TECHNISCHE UNIVERSITÄT MÜNCHEN
Fachgebiet Methoden der Signalverarbeitung

Efficient Near Optimum Algorithms for Multiuser Multiantenna Systems

Christian A. Guthy

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. sc. Samarjit Chakraborty
Prüfer der Dissertation:
1. Univ.-Prof. Dr.-Ing. Wolfgang Utschick
2. Prof. Michael L. Honig
   Northwestern University Illinois / USA

Die Dissertation wurde am 21.06.2011 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 20.07.2012 angenommen.

Contents

Acknowledgments

This thesis would not have been possible without the inspiring and creative environment at the Fachgebiet Methoden der Signalverarbeitung at TUM. I am therefore deeply thankful to Prof. Wolfgang Utschick for giving me the opportunity to be part of his group and supervising my thesis. This has been a great and unique experience for me. He posed the right questions at the right time and always encouraged me during my work when needed. Furthermore he made it possible for me to collaborate with Prof. Michael L. Honig. I am extremely grateful to Mike for introducing the methods and concepts of large system analysis to me, the fruitful ideas and discussions on the asymptotic analysis of SESAM and the sum capacity, and finally for being my second examiner. Regarding the algorithmic part of the thesis, it has been a pleasure for me to exchanging ideas and writing papers together with Pedro Tejera and Michael Joham. Pedro furthermore served as a wonderful office mate for two and a half years. I learned some Spanish words from him and he often challenged my German grammar knowledge. I would also like to thank the other colleagues, with which I shared some time in the same group, notably Johannes Brehmer, Peter Breun, Guido Dietl, Frank Dietrich, Andreas Dotzler, Lennart Gerdes, Andreas Gründinger, Christoph Hellings, Raphael Hunger, Alex Krebs, Max Riemensberger and David Schmidt. Explicitly mentioning each hint given by or problem solved together with them during my time would not fit on this page and they all contributed a lot to the great time I had at the institute. Furthermore I appreciated working together as a teaching assistant with Dr. Rainer Pauli. From the administrative staff of the institute, I have been glad to receive support from Lida Barth, Ulrike Heer, Sergey Fedorov, Hartmut Peters, Elizabeth Soeder and Ali Yilmazcan. Special thanks also go to the students Alex Krebs, Tobias Lutz, and Christian Neumair, who I have supervised.

Prof. Samarjit Chakraborty and his secreatary, Mrs. Dippold, had the workload associated with heading the committee and Lennart Gerdes accomplished the circulation of the thesis amongst professors. I am therefore also very grateful to them.

During the first three years my work has been supported by DOCOMO Eurolabs. I owe many thanks to Gerhard Bauch and Guido Dietl for a very good collaboration. Furthermore I would like to thank the Deutsche Forschungsgemeinschaft (DFG) for giving a grant for my work on large system analysis.

Finally, I would like to thank my parents for their sustained and strong support not only during my time as a PhD candidate but throughout my whole life and my girlfriend Anita for her backing and sympathy during writing this thesis.

Abstract

The use of multiple antennas at the transmitter and the receivers in a wireless communication system enables an efficient use of resources such as bandwidth and transmit power. On the other hand this advantage comes along with an increased complexity of the signal processing algorithms compared to single-antenna systems. In this thesis four popular optimization criteria for the design of transmit and receive filters in Multiple-Input Multiple-Output (MIMO) systems are covered. Those include the problem of maximizing the weighted sum of the transmission rates under a sum power constraint. Quality of Service constraints in terms of minimum and relative rate requirements for each user are additionally taken into account and furthermore the problem of minimizing the transmit power required to satisfy guaranteed rates for each user is treated. The optimum solutions to these problems are state-of-the-art, work iteratively and are numerically very complex. For practical purposes it is therefore desired to use algorithms that are less complex but have some acceptable performance losses. Those kind of methods are presented in this thesis along with a summary of the optimum algorithms. In order to show the near optimum performance of these algorithms not only by simulations, analytical results are derived for systems with infinite number of transmit antennas and infinite number of users or infinite number of receive antennas. In these large system limits, the rates achievable by the algorithms become deterministic, although the characteristics of the channel are random. The results obtained this way serve as a good approximation of the average performance in systems with finite parameters and are therefore valuable for the analysis of those systems.

1. Introduction

The use of multiple antennas in wireless communication systems enables strong performance improvements compared to single antenna systems. While a single-antenna transmitter emits its signals omni-directional, with multiple antennas the transmit signals can be designed so that the signals emitted from different antennas superpose constructively at desired spatial directions, whereas they overlap destructively in other spatial directions. Additionally, receivers with multiple antennas can combine certain signals constructively and better suppress interference from other transmitters this way. Thus, significantly higher transmission rates than in single-antenna systems can be achieved with the same amount of transmit power [1] or the same performance can be maintained at reduced transmit power, which diminishes interference to other users. Thereby, these Multiple-Input Multiple-Output (MIMO) systems ease spatial multiplexing. This implies that a multi-antenna transmitter transmits several different data symbols at the same time and on the same frequency, which enables an efficient use of the rare resource bandwidth. Those data symbols can be dedicated to different receivers or to the same receiver, where they may be combined for error correction or used for different data symbols. Vice versa, a multi-antenna receiver can better detect data symbols from different transmitters. However, these performance improvements come at the price of increased complexity. Besides the increased hardware requirements of multiple antennas, the numerical complexity of the signal processing algorithms in the physical layer rises drastically. While in single-antenna systems, the transmit power is the only scalar parameter to optimize at the physical layer, in multi-antenna systems the transmit and receive filters, that describe how the signals for or from the antennas are combined, become vector-valued. When only one transmitter and one receiver are involved in the communication system, i.e., in single-user MIMO systems, this fact is still manageable. There, the sum capacity, i.e., the maximum theoretically achievable sum of the error-free transmission rates, can be computed by a Singular Value Decomposition (SVD) [2]. However, in order to fully exploit the increased performance of spatial multiplexing, it is inevitable to work with multi-user MIMO systems for most scenarios. In [3] for example, the gains of multi-user MIMO compared to simpler transmission schemes based single-user MIMO are evaluated. However, the computation of the sum capacity achieving filters in multi-user MIMO systems is only possible with iterative, numerically involved algorithms. First, the optimum transmit and receive strategy achieving sum capacity has been solved for the Multiple-Access Channel (MAC) in [4], where several transmitters send different symbols to one common receiver. The determination of optimum signal processing algorithms for the MIMO broadcast channel, where one transmitter emits different signals for several users, has been enabled by the duality between the broadcast channel and the dual MAC established in [5]. In contrast to the physical MAC described above, where individual power constraints have to be met at each transmitter, in the dual MAC a sum power constraint, which is the same as in the broadcast channel, is considered over all transmitters. This duality states that the same rates are achievable in the dual MAC as in the broadcast channel. That is why the problem of determining the optimum signal processing filters can be solved in the dual MAC, where the problem of sum rate maximization becomes convex as exploited in [6]. Nevertheless, that problem still has to be solved iterative and in a numerically complex manner. Additionally, the duality implies a numerically complex transformation of the filters optimum in the dual MAC to the broadcast channel after convergence of the algorithm in the dual MAC. In [7] the more general problem of weighted sum rate maximization, where different priorities can be assigned to the data streams, has first been solved for the broadcast channel. Further computational burdens are introduced with additional constraints on the rates, like minimum rate requirements for each user as in [8], or relative rate constraints as considered in [9], as those problems are tackled by iteratively solving weighted sum rate maximization problems. All those algorithms are derived for frequency-flat channels, i.e., there is only one temporal propagation path between each transmitter and receiver. In case of multi-path propagation, additional measures have to be taken to mitigate the inter-symbol interference. Orthogonal Frequency Multiplexing (OFDM) can for example be applied to decompose each multi-path channel into a system of frequency-flat channels. Although the algorithms mentioned above are still applicable in OFDM systems, the dimensions of the matrices describing the effect of channel propagation increase making these algorithms even more complex.

