resource assignment are defined. In Section 4.2.2.2, the CAP-BFA of Table 4.1 is extended to perform joint SDMA grouping and resource assignment over the multiple resources altogether.

In Section 4.2, all the GA algorithms resulting from the combination of an SDMA algorithm, an MS or group priority, and an assignment algorithm considered in this chapter are possible, which results in a large number of GA algorithms. Due to this, the GA algorithms considered in this chapter will be defined only in Section 4.4, together with the definition of the RA strategies. Therefore, the convention adopted in this chapter differs from the one that has been adopted in Chapter 3.

In Section 4.3, the precoding and power allocation algorithms for the scenario with multiple resources are introduced. In Section 4.3.1, some remarks regarding the precoding algorithm considered in this chapter are made. In Section 4.3.2, the WFA of Section 3.3.2 is extended to take into account multiple resources. In Section 4.3.3, the considered precoding and power allocation algorithms are defined.

In Section 4.4, the GA algorithms described in Section 4.2 are combined with the precoding and power allocation algorithms of Section 4.3 in order to define the RA strategies investigated in this chapter.

In Section 4.5, the performance of the RA strategies defined in Section 4.4 is investigated and a short summary of the results of this chapter is provided.

## 4.2 Grouping & Assignment Algorithm

### 4.2.1 Group Priority

In this section, the concept of MS priorities is extended to group priorities. The group priority defined here quantifies the efficiency of allocating a resource to a given SDMA group. An SDMA group may have a different priority for each of the resources of the system and, consequently, group priorities can be used to determine which resources should be allocated to which group. On resource $b$, let the $g^{\text{th}}$ SDMA group be denoted by $\mathcal{G}_{g,b}$ and let $\mathfrak{v}_{g,b}$ denote its group priority. One simple form of defining the group priority corresponds to sum up the priorities $\mathfrak{u}_{i,b}$ of the MSs in $\mathcal{G}_{g,b}$ on resource $b$, i.e., to define

$$\mathfrak{v}_{g,b} = \sum_{i \in \mathcal{G}_{g,b}} \mathfrak{u}_{i,b}. \tag{4.1}$$

[MK07b, MK08]. Using (4.1) and considering the CM, the WRR, and the WPF criteria defined in Section 3.5.9, the group priorities can be defined as

$$\mathfrak{v}_{g,b} = f_{\text{CAP}}(\mathcal{G}_{g,b}) \tag{4.2a}$$

for the CM criterion,

$$\mathfrak{v}_{g,b} = \sum_{i \in \mathcal{G}_{g,b}} \mathfrak{u}_i^{(\tilde{p})} \tag{4.2b}$$

for the WRR criterion, and

$$\mathfrak{v}_{g,b} = \sum_{i \in \mathcal{G}_{g,b}} \frac{R_i^{(0)}}{\bar{R}_i^{(\tilde{p})}} \log_2 \left( 1 + \frac{p_{i,b} \left\| \hat{\mathbf{g}}_i \mathbf{w}_k \right\|_2^2}{\sigma^2 + \sum\limits_{\substack{j \in \mathcal{G}_{g,b} \\ j \neq i}} p_{j,b} \left\| \hat{\mathbf{g}}_{i,b} \mathbf{w}_{j,b} \right\|_2^2} \right) \tag{4.2c}$$

for the WPF criterion.

In the following, the group priorities defined in (4.2) are shortly discussed. For the CM criterion, the higher the group capacity of an SDMA group $\mathcal{G}_{g,b}$ on a resource $b$ is, the higher the group priority on this resource is and, consequently, the higher the chances of assigning the resource $b$ to the SDMA group $\mathcal{G}_{g,b}$.

For the WRR criterion, the MS priorities $\mathfrak{u}_i^{(\tilde{p})}$ of the all MS in an SDMA group $\mathcal{G}_{g,b}$ are summed up. Thus, according to (3.56), SDMA groups containing MSs which got no access to the channel for a long period will have high group priority. Consequently, there will be high chances of assigning resources to these SDMA groups.

For the WPF criterion, the rate of the MSs in an SDMA group are scaled by the throughput ratio $\frac{R_k^{(0)}}{\bar{R}_k^{(\tilde{p})}}$, which measures how well the MS has met its QoS requirements [SWJO07]. Thus, SDMA groups containing MSs achieving a high rate or MSs whose QoS requirements have not been fulfilled will have high group priority. Consequently, there will be high chances of assigning resources to these SDMA groups.

Differently from (3.54) and (3.58), in which the rate of a single MS could be replaced by the channel gain, replacing the group capacity in (4.2a) and (4.2c) by the norm of the group channel matrix may be not an adequate solution, since it would not efficiently represent neither the effective channel matrix of the SDMA group, which depends on the precoding vectors, nor the individual products between the priority and rate of each MS in the SDMA group. However, whenever the SDMA grouping and the resource assignment problem are solved separately, the composition of the candidate SDMA group is known and, consequently, the group priority can be calculated for each of candidate group using (4.2). Nevertheless, this leads to additional complexity compared to the case in which group priorities are not considered, since precoding and power allocation have to be taken into account in the computation of the group priorities. It must be noted that this applies for the CM and WPF criteria, but this does not apply to the WRR criterion in which the group priority does not depend neither on precoding nor on power allocation, cf. (4.2b).

### 4.2.2 SDMA and Assignment Algorithms

#### 4.2.2.1 Separated Grouping and Assignment

In this section, the different algorithms for performing Separated Grouping and Assignment (SGA) are described. In order to perform SGA, an a-priori distribution of the power among the system resources has to be assumed. In this chapter, an EPA among the $B$ resources of the system is considered by the SGA algorithms, so that $P_b = \frac{P}{S}, b = 1, \ldots, B$. In this section, the following SGA algorithms are defined:

1. The SGA Algorithm 1 (SGA-A1), which is used by the RA strategies that work on a resource-by-resource basis, so that the resources in the system are considered sequentially.
2. The SGA Algorithm 2 (SGA-A2), which first solves the resource assignment problem by assigning each resource to an initial MS and which solves the SDMA grouping problem afterwards by applying the SDMA algorithms to build groups on each resource.
3. The SGA Algorithm 3 (SGA-A3), which first solves the SDMA grouping problem by building a set of candidate SDMA groups for each resource and which solves the resource assignment problem afterwards by assigning the resources to a subset of the candidate SDMA groups.

The three SGA algorithms mainly differ in the form and order in which the SDMA grouping and the resource assignment problems are solved.

In the following, the SGA-A1 is described. The SGA-A1 is a simple SGA algorithm which can be used by RA strategies that do not really take into account the existence of multiple resources into the system, but deal with each of the $B$ resources of the system individually. An illustration of the SGA-A1 is shown in Fig. 4.1.

Considering the SGA-A1, the first resource is assigned to an initial MS according to the MS priority criteria defined in Section 3.5.9. This is the resource assignment step indicated by ① in Fig. 4.1. Considering this initial MS, the SDMA algorithm of SGA-A1 builds an SDMA group on the considered resource and other MSs are put together with the initial MS into the SDMA group. This is the SDMA grouping step indicated by ② in Fig. 4.1. After that, the precoding, power allocation, and the SRA are applied for the SDMA group while considering the available power $P_b = \frac{P}{S}$ on the resource. This is the precoding and power allocation step indicated by ③ in Fig. 4.1. Due to the SRA, MSs admitted to the SDMA group might be removed, thus resulting in a smaller group, as illustrated in Fig. 4.1. After precoding and power allocation, the RA strategy is finished with this particular resource and the process is repeated for the next resource, as indicated by step ④ in Fig. 4.1.

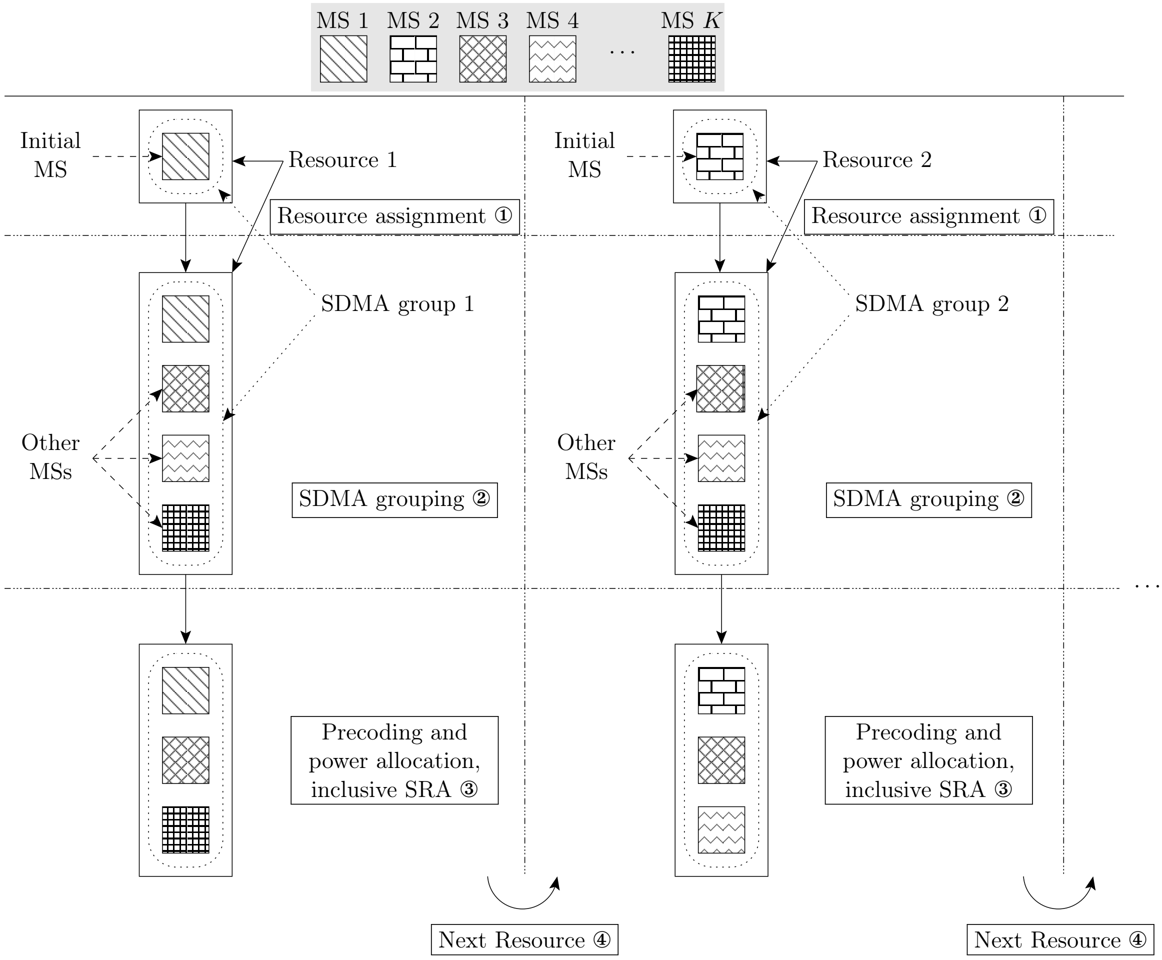



Figure 4.1. Illustration of the SGA Algorithm 1 (SGA-A1).

For each of the RA strategies discussed in Chapter 3, in which a single resource has been considered, the use of the SGA-A1 can be interpreted as a sequential application of the RA strategy itself for each resource $b, b = 1, \ldots, B$.

In the following, the SGA-A2 is described. Differently from the SGA-A1, in which resources are assigned one-by-one to initial MSs, the SGA-A2 assigns all the resources to a set of initial MSs at once. An illustration of the SGA-A2 algorithm is shown in Fig. 4.2.

Considering the SGA-A2, first an initial MS is assigned to each resource. This is the resource assignment step indicated by ① in Fig. 4.2. Then, on each resource $b, b = 1, \ldots, B$, an SDMA group is built by the SDMA algorithm of the SGA-A2 and other MSs are grouped with the initial MSs to which the resources have been assigned. This is the SDMA grouping step indicated by ② in Fig. 4.2. Considering EPA among resources, the precoding and power allocation algorithm is applied to each resource. This is the precoding and power allocation step indicated by ③ in Fig. 4.2. Due to the SRA, MSs admitted to the SDMA groups might be removed, thus resulting in a smaller groups, as illustrated in Fig. 4.2.

In order to decide which resources should be assigned to which initial MSs, a standard assignment problem [NW99] considering the MS priorities defined in Section 3.5.9 is for-

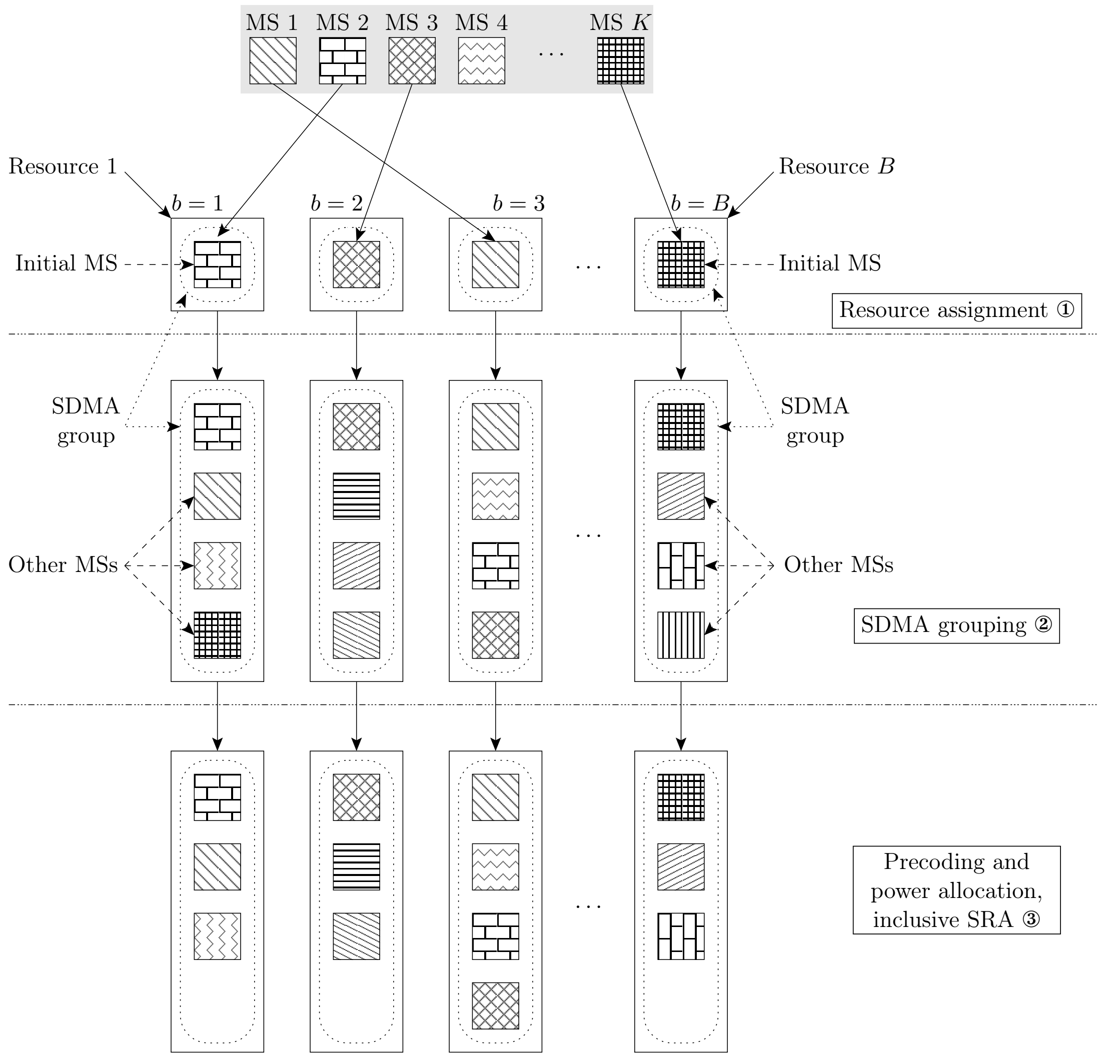



Figure 4.2. Illustration of the SGA Algorithm 2 (SGA-A2).