In order to overcome this drawback of high computational complexity, but still profit from the benefits of multi-user MIMO, efficient and low complex algorithms need to be developed, that achieve the optimum solutions as close as possible. A procedure of this kind, named Successive Encoding Successive Allocation Method (SESAM), has been proposed for the problem of sum rate maximization in [10], where it is shown by simulation results that it is able to achieve the sum capacity at negligible performance losses. An extension to the problem of sum rate maximization with relative rate requirements is presented in [11]. Both methods are based on the concepts of spatial zero-forcing and successive resource allocation, i.e., interference between spatially multiplexed data streams is suppressed completely, and data streams are successively allocated to the users. As the respective optimum algorithms, this method is still based on the principle of Dirty Paper Coding (DPC). While its theoretical concept stating that parts of the interference can be cancelled perfectly without affecting the transmission rates, is relatively simple, its near-optimum implementation involves huge computational complexity. Giving up DPC introduces however additional performance losses but further reduces the numerical load. In [12] the concepts of spatial zero-forcing and successive resource allocation are therefore incorporated into an efficient algorithm for sum rate maximization without DPC leading to acceptable performance losses. However, it has been designed for Multiple-Input Single-Output (MISO) systems, where the receivers have only one antenna.

In this book these concepts of successive resource allocation and spatial zero-forcing are used to derive efficient near-optimum algorithms for a wide range of optimization problems in the MIMO broadcast channel. Chapter 3 focuses on the weighted sum rate maximization. Starting with the original optimization problem, for which an optimum algorithm is reviewed in the first section of Chapter 3, several simplifying steps are explained and introduced to this problem until it can be solved efficiently and near-optimally. For this purpose, first the case is covered that it is still affordable to implement DPC, before further simplifications are made to consider the absence of DPC. Quality of Service constrained optimization problems in the MIMO broadcast channel are handled in Chapter 4. Those include the weighted sum rate maximization under a transmit sum power constraint and minimum rate requirements for each user and the sum power minimization to achieve target user rates. While the optimum algorithms, which are explained in the first section of Chapter 4, rely on a iterative weighted sum rate maximization, the proposed efficient methods work differently and are therefore only a little bit more complex than the efficient algorithms for weighted sum rate maximization. As in Chapter 3, two variants with and without DPC are presented. Both Chapters 3 and 4 conclude with presenting procedures for further reducing the computational complexity of the proposed algorithms and numerical results. To underline the close-to-optimum performance not only by simulation results, some of the efficient algorithms will be analyzed in the large system limit in Chapter 5. Thereby, in each case two system parameters grow towards infinity, where their ratio remains fixed and finite. In contrast to analysis methods, where only one parameter, as the transmit power or the number of users, grows towards infinity, the obtained results also serve as a good approximation of the average performance in systems with finite parameters having the same ratio. First, the number of transmit antennas and the number of users become infinite, where for simplicity MISO systems are analyzed. Afterwards MIMO systems having infinite number of transmit and receive antennas are analyzed. The notation and system model used throughout this book are explained in Chapter 2 and some concluding remarks are given in Chapter 6.

2. Notation and System Model

Bold lower and uppercase letters denote vectors and matrices, respectively, where throughout this book column vectors are used. $(\bullet)^{\mathsf{T}}$ and $(\bullet)^{\mathrm{H}}$ describe the transpose and the Hermitian of a vector or a matrix, respectively. $\rho_i(\boldsymbol{A})$, $\mathrm{tr}(\boldsymbol{A})$, $|\boldsymbol{A}|$, $\|\boldsymbol{A}\|_{\mathrm{F}}$, and $[\boldsymbol{A}]_{i,j}$ are the ith eigenvalue, the trace, the determinant, the Frobenius norm, and the element in row i and column j of the matrix $\boldsymbol{A}$, respectively. $\boldsymbol{A}^{+}$ denotes the Moore-Penrose pseudo-inverse of the matrix $\boldsymbol{A}$, and $\mathrm{vec}(\boldsymbol{A})$ stacks the columns of the matrix $\boldsymbol{A}$ in one vector.

$$\mathrm{diag}(a_1,\ldots,a_n) = \begin{bmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{bmatrix}$$

denotes a diagonal matrix with the elements $a_1,\ldots,a_n$ on its diagonal and

$$\mathrm{blockdiag}\left(\boldsymbol{A}_1,\ldots\boldsymbol{A}_n\right) = \begin{bmatrix} \boldsymbol{A}_1 & & \\ & \ddots & \\ & & \boldsymbol{A}_n \end{bmatrix}$$

is the notation for a block-diagonal matrix. $\boldsymbol{I}_i$ is the $i \times i$ identity matrix, $\mathbf{0}_{i,j}$ is the $i \times j$ zero matrix, $\mathbf{1}_i$ is the i-dimensional all-ones vector and $\boldsymbol{e}_j$ denotes the j-th canonical unit vector.

$$\mathrm{span}\{\boldsymbol{A}\} = \{\boldsymbol{y} \in \mathbb{C}^m | \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}, x \in \mathbb{C}^n\}, \quad \mathrm{null}\{\boldsymbol{A}\} = \{\boldsymbol{x} \in \mathbb{C}^n | \boldsymbol{A}\boldsymbol{x} = \mathbf{0}_{m,1}\}$$

and $(\mathrm{null}\{\boldsymbol{A}\})^{\perp}$ denote the range, the nullspace of the matrix $\boldsymbol{A} \in \mathbb{C}m \times n$ and the orthogonal complement to this nullspace, respectively.
For n-dimensional vectors $\boldsymbol{a} \in \mathbb{R}^n$ and $\boldsymbol{b} \in \mathbb{R}^n$ inequalities hold element-wise, i.e., $\boldsymbol{a} \le \boldsymbol{b}$ is the short notation for the system of inequalities

$$\boldsymbol{e}_j^{\mathsf{T}}\boldsymbol{a} \le \boldsymbol{e}_j^{\mathsf{T}}\boldsymbol{b}, \quad \forall j = 1,\ldots,n$$