mulated and solved using Munkres' algorithm [BL71, MK07b, MK08]. In the following, the assignment problem considered by the SGA-A2 is formulated. Let $\mathfrak{U} \in \mathbb{R}_+^{B\times K}$ denote an MS priority matrix containing the MS priorities $\mathfrak{u}_{b,k}$ of each MS $k, k = 1, \ldots, K$, on each resource $b, b = 1, \ldots, B$. Let $\mathbf{U} \in \mathbb{B}^{B\times K}$ denote a resource-to-MS assignment matrix whose binary entries $[\mathbf{U}]_{b,k} = u_{b,k} \in \{0, 1\}$ indicate whether a resource $b$ is assigned to an MS $k$. The matrices $\mathfrak{U}$ and $\mathbf{U}$ can be written as

$$\mathfrak{U} = \begin{bmatrix} \mathfrak{u}_{1,1} & \mathfrak{u}_{1,2} & \ldots & \mathfrak{u}_{1,K} \\ \mathfrak{u}_{2,1} & \mathfrak{u}_{2,2} & \ldots & \mathfrak{u}_{2,K} \\ \vdots & \vdots & \ddots & \vdots \\ \mathfrak{u}_{B,1} & \mathfrak{u}_{B,2} & \ldots & \mathfrak{u}_{B,K} \end{bmatrix} \tag{4.3a}$$

and

$$\mathbf{U} = \begin{bmatrix} u_{1,1} & u_{1,2} & \dots & u_{1,K} \\ u_{2,1} & u_{2,2} & \dots & u_{2,K} \\ \vdots & \vdots & \ddots & \vdots \\ u_{B,1} & u_{B,2} & \dots & u_{B,K} \end{bmatrix}, \tag{4.3b}$$

respectively. Let $\odot$ denote the Hadamard product between two vectors or two matrices [GL96]. Then, using (4.3) the assignment of the resources to the initial MSs can be formulated as

$$\mathbf{U}^{\star} = \arg\max_{\mathbf{U}} \left\{ \mathbf{1}_B^{\mathrm{T}} \left( \mathfrak{U} \odot \mathbf{U} \right) \mathbf{1}_K \right\} \tag{4.4a}$$

$$\text{subject to: } \mathbf{U}\mathbf{1}_K = \mathbf{1}_B \tag{4.4b}$$

$$\mathbf{U}^{\mathrm{T}}\mathbf{1}_B = \mathbf{1}_K, \tag{4.4c}$$

which is a standard assignment problem that can be efficiently solved with Munkres' algorithm [BL71]. After solving (4.4), the resource assignment problem is solved and the resources assigned to the initial MSs are determined by the non-zero entries of $\mathbf{U}^{\star}$. The RA strategies employing the SGA-A2 work similarly to those employing the SGA-A1 and starting with the initial MSs to which the resources have been assigned, the SDMA algorithm builds an SDMA group on each resource. After the resource assignment and the SDMA grouping problems are solved, the precoding and power allocation algorithm of the RA strategy can be applied to the SDMA groups associated with each resource.

In the following, the SGA-A3 is described. Differently from SGA-A1 and SGA-A2, the SGA-A3 solves first the SDMA grouping problem and builds a set of candidate SDMA groups on each resource. After that, the assignment algorithm assigns each resource to one of the candidate SDMA groups. An illustration of the SGA-A3 algorithm is shown in Fig. 4.3.

Considering the SGA-A3, first a set of $K$ candidate SDMA groups are created on each resource $b, b = 1, \dots, B$. On a given resource $b$, the $k^{\text{th}}$ SDMA group is created by selecting the $k^{\text{th}}$ MS as initial MS and applying the SDMA algorithm to build the SDMA group considering this initial MS. Consequently, $G = K \cdot B$ candidate groups are created. These steps are indicated by ① and ② in Fig. 4.3, respectively. For each candidate SDMA group $g, g = 1, \dots, G$, the precoding and power allocation algorithm is applied considering EPA among resources. This is the precoding and power allocation step indicated by ③ in Fig. 4.3. Then, similarly to the SGA-A2 the assignment of resources to SDMA groups is formulated as a standard assignment problem based on the group priorities [NW99], which is solved using Munkres' algorithm [BL71, MK07b, MK08], so that the $B$ resources are assigned to $B$ of the $G$ candidate SDMA groups. This is the resource assignment step indicated by ④ in Fig. 4.2.

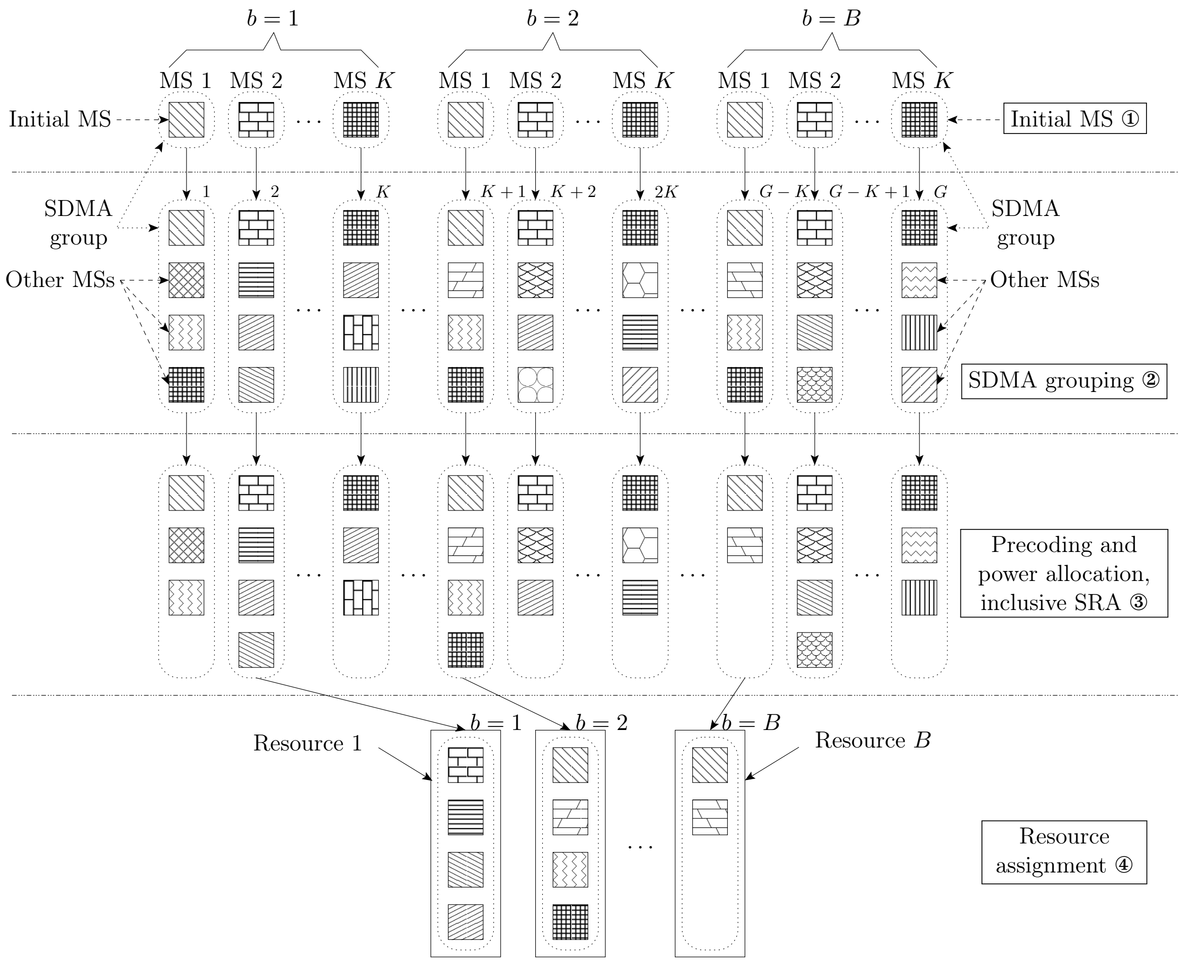



Figure 4.3. Illustration of the SGA Algorithm 3 (SGA-A3).

In the following, the assignment problem considered by the SGA-A3 is formulated. Differently from the SGA-A2, resources are assigned to whole SDMA groups by the SGA-A3 and the group priorities defined in Section 4.2.1 are considered in the problem formulation instead of the MS priorities considered by the SGA-A2. Let $\mathfrak{V} \in \mathbb{R}_{+}^{B\times G}$ denote a group priority matrix containing the group priorities $\mathfrak{v}_{b,g}$ of each SDMA group $g, g = 1, \ldots, G$, on each resource $b, b = 1, \ldots, B$. Let $\mathbf{V} \in \mathbb{B}^{B\times G}$ denote a resource-to-group assignment matrix whose binary entries $[\mathbf{V}]_{b,g} = v_{b,g} \in \{0, 1\}$ indicate whether a resource $b$ is assigned to an SDMA $g$. The matrices $\mathfrak{V}$ and $\mathbf{V}$ can be written as

$$\mathfrak{V} = \begin{bmatrix} \mathfrak{v}_{1,1} & \mathfrak{v}_{1,2} & \ldots & \mathfrak{v}_{1,G} \\ \mathfrak{v}_{2,1} & \mathfrak{v}_{2,2} & \ldots & \mathfrak{v}_{2,G} \\ \vdots & \vdots & \ddots & \vdots \\ \mathfrak{v}_{B,1} & \mathfrak{v}_{B,2} & \ldots & \mathfrak{v}_{B,G} \end{bmatrix} \tag{4.5a}$$

and

$$\mathbf{V} = \begin{bmatrix} v_{1,1} & v_{1,2} & \dots & v_{1,G} \\ v_{2,1} & v_{2,2} & \dots & v_{2,G} \\ \vdots & \vdots & \ddots & \vdots \\ v_{B,1} & v_{B,2} & \dots & v_{B,G} \end{bmatrix}, \tag{4.5b}$$

respectively, and the assignment of resources to SDMA groups can be formulated as

$$\mathbf{V}^{\star} = \arg\max_{\mathbf{V}} \left\{ \mathbf{1}_B^{\mathrm{T}} \left( \boldsymbol{\mathfrak{V}} \odot \mathbf{V} \right) \mathbf{1}_G \right\} \tag{4.6a}$$

$$\text{subject to: } \mathbf{V}\mathbf{1}_G = \mathbf{1}_B \tag{4.6b}$$

$$\mathbf{V}^{\mathrm{T}}\mathbf{1}_B = \mathbf{1}_G, \tag{4.6c}$$

which is also a standard assignment problem and can be efficiently solved with Munkres' algorithm [BL71]. After solving (4.6), the resource assignment problem is solved and the resources assigned to SDMA groups are determined by the non-zero entries of $\mathbf{V}^{\star}$. After that, the RA is completed.

In the following, some additional remarks are made regarding the SGA algorithms described in this section. Observing the constraints (4.4b) and (4.4c), it can be noted that, differently from the SGA-A1, the SGA-A2 does not allow more than one of the $B$ resources to be allocated to the same initial MS. Considering EPA among resources, this is the basic difference between SGA-A1 and SGA-A2. For example, consider two MSs, two resources, the CM criterion, and consider also that the channel gains of MS $k = 1$ are higher than the channels gains of MS $k = 2$ on both resources. Then, if the SGA-A1 is employed by the RA strategy, MS $k = 1$ will be chosen as initial MS for both resources. However, if the SGA-A2 is employed by the RA strategy, MS $k = 1$ will be chosen as initial MS for the resource on which its channel gain is maximum and the other resource will be assigned to MS $k = 2$. Consequently, the SGA-A2 leads to a more fair distribution of the resources among the initial MSs. In order to allow the SGA-A2 to assign multiple resources to a given initial MS $k$, it is enough to replicate the $k^{\text{th}}$ column of $\boldsymbol{\mathfrak{U}}$ as many times as needed, correspondingly extend the matrix $\mathbf{U}$, and solve (4.4) considering such extended versions of $\boldsymbol{\mathfrak{U}}$ and $\mathbf{U}$ while keeping track of the correspondence between the added columns and the considered MS $k$. In this way, an MS $k$ might get as many resources as the number of columns of $\boldsymbol{\mathfrak{U}}$ and $\mathbf{U}$ that are associated with it and the SGA-A2 can perfectly emulate the SGA-A1. Anyway, because the composition of the SDMA groups is decided afterwards, multiple resources might be allocated to SDMA groups of same composition.

Similarly to the SGA-A2, extended matrices $\boldsymbol{\mathfrak{V}}$ and $\mathbf{V}$ can be used with the SGA-A3 to allow more resources to be allocated to a given SDMA group. Anyway, because SDMA groups having the same composition might be built on different resources, multiple resources might be allocated to the same SDMA group.

On the one hand, extending the matrices $\mathfrak{U}$ and $\mathbf{U}$, or $\mathfrak{V}$ and $\mathbf{V}$, might allow to allocate multiple resources to the same MS or SDMA group, respectively. On the other hand, they increase the dimensions of problems (4.4) and (4.6) and increase their complexity. Moreover, if fairness aspects are taken into account, extending the matrices $\mathfrak{U}$ and $\mathbf{U}$, or $\mathfrak{V}$ and $\mathbf{V}$, might reduce the throughput fairness among the MSs due to the possible concentration of resources by a few MSs or SDMA groups.

The assignment algorithm of the SGA-A3 is structurally analogous to that of the SGA-A2, but assigns whole SDMA groups to resources instead of only initial MSs. In order to avoid that an unsuitable resource be assigned to an SDMA group, the group priorities of the $K$ SDMA groups created on each resource $b$ are computed only for this specific resource and are set to zero on the other $B-1$ resources. This choice has been made by the author in [MK07b, MK08] and is reasonable since spatial compatibility among MSs is resource-dependent. Consequently, for each resource $b$, the solution of (4.6) reduces to assigning each resource to the highest priority group among the $K$ SDMA groups created on the corresponding resource.

#### 4.2.2.2 Joint Grouping and Assignment

In this section, the Joint Grouping and Assignment (JGA) is discussed, which can be performed if the SDMA algorithm is aware of the precoding and power allocation. From the SDMA algorithms listed in Section 4.1, only the CAP-BFA is aware of the precoding and power allocation and, consequently, it is the unique SDMA algorithm considered in this section.

In the following, the CAP-BFA of Table 3.2 is extended to consider multiple resources and perform JGA. In order to extend the CAP-BFA to perform JGA, the following aspects have to be taken into account:

- The initial MS $k'$ corresponds now to the one with the highest channel gain $\|\hat{\mathbf{h}}_{k,b}\|_2^2$ over all the resources $b, b = 1, \ldots, B$, and the initial resource $b'$ is the resource on which the channel gain $\|\hat{\mathbf{h}}_{k',b}\|_2^2$ of the initial MS $k'$ is maximum. Then, an SDMA group $\mathcal{G}_{b'}$ containing the initial MS $k'$ is build on the initial resource $b'$.
- The next MS to be assigned a resource can now be a new MS admitted to an SDMA group to which a resource has already been allocated or an MS building a new SDMA group on a resource which has not been allocated yet.
- The CAP-BFA stops adding new MSs to the SDMA group $\mathcal{G}_b$ to which the resource $b$ is allocated whenever the target group size $K^{\star}_{\mathcal{G}_b}$ on that resource has been reached.
- The CAP-BFA ends whenever the admission of a new MS to an SDMA group on any of the resources does not increase the sum rate anymore.

Modifying the CAP-BFA of Table 3.2 to take into account the above aspects is straightforward. Let $\mathcal{G}_b$ denote the SDMA group currently allocated on resource $b$, let $K_{\mathcal{G}_b}$

denote the size of the SDMA group $\mathcal{G}_b$, and let $\mathscr{G}$ denote the set of the SDMA groups to which the resources are allocated. Considering these definitions, the CAP-BFA extended for JGA is presented in Table 4.1.

Table 4.1. CAP-BFA extended for JGA.

1. Select the initial MS $k'$ and resource $b'$ as $\{k', b'\} = \arg\max_{\{k,b\}} \left\{ \|\hat{\mathbf{h}}_{k,b}\|_2^2 \right\}$ and build $\mathcal{G}_{b'} = \{k'\}$.
2. Set $\mathscr{G} = \{\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_B\}$, where $\mathcal{G}_b$ are initially empty $\forall b \neq b'$.
3. While $K_{\mathcal{G}_b} < K^\star_{\mathcal{G}_b}$ for any resource $b$
   a. For each resource $b$, set $\mathcal{G}'_b = \mathcal{G}_b \cup \left\{ \arg\max_{k \notin \mathcal{G}_b} \{ f_{\text{CAP}}(\mathcal{G}_b \cup \{k\}) \} \right\}$.
   b. Set $b^\star = \arg\max_b \left\{ \sum_{\mathcal{G}_i} f_{\text{CAP}}(\mathcal{G}_i) \right\}$, with $\mathcal{G}_i \in \{(\mathscr{G} \setminus \mathcal{G}_b) \cup \mathcal{G}'_b\}$.
   c. Set $\mathscr{G}' = \{(\mathscr{G} \setminus \mathcal{G}_{b^\star}) \cup \mathcal{G}'_{b^\star}\}$.
   d. If $\sum_{\mathcal{G}_i \in \mathscr{G}'} f_{\text{CAP}}(\mathcal{G}_i) > \sum_{\mathcal{G}_i \in \mathscr{G}} f_{\text{CAP}}(\mathcal{G}_i)$, set $\mathscr{G} = \mathscr{G}'$, otherwise stop.
4. Define the set of the best SDMA groups as $\mathscr{G}^\star = \mathscr{G}$.

The algorithm in Table 4.1 remains unchanged if a grouping metric other than $f_{\text{CAP}}(\mathcal{G})$ is used. However, for any grouping metric not involving the actual precoding and power allocation, the algorithm shown in Table 4.1 leads to the same result than applying the BFA from Table 3.2 on a resource-by-resource basis.

# 4.3 Precoding and Power Allocation Algorithms

## 4.3.1 Precoding Algorithm

As mentioned in Section 4.1, only the ZF of Section 3.3.1 is considered in this chapter. Because ZF precoding diagonalizes the channel matrix of the SDMA groups assigned on each resource, cf. (3.31), the precoding and the power allocation are decoupled from each other, as discussed in Section 3.3, and precoding vectors can be determined on a resource basis. Thus, if the composition of the SDMA groups allocated on the different resources of the system does not change, the efficiency of the precoding algorithm is not affected by any changes in the distribution of the power among the resources allocated to the MSs. Therefore, no additional considerations are required in this chapter regarding the ZF precoding of Section 3.3.1.

## 4.3.2 Power Allocation Algorithm

In this section, the WFA of Table 3.5 is extended to cope with multiple resources. The extension of the WFA and of the SDA is straightforward. However, due to higher complexity of the SDA compared to the WFA and of the GEP algorithm of Table 3.9 compared to the ZF algorithm of Table 3.8, the SDA is not considered in this section.

Let the indices $i$ and $b$ be used to indicate the MS $i$ in the SDMA group $\mathcal{G}_b$ to which the resource $b$ has been allocated. Let $K_{\mathcal{G}_b}$ be the size of the SDMA group $\mathcal{G}_b$. Then, the total number $K_A$ of MSs to which resources have been allocated can be written as $K_A = \sum_{b=1}^{B} K_{\mathcal{G}_b}$. Let $j, j = 1, \ldots, K_A$, indicate the $j^{\text{th}}$ MS to which resources have been allocated, and let $\mathcal{M}(b,i)$ be a function mapping the indexes $b$ and $i$ of an MS to its corresponding index $j$, and $\mathcal{M}^{-1}(j)$ denote the function mapping the index $j$ back to the indexes $b$ and $i$. Considering these definitions, the WFA for power allocation over multiple resources is presented in Table 4.2.

Table 4.2. WFA for power allocation over multiple resources.

1. Get the index $j$ of each MSs by using $\mathcal{M}(i,b)$, for $b = 1, \ldots, B$, and $i = 1, \ldots, K_{\mathcal{G}_b}$.
2. Sort the $K_A$ MSs so that $\|\hat{\mathbf{g}}_1\mathbf{w}_1\|_2^2 \geq \|\hat{\mathbf{g}}_2\mathbf{w}_2\|_2^2 \geq \ldots \geq \|\hat{\mathbf{g}}_j\mathbf{w}_j\|_2^2 \ldots \geq \|\hat{\mathbf{g}}_{K_A}\mathbf{w}_{K_A}\|_2^2$.
3. For $\tilde{K} = K_A$ to 1,

   If $\left(\frac{\sigma^2}{\|\hat{\mathbf{g}}_{\tilde{K}}\mathbf{w}_{\tilde{K}}\|_2^2}\right)\tilde{K} - \left(P + \sum_{j'=1}^{\tilde{K}} \frac{\sigma^2}{\|\hat{\mathbf{g}}_{j'}\mathbf{w}_{j'}\|_2^2}\right) < 0$, go to step 4.
4. Compute the water level $\mu^\star = \frac{1}{\tilde{K}^\star}\left(P + \sum_{j'=1}^{\tilde{K}^\star} \frac{\sigma^2}{\|\hat{\mathbf{g}}_{j'}\mathbf{w}_{j'}\|_2^2}\right)$.
5. For $j = 1$ to $K_A$

   If $j \leq \tilde{K}$, set $p_j = \mu^\star - \frac{\|\hat{\mathbf{g}}_j\mathbf{w}_j\|_2^2}{\sigma^2}$, else set $p_{\mathtt{m}(j)} = 0$.
6. Sort the MSs back to their original order.
7. Get the indexes $b$ and $i$ of each MS by using $\mathcal{M}^{-1}(j)$, for $j = 1, \ldots, K_A$.

The only difference between the WFA in Table 4.2 and the WFA in Table 3.5 is that the former performs water-filling in both frequency and space dimensions while the latter performs it only in the space dimension since a single frequency resource is considered.

### 4.3.3 Precoding and Power Allocation Algorithm Definition

In this section, the precoding and power allocation algorithms considered in this chapter are defined. Two precoding and power allocation algorithms are considered in this chapter.

If EPA among resources is considered, no changes are required to WFA of Section 3.3.2 or for the SRA of Section 3.3.3 to be applied in this chapter. EPA among resources is considered by the RA strategies performing SGA and which employ the ZF algorithm of Table 3.8.

If an EPA among resources is not considered, the precoding and power allocation algorithm combines the ZF precoding, cf. Section 3.3.1, and the WFA of Table 4.2, which performs power allocation to the MSs over the resources altogether. These are the precoding algorithm and the power allocation algorithm considered by the RA strategies performing JGA.

Note that the precoding and power allocation algorithms adopted by an RA strategy is clearly identified by the assignment algorithm employed by the same RA strategy and, consequently, no additional denominations for the precoding and power allocation algorithms are required. The association between assignment algorithms and the precoding and power allocation algorithms will become more clear when defining the RA strategies in the next section.

Moreover, because the SRA of Table 3.7 is used without changes by the RA strategies performing SGA and because the CAP-BFA of Table 4.1 considered by the RA strategy performing JGA does not ask for an SRA, no additional section concerning an extension of the SRA of Table 3.7 to multiple resources is introduced in this chapter.

## 4.4 Resource Allocation Strategy Definition

In this section, the RA strategies considered in this chapter are defined. They consist of combining a GA algorithm of Section 4.2 with a precoding and power allocation algorithm of Section 4.3.3. The same simulation parameters provided in Table 3.11 and in Table 3.12 are considered in this chapter. The RA strategies considered in this chapter corresponds to the different combinations allowed by the elements presented in Table 4.3.

Table 4.3. RA strategies: multiple resources.

<table>
<tr><th colspan="3">Grouping & Assignment algorithm</th><th rowspan="2">Precoding and power allocation algorithm</th></tr>
<tr><th>SDMA algorithm</th><th>Priority criterion</th><th>Assignment algorithm</th></tr>
<tr><td>CAP-BFA</td><td rowspan="4">CM,<br>WRR,<br>or<br>WPF</td><td rowspan="4">SGA-A1,<br>SGA-A2,<br>or<br>SGA-A3</td><td rowspan="4">ZF, cf. Section 3.3.1, WFA, cf. Table 3.5,<br>SRA, cf. Table 3.7, and EPA among resources</td></tr>
<tr><td>SP-BFA</td></tr>
<tr><td>CC-BFA</td></tr>
<tr><td>CC-CGA</td></tr>
<tr><td>CAP-BFA</td><td>CM</td><td>JGA</td><td>ZF, cf. Section 3.3.1, and WFA, cf. Table 4.2</td></tr>
</table>

In Table 4.3, note that there is a clear distinction between the precoding and power allocation algorithms employed by RA strategies performing SGA and JGA, as discussed in Section 4.3.3. In this chapter, the RA strategies could be named after the SDMA algorithm, MS or group priority criterion, and assignment algorithm that they employ. For example, the RA strategy combining the SP-BFA as SDMA algorithm, the CM criterion for MS priorities, and the SGA-A1 as assignment algorithm would be named SP-BFA-CM-SGA-A1. However, such long denominations for the RA strategies will be avoided in this chapter and the distinction among the different RA strategies will be made explicitly in the text.

## 4.5 Performance of the Resource Allocation Strategies

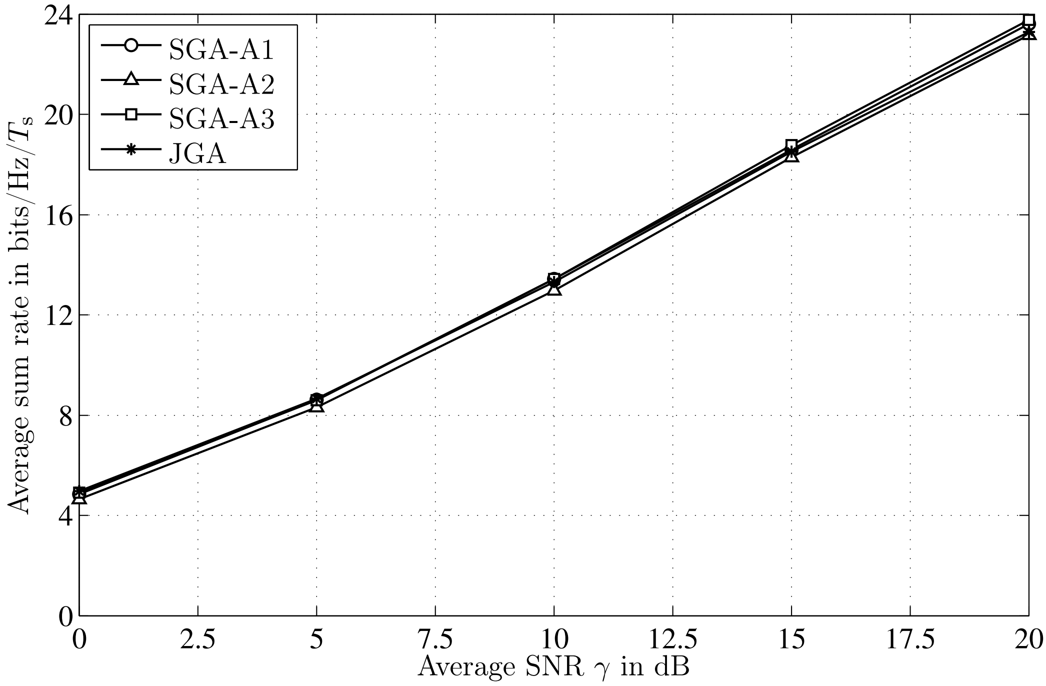
In this section, the performance of the RA strategies of Section 4.4 is investigated. Only the B-CSIT model of Section 2.2.4 is considered in this section. Moreover, the RA is assumed to be performed at each TS in order to cope with the different criteria for MS or group priority.

Because the assignment algorithms are a new element considered in this chapter, it is interesting to initially compare the performance of RA strategies employing the same SDMA algorithm but different assignment algorithms. In order to do this, only the RA strategies employing the CAP-BFA as SDMA algorithm will be considered, because they provided the highest average sum rate figures in Chapter 3 and because the CAP-BFA is the only SDMA algorithm considered in the JGA, cf. Section 4.2.2.2. Fig. 4.4(a) shows the average sum rate achieved by RA strategies considering the CAP-BFA as SDMA algorithm, the CM criterion, and the different assignment algorithms, cf. Section 4.2.2, for a varying average SNR $\gamma$. Fig. 4.4(b) shows Jain's Index of Fairness (JIF) obtained with the same RA strategies after a varying number of frames and considering an average SNR $\gamma = 10$ dB.

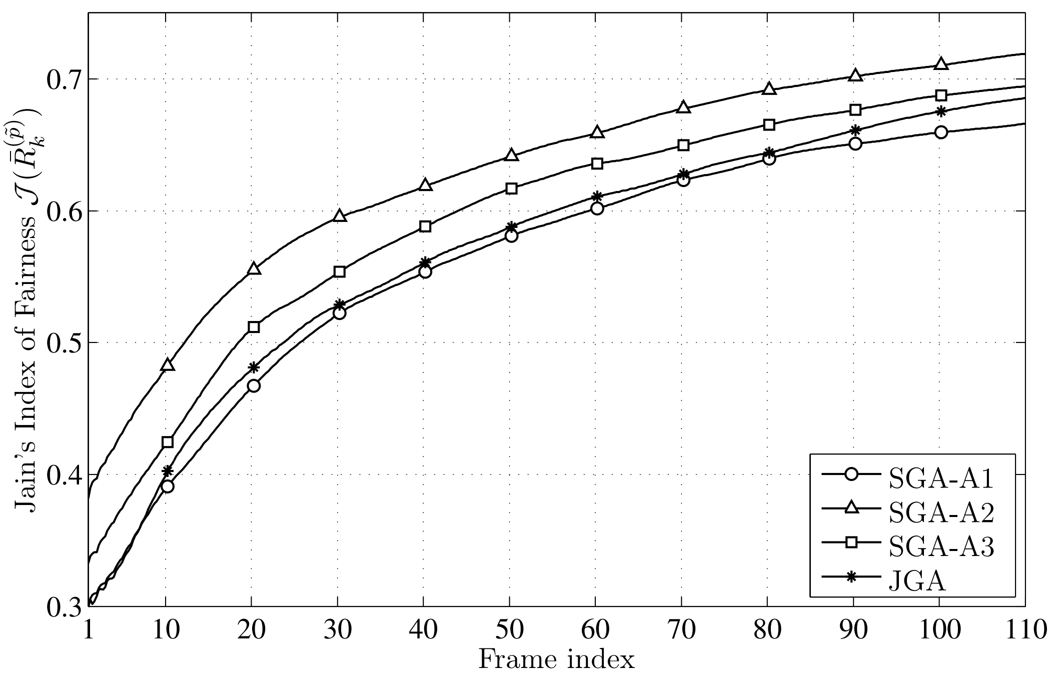
In Fig. 4.4(a), it can be seen that the RA strategies with the different assignment algorithms achieve almost the same average sum rate. The slightly higher average sum rates of the SGA-A1, SGA-A3, and JGA compared to that of the SGA-A2 are due to the less fair assignment of resources to MSs. Considering SGA-A2, the same initial MS is not assigned on multiple resources. For the number $K = 16$ of MSs and the number $B = 8$ of resources considered in this work, eight distinct initial MSs will be assigned resources. Consequently, during each allocation period, half of the MSs are assigned resources and, consequently, one can say that MSs with low channel gain also have relatively higher chances of being allocated resources considering the SGA-A2 than considering SGA-A1, SGA-A3, and JGA.