and for matrices the expression "$\succeq 0$" implies positive semi-definiteness, i.e., $\boldsymbol{A} \succeq 0$ implies that all eigenvalues of $\boldsymbol{A}$ are larger or equal to zero. $[\boldsymbol{a}]^{+}$, where $\boldsymbol{a} \in \mathbb{R}^n$, is the short notation for $\max(\mathbf{0}_{n,1}, \boldsymbol{a})$, where the maximum operator is applied element-wise, i.e., $[\boldsymbol{a}]^{+}$ sets all negative elements in $\boldsymbol{a}$ to zero. For a complex vector $\boldsymbol{a} \in \mathbb{C}^n$ $\mathrm{Re}\{\boldsymbol{a}\}$ and $\mathrm{Im}\{\boldsymbol{a}\}$ returns the real and the imaginary part of this vector, respectively.
Following [2], a random variable $\boldsymbol{x}$ is Gaussian and circularly symmetric with mean $\bar{\boldsymbol{x}}$ and covariance matrix $\boldsymbol{Q}$, which will be denoted as

$$\boldsymbol{x} \sim \mathcal{CN}(\bar{\boldsymbol{x}}, \boldsymbol{Q})$$

in the following, if the vector $\begin{bmatrix} \mathrm{Re}\{\boldsymbol{x}\} \\ \mathrm{Im}\{\boldsymbol{x}\} \end{bmatrix}$ is a multivariate Gaussian variable with mean $\begin{bmatrix} \mathrm{Re}\{\bar{\boldsymbol{x}}\} \\ \mathrm{Im}\{\bar{\boldsymbol{x}}\} \end{bmatrix}$ and covariance matrix

$$\frac{1}{2}\begin{bmatrix} \mathrm{Re}\{\boldsymbol{Q}\} & -\mathrm{Im}\{\boldsymbol{Q}\} \\ \mathrm{Im}\{\boldsymbol{Q}\} & \mathrm{Re}\{\boldsymbol{Q}\} \end{bmatrix}.$$

In this book a multi-user MIMO system with one base station or access point and K users is considered. The main focus is put on the downlink transmission, i.e., the base station acts as transmitter and the user terminals as receivers. The number of antennas at the transmitter is denoted by N_T and r_k is the number of antennas at user k. The messages intended for each user are different. Thus, from an information theoretic point of view, a broadcast channel is analyzed [13]. Orthogonal Frequency Division Multiplexing (OFDM) is used to mitigate the interference caused by multi-path propagation, commonly known as Intersymbol Interference (ISI), where the number of carrier frequencies is denoted by C. An overview of the MIMO OFDM broadcast channel is depicted in Figure 2.1. The source emits discrete symbols $\boldsymbol{s}_k[n] \in \mathbb{C}^{d_k}$ for each user k every T_s

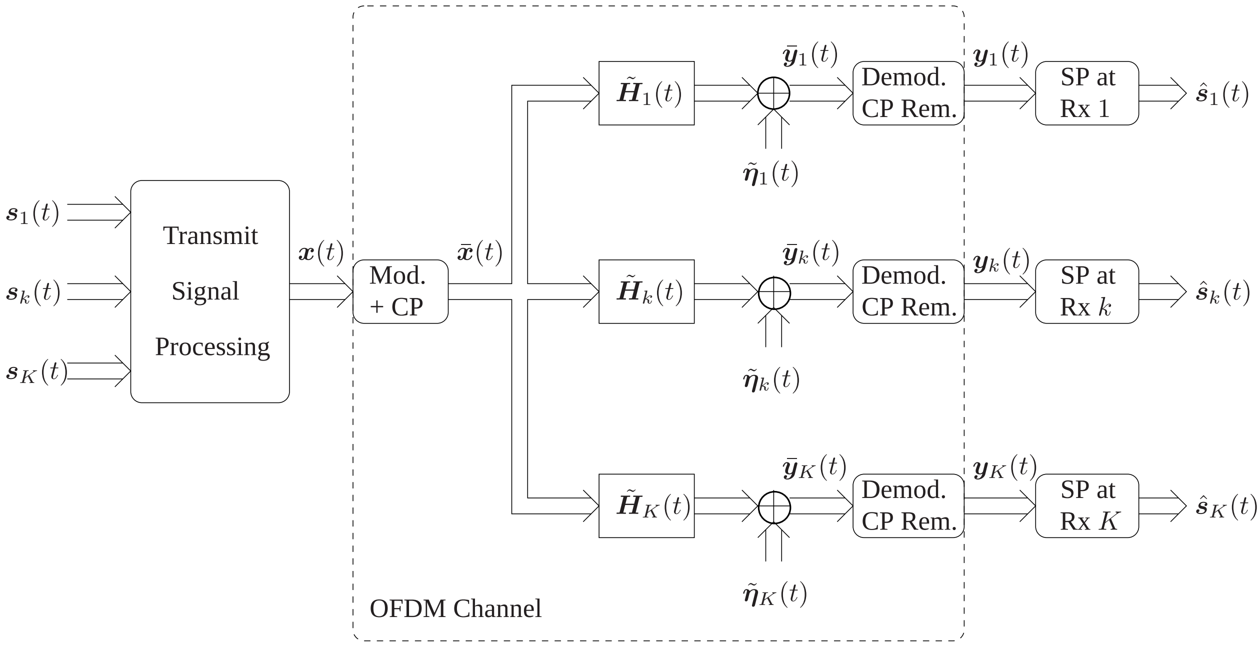

Figure 2.1: MIMO OFDM Broadcast Channel

seconds, so that

$$\boldsymbol{s}_k(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT_s)\boldsymbol{s}_k[n],$$

where n, d_k, and $\delta(x)$ denote the index of the channel access, the number of symbols transmitted to user k during one channel access, and the Dirac function respectively. The symbols $\boldsymbol{s}_k[n]$ are already preprocessed by source and channel coding, where it is assumed that Gaussian codebooks have been used for this purpose. This idealized choice is motivated by the fact that those kind of codebooks achieve the capacity of the MIMO broadcast as shown in [14]. Thus, the entries of the $\boldsymbol{s}_k[n]$ are circularly symmetric Gaussian with zero mean and unit variance, so that $\boldsymbol{s}_k[n] \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{I}_{d_k})$. Furthermore the transmission symbols for different users are uncorrelated. In the following the different parts will be explained in detail, before a compact system model can be introduced at the end, which will be used in the remaining chapters.