Considering the JGA, it can be seen in Fig. 4.4(a) that the application of the precoding and power allocation algorithm over the resources altogether does not improve the average sum rate compared to the SGA-A1 and SGA-A3. Because the CAP-BFA is aware of the actual precoding and power allocation, the power redistribution performed by the WFA of Table 4.2 has only a marginal impact on the performance of the RA strategy. According to [Cal04], the performance of the RA strategies considering the WFA of Table 3.5 and EPA among resources or considering the WFA of Table 4.2 is expected to be the same for high average SNR values and/or low number of MSs per SDMA group. Thus, the JGA offers no benefit compared to the SGA-A1 and SGA-A3, but it leads to undesired additional complexity, as it can be deduced from results presented in Section 3.5.8 and the formulation of the CAP-BFA in Table 4.1.

In Fig. 4.4(b), it can be seen that the throughput fairness among the MSs obtained with the SGA-A2 is higher than that obtained with SGA-A1, SGA-A3, or JGA. As




(a) Average sum rate.




(b) Jain's Index of Fairness. Average SNR $\gamma = 10$ dB.

Figure 4.4. Average sum rate and throughput fairness achieved by the RA strategies considering CAP-BFA, CM criterion, and the different assignment algorithms of Section 4.2.2.

discussed for Fig. 4.4(a), the resources are always allocated to different initial MSs with the SGA-A2, which leads to a more fair distribution of resources compared to SGA-A1, SGA-A3, or JGA.

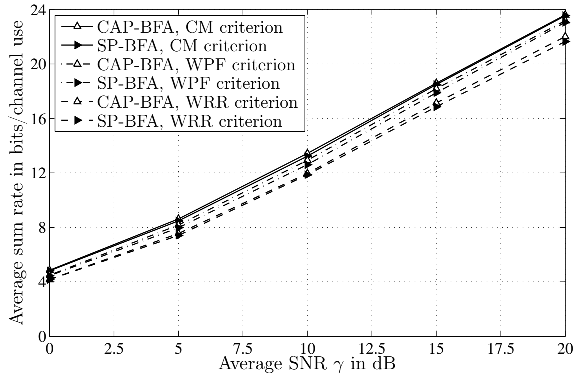
In Fig. 4.4(b), it can also be seen that the SGA-A3 achieves a slightly higher throughput fairness than the SGA-A1 and JGA. This result is explained in the sequel. With the

SGA-A1, the MS with the highest channel gain will be always selected as initial MS due to the considered CM criterion. Similarly, the SDMA groups built on a given resource by JGA will always start with the MS having the highest channel gain on this resource. Thus, the MSs with high channel gain concentrate the resources when the the SGA-A1 or JGA are considered by the RA strategy. Differently from SGA-A1 and JGA, $K$ candidate SDMA groups considering different initial MSs are created on each resource with the SGA-A3, cf. Section 4.2.2.1. The group priority of each candidate group is computed using (4.1) and the resources are assigned to the best SDMA groups. Because there is no warranty that the best SDMA group on each resource contains the MS with the highest channel gain, the SGA-A3 might allocate resources to SDMA groups that differ from those built by the SGA-A1 and JGA. Consequently, the resources are not so concentrated by the MS with high channel gains. Moreover, because the group priority takes into account the whole SDMA group, the resource assignment performed by the SGA-A3 is efficient and no losses in terms of average sum rate are observed in Fig. 4.4(a) compared to SGA-A1 and JGA.

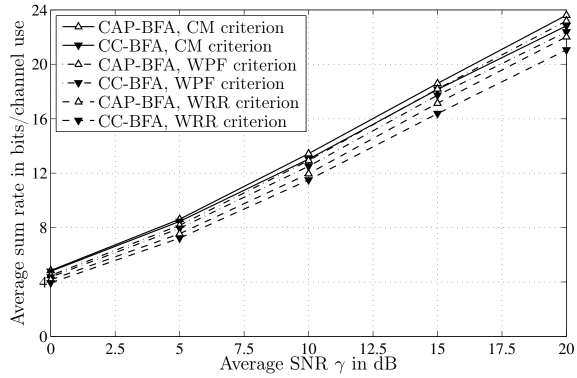
In the following, some comments regarding the complexity of the RA strategies considered in the analysis of Fig. 4.4 are made. In Fig. 4.4(a), the RA strategies presented almost the same performance in terms of average sum rate independent of the adopted assignment algorithm. However, the RA strategies considerably differ in terms of complexity. It is easy to conclude that the least complex RA strategy is the one employing SGA-A1, which just scales roughly by $B$ the complexity presented in Table 3.15 for CAP-BFA-ZF. The next least complex strategy is the one employing SGA-A2, which incorporates the assignment of the initial MS according to Munkres' algorithm. Then, the RA strategy employing SGA-A3 comes as next and its complexity is roughly $K \cdot B$ times larger than that presented in Table 3.15 for CAP-BFA-ZF. As last comes the RA strategy considering JGA, whose complexity is hard to estimate. However, since a number of groups larger than $K \cdot B$ might be tested when considering JGA, this strategy is considered as the most complex. The same comments hold for all the RA strategies that can be obtained by combining the elements presented in Table 4.3. Due to its relatively bad performance-complexity trade-off, JGA will not be considered in the sequel.

In the following, the average sum rate achieved by the RA strategies considering the different criteria for the MS or group priority and the assignment algorithms is investigated. The same analysis done for Fig. 4.4(a) considering CAP-BFA has been performed for the RA strategies considering the other SDMA algorithms. In Fig. 4.5, the performance in terms of average sum rate of RA strategies that use SP-BFA and CC-BFA is compared to that of the RA strategies that use the CAP-BFA considering the different SGA algorithms of Section 4.2.2.1 and the different criteria for the MS or group priority. Results related to the SP-BFA and CC-BFA are shown on the left and right sides of Fig. 4.5, respectively.

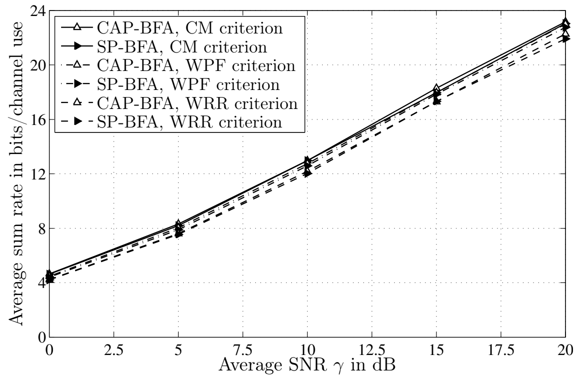
In Fig. 4.5, no results are presented for the CC-CGA because it has almost the same performance in terms of average sum rate and fairness as the CC-BFA. The CC-BFA has been preferred here because it is less complex than CC-CGA, cf. Section 3.5.8.



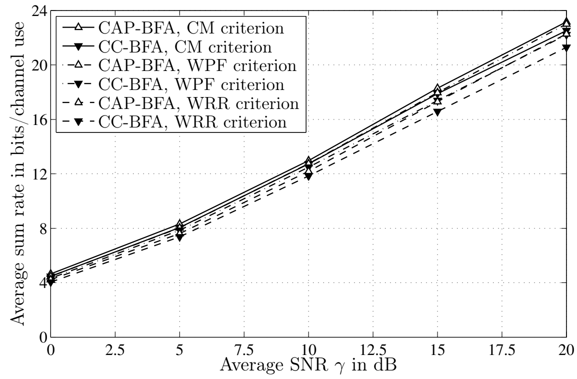

(a) SP-BFA, SGA-A1.



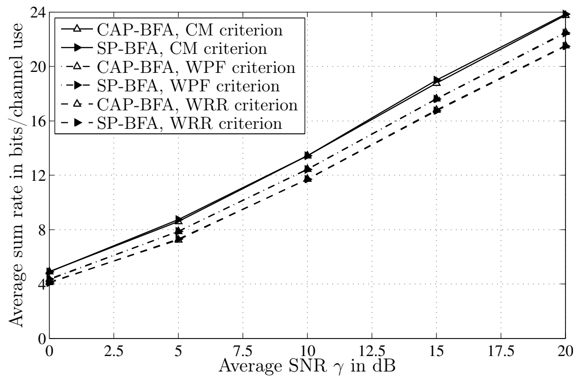

(b) CC-BFA, SGA-A1.



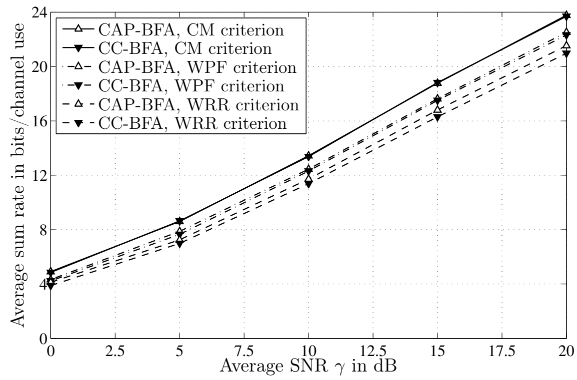

(c) SP-BFA, SGA-A2.




(d) CC-BFA, SGA-A2.




(e) SP-BFA, SGA-A3.




(f) CC-BFA, SGA-A3.

Figure 4.5. Average sum rate achieved using SP-BFA (left) and CC-BFA (right) compared to that of CAP-BFA (left/right). The different SGA algorithms and criterion for MS or group priority are considered.

In Fig. 4.5, it can be seen that the SP-BFA and CC-BFA have obtained almost the same performance than the CAP-BFA independent of the priority criterion being considered. It can also be seen that the average sum rates achieved with the SP-BFA in Fig. 4.5(a), Fig. 4.5(c), and Fig. 4.5(e) are slightly higher than that achieved with CC-BFA in Fig. 4.5(b), Fig. 4.5(d), and Fig. 4.5(f), respectively. This relative performance is similar to that obtained in Chapter 3 and has already been discussed in Section 3.5. Indeed, the RA strategies using the SP-BFA and CC-BFA achieve over 95% of the average sum rate achieved with the corresponding RA strategies using CAP-BFA, as it can be seen in Fig. 4.5. However, as it has been discussed in Section 3.5, the SP-BFA and especially the CC-BFA are considerably less complex than the CAP-BFA.

Also similarly to the results discussed in Section 3.5.9, the adoption of the WPF and WRR criteria leads only to small reductions in the average sum rate achieved considering the CM criterion. Considering an average SNR $\gamma$ of 10 dB, the average sum rates achieved with SGA-A1, SGA-A2, and SGA-A3 considering the WPF criterion are 4%, 2%, and 7% lower than considering the same assignment algorithms and the CM criterion, respectively. Considering the WRR criterion, the average sum rates achieved with SGA-A1, SGA-A2, and SGA-A3 are 12%, 7%, and 15% lower than considering the same assignment algorithms and the CM criterion, respectively.

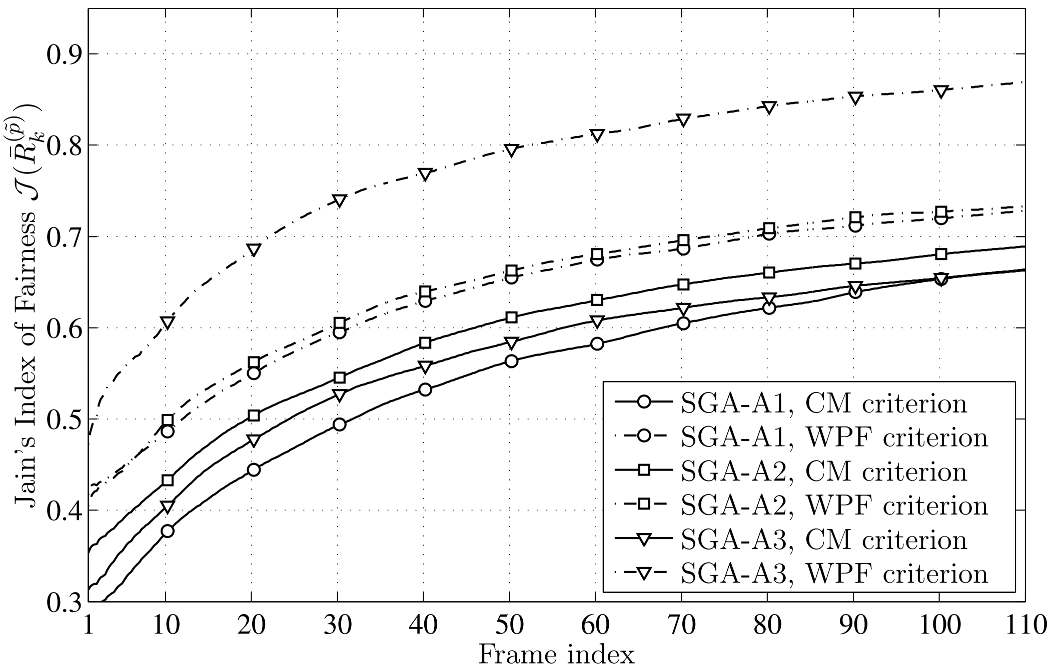
According to the results shown in Fig. 4.4 and Fig. 4.5, it can be noted that the RA strategies employing either SGA-A1, SGA-A2, or SGA-A3 have quite similar performances in terms of the achieved average sum rate. It can also be noted that SGA-A3 presents the highest reductions in the average sum rate when considering the WPF and WRR criteria. These somehow higher reductions in the average sum rate are a consequence of the use of group priorities by the assignment algorithm. Because the priority of all MSs in the SDMA groups are considered in the resource assignment, an assignment of resources oriented to the QoS provision is more effectively performed by the SGA-A3, as it is discussed in the sequel.

In order to illustrate the effectiveness of SGA-A3 to improve the throughput fairness among the MSs compared to the other assignment algorithms, Fig. 4.6 shows JIF achieved by the RA strategies considering the SP-BFA and CC-BFA, the CM and the WPF criteria, and the different assignment algorithms for an average SNR $\gamma$ of 10 dB.

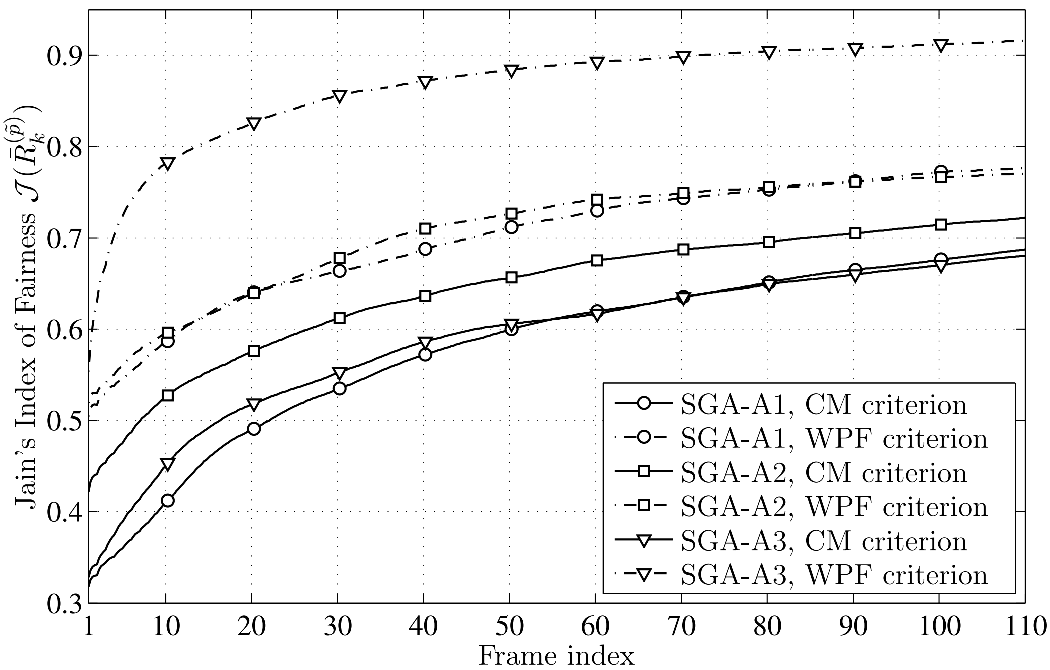
Before discussing the results in Fig. 4.6, some remarks are made. No results are presented for the CC-CGA for the same reasons discussed regarding Fig. 4.5. Additionally, the throughput fairness results obtained with RA strategies considering the WRR criterion are located in between those obtained with the CM and the WPF criteria, thus being omitted in Fig. 4.6.

Considering the WPF criterion, it can be seen in Fig. 4.6 that the SGA-A3 considerably improves the degree of throughput fairness compared to the other assignment algorithms. In Fig. 4.6(a), it can be observed that after a short time (less than 100 ms), the RA strategy considering the SP-BFA and the SGA-A3 achieves a fairness index of 0.85. As discussed in Section 3.5.9, a JIF value of $0.85 \cdot 100$ can be interpreted as having 85% of the MSs being fairly served and 15% of the MSs getting absolutely no throughput. Moreover, considering the WPF criterion the RA strategies can still attain about 85% of the sum rate achievable considering the CM criterion. For the CC-BFA in Fig. 4.6(b), the values assumed by JIF considering the WPF criterion are even higher. The reason for a lower degree of fairness for the SP-BFA compared to the CC-BFA is that the SP-BFA is more greedy than the CC-BFA and favors MSs with high channel gain, while the CC-BFA looks for a trade-off between spatial correlation and channel gains.

The higher degree of fairness achieved with the SGA-A3 is a consequence of the use of group priorities instead of MS priorities, which makes this assignment algorithm




(a) SP-BFA.




(b) CC-BFA.

Figure 4.6. JIF for the RA strategies employing the SP-BFA and the CC-BFA, the CM and WPF criteria, and the different assignment algorithms. Average SNR $\gamma = 10$ dB

capable of following the priority criteria more effectively than the other assignment algorithms. However, this higher degree of fairness is reached at the expense of additional complexity and reduced sum rate, as already discussed in this section.

Because the SGA-A3 is the assignment algorithm providing the best degree of fairness, it is of interest to see how the system throughput is divided among low- and high-priority MSs when considering this assignment algorithm. Analogously to Fig. 3.19 presented in Chapter 3, Fig. 4.7 shows the percentual difference between the throughput of high- and low-priority MSs for RA strategies combining SP-BFA, CC-BFA, or CC-CGA with SGA-A3 and considering the WPF and WRR criteria. An average SNR $\gamma$ of 10 dB is considered.

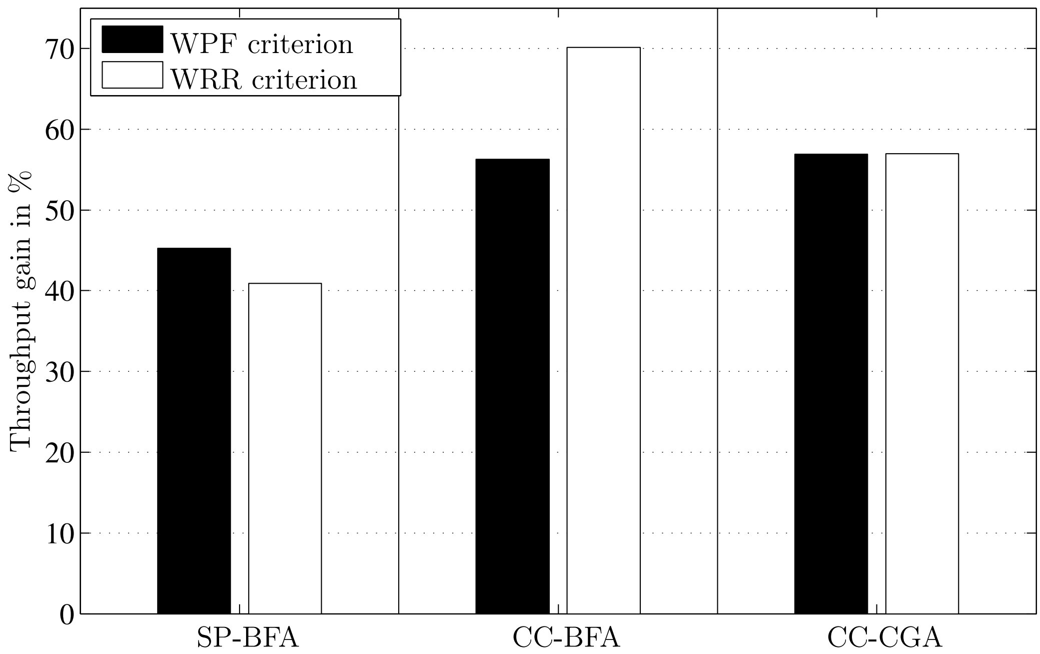



Figure 4.7. Percentual throughput difference for high-priority MSs compared to low-priority MSs considering SP-BFA, CC-BFA, or CC-CGA with SGA-A3 and the WPF or WRR criteria. Average SNR $\gamma = 10$ dB.

Compared to the results in Fig. 3.19, which can be associated to the SGA-A1, a significant improvement in the throughput of high-priority MSs is observed for both WPF and WRR criteria in Fig. 4.7 and high-priority MSs perceive throughputs at least 40% higher than low-priority MSs. Again, the reason for the improved throughput gains for high-priority MSs is the better management of priorities performed by the SGA-A3.

Comparing the results for the WRR criterion in Fig. 3.19 and Fig. 4.7, it can also be noted that the throughput gains obtained with the WRR criterion and the SGA-A3 are considerably higher than in Fig. 3.19. Because the priorities given by (3.56) of all the MSs in an SDMA group are considered by the SGA-A3 and not only that of the initial MS $k'$, the groups containing high-priority MSs get a considerably higher fraction of the resources than the low-priority MSs. Consequently, the fraction of resources assigned to the MSs is distributed more proportionally to their static priorities in (3.56) and the adoption of the WRR criterion effectively gives high-priority MSs a meaningful throughput advantage over the low-priority ones, as it is desired for this criterion. However, because the priorities are not taken into account by the SDMA algorithms, MSs might still get higher amount of resources than that expected from its static priority. It must also be noted that considering the WRR criterion the throughput fairness among MSs might be reduced since groups containing oversatisfied high-priority MSs will continue being allocated resources more often than groups containing low-priority MSs.

An improved QoS control and, consequently, improved throughput fairness among MSs may be achieved by employing precoding and power allocation algorithms oriented towards this objective, such as those in [SB04, MK07c, MK07c]. However, the investigation of this topic is out of the scope of this thesis.

In the following, some results presented in this chapter are shortly summarized.

In this chapter, different suboptimal RA strategies have been studied considering multiple radio resources. The studied RA strategies aim at maximizing the sum rate of the system and are composed of a GA algorithm and a precoding and power allocation algorithm.

Only a subset of the SDMA algorithms considered in Chapter 3, as well as only linear ZF precoding, have been considered in this chapter, which has focused more on the assignment algorithms, on the criteria for MS or group priorities, and on their implications on the fairness among the MSs.

Both SGA and JGA have been investigated and it has been shown that, considering an EPA among resources and the SGA-A1, a sequential assignment of resources to SDMA groups provides the same performance in terms of average sum rate as a joint SDMA grouping and resource assignment, while being however considerably less complex than the JGA.

Different assignment algorithms performing SGA have been investigated. All of them achieved quite similar performances in terms of the average sum rate. However, these assignment algorithms differ in the way in which MS or group priorities are used to perform the resource assignment and in terms of complexity.

In particular, it has been shown that assigning a different initial MS to each resource with the SGA-A2 can already enhance the fairness of the RA strategies without compromising substantially the average sum rate. Additionally, this fairness enhancement has no negative impact in terms of complexity since Munkres' algorithm involves only simple additions, subtractions, and logical operations, but no multiplications, and has consequently low complexity.

For the SGA-A3, where priorities are computed considering the whole SDMA groups to which the resources are assigned, it has been shown that considerable fairness improvements can be achieved compared to the other assignment algorithms. A higher degree of fairness has been observed for the WPF criterion and arises from the fact that group priorities are employed instead of MS priorities in order to decide which group to assign to each resource. Nevertheless, it must be noted that MS or group priorities and the grouping metrics are kept relatively decoupled so that the spatial compatibility in an SDMA group is not compromised by the QoS control mechanisms. In particular, it has been observed that for the assignment algorithm SGA-A3, the WRR criterion provides a considerable improvement on the throughput of high-priority MSs. However, this improvement comes at the expense of reduced average sum rate.

The complexity of the considered RA strategies, and in particular of the assignment algorithms, has not been explored in details. Nevertheless, one can simply deduce that the complexity the RA strategies as a function of the selected assignment algorithm increases in the sequence SGA-A1, SGA-A2, SGA-A3, and JGA, while the dependency of the complexity on the SDMA algorithms can be derived from the results presented in Chapter 3.

From the analyses conducted in this chapter, the consideration of group priorities directly in the SDMA algorithms can allow to design RA strategies merging the benefits of the SGA-A1 and of the SGA-A3, i.e., the simplicity of the SGA-A1 and the effectiveness of the SGA-A3 in the priority management. Such an RA strategy is not considered in this work and might be topic of future investigations.

# Chapter 5

# Conclusions

This thesis deals with suboptimal RA strategies in the DL of MIMO-OFDMA systems aiming at the maximization of the sum rate. In order to solve the problem of maximizing the sum rate with affordable complexity, a new formulation for the problem is proposed in Chapter 1 which divides it into four subproblems, namely the SDMA grouping problem, the precoding problem, the power allocation problem, and the resource assignment problem. This model is proposed as a framework for suboptimal RA strategies into which several RA strategies available in the literature suitably fit into. The mathematical formulation of the problem is presented in Chapter 2, in which some insight on the nature and interdependency of the four mentioned subproblems is provided. Then, for each subproblem, several existing or newly proposed algorithms oriented towards the maximization of the sum rate of the system are applied. Through the combination of these algorithms, new suboptimal rather efficient RA strategies are obtained.

Considering a scenario with a single resource, or in which resources are considered one-by-one, various suboptimal RA strategies are proposed in Chapter 3 and their performance in terms of achieved average sum rate is investigated. In particular, four new SDMA algorithms, namely the CC-CGA, CC-BFA, WN-BFA, and WN-FFA are proposed to solve the SDMA grouping problem. The proposed SDMA algorithms are shown to perform as good as the CAP-BFA and SP-BFA and are shown to achieve over 90% of the average sum rate obtained by the CAP-ESA-ZF strategy, which performs an ES for the SDMA group maximizing the sum rate of the system. Expressions for the complexity of the RA strategies considering ZF have been derived and it has been shown that the proposed RA strategies considering the CC-CGA and especially the CC-BFA have in many cases a considerably lower computational complexity than those employing CAP-BFA and SP-BFA as SDMA algorithm. RA strategies employing the WN-BFA and WN-FFA proposed in Chapter 3 offer a suboptimal solution to the RA problem with even lower complexity than those employing the CC-CGA and CC-BFA as SDMA algorithm, but with slightly worse performance in terms of average sum rate. The performance of the RA strategies has also been investigated considering imperfect CSI, to which the referred SDMA algorithms are as sensitive as the other algorithms considered for comparison. Thus, the proposed SDMA algorithms are shown to provide a good performance-complexity trade-off compared to the existing SDMA algorithms selected for benchmarking.

In Chapter 3, a new iterative SDA is proposed for the power allocation problem. This algorithm is subsequently combined with GEP into a new iterative joint precoding and power allocation algorithm, for which a proof of convergence is also provided. The proposed iterative joint precoding and power allocation algorithm is flexible and can pursue either the maximization of the sum rate or the provision of QoS to the MSs by means of a simple parameter setting. Anyway, only the maximization of the sum rate has been considered in this thesis and, in this case, separated precoding and power allocation is shown to perform as well as the joint precoding and power allocation, but with lower complexity.

Also as part of the precoding and power allocation algorithm, a new SRA is proposed, which removes MSs from SDMA groups as to enhance the sum rate. When considered by RA strategies whose SDMA algorithms are unaware of the actual precoding and power allocation, the proposed SRA is able to improve by 10% to 30% the average sum rate achieved by the RA strategies.

For the resource assignment problem, algorithms performing either SGA or JGA are proposed and compared in Chapter 4. For the maximization of the sum rate of the system, it is shown that a sequential assignment of resources to SDMA groups performs as good as JGA while being considerably more simple. Moreover, different criteria to prioritize MSs or SDMA groups are considered by the assignment algorithms proposed in Chapter 4, and it is shown that by adopting a suitable priority criterion, the throughput fairness among the MSs can be considerably improved at the expense of only small reductions in the average sum rate of the system.

The SGA-A2, which assigns resources to different initial MSs, has been shown to enhance the fairness of the RA strategies without compromising substantially the average sum rate of the system or incurring increased complexity. The SGA-A3, which employs group priorities and assigns resources directly to whole SDMA groups, has also been shown to achieve good average sum rate figures and considerably higher throughput fairness than the other assignment algorithms. This result arises from selecting a subset of SDMA groups from a set of candidate SDMA groups and from the fact of employing group priorities, both aspects which are not considered by the other algorithms. However, considering the maximization of the sum rate of the system, the other SGA algorithms represent solutions with lower complexity and almost the same performance.

# Appendix

## A.1 Complexity of some mathematical operations and functions

In this section, the complexity of some general mathematical operations is presented, which is required for determining the complexity of the RA strategies of Section 3.5.8. Table A.1 shows the complexity of several mathematical operations involving scalars, vectors, and matrices. The complexity of the operations is expressed in terms of the number of required complex multiplications, and the complexity order $\mathcal{O}(\cdot)$ follows the big O notation [GL96]. Part of the complexity values shown in Table A.1 have been obtained from [Hun07]. Therein, a similar table is provided, which includes the summations as well. In Table A.1, the Eigenvalue Decomposition (EVD) is assumed to be computed by the QR algorithm [GL96], and the generalized EVD is assumed to be computed by performing a matrix inversion and a standard EVD.