Transmit Signal Processing

The signals $\boldsymbol{s}_k(t)$ are processed at the transmitter as depicted in Figure 2.2. At first Dirty Paper Coding (DPC) can be applied to the transmit signals. During each channel access DPC further

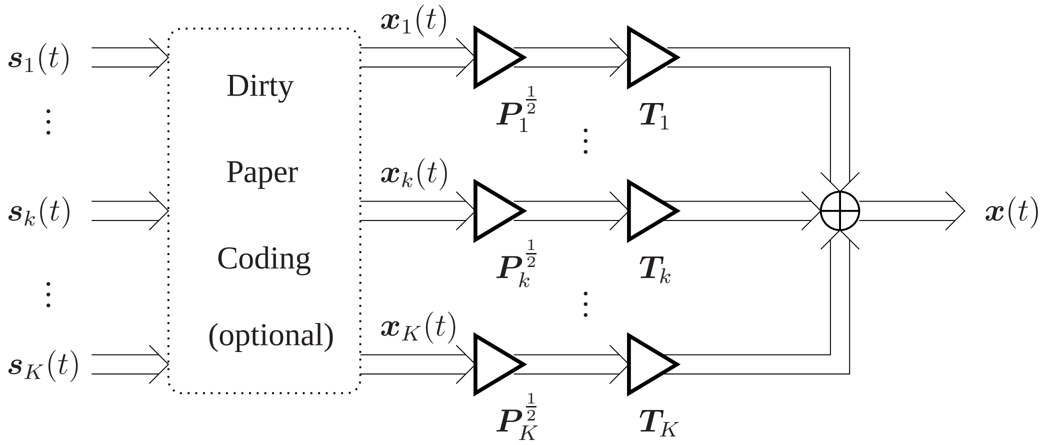

Figure 2.2: Transmit Signal Processing

encodes the signals $\boldsymbol{s}_k[n]$ leading to the symbols $\boldsymbol{x}_k[n]$, which are emitted every T_s seconds so that

$$\boldsymbol{x}_k(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT_s)\boldsymbol{x}_k[n].$$

The DPC principle has been proofed by Costa [15]. It does not change the statistics of the signals, i.e., $\boldsymbol{x}_k[n] \sim \mathcal{CN}(\mathbf{0}_{d_k,0}, \boldsymbol{I}_{d_k})$, and it states that interference that is known when a certain symbol $[\boldsymbol{x}_k[n]]_i$ is encoded does not affect that user's rate. That rate is therefore the same as if this interference was not present. The interference for the symbol encoded first on each carrier is unknown and can therefore not be taken into account during encoding. When another symbol is encoded at second place on the same carrier, the interference from the symbol encoded first is known and can be cancelled by DPC, while the remaining interference has to be mitigated by other measures. Correspondingly, the interference for the symbol encoded last can be eliminated completely by DPC. Interference between symbols on different carriers is eliminated by OFDM, as shown in the next section. Costa however only proofed the theoretical concept of DPC. Practical near optimum implementations for DPC are vector precoding [16] or the coding scheme from [17]. In the remainder of this book the common idealized assumption will be made that DPC can be implemented perfectly. As this implies a huge numerical effort, algorithms without DPC will be introduced in Sections 3.3 and 4.4, where $\boldsymbol{x}_k[n] = \boldsymbol{s}_k[n], \forall k = 1, \ldots, K$ holds.
The linear precoding for each vector $\boldsymbol{s}_k(t)$ is split into the power allocation matrices $\boldsymbol{P}_k \in \mathbb{C}^{d_k \times d_k}$ and the beamforming matrices $\boldsymbol{T}_k \in \mathbb{C}^{CN_\mathrm{T} \times d_k}$. The jth column of the matrix $\boldsymbol{T}_k$ corresponds to the beamforming vector for the jth data stream of user k normalized to one, i.e., $\boldsymbol{e}_j^\mathsf{T} \boldsymbol{T}_k^\mathrm{H} \boldsymbol{T}_k \boldsymbol{e}_j = 1$. The $\boldsymbol{P}_k$ are diagonal and contain the power allocations, i.e., $\boldsymbol{e}_j^\mathsf{T} \boldsymbol{P}_k \boldsymbol{e}_j$ denotes the power allocated to the jth data stream of user k. Thus the $\boldsymbol{P}_k$ must be positive semidefinite. Additionally an average transmit power constraint P_Tx is considered, which implies that

$$\sum_{k=1}^{K} \mathrm{tr}\left(\boldsymbol{P}_k\right) \leq P_\mathrm{Tx}$$

must be fulfilled. Throughout this book the figures of merit are analyzed for a specific transmit and receive strategy with fixed channel properties. For notational convenience the transmit filters do therefore not depend on the channel access index n. The signal $\boldsymbol{x}(t) \in \mathbb{C}^{CN_\mathrm{T}}$ after transmit signal processing is then given by

$$\boldsymbol{x}(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT_s) \sum_{k=1}^{K} \boldsymbol{T}_k \boldsymbol{P}_k^{\frac{1}{2}} \boldsymbol{x}_k[n] = \sum_{n=-\infty}^{\infty} \delta(t - nT_s)\boldsymbol{x}[n].$$

OFDM Channel

The kth user's channel is modeled by a $r_k \times N_\mathrm{T}$ channel matrix $\tilde{\boldsymbol{H}}_k(t)$ and a tapped delay line model is used to describe the effect of multi-path propagation. Thus, the kth user's channel matrix is given by

$$\tilde{\boldsymbol{H}}_k(t) = \sum_{\ell=1}^{L_k} \tilde{\boldsymbol{H}}_{k,\ell}\delta(t - \tau_{k,\ell}),$$

where $\tilde{\boldsymbol{H}}_{k,\ell} \in \mathbb{C}^{r_k \times N_\mathrm{T}}$ contains the complex attenuation coefficients of the channel of the propagation path with delay $\tau_{k,\ell}$ and L_k denotes the number of propagation paths of user k. The properties of the channels are assumed to be constant, which is why the $\tilde{\boldsymbol{H}}_{k,\ell}$ and the delays $\tau_{k,\ell}$ are independent of t. The additive noise $\tilde{\boldsymbol{\eta}}_k(t) \in \mathbb{C}^{r_k}$ experienced by user k is Gaussian circularly symmetric with zero mean and power spectral density $\tilde{\boldsymbol{R}}_k$, i.e., noise vectors $\tilde{\boldsymbol{\eta}}_k(t_1)$ and $\tilde{\boldsymbol{\eta}}_k(t_2)$ observed at different time instances $t_1 \neq t_2$ are uncorrelated. Additionally $\tilde{\boldsymbol{\eta}}_k(t)$ is uncorrelated to the noise vectors of other users. The effect of modulation, demodulation and the cyclic prefix can be included in the channel model to obtain an effective frequency representation of the OFDM channel as shown in Appendix A1. Thus, the kth user's channel is modeled by the $Cr_k \times CN_\mathrm{T}$ matrix

$$\begin{aligned}\boldsymbol{H}_k &= \sum_{c=1}^{C} \boldsymbol{S}_{c,k}^{\mathsf{T}} \sum_{\ell=1}^{L_k} \tilde{\boldsymbol{H}}_{k,\ell} \exp(j2\pi f_c(\tau_{k,1} - \tau_{k,\ell}))\boldsymbol{S}_{c,\mathrm{T}} = \\ &= \operatorname{blockdiag}\left(\sum_{\ell=1}^{L_k} \tilde{\boldsymbol{H}}_{k,\ell} \exp(j2\pi f_c(\tau_{k,1} - \tau_{k,\ell}))\right), \qquad (2.1)\end{aligned}$$

where

$$\boldsymbol{S}_{c,k} = \left[\boldsymbol{0}_{r_k,(c-1)r_k}, \boldsymbol{I}_{r_k}, \boldsymbol{0}_{r_k,(C-c)r_k}\right], \qquad \boldsymbol{S}_{c,\mathrm{T}} = \left[\boldsymbol{0}_{N_\mathrm{T},(c-1)N_\mathrm{T}}, \boldsymbol{I}_{N_\mathrm{T}}, \boldsymbol{0}_{N_\mathrm{T},(C-c)N_\mathrm{T}}\right], \qquad (2.2)$$

and the received signal $\boldsymbol{y}_k(t)$ is a sum of Dirac impulses, which is given by

$$\boldsymbol{y}_k(t) = \sum_{n=-\infty}^{\infty} \delta(t - ((n+1)T_s + T_\mathrm{cp} + \tau_{k,1}))\boldsymbol{y}_k[n]$$

with T_cp being the length of the cyclic prefix in OFDM. The symbols $\boldsymbol{y}_k[n]$ compute according to

$$\boldsymbol{y}_k[n] = \boldsymbol{H}_k\boldsymbol{x}[n] + \boldsymbol{\eta}_k[n] = \boldsymbol{H}_k \sum_{m=1}^{K} \boldsymbol{T}_m \boldsymbol{P}_m^{\frac{1}{2}} \boldsymbol{x}_m[n] + \boldsymbol{\eta}_k[n], \qquad (2.3)$$

i.e., the symbols $\boldsymbol{y}_k[n]$ experience no interference from the pervious transmit symbols $\boldsymbol{x}_m[n-1]$, $m = 1, \ldots, K$. The effective noise $\boldsymbol{\eta}_k[n] \in \mathbb{C}^{Cr_k}$ contains circularly symmetric Gaussian variables with zero mean and covariance matrix

$$\boldsymbol{R}_k = \mathrm{E}\left[\boldsymbol{\eta}_k[n]\boldsymbol{\eta}_k[n]^\mathrm{H}\right] = \frac{1}{T}\sum_{c=1}^{C} \boldsymbol{S}_{c,k}^{\mathsf{T}} \tilde{\boldsymbol{R}}_k \boldsymbol{S}_{c,k} = \frac{1}{T}\operatorname{blockdiag}\left(\tilde{\boldsymbol{R}}_k, \ldots, \tilde{\boldsymbol{R}}_k\right), \qquad (2.4)$$

where T is defined via the system's bandwidth

$$B = \frac{C}{T}.$$

Throughout this book it is assumed that the channel matrices $\boldsymbol{H}_k$ or alternatively the matrices $\tilde{\boldsymbol{H}}_{k,\ell}$ and the delays $\tau_{k,\ell}$, $\ell = 1, \ldots, L_k$ are perfectly known at user k and that the transmitter has complete and perfect knowledge of the channel matrices $\boldsymbol{H}_k$ of all users. Clearly, this assumption is idealized and most realistic in slowly time-varying scenarios, like for instance indoor office environments. For an overview of finite-rate feedback schemes, where the channel matrices $\boldsymbol{H}_k$ are perfectly known only at user k and fed back at a fixed and finite rate, the reader is referred to [18]. References [19] and [20] focus on finite feedback schemes in OFDM systems, where correlations between channel matrices of adjacent subcarriers are exploited for feedback reductions. The performance of the algorithm presented in Section 3.2 with several finite rate feedback schemes has been analyzed in [21], a new low complexity feedback method for this algorithm has been proposed in [22].

Receive Signal Processing

The signal $\boldsymbol{y}_k(t)$ is finally processed by a linear filter $\boldsymbol{G}_k \in \mathbb{C}^{Cr_k \times d_k}$ as depicted in Figure 2.3, so that the estimation $\hat{\boldsymbol{s}}_k(t)$ of the transmit signal $\boldsymbol{s}_k(t)$ is given by

$$\hat{\boldsymbol{s}}_k(t) = \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{y}_k(t).$$

The receive filters $\boldsymbol{G}_k$ are usually determined at the transmitter, as each user is only aware of its

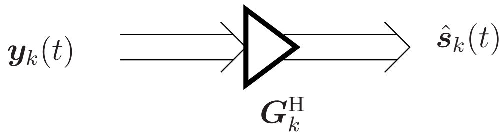

Figure 2.3: Signal Processing at Receiver k

own channel matrix, but a good filter design must be based on the knowledge of all users' channels. The filters are therefore communicated in a signaling phase before transmission. Similarly to [23] or [24], first common pilot symbols are sent to the users, where the pilot symbols are precoded so that the estimate is equal to the filters to be applied at the receivers. To detect which of the estimated filters should be used by a certain user, user identifiers are sent over the precoded channels in a second phase. Alternatively, [25] proposes signaling schemes with quantized feedforward of the terminals' filters.

Summarizing the results of the previous sections a compact system model can be established as depicted in Figure 2.4. where a discrete signal model for input, output and noise symbols is used. The system parameters are summarized in Table 2.1. In the remainder of this book noise covariance matrices

$$\tilde{\boldsymbol{R}}_k = N_0 \boldsymbol{I}_{r_k}$$

will be used, which implies that

$$\boldsymbol{R}_k = \frac{N_0}{T} \boldsymbol{I}_{Cr_k} = \frac{\sigma_{\mathrm{n}}^2}{C} \boldsymbol{I}_{Cr_k},$$

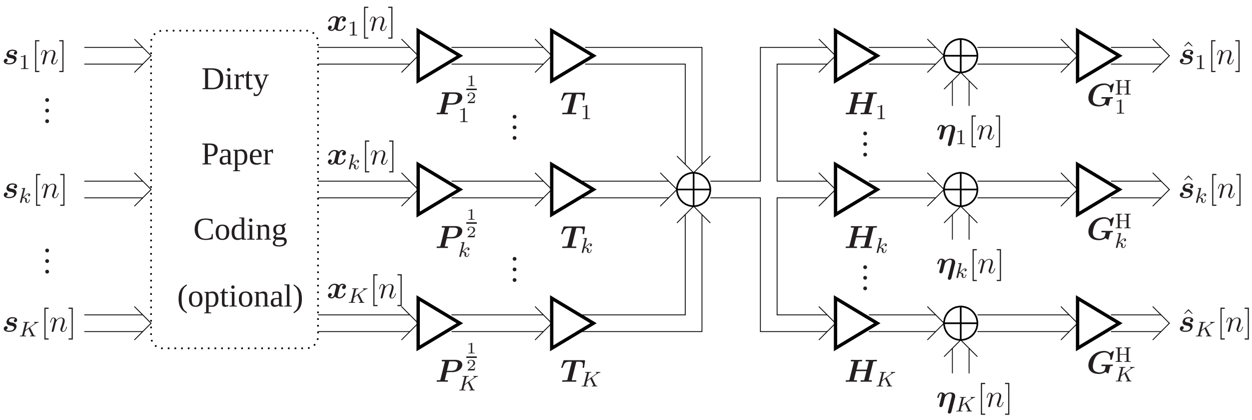

Figure 2.4: Compact OFDM System Model

Parameter	Meaning
N_{T}	Number of transmit antennas
r_k	Number of receive antennas at user k
K	Number of users
C	Number of carriers
P_{Tx}	Transmit power constraint
d_k	Number of data streams allocated to user k
$\boldsymbol{P}_k \in \mathbb{C}^{d_k \times d_k}$	Diagonal power allocation matrix of user k
$\boldsymbol{T}_k \in \mathbb{C}^{CN_{\mathrm{T}} \times d_k}$	Precoding matrix of user k with normalized columns
$\boldsymbol{H}_k \in \mathbb{C}^{Cr_k \times CN_{\mathrm{T}}}$	Block diagonal OFDM channel matrix of user k
$\boldsymbol{G}_k \in \mathbb{C}^{Cr_k \times d_k}$	Receive filter at user k
$\boldsymbol{x}_k[n] \sim \mathcal{CN}(\mathbf{0}_{d_k,1}, \boldsymbol{I}_{d_k})$	Transmit symbol for user k during nth transmission slot
$\boldsymbol{\eta}_k[n] \sim \mathcal{CN}(\mathbf{0}_{Cr_k}, \frac{N_0 B}{C}\boldsymbol{I}_{Cr_k})$	Additive noise for user k during nth transmission slot
$\hat{\boldsymbol{s}}_k[n]$	Estimated symbol at user k during nth transmission slot