Table A.1. Complexity of mathematical operations.

| **Operation/function** | **Expression** | **# of multiplications** | $\mathcal{O}(\cdot)$ |
|---|---|---|---|
| Product | $xy$ | 1 | $\mathcal{O}(1)$ |
| Division | $x/y$ | 1 | $\mathcal{O}(1)$ |
| Square root | $\sqrt{x}$ | 1 | $\mathcal{O}(1)$ |
| Vector scaling | $x\mathbf{y}, \quad \mathbf{y}: a \times 1$ | $a$ | $\mathcal{O}(a)$ |
| Inner product | $\mathbf{x}^{\mathrm{H}}\mathbf{y}, \quad \mathbf{x}, \mathbf{y}: a \times 1$ | $a$ | $\mathcal{O}(a)$ |
| Self outer product | $\mathbf{x}\mathbf{x}^{\mathrm{H}}, \quad \mathbf{x}: a \times 1$ | $(a^2 + a)/2$ | $\mathcal{O}(a^2/2)$ |
| Outer product | $\mathbf{x}\mathbf{y}^{\mathrm{H}}, \quad \mathbf{x}: a \times 1, \mathbf{y}: b \times 1$ | $ab$ | $\mathcal{O}(ab)$ |
| Matrix scaling | $x\mathbf{Y}, \quad \mathbf{Y}: a \times b$ | $ab$ | $\mathcal{O}(ab)$ |
| Matrix-vector product | $\mathbf{X}\mathbf{y}, \quad \mathbf{X}: a \times b, \mathbf{y}: b \times 1$ | $ab$ | $\mathcal{O}(ab)$ |
| Matrix-matrix product | $\mathbf{X}\mathbf{Y}, \quad \mathbf{X}: a \times b, \mathbf{Y}: b \times c$ | $abc$ | $\mathcal{O}(abc)$ |
| Gram matrix generation | $\mathbf{X}\mathbf{X}^{\mathrm{H}}, \quad \mathbf{X}: a \times b$ | $(a^2 b + ab)/2$ | $\mathcal{O}(a^2 b/2)$ |
| Frobenius norm | $\lVert\mathbf{X}\rVert_{\mathrm{F}}, \quad \mathbf{X}: a \times b$ | $ab + 1$ | $\mathcal{O}(ab)$ |
| Matrix-diagonal matrix product | $\mathbf{X}\mathcal{D}\{\mathbf{y}\}, \quad \mathbf{X}: a \times b, \mathbf{y}: b \times 1$ | $ab$ | $\mathcal{O}(ab)$ |
| Matrix inversion | $\mathbf{X}^{-1}, \quad \mathbf{X}: a \times a$ | $(a^3 + 3a^2)/2$ | $\mathcal{O}(a^3/2)$ |
| Matrix pseudo-inversion[a] | $\mathbf{X}^{\dagger} = \mathbf{X}^{\mathrm{H}}(\mathbf{X}\mathbf{X}^{\mathrm{H}})^{-1}, \quad \mathbf{X}: a \times b$ | $(a^3 + 3a^2 b + 3a^2 + ab)/2$ | $\mathcal{O}(a^3/2)$ |
| EVD[b] | $\mathbf{X}\mathbf{y} = \lambda\mathbf{y}, \mathbf{X}: a \times a, \mathbf{y}: a \times 1$ | $2a^3/3 + \mathcal{O}(a^2)$ | $\mathcal{O}(2a^3/3)$ |
| Generalized EVD[b] | $\mathbf{X}\mathbf{y} = \lambda\mathbf{Z}\mathbf{y}, \quad \mathbf{X}, \mathbf{Z}: a \times a, \mathbf{y}: a \times 1$ | $(7a^3 + 9a^2)/6 + \mathcal{O}(a^2)$ | $\mathcal{O}(7a^3/6)$ |

[a] $\mathbf{X}$ is assumed to be full row-rank.
[b] $\mathbf{X}$ and $\mathbf{Z}$ are assumed to be symmetric and full rank.

# List of Acronyms

| | |
|---|---|
| **3GPP** | 3rd Generation Partnership Project |
| **AA** | Antenna Array |
| **AoA** | Angle of Arrival |
| **AoD** | Angle of Departure |
| **AWGN** | Additive White Gaussian Noise |
| **B-CSIT** | Block Channel State Information at the Transmitter |
| **BC** | Broadcast Channel |
| **BD** | Block Diagonalization |
| **BFA** | Best Fit Algorithm |
| **BS** | Base Station |
| **CAP** | Capacity |
| **CC** | Convex Combination |
| **CGA** | Convex Grouping Algorithm |
| **CM** | Capacity Maximization |
| **COA** | Compatibility Optimization Algorithm |
| **CP** | Cyclic Prefix |
| **CSI** | Channel State Information |
| **CSIT** | Channel State Information at the Transmitter |
| **DL** | Downlink |
| **DPC** | Dirty Paper Coding |
| **EP** | Eigen-Precoding |
| **EPA** | Equal Power Allocation |
| **ES** | Exhaustive Search |
| **ESA** | Exhaustive Search Algorithm |
| **EVD** | Eigenvalue Decomposition |
| **FDD** | Frequency Division Duplexing |
| **FDMA** | Frequency Division Multiple Access |
| **FFA** | First Fit Algorithm |
| **FFT** | Fast Fourier Transform |
| **GA** | Grouping & Assignment |
| **GEP** | Generalized Eigen-Precoding |
| **GI** | Guard Interval |
| **HPF** | Highest Power First |
| **ICI** | Inter-Carrier Interference |
| **ISI** | Inter-Symbol Interference |
| **JGA** | Joint Grouping and Assignment |
| **JIF** | Jain's Index of Fairness |
| **KKT** | Karush-Kuhn-Tucker |
| **LGF** | Lowest Gain First |
| **LOS** | Line Of Sight |

| | |
|---|---|
| **MAC** | Multiple Access Channel |
| **MIMO** | Multiple Input Multiple Output |
| **MISO** | Multiple Input Single Output |
| **MMSE** | Minimum Mean Square Error |
| **MS** | Mobile Station |
| **MU** | Multi-User |
| **NLOS** | Non Line Of Sight |
| **NP-C** | Non-deterministic Polynomial time Complete |
| **OFDM** | Orthogonal Frequency Division Multiplexing |
| **OFDMA** | Orthogonal Frequency Division Multiple Access |
| **P-CSIR** | Perfect Channel State Information at the Receiver |
| **P-CSIT** | Perfect Channel State Information at the Transmitter |
| **PC** | Power Control |
| **PRB** | Physical Resource Block |
| **QoS** | Quality of Service |
| **RA** | Resource Allocation |
| **RGA** | Random Grouping Algorithm |
| **RR** | Round Robin |
| **RV** | Random Variable |
| **S-CSIT** | Second-order Channel State Information at the Transmitter |
| **SDA** | Soft Dropping Algorithm |
| **SDMA** | Space Division Multiple Access |
| **SDP** | Semidefinite Programming |
| **SDPC** | Soft Dropping Power Control |
| **SGA** | Separated Grouping and Assignment |
| **SGA-A1** | SGA Algorithm 1 |
| **SGA-A2** | SGA Algorithm 2 |
| **SGA-A3** | SGA Algorithm 3 |
| **SIC** | Successive Interference Cancellation |
| **SINR** | Signal-to-Interference plus Noise Ratio |
| **SISO** | Single Input Single Output |
| **SNR** | Signal-to-Noise Ratio |
| **SP** | Successive Projection |
| **SRA** | Sequential Removal Algorithm |
| **SU** | Single-User |
| **SVD** | Singular Value Decomposition |
| **TDD** | Time Division Duplexing |
| **TDMA** | Time Division Multiple Access |
| **TS** | Time-Slot |
| **UL** | Uplink |
| **ULA** | Uniform Linear Array |
| **WFA** | Water Filling Algorithm |

| | |
|---|---|
| WIM | WINNER Phase I Channel Model |
| WINNER | Wireless World Initiative New Radio |
| WN | Weighted Norm |
| WPF | Weighted Proportional Fair |
| WRR | Weighted Round Robin |
| ZF | Zero-Forcing |
| ZMCSCG | Zero Mean Circularly Symmetric Complex Gaussian |

# List of Symbols

| | |
|---|---|
| $\mathbf{1}_M$ | $M \times 1$ vector of ones |
| $\mathbf{a}$ | Attenuation vector containing the channel attenuation of all MSs |
| $\mathbf{a}_\mathcal{G}$ | Attenuation vector containing the channel attenuation for the MSs in the SDMA group $\mathcal{G}$ |
| $\arg\max_x \{f(x)\}$ | Maximum argument $x$ of the function $f(x)$ |
| $\arg\min_x \{f(x)\}$ | Minimum argument $x$ of the function $f(x)$ |
| $B$ | Number of resources, i.e., frequency blocks |
| $b$ | Resource index |
| $B_\mathrm{c}$ | Coherence bandwidth |
| $\beta$ | Parameter controlling the relevance given to spatial correlation and channel gain in the CGA |
| $B_\mathrm{sys}$ | System bandwidth |
| $\mathbf{C}$ | Correlation matrix containing the spatial correlation among each pair of MSs |
| $\mathbf{C}_\mathcal{G}$ | Correlation matrix containing the spatial correlation among each pair of MSs in the SDMA group $\mathcal{G}$ |
| $\tilde{\mathbf{C}}$ | Diagonally loaded positive semidefinite version of the spatial correlation matrix $\mathbf{C}$ |
| $\Delta f$ | Subcarrier bandwidth |
| $\Delta_R$ | Distance between adjacent elements of the AA of an MS |
| $\Delta_T$ | Distance between adjacent elements of the AA of the BS |
| $\mathbf{E}_{k,b}$ | Additive ZMCSCG channel estimation error term for the $k^\mathrm{th}$ MS on frequency block $b$ |
| $\mathbf{e}_m$ | $m^\mathrm{th}$ canonical base vector |
| $\epsilon$ | Diagonal loading factor of $\tilde{\mathbf{C}}$ |
| $\epsilon_\mathrm{p}$ | Power precision for the convergence of the SDA |
| $\varepsilon$ | Upper bound on $\epsilon$ |
| $\boldsymbol{\eta}$ | Noise vector for the DL power allocation of the SDA |
| $\mathcal{E}\{\cdot\}$ | Statistical expectation |
| $f_{(\cdot)}(\mathcal{G})$ | A grouping metric for an SDMA group $\mathcal{G}$ |
| $f_\mathrm{CAP}(\mathcal{G})$ | Group capacity of the SDMA group $\mathcal{G}$ |
| $f_\mathrm{CC}(\mathcal{G})$ | Convex combination of the total spatial correlation and channel gains |
| $\tilde{\mathbf{W}}_{k,b}$ | Receive beamformer matrix of MS $k$ on frequency block $b$ |
| $\check{\mathbf{w}}_{k,b,n}$ | Receive beamformer of MS $k$ on frequency block $b$ and spatial layer $n$ |
| $\tilde{\mathbf{w}}_{k,b,n}$ | Approximate receive beamformer of MS $k$ on frequency block $b$ and spatial layer $n$ |
| $\lfloor\cdot\rfloor$ | Nearest integer smaller than or equal to the argument |
| $f_\mathrm{SP}(\mathcal{G})$ | Sum of channel gains with null space SPs for the SDMA group $\mathcal{G}$ |
| $\tilde{f}$ | Frame index |

| | |
|---|---|
| $f_{\text{WN}}(\mathcal{G})$ | Weighted norm of the total spatial correlation |
| $\hat{\mathbf{G}}$ | Channel matrix of the SDMA group $\mathcal{G}$ |
| $G$ | Number of SDMA groups |
| $g$ | SDMA group index |
| $\gamma$ | Average SNR measured at the output of the AA of the BS |
| $\gamma_{\text{CSI}}$ | Measure of the quality of the CSI available at the BS |
| $\gamma_i$ | DL SINR of MS $i$ in an SDMA group |
| $\gamma_i'$ | UL SINR of MS $i$ in an SDMA group |
| $\gamma_{k,b}$ | DL SINR of MS $k$ on frequency block $b$ |
| $\bar{\gamma}_i^{(\tilde{t})}, \bar{\Gamma}_i^{(\tilde{t})}$ | Target SINR in linear and dB scales, respectively, for MS $i$ at iteration $\tilde{t}$ of the SDA |
| $\check{\gamma}, \check{\Gamma}$ | Minimum target SINR in linear and dB scales, respectively, for the SDA |
| $\hat{\gamma}, \hat{\Gamma}$ | Maximum target SINR in linear and dB scales, respectively, for the SDA |
| $\mathcal{G}$ | An SDMA group |
| $\mathcal{G}^\star$ | The best SDMA group |
| $\odot$ | Hadamard product between vectors/matrices |
| $\hat{\mathbf{H}}_b$ | Estimated MIMO channel matrix of all MSs on frequency block $b$ |
| $\hat{\mathbf{H}}_{k,b}$ | Estimated MIMO channel matrix of MS $k$ on frequency block $b$ |
| $\mathbf{H}_{k,s}$ | MIMO channel matrix of MS $k$ on subcarrier $s$ |
| $\hat{\mathbf{H}}_{k,s}$ | Estimated MIMO channel matrix of MS $k$ on subcarrier $s$ |
| $h_{n,m}$ | Channel coefficient between the $m^{\text{th}}$ antenna of the BS and the $n^{\text{th}}$ antenna of an MS on a subcarrier $s$ |
| $\mathbf{H}_s$ | MIMO channel matrix of all MSs on subcarrier $s$ |
| $\hat{\mathbf{H}}_s$ | Estimated MIMO channel matrix of all MSs on subcarrier $s$ |
| $\mathcal{I}(\mathbf{p}), \mathcal{I}(\mathbf{q})$ | Standard interference functions |
| $\mathbf{I}_M$ | $M \times M$ identity matrix |
| $I_{\text{CGA}}$ | Number of iterations required by the CGA to converge |
| $\bar{I}_{\text{CGA}}$ | Estimated average number of iterations required by the CGA to converge |
| $(\cdot)^{-1}$ | Inversion of a matrix |
| $I_{\text{SDA}}$ | Number of iterations required by the SDA to converge |
| $\mathcal{J}(\bar{R}_k^{(\tilde{p})})$ | Jain's Index of Fairness |
| $K$ | Number of MSs |
| $k$ | MS index |
| $K_A$ | Total number of MSs to which resources have been allocated |
| $K_{\mathcal{G}}$ | Number of MSs in $\mathcal{G}$ |
| $K_{\mathcal{G}}^\star$ | Target SDMA group size |
| $k'$ | Initial MS selected to build SDMA groups |
| $L$ | Number of clusters in the WIM |

| | |
|---|---|
| $l$ | Clusters index in the WIM |
| $\ell_1$ | Additional term in the recursive form $\mu_{\mathcal{G}'}$ of $\mu_{\mathcal{G}}$ |
| $\ell_2$ | Additional term in the recursive form $\\|\mathbf{C}_{\mathcal{G}'}\\|_{\mathrm{F}}$ of $\\|\mathbf{C}_{\mathcal{G}}\\|_{\mathrm{F}}$ |
| $\ell_3$ | Additional term in the recursive form $f_{\mathrm{CC}}(\mathcal{G}')$ of $f_{\mathrm{CC}}(\mathcal{G})$ |
| $\mathbf{\Lambda}_{k,b}$ | Eigenvalue matrix of $\mathbf{R}_{k,b}$ |
| $\lambda^{-}(\mathbf{A})$ | Smallest eigenvalue of matrix $\mathbf{A}$ |
| $\lambda_i(\mathbf{A})$ | $i^{\mathrm{th}}$ eigenvalue of matrix $\mathbf{A}$ |
| $M$ | Number of elements of the AA of the BS |
| $\max\{\ldots\}$ | Maximum among the arguments |
| $\mathcal{M}(b,i), \mathcal{M}^{-1}(j)$ | Index mapping functions for the WFA over multiple resources |
| $\min\{\ldots\}$ | Minimum among the arguments |
| $\lvert\cdot\rvert$ | Absolute value of a complex number |
| $\mu$ | Optimal water-level of the WFA |
| $\mu_{\mathcal{G}}$ | Geometric mean of the square root of the channel gains of the MSs in the SDMA group $\mathcal{G}$ |
| $N$ | Total number of receive antennas |
| $N_0$ | AWGN power spectral density |
| $N_k$ | Number of elements of the AA of MS $k$ |
| $\\|\cdot\\|_1$ | 1-norm of a vector |
| $\\|\cdot\\|_2$ | 2-norm of a vector |
| $\\|\cdot\\|_{\mathrm{F}}$ | Frobenius norm of a vector/matrix |
| $\nu$ | Erroneous CSI model parameter |
| $\mathcal{N}$ | Null space |
| $\mathcal{O}(\cdot)$ | Complexity order of an algorithm |
| $\mathbf{p}$ | DL power vector containing the powers allocated to the MSs in an SDMA group |
| $P$ | Available transmit power of the BS |
| $\bar{\phi}_l$ | Average AoD of the $l^{\mathrm{th}}$ cluster in the WIM |
| $\bar{\varphi}_l$ | Average AoA of the $l^{\mathrm{th}}$ cluster in the WIM |
| $\phi_{l,j}$ | AoD of $j^{\mathrm{th}}$ scatterer/multipath component of the $l^{\mathrm{th}}$ cluster in the WIM |
| $\varphi_{l,j}$ | AoD of $j^{\mathrm{th}}$ scatterer/multipath component of the $l^{\mathrm{th}}$ cluster in the WIM |
| $p_i$ | DL power allocated to MS $i$ in an SDMA group |
| $p_{k,b}$ | Power allocated to MS $k$ on frequency block $b$ |
| $p^{\star}_{k,b}$ | Optimal power allocated to MS $k$ on frequency block $b$ |
| $p_i^{(\tilde{t})}, P_i^{(\tilde{t})}$ | Allocated power in W and dBW scales, respectively, for MS $i$ at iteration $\tilde{t}$ of the SDA |
| $\check{p}, \check{P}$ | Minimum allocable power in W and dBW scales, respectively, for the SDA |
| $\hat{p}, \hat{P}$ | Maximum allocable power in W and dBW scales, respectively, for the SDA |

| | |
|---|---|
| $(\cdot)^{\dagger}$ | Pseudo-inversion of a matrix |
| $\tilde{p}$ | Resource allocation period |
| $\mathbf{q}$ | UL power vector containing the powers allocated to the MSs in an SDMA group |
| $q_i$ | UL power allocated to MS $i$ in an SDMA group |
| $Q_{\text{sub}}$ | Number of adjacent subcarriers per frequency block |
| $Q_{\text{sym}}$ | Number of OFDMA symbols per TS |
| $\text{Re}\{\cdot\}$ | Real part of a complex number |
| $\rho_{j,k}$ | Maximum normalized scalar product |
| $\varrho_{k,j}$ | An arbitrary pairwise spatial correlation for combination with the CGA |
| $R_k^{(0)}$ | Contracted average throughput of MS $k$ with the WPF priority criterion |
| $\mathbf{R}_{k,b}$ | Spatial covariance matrix of MS $k$ on frequency block $b$ |
| $\bar{R}_k^{(\tilde{p})}$ | Perceived average throughput of MS $k$ at the allocation period $\tilde{p}$ with the WPF priority criterion |
| $S$ | Number of subcarriers |
| $s$ | Subcarrier index |
| $s_b$ | First subcarrier of the $b^{\text{th}}$ frequency block |
| $\overline{s}_b$ | Middle subcarrier of the $b^{\text{th}}$ frequency block |
| $\sigma^2$ | Average AWGN power per subcarrier |
| $\mathbf{T}_i$ | Null space projection matrix for the $i^{\text{th}}$ MS in an SDMA group |
| $T$ | Number of TSs per frame |
| $\tilde{\tau}$ | Feedback delay |
| $J_0\left(\frac{2\pi v_{\text{MS}}\tilde{\tau}}{\lambda}\right)$ | Spaced-time correlation function |
| $T_{\text{c}}$ | Coherence time |
| $T_{\text{d}}$ | Duplexing time |
| $T_{\text{FRM}}$ | Frame duration |
| $(\cdot)^{\text{T}}$ | Transpose of a vector/matrix |
| $(\cdot)^{\text{H}}$ | Transpose conjugate of a vector/matrix |
| $T_{\text{s}}$ | OFDMA symbol duration |
| $\tilde{t}$ | Iteration index of the SDA |
| $T_{\text{TS}}$ | TS duration |
| $\mathbf{U}$ | Resource-to-MS binary assignment matrix |
| $\tilde{\mathbf{u}}$ | Continuous relaxed version of $\mathbf{u}$ |
| $\mathbf{u}$ | Binary selection vector for SDMA grouping |
| $\boldsymbol{\mathfrak{U}}$ | MS priority matrix |
| $u_k$ | The $k^{\text{th}}$ component of $\mathbf{u}$ |
| $u_{k,b}$ | Binary variable indicating whether MS $k$ is allocated on frequency block $b$ |
| $u_{k,b}^{\star}$ | Optimal binary variable indicating whether MS $k$ is allocated on frequency block $b$ |

| | |
|---|---|
| $\mathfrak{u}_k$ | Priority of MS $k$ |
| $\mathfrak{u}_k^{(0)}$ | Static priority of MS $k$ for the WRR and WPF priority criteria |
| $\mathfrak{u}_k^{(\tilde{p})}$ | Priority of MS $k$ at the allocation period $\tilde{p}$ |
| $\tilde{u}_k$ | Continuous relaxed version of $u_k$ |
| $\boldsymbol{\Upsilon}$ | Gain matrix for the DL power allocation of the SDA |
| $\mathbf{u}^\star$ | Optimal binary selection vector associated with the best SDMA group $\mathcal{G}^\star$ |
| $\tilde{\mathbf{u}}^\star$ | Continuous relaxed version of $\mathbf{u}^\star$ |
| $\mathbf{V}$ | Resource-to-group binary assignment matrix |
| $\mathfrak{V}$ | Group priority matrix |
| $v_{g,b}$ | Binary variable indicating whether SDMA group $g$ is allocated on frequency block $b$ |
| $\mathfrak{v}_{g,b}$ | Priority of the SDMA group $\mathcal{G}_{g,b}$ on resource $b$ |
| $v^\star_{g,b}$ | Optimal binary variable indicating whether SDMA group $g$ is allocated on frequency block $b$ |
| $\mathbf{V}_{k,b}$ | Eigenvector matrix of $\mathbf{R}_{k,b}$ |
| $v_{\text{MS}}$ | Average speed of the MSs |
| $\mathbf{v}_i(\mathbf{A})$ | $i^{\text{th}}$ eigenvector of matrix $\mathbf{A}$ |
| $\mathbf{W}$ | Precoding matrix for SDMA group $\mathcal{G}$ |
| $W$ | Window size for S-CSIT |
| $\mathbf{w}_{k,b}$ | Precoding vector of MS $k$ on frequency block $b$ |
| $\mathbf{w}^\star_{k,b}$ | Optimal precoding vector of MS $k$ on frequency block $b$ |
| $\xi$ | Power backoff parameter of SDA |
| $\zeta$ | SDA parameter relating allocated powers and target SINRs |

# Bibliography

[3GP06a] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); radio interface protocol aspects," 3rd. Generation Partnership Project, Tech. Rep. TR 25.813 V7.1.0, Sept. 2006. [Online]. Available: http://www.3gpp.org

[3GP06b] ——, "Physical Layer Aspects for Evolved Universal Terrestrial Radio Access (UTRA)," 3rd. Generation Partnership Project, Tech. Rep. TR 25.814 V7.1.0, Sept. 2006. [Online]. Available: http://www.3gpp.org

[3GP07] ——, "Spatial channel model for Multiple Input Multiple Output (MIMO) simulations," 3rd. Generation Partnership Project, Tech. Rep. TR 25.996 V7.0.0, Jun. 2007.

[AEW94] M. Almgren, H. Eriksson, and K. Wallstedt, "Power control in a cellular system," in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, vol. 2, Jun. 1994, pp. 833–837.

[AHSB$^+$06] C. Antón-Haro, P. Svedman, M. Bengtsson, A. Alexiou, and A. Gameiro, "Cross-layer scheduling for multi-user MIMO systems," *IEEE Communications Magazine*, vol. 44, no. 9, pp. 39–45, Sept. 2006.

[B$^+$06] G. Bauch *et al.*, "Aspects of multiuser MIMO for cell throughput maximization," in *Proceedings of the Internatational OFDM Workshop*, Aug. 2006.

[Bal05] C. A. Balanis, *Antenna theory: analysis and design*, 3rd ed. John Wiley & Sons, Apr. 2005.

[BGWM07] M. Bohge, J. Gross, A. Wolisz, and M. Meyer, "Dynamic resource allocation in OFDM systems: an overview of cross-layer optimization principles and techniques," *IEEE Network*, pp. 53–59, Jan. 2007.

[BHV06] B. Bandemer, M. Haardt, and S. Visuri, "Linear MMSE multi-user MIMO downlink precoding for users with multiple antennas," in *Proceedings of the IEEE Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sept. 2006.

[BL71] F. Bourgeois and J.-C. Lassalle, "An extension of the Munkres algorithm for the assignment problem to rectangular matrices," *Communications of the ACM*, vol. 14, no. 12, pp. 802–804, Dec. 1971.

[BO02] M. Bengtsson and B. Ottersten, *Handbook of antennas in wireless communications*. CRC Press, 2002, ch. Optimum and suboptimum transmit beamforming.

[BS05] H. Boche and M. Schubert, *Smart antennas: state of the art*. Hindawi Publishing Coorporation, 2005, ch. Duality theory for uplink and downlink multiuser beamforming, pp. 545–576.

[BV04] S. Boyd and L. Vandenberghe, *Convex optimization*, 1st ed. Cambridge University Press, 2004.

[Cal04] D. B. Calvo, "Fairness analysis of wireless beamforming schedulers," Ph.D. dissertation, Technical University of Catalonia, Spain, Nov. 2004.

[CC05] P. W. C. Chan and R. S. K. Cheng, "Optimal power allocation in zero-forcing MIMO-OFDM downlink with multiuser diversity," in *Proceedings of the IST Mobile and Wireless Communications Summit*, Jun. 2005.

[CC07] P. W. C. Chan and R. S. Cheng, "Capacity maximization for zero-forcing MIMO-OFDMA downlink systems with multiuser diversity," *IEEE Transactions on Wireless Communications*, vol. 6, no. 5, pp. 1880–1889, May 2007.

[Cos83] M. Costa, "Writing on dirty paper," *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[CPIPN03] D. B. Calvo, A. Pascual-Iserte, and A. I. Perez-Neira, "Spatial scheduling algorithms for wireless systems," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Apr. 2003.

[CS03] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.

[CT06] T. M. Cover and J. A. Thomas, *Elements of information theory*, 2nd ed. John Wiley & Sons, 2006.

[CTRF02] J.-H. Chang, L. Tassiulas, and F. Rashid-Farrokhi, "Joint transmitter receiver diversity for efficient space division multiaccess," *IEEE Transactions on Wireless Communications*, vol. 1, no. 1, pp. 16–27, Jan. 2002.

[DH03] G. Del Galdo and M. Haardt, "Geometry-based channel modelling of MIMO channels in comparison with channel sounder measurements," *Advances in Radio Science - Kleinheubacher Berichte*, vol. 50, no. 5, pp. 117–126, Oct. 2003.

[DH04] ——, "Comparison of zero-forcing methods for downlink spatial multiplexing in realistic multi-user MIMO channels," in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, May 2004.

[DS04] G. Dimic and N. D. Sidiropoulos, "Low-complexity downlink beamforming for maximum sum capacity," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 4, May. 2004, pp. 701–704.

[DS05] ——, "On downlink beamforming with greedy user selection: performance analysis and a simple new algorithm," *IEEE Transactions on Signal Processing*, vol. 53, no. 10, pp. 3857–3868, Oct. 2005.

[DSK+06] M. Doettling, M. Sternad, G. Klang, J. von Hafen, and M. Olsson, "Integration of spatial processing in the WINNER B3G air interface design," in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, vol. 1, May 2006, pp. 246–250.

[FDH05] M. Fuchs, G. Del Galdo, and M. Haardt, "A novel tree-based scheduling algorithm for the downlink of multi-user MIMO systems with ZF beamforming," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 3, Mar. 2005, pp. 1121–1124.

[FDH07] ——, "Low-complexity space-time-frequency scheduling for MIMO systems with SDMA," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 5, pp. 2775–2784, Sept. 2007.

[FG98] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[FKC06] T. Frank, A. Klein, and E. Costa, "Low complexity multi carrier multiple access with cyclic delay diversity," in *Proceedings of the IEEE Workshop on Smart Antennas*, March 2006.

[FKKC06] T. Frank, A. Klein, A. Kuehne, and E. Costa, "Low complexity and power efficient space-time-frequency coding for OFDMA," in *Proceedings of the IST Mobile and Wireless Communications Summit*, Jun. 2006.

[FL02] H. Fattah and C. Leung, "An overview of scheduling algorithms in wireless multimedia networks," *IEEE Wireless Communications Magazine*, vol. 9, no. 5, pp. 76–83, Oct. 2002.

[FLFV00] F. R. Farrokhi, A. Lozano, G. J. Foschini, and R. A. Valenzuela, "Spectral efficiency of wireless systems with multiple transmit and receive antennas," in *Proceedings of the IEEE Personal, Indoor and Mobile Radio Communications (PIMRC)*, vol. 1, Sept. 2000, pp. 18–21.

[FN96] C. Farsakh and J. A. Nossek, "A real time downlink channel allocation scheme for an SDMA mobile radio system," in *Proceedings of the IEEE Personal, Indoor and Mobile Radio Communications (PIMRC)*, vol. 3, Oct. 1996, pp. 1216–1220.

[Ger05] A. B. Gershman, *Smart antennas: state of the art.* Hindawi Publishing Coorporation, 2005, ch. Uplink robust adaptive beamforming, pp. 173–188.

[GJ79] M. R. Garey and D. S. Johnson, *Computers and intractability - a guide to the theory of NP-completeness*, 1st ed. W. H. Freeman, Jan. 1979.

[GJJV03] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 684–702, Jun. 2003.

[GL96] G. H. Golub and C. F. V. Loan, *Matrix computations*, 3rd ed. Johns Hopkins University Press, 1996.

[Gun00] F. Gunarsson, "Power control in cellular systems: analysis, design and estimation," Ph.D. dissertation, Linköping University, Linköping, Sweden, Mar. 2000.

[Hay02] S. Haykin, *Adaptive filter theory*, 4th ed., T. Kailath, Ed. Prentice Hall, 2002.

[HH03] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Transactions on Information Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.

[HPS05] B. M. Hochwald, C. B. Peel, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication - part II: perturbation," *IEEE Transactions on Communications*, vol. 53, no. 3, pp. 537–544, Mar. 2005.

[HRM02] T. Halonen, J. Romero, and J. Melero, *GSM, GPRS and EDGE performance: evolution towards 3G/UMTS*, 1st ed. John Wiley & Sons, 2002.

[HT02] H. Holma and A. Toskala, *WCDMA for UMTS*, 2nd ed. John Wiley & Sons, Sept. 2002.

[Hun07] R. Hunger, "Floating point operations in matrix-vector calculus," Munich University of Technology, Tech. Rep. v1.3, Sept. 2007.

[IEE04] IEEE, "Air interface for fixed broadband wireless access systems," Institute of Electrical and Electronics Engineers, Tech. Rep. IEEE 802.16-2004, Oct. 2004.

[Jai91] R. Jain, *The art of computer systems performance analysis: techniques for experimental design, measurement, simulation, and modeling*, 1st ed. John Wiley & Sons, 1991.

[Joh04] M. Joham, "Optimization of linear and nonlinear transmit signal processing," Ph.D. dissertation, Munich University of Technology, Munich, Germany, Jun. 2004.

[JRV+05] N. Jindal, W. Rhee, S. Vishwanath, S. A. Jafar, and A. Goldsmith, "Sum power iterative water-filling for multi-antenna gaussian broadcast channels," *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1570–1580, Apr. 2005.

[JUN05] M. Joham, W. Utschick, and J. A. Nossek, "Linear transmit processing in MIMO communications systems," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 2700–2712, Aug. 2005.

[KK07a] A. Kühne and A. Klein, "Adaptive subcarrier allocation with imperfect channel knowledge versus diversity techniques in a multi-user OFDM-system," in *Proceedings of the IEEE Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sept. 2007.

[KK07b] ——, "An analytical consideration of imperfect CQI feedback on the performance of a multi-user OFDM-system," in *Proceedings of the Internatational OFDM Workshop*, Aug. 2007.

[KT02] I. Koutsopoulos and L. Tassiulas, "Adaptive resource allocation in SDMA-based wireless broadband networks with OFDM signaling," in *Proceedings of the IEEE Conference on Computer Communications (INFOCOM)*, vol. 3, Jun. 2002, pp. 1376–1385.

[KT08] ——, "The impact of space division multiplexing on resource allocation: a unified treatment of TDMA, OFDMA and CDMA," *IEEE Transactions on Communications*, vol. 56, no. 2, pp. 260–269, Feb. 2008.

[KTSK01] R. Kuehner, T. D. Todd, F. Shad, and V. Kezys, "Foward-link capacity in smart antenna base stations with dynamic slot allocation," *IEEE Transactions on Vehicular Technology*, vol. 50, no. 4, pp. 1024–1038, Jul. 2001.