Table 2.1: Overview of System Parameters

with

$$\sigma_{\mathrm{n}}^2 = N_0 B.$$

Using this simplification and the fact that the transmit symbols $\boldsymbol{x}_k[n]$ and the effective noise vectors $\boldsymbol{\eta}_k[n]$ are circularly symmetric Gaussian random variables, the kth user can at most transmit error-free at a rate

$$
\begin{aligned}
R_k &= \log_2 \frac{\left| \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{G}_k \frac{\sigma_{\mathrm{n}}^2}{C} + \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{T}_k \boldsymbol{P}_k \boldsymbol{T}_k^{\mathrm{H}} \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{G}_k + \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{H}_k \left[\sum_{\substack{m=1\\ m\in\mathcal{I}_k}}^{K} \boldsymbol{T}_m \boldsymbol{P}_m \boldsymbol{T}_m^{\mathrm{H}} \right] \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{G}_k \right|}{\left| \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{G}_k \frac{\sigma_{\mathrm{n}}^2}{C} + \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{H}_k \left[\sum_{\substack{m=1\\ m\in\mathcal{I}_k}}^{K} \boldsymbol{T}_m \boldsymbol{P}_m \boldsymbol{T}_m^{\mathrm{H}} \right] \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{G}_k \right|} \\
&= \log_2 \left| \boldsymbol{I}_{d_k} + \left(\boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{G}_k \frac{\sigma_{\mathrm{n}}^2}{C} + \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{H}_k \left[\sum_{\substack{m=1\\ m\in\mathcal{I}_k}}^{K} \boldsymbol{T}_m \boldsymbol{P}_m \boldsymbol{T}_m^{\mathrm{H}} \right] \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{G}_k \right)^{-1} \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{T}_k \boldsymbol{P}_k \boldsymbol{T}_k^{\mathrm{H}} \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{G}_k \right|
\end{aligned}
\tag{2.5}
$$

where it must be ensured that each matrix $\boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{G}_k$ is invertible, i.e., has rank d_k. In case $\boldsymbol{G}_k$ contains linearly dependent columns one has to combine those columns to one data stream, so that the rate expressions in (2.5) can be stated. $\mathcal{I}_k$ denotes the set of users that interfere with user k. In case DPC is not used, $\mathcal{I}_k$ contains all active users except user k. Otherwise, all users that are encoded after user k can be found in $\mathcal{I}_k$, where for that case it is assumed in (2.5) that the encoding order is the same on all carriers and data streams of the same user are encoded simultaneously. Note that the case of colored noise, i.e., the $\boldsymbol{R}_k$ are not diagonal, can be easily considered by using channel matrices $\boldsymbol{H}_k = \boldsymbol{R}_k^{-\frac{1}{2}}\hat{\boldsymbol{H}}_k$ and receive filters $\boldsymbol{G}_k = \boldsymbol{R}_k^{\frac{1}{2}}\hat{\boldsymbol{G}}_k$in (2.5), where $\hat{\boldsymbol{H}}_k$ and $\hat{\boldsymbol{G}}_k$denote the actual channel matrices and receive filters in the system with colored noise, respectively.

3. Weighted Sum Rate Maximization in the MIMO Broadcast Channel

The problem of weighted sum rate maximization is treated in this chapter. First, an optimum algorithm will be presented in Section 3.1. The concepts of spatial zero-forcing and successive resource allocation will be introduced in Section 3.2 and used to derive an efficient near-optimum algorithm, when DPC can be applied at the transmitter. Due to the complexity associated with practical implementations of DPC, an algorithm without DPC will be presented in Section 3.3. Although these algorithms are already able to reduce the complexity of the optimum algorithm drastically at little performance losses, further complexity reductions for these methods will be presented in Section 3.4 before the chapter is concluded with numerical results in Section 3.5.

3.1 Optimum Algorithm

Mathematically, the maximization of a weighted sum of the users' rates under a sum transmit power constraint reads as

$$\max_{\{\boldsymbol{T}_k,\boldsymbol{P}_k,\boldsymbol{G}_k,\hat{\pi}(k)\}_{k=1,\ldots,K}} \sum_{k=1}^{K} \mu_k R_k,$$

$$\text{s.t.} \sum_{k=1}^{K} \mathrm{tr}(\boldsymbol{P}_k) \leq P_{\mathrm{Tx}}, \quad \boldsymbol{e}_j^{\mathrm{T}} \boldsymbol{T}_k^{\mathrm{H}} \boldsymbol{T}_k \boldsymbol{e}_j = 1, \forall j = 1,\ldots,Cr_k, \forall k, \quad \boldsymbol{P}_k \succeq 0, \boldsymbol{P}_k \text{ diagonal }, \forall k, \tag{3.1}$$

where the rates R_k compute according to (2.5). The weights μ_k are given a priori and reflect the priorities assigned to the corresponding user from higher layers of the communication system. μ_k can for example be proportional to the queue length of data packets waiting for transmission to user k (e.g. [26] and references therein). Solving (3.1) leads to a rate vector on the boundary of the capacity region [7, 14] and for this purpose DPC has to be applied [14]. Hence, an optimum encoding order has also to be found in (3.1), which is reflected by the variables $\hat{\pi}(i)$, $i = 1,\ldots,K$. The user $\hat{\pi}(i)$ is encoded at ith place, i.e., the function

$$\hat{\pi} : \{1,\ldots,K\} \mapsto \{1,\ldots,K\}, i \mapsto \hat{\pi}(i). \tag{3.2}$$

maps the encoding position i for DPC to a user index k. According to the principle of DPC explained in the last chapter, the sets $\mathcal{I}_k$ in (2.5) are given by

$$\mathcal{I}_k = \{\hat{\pi}(j) \in \{1,\ldots,K\} | j > i, \hat{\pi}(i) = k\}. \tag{3.3}$$