[Lau02] V. K. N. Lau, "Proportional fair spatial scheduling for wireless access point with multiple antenna - reverse link with scalar feedback," in *Proceedings of the IEEE Global Telecommunications Conference*, vol. 1, Nov. 2002, pp. 763–767.

[Lau04] ——, "Optimal downlink space-time scheduling design with convex utility functions - multiple antenna systems with orthogonal transmit beamforming," in *Proceedings of the IEEE International Conference on Communications (ICC)*, vol. 5, Jun. 2004, pp. 3040–3045.

[Lau05] ——, "Optimal downlink space-time scheduling design with convex utility functions-multiple-antenna systems with orthogonal spatial multiplexing," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 4, pp. 1322–1333, Jul. 2005.

[Leu04] J. Y.-T. Leung, *Handbook of scheduling: algorithms, models and performance analysis.* Chapman & Hall/CRC, 2004, ch. A tutorial on complexity.

[Liu04] Y. Liu, "Modeling and simulation of service area based OFDM air interfaces for beyond 3G mobile radio systems," Ph.D. dissertation, Technische Universität Kaiserslautern, Dec. 2004.

[LL05] H. Liu and G. Li, *OFDM-based broadband wireless networks - design and optimization*, 1st ed. John Wiley & Sons, 2005.

[LR99] J. C. Liberti, Jr. and T. S. Rappaport, *Smart antennas for wireless communications: IS-95 and third generation CDMA applications*, 1st ed., T. S. Rappaport, Ed. Prentice Hall, 1999.

[LZ06] K. B. Letaief and Y. J. Zhang, "Dynamic multiuser resource allocation and adaptation for wireless systems," *IEEE Wireless Communications Magazine*, vol. 13, no. 4, pp. 38–47, Aug. 2006.

[MBQ04] M. Meurer, P. W. Baier, and W. Qiu, "Receiver orientation versus transmitter orientation in linear MIMO transmission systems," *EURASIP Journal on Applied Signal Processing*, vol. 9, pp. 1191–1198, 2004.

[Mey01] C. D. Meyer, *Matrix analysis and applied linear algebra*, 1st ed. Society for Industrial and Applied Mathematics (SIAM), Feb. 2001.

[MK06] T. F. Maciel and A. Klein, "A low-complexity SDMA grouping strategy for the downlink of Multi-User MIMO systems," in *Proceedings of the IEEE Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sept. 2006.

[MK07a] ——, "A convex quadratic SDMA grouping algorithm based on spatial correlation," in *Proceedings of the IEEE International Conference on Communications (ICC)*, Jun. 2007.

[MK07b] ——, "A low-complexity resource allocation strategy for SDMA/OFDMA systems," in *Proceedings of the IST Mobile and Wireless Communications Summit*, Jul. 2007.

[MK07c] ——, "On the performance of SDMA with soft dropping and SINR balancing power control in the downlink of multi-user MIMO systems," in *Proceedings of the IEEE Workshop on Smart Antennas*, Feb. 2007.

[MK08] ——, *Advances in mobile and wireless communications - views of the 16th IST Mobile and Wireless Communications Summit*, 1st ed., ser. Lecture Notes in Electrical Engineering. Springer, Jun. 2008, vol. 16, ch. Resource allocation strategies for SDMA/OFDMA systems, pp. 115–132.

[MSLT07] E. Matskani, N. D. Sidiropoulos, Z.-Q. Luo, and L. Tassiulas, "Joint multiuser downlink beamforming and admission control: a semidefinite relaxation approach," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 3, no. 585-588, Apr. 2007.

[NLTW98] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1423–1436, Oct. 1998.

[NP00] R. V. Nee and R. Prasad, *OFDM for wireless multimedia communications.* Artech House Publishers, 2000.

[NW99] G. Nemhauser and L. Wosley, *Integer and combinatorial optimization.* John Wiley & Sons, 1999.

[OOI97] T. Ohgane, Y. Ogawa, and K. Itoh, "A study on a channel allocation scheme with an adaptive array in SDMA," in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, vol. 2, May 1997, pp. 725–729.

[PED04] J. Papandriopoulos, J. Evans, and S. Dey, "Distributed power control for cellular MIMO systems with temporal and spatial filtering," in *Proceedings of the Australian Communication Theory Workshop*, Feb. 2004, pp. 164–175.

[PF05] D. P. Palomar and J. R. Fonollosa, "Practical algorithms for a family of waterfilling solutions," *IEEE Transactions on Signal Processing*, vol. 53, no. 2, pp. 687–695, Feb. 2005.

[PHS05] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication - part I: channel inversion and regularization," *IEEE Transactions on Communications*, vol. 53, no. 1, pp. 195–202, Jan. 2005.

[PNG03] A. Paulraj, R. Nabar, and D. Gore, *Introduction to space-time wireless communications*, 1st ed. Cambridge University Press, 2003.

[PP02] A. Papoulis and S. U. Pillai, *Probability, random variables and stochastic processes.* McGraw-Hill, 2002.

[Pro95] J. G. Proakis, *Digital communications*, 3rd ed., ser. Electrical Engineering Series. McGraw-Hill, 1995.

[Qiu04] W. Qiu, "Transmit power reduction in MIMO multi-user mobile radio downlinks by the rationale receiver orientation," Ph.D. dissertation, Kaiserslautern University of Technology, Kaiserslautern, Germany, Dec. 2004.

[Rap99] T. S. Rappaport, *Wireless communications: principles and practice*, 1st ed., ser. Prentice Hall Communications Engineering and Emerging Technologies. Prentice Hall, Jul. 1999.

[RFH08] F. Römer, M. Fuchs, and M. Haardt, "Distributed MIMO systems with spatial reuse for high-speed-indoor mobile radio access," in *Proceedings of the Wireless World Research Forum (WWRF)*, Apr. 2008.

[Sat78] H. Sato, "An outer bound to the capacity region of of broadcast channels," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 374–377, May 1978.

[SB04] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 1, pp. 18–28, Jan. 2004.

[SBO06] R. Stridh, M. Bengtsson, , and B. Ottersten, "System evaluation of optimal downlink beamforming with congestion control in wireless communication," *IEEE Transactions on Wireless Communications*, vol. 5, no. 4, pp. 743–751, Apr. 2006.

[SDS+05] J. Salo, G. Del Galdo, J. Salmi, P. Kyösti, L. Hentilä, M. Milojevic, D. Laselva, P. Zetterberg, and C. Schneider, "Matlab implementation of the WINNER phase I channel model ver1.5," WINNER Project, Tech. Rep., Dec. 2005. [Online]. Available: https://www.ist-winner.org/

[Set04] J. Sethuraman, *Handbook of scheduling: algorithms, models, and performance analysis.* Chapman & Hall/CRC, 2004, ch. Convex quadratic relaxations in scheduling.

[SH05] V. Stankovic and M. Haardt, "Successive optimization Tomlinson-Harashima precoding (SO THP) for multi-user MIMO systems," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 3, Mar. 2005, pp. 1117–1120.

[SH08] ——, "Generalized design of multi-user MIMO precoding matrices," *IEEE Transactions on Wireless Communications*, vol. 7, no. 3, pp. 953–961, Mar. 2008.

[Skl97] B. Sklar, "Rayleigh fading channels in mobile digital communication systems I - characterization," *IEEE Communications Magazine*, vol. 35, pp. 90–100, Jul. 1997.

[SPSH04] Q. H. Spencer, C. B. Peel, A. L. Swindlehurst, and M. Haardt, "An introduction to the multi-user MIMO downlink," *IEEE Communications Magazine*, vol. 42, no. 10, pp. 60–67, Oct. 2004.

[SS04a] Q. H. Spencer and A. L. Swindlehurst, "Channel allocation in multi-user MIMO wireless communications systems," in *Proceedings of the IEEE International Conference on Communications (ICC)*, vol. 5, Jun. 2004, pp. 3035–3039.

[SS04b] ——, "A hybrid approach to spatial multiplexing in multiuser MIMO downlinks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2, no. 2, pp. 236–247, 2004.

[SSH04] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 461–471, Feb. 2004.

[Sta06] V. Stankovic, "Multi-user MIMO wireless communications," Ph.D. dissertation, Technische Universität Ilmenau, Ilmenau, Germany, Nov. 2006.

[STKL01] F. Shad, T. D. Todd, V. Kezys, and J. Litva, "Dynamic Slot Allocation (DSA) in indoor SDMA/TDMA using a smart antenna basestation," *IEEE/ACM Transactions on Networking*, vol. 9, no. 1, pp. 69–81, Feb. 2001.

[STSK98] R. Sinha, T. D. Todd, F. Shad, and V. Kezys, "Forward link capacity in smart antenna base stations with dynamic slot allocation," in *Proceedings of the IEEE Personal, Indoor and Mobile Radio Communications (PIMRC)*, vol. 2, Sept. 1998, pp. 942–946.

[SWJO07] P. Svedman, S. K. Wilson, L. J. C. Jr., and B. Ottersten, "Opportunistic beamforming and scheduling for OFDMA systems," *IEEE Transactions on Communications*, vol. 55, no. 5, pp. 941–952, May 2007.

[TC04a] Y. M. Tsang and R. S. Cheng, "Optimal resource allocation in SDMA/Multi-Input-Single-Output/OFDM systems under QoS and power constraints," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, vol. 3, Mar. 2004, pp. 1595–1600.

[TC04b] ——, "Rate maximization in Space-Division Multiple Access (SDMA)/Multi-Input-Single-Output (MISO)/OFDM systems under QoS and power constraints," in *Proceedings of the IEEE International Conference on Communications (ICC)*, vol. 5, Jun. 2004, pp. 2970–2974.

[Tel95] I. E. Telatar, "Capacity of multi-antenna gaussian channels," Lucent Technologies / Bell Labs Innovations, Technical Memorandum, Oct. 1995. [Online]. Available: http://mars.bell-labs.com/cm/ms/what/mars/papers/proof/

[TJ05] A. Tölli and M. Juntti, "Scheduling for multiuser MIMO downlink with linear processing," in *Proceedings of the IEEE Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2005.

[TUBN05] P. Tejera, W. Utschick, G. Bauch, and J. A. Nossek, “Sum-rate maximizing decompositon approaches for multiuser MIMO-OFDM,” in *Proceedings of the IEEE Personal, Indoor and Mobile Radio Communications (PIMRC)*, vol. 1, Sept. 2005, pp. 231–235.

[TUBN06] ——, “Subchannel allocation in multiuser Multiple-Input-MultipleOutput systems,” *IEEE Transactions on Information Theory*, vol. 52, no. 10, pp. 4721–4733, Oct. 2006.

[Van02] H. L. Van Trees, *Optimum array processing.* John Wiley & Sons, 2002.

[VJG03] S. Vishwanath, N. Jindal, and A. Goldsmith, “Duality, achievable rates, and sum-rate capacity of Guassian MIMO broadcast channels,” *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.

[VM01] E. Visotsky and U. Madhow, “Space-time transmit precoding with imperfect feedback,” *IEEE Transactions on Information Theory*, vol. 47, no. 6, pp. 2632–2639, Sept. 2001.

[VT03] P. Viswanath and D. N. C. Tse, “Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality,” *IEEE Transactions on Information Theory*, vol. 49, no. 8, pp. 1912–1921, Aug. 2003.

[VVE+00] P. Vandenameele, L. Van Der Perre, M. G. E. Engels, B. Gyselinckx, and H. J. De Man, “A combined OFDM/SDMA approach,” *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 11, pp. 2312–2321, Nov. 2000.

[Wil06] T. Wild, “Successive user insertion for long-term adaptive beams with SDMA using short-term CQI,” in *Proceedings of the Internatational OFDM Workshop*, Aug. 2006.

[WIN05a] WINNER Project, “Assessment of advanced beamforming and MIMO techniques,” IST, Tech. Rep. D2.7, Feb. 2005. [Online]. Available: https://www.ist-winner.org/DeliverableDocuments/D2.7.pdf

[WIN05b] ——, “Final report on identified RI key technologies, system concept, and their assessment,” IST, Tech. Rep. D2.10, Dec. 2005. [Online]. Available: https://www.ist-winner.org/DeliverableDocuments/D2.10.pdf

[WIN05c] ——, “Final report on link level and system level channel models,” IST, Tech. Rep. D5.4, Nov. 2005. [Online]. Available: https://www.ist-winner.org/DeliverableDocuments/D5.4.pdf

[Yat95] R. D. Yates, “A framework for uplink power control in cellular radio systems,” *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 7, pp. 1341–1347, Sept. 1995.

[YG05] T. Yoo and A. Goldsmith, “Optimality of zero-forcing beamforming with multiuser diversity,” in *Proceedings of the IEEE International Conference on Communications (ICC)*, vol. 1, May 2005, pp. 542–546.

[YG06] ——, "Capaciy and power allocation for fading MIMO channels with channel estimation error," *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 2203–2214, May 2006.

[YGRS97] R. D. Yates, S. Gupta, C. Rose, and S. Sohn, "Soft dropping power control," in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, vol. 3, May 1997, pp. 1694–1698.

[YO02] K. Yu and B. Ottersten, "Models for MIMO propagation channels, A review," *Wireless Communications and Mobile Computing*, vol. 2, no. 7, pp. 653–666, Nov. 2002.

[Zet95] P. Zetterberg, "A comparison of two systems for down link communication with antennas at the base," in *Proceedings of the IEEE Wireless Communications System Symposium*, Nov. 1995, pp. 15–20.

[Zet99] ——, "A comparison of two systems for downlink communication with base station antenna arrays," *IEEE Transactions on Vehicular Technology*, vol. 48, no. 5, pp. 1356–1370, Sept. 1999.

[Zha95] H. Zhang, "Service disciplines for guaranteed performance service in packet-switching networks," *Proceedings of the IEEE*, vol. 83, no. 10, pp. 1374–1396, Oct. 1995.

[ZKAQ01] J. Zander, S.-L. Kim, M. Almgren, and O. Queseth, *Radio resource management for wireless networks*, ser. Mobile Communications. London, England: Artech House Publishers, 2001.

[ZL05] Y. J. Zhang and K. B. Letaief, "An efficient resource-allocation scheme for spatial multiuser access in MIMO/OFDM systems," *IEEE Transactions on Communications*, vol. 53, no. 1, pp. 107–116, Jan. 2005.

[ZTC06] Y. Zhang, A. Totaro, and E. Costa, "Adaptive resource allocation in chunk-based OFDMA/SDMA systems," in *Proceedings of the Internatational OFDM Workshop*, Aug. 2006.

# Lebenslauf

| | |
|---|---|
| Name: | Tarcisio Ferreira Maciel |
| Anschrift: | Lichtenbergstrasse 39, 64239 Darmstadt |
| Geburtsdatum: | 15. Mai 1977 |
| Geburtsort: | Fortaleza, Brasilien |
| Familienstand: | verheiratet |

**Schulausbildung**

| | |
|---|---|
| 1983-1987 | Grundschule in Fortaleza, Brasilien |
| 1988-1994 | Gymnasium in Fortaleza, Brasilien |

**Studium**

| | |
|---|---|
| 1996-2001 | Studium der Elektrotechnik an der<br>Federal University of Ceará, Fortaleza, Brasilien,<br>Studienabschlussß: Bachelor of Engineering |
| 2002-2003 | Studium der Elektrotechnik an der<br>Federal University of Ceará, Fortaleza, Brasilien,<br>Studienabschluß: Master of Science |

**Berufstätigkeit**

| | |
|---|---|
| 2001-2004 | Forschungstätigkeit als Mitarbeiter am<br>Wireless Telecommunications Research Group (GTEL),<br>Federal University of Ceará, Fortaleza, Brasilien |
| seit 2005 | wissenschaftlicher Mitarbeiter am<br>Fachgebiet Kommunikationstechnik,<br>Institut für Nachrichtentechnik,<br>Technische Universität Darmstadt |