By using receive filters

$$\boldsymbol{G}_k = \left(\boldsymbol{I}_{Cr_k} \frac{\sigma_{\mathrm{n}}^2}{C} + \boldsymbol{H}_k \left[\sum_{m \in \mathcal{I}_k, m=k}^{K} \boldsymbol{T}_m \boldsymbol{P}_m \boldsymbol{T}_m^{\mathrm{H}} \right] \boldsymbol{H}_k^{\mathrm{H}} \right)^{-1} \boldsymbol{H}_k \boldsymbol{T}_k \boldsymbol{P}_k^{\frac{1}{2}}, \tag{3.4}$$

which minimize the Mean Square Error (MSE) between the received and the desired signal, the rates R_k in (2.5) compute according to

$$R_k = \log_2 \frac{\left| \boldsymbol{I}_{Cr_k} \frac{\sigma_\mathrm{n}^2}{C} + \boldsymbol{H}_k \boldsymbol{T}_k \boldsymbol{P}_k \boldsymbol{T}_k^\mathrm{H} \boldsymbol{H}_k^\mathrm{H} + \boldsymbol{H}_k \left[\sum_{\substack{m=1 \\ m \in \mathcal{I}_k}}^{K} \boldsymbol{T}_m \boldsymbol{P}_m \boldsymbol{T}_m^\mathrm{H} \right] \boldsymbol{H}_k^\mathrm{H} \right|}{\left| \boldsymbol{I}_{Cr_k} \frac{\sigma_\mathrm{n}^2}{C} + \boldsymbol{H}_k \left[\sum_{\substack{m=1 \\ m \in \mathcal{I}_k}}^{K} \boldsymbol{T}_m \boldsymbol{P}_m \boldsymbol{T}_m^\mathrm{H} \right] \boldsymbol{H}_k^\mathrm{H} \right|}, \tag{3.5}$$

i.e., become independent of the receive filters $\boldsymbol{G}_k$. The rates in (3.5) correspond to the rates achievable before receive filtering. As receive filtering can only lower the rates and additionally the constraints in (3.1) are independent of the receive filters, the optimum $\boldsymbol{G}_k$ are given by (3.4) and Problem (3.1) reads as

$$\max_{\{\boldsymbol{T}_k, \boldsymbol{P}_k, \hat{\pi}(k)\}_{k=1,\ldots,K}} \sum_{k=1}^{K} \mu_k R_k =$$

$$\max_{\{\boldsymbol{T}_k, \boldsymbol{P}_k, \hat{\pi}(k)\}_{k=1,\ldots,K}} \sum_{k=1}^{K} \mu_k \log_2 \frac{\left| \boldsymbol{I}_{Cr_k} + \frac{C}{\sigma_\mathrm{n}^2} \boldsymbol{H}_k \boldsymbol{T}_k \boldsymbol{P}_k \boldsymbol{T}_k^\mathrm{H} \boldsymbol{H}_k^\mathrm{H} + \frac{C}{\sigma_\mathrm{n}^2} \boldsymbol{H}_k \left[\sum_{\substack{m=1 \\ m \in \mathcal{I}_k}}^{K} \boldsymbol{T}_m \boldsymbol{P}_m \boldsymbol{T}_m^\mathrm{H} \right] \boldsymbol{H}_k^\mathrm{H} \right|}{\left| \boldsymbol{I}_{Cr_k} + \frac{C}{\sigma_\mathrm{n}^2} \boldsymbol{H}_k \left[\sum_{\substack{m=1 \\ m \in \mathcal{I}_k}}^{K} \boldsymbol{T}_m \boldsymbol{P}_m \boldsymbol{T}_m^\mathrm{H} \right] \boldsymbol{H}_k^\mathrm{H} \right|},$$

$$\text{s.t.} \sum_{k=1}^{K} \mathrm{tr}(\boldsymbol{P}_k) \leq P_\mathrm{Tx}, \quad \boldsymbol{e}_j^\mathsf{T} \boldsymbol{T}_k^\mathrm{H} \boldsymbol{T}_k \boldsymbol{e}_j = 1, \forall j = 1, \ldots, Cr_k, \forall k, \quad \boldsymbol{P}_k \succeq 0, \boldsymbol{P}_k \text{ diagonal}, \forall k, \tag{3.6}$$

As the objective function in (3.6) depends only on the covariance matrices $\boldsymbol{Q}_k = \boldsymbol{T}_k \boldsymbol{P}_k \boldsymbol{T}_k^\mathrm{H} \in \mathbb{C}^{CN_\mathrm{T} \times CN_\mathrm{T}}$, the maximization can also be conducted with respect to the $\boldsymbol{Q}_k$ and can be stated as

$$\max_{\{\boldsymbol{Q}_k, \hat{\pi}(k)\}_{k=1,\ldots,K}} \sum_{k=1}^{K} \mu_k \log_2 \frac{\left| \boldsymbol{I}_{Cr_k} + \frac{C}{\sigma_\mathrm{n}^2} \boldsymbol{H}_k \boldsymbol{Q}_k \boldsymbol{H}_k^\mathrm{H} + \frac{C}{\sigma_\mathrm{n}^2} \boldsymbol{H}_k \left[\sum_{\substack{m=1 \\ m \in \mathcal{I}_k}}^{K} \boldsymbol{Q}_m \right] \boldsymbol{H}_k^\mathrm{H} \right|}{\left| \boldsymbol{I}_{Cr_k} + \frac{C}{\sigma_\mathrm{n}^2} \boldsymbol{H}_k \left[\sum_{\substack{m=1 \\ m \in \mathcal{I}_k}}^{K} \boldsymbol{Q}_m \right] \boldsymbol{H}_k^\mathrm{H} \right|},$$

$$\text{s.t.} \sum_{k=1}^{K} \mathrm{tr}(\boldsymbol{Q}_k) \leq P_\mathrm{Tx}, \quad \boldsymbol{Q}_k \succeq 0, \forall k. \tag{3.7}$$

Note that the optimum transmit filter vectors $\boldsymbol{T}_k$ and $\boldsymbol{P}_k$ can be obtained from any decomposition of the optimum $\boldsymbol{Q}_k = \boldsymbol{T}_k \boldsymbol{P}_k \boldsymbol{T}_k^\mathrm{H}$, as long as the constraints on the $\boldsymbol{T}_k$ in (3.6) are fulfilled. Choosing the $\boldsymbol{T}_k$ as the eigenvectors of $\boldsymbol{Q}_k$ and $\boldsymbol{P}_k$ as the diagonal matrix containing the corresponding non-zero eigenvalues for example leads to a set of transmit filter vectors fulfilling these constraints. Unfortunately, Problem (3.7) is non-convex and can therefore not be solved straightforwardly. However, the duality between the broadcast channel and the dual Multiple Access Channel (MAC) from [5] can be applied to obtain the optimum covariances and encoding order from a convex optimization problem as described in the following. The duality implies that the transmitter in the broadcast channel becomes the receiver in the dual MAC and the kth user terminal sends the same data symbols $\boldsymbol{s}_k[n]$ as dedicated to it in the broadcast channel to the receiver. The additive noise in the dual MAC is also additive Gaussian with zero mean and covariance matrix $\frac{\sigma_\text{n}^2}{C}\boldsymbol{I}_{CN_\text{T}}$. In contrast to the physical MAC, the channel between user k and the receiver is given by $\boldsymbol{H}_k^\text{H}$ and a sum power constraint P_Tx across all users has to be fulfilled in the dual MAC. Additionally, instead of DPC in the broadcast channel, the receiver applies Successive Interference Cancellation (SIC). This implies that interference from symbols decoded before a certain symbol can be subtracted before decoding that symbol. This interference does therefore not reduce the rate for that symbol. Under these conditions a rate vector $[R_1, \ldots, R_K]^\mathsf{T}$ achievable with an encoding order $\hat{\pi}(i)$ and covariance matrices $\boldsymbol{Q}_1, \ldots, \boldsymbol{Q}_K$ fulfilling the power constraint in the broadcast channel, is also achievable in the dual MAC under the same power constraint with covariance matrices $\boldsymbol{W}_k \in \mathbb{C}^{Cr_k \times Cr_k}$ and a decoding order $\tilde{\pi}(i)$. The decoding order in the dual MAC is the reverse encoding order of the broadcast channel, i.e.,

$$\tilde{\pi}(i) = \hat{\pi}(K - i + 1), \tag{3.8}$$

which implies that the user encoded last in the broadcast channel is decoded first in the dual MAC [5]. Note that thereby counterintuitively users that have to be considered as interferers in the broadcast channel are suppressed by SIC in the dual MAC and vice versa. Thus, in the dual MAC the users interfering with user k are given by those users not contained in the interference set $\mathcal{I}_k$ of the broadcast channel, except user k, which does not interfere with itself. The relation between the covariances $\boldsymbol{W}_k$ in the dual MAC and the $\boldsymbol{Q}_k$ in the broadcast channel is given by the duality transformations in [5]. As a consequence, the encoding order and covariance matrices maximizing the weighted sum rate in (3.7), can also be obtained from a weighted sum rate maximization in the dual MAC, which is given by

$$\max_{\{\boldsymbol{W}_k, \hat{\pi}(k)\}_{k=1,\ldots,K}} \sum_{k=1}^{K} \mu_k \log_2 \frac{\left| \boldsymbol{I}_{CN_\text{T}} + \frac{C}{\sigma_\text{n}^2}\boldsymbol{H}_k^\text{H}\boldsymbol{W}_k\boldsymbol{H}_k + \frac{C}{\sigma_\text{n}^2}\left[\sum_{\substack{m=1\\ m\notin\mathcal{I}_k, m\neq k}}^{K} \boldsymbol{H}_m^\text{H}\boldsymbol{W}_m\boldsymbol{H}_m\right]\right|}{\left| \boldsymbol{I}_{CN_\text{T}} + \frac{C}{\sigma_\text{n}^2}\left[\sum_{\substack{m=1\\ m\notin\mathcal{I}_k, m\neq k}}^{K} \boldsymbol{H}_m^\text{H}\boldsymbol{W}_m\boldsymbol{H}_m\right]\right|},$$

$$\text{s.t.} \sum_{k=1}^{K} \operatorname{tr}(\boldsymbol{W}_k) \leq P_\text{Tx}, \quad \boldsymbol{W}_k \succeq 0, \forall k, \tag{3.9}$$

where the maximization is conducted directly over the broadcast encoding order $\hat{\pi}(k)$, which is related to the decoding order in the dual MAC via (3.8). The optimum decoding order in the dual MAC is given as the inverse order of the weights [27], i.e., the user with the smallest weight is decoded first, the user with the second smallest weight decoded second and so forth until the user with the highest weight is decoded last so that

$$\mu_{\tilde{\pi}(1)} \leq \mu_{\tilde{\pi}(2)} \leq \ldots \leq \mu_{\tilde{\pi}(K)}, \tag{3.10}$$

which is equivalent to

$$\mu_{\hat{\pi}(1)} \geq \mu_{\hat{\pi}(2)} \geq \ldots \geq \mu_{\hat{\pi}(K)} \tag{3.11}$$

using (3.8). The optimality of this decoding order can be proofed similarly to Theorem 2 in [28], where the MAC with single antenna users and a single antenna receiver is considered. If a group of users has equal weights, any decoding order within this group will lead to the same optimum of the objective function. Taking the optimum decoding order into account, Problem (3.9) can be simplified to

$$\begin{aligned}\max_{\{\boldsymbol{W}_k\}_{k=1,\ldots,K}} \sum_{k=1}^{K} \Delta\mu_k \log_2 \left| \boldsymbol{I}_{CN_{\mathrm{T}}} + \frac{C}{\sigma_{\mathrm{n}}^2} \left[\sum_{m=1}^{k} \boldsymbol{H}_{\hat{\pi}(m)}^{\mathrm{H}} \boldsymbol{W}_{\hat{\pi}(m)} \boldsymbol{H}_{\hat{\pi}(m)} \right] \right| = \\ = \max_{\{\boldsymbol{W}_k\}_{k=1,\ldots,K}} R_{\mathrm{wsr}}(\boldsymbol{W}_1, \ldots \boldsymbol{W}_K), \\ \text{s.t.} \sum_{k=1}^{K} \mathrm{tr}(\boldsymbol{W}_k) \leq P_{\mathrm{Tx}}, \quad \boldsymbol{W}_k \succeq 0, \forall k, \end{aligned} \tag{3.12}$$

where

$$\Delta\mu_k = \begin{cases} \mu_{\hat{\pi}(k)}, & k = K, \\ \mu_{\hat{\pi}(k)} - \mu_{\hat{\pi}(k+1)}, & k = 1, \ldots, K-1 \end{cases}.$$

Note that due to (3.11), the $\Delta\mu_k$ are greater or equal to zero. Finally Problem (3.9) consists of an objective function concave in the $\boldsymbol{W}_k$ and convex constraints sets, which make the whole problem convex. It could therefore be solved by interior point methods (e.g. [29]). Nevertheless, some more efficient algorithms have been published since the discovery of the duality between the MIMO broadcast channel and the dual MAC. Viswanathan et al. have first proposed an iterative algorithm in [7] to solve (3.12). In each iteration a rank-one update of one user's covariance matrix is performed, convergence is however very poor [30]. The algorithms in [31] and [32] use conjugate gradient methods with projections, where the first algorithm works with the precoding matrices in the dual uplink instead of the covariance matrices $\boldsymbol{W}_k$. Hunger et al. propose an algorithm in [30] which also relies on a projected gradient, but uses different update rules for the covariances matrices in each step than [32]. It exhibits a faster convergence than the other mentioned algorithms and will therefore be reviewed in the following. Beginning with initial covariance matrices $\boldsymbol{W}_k^{(0)} = \frac{P_{\mathrm{Tx}}}{C\sum_{k=1}^{K} r_k} \boldsymbol{I}_{Cr_k}$ the algorithm iteratively updates the covariance matrices so that an increase in weighted sum rate occurs until the improvement from one iteration to another is less than a predefined threshold ε. Each iteration consists of two main steps, an unconstrained gradient update and an orthogonal projection. The first step implies that the covariance matrices are changed in the direction of the steepest ascent of the cost function according to

$$\hat{\boldsymbol{W}}_k^{(i)} = \boldsymbol{W}_k^{(i-1)} + \frac{1}{d^{(i-1)}} \frac{P_{\mathrm{Tx}}}{\sum_{k=1}^{K} \mathrm{tr}\left(\boldsymbol{\Phi}_k \left(\boldsymbol{W}_1^{(i-1)}, \ldots, \boldsymbol{W}_K^{(i-1)}\right)\right)} \boldsymbol{\Phi}_k \left(\boldsymbol{W}_1^{(i-1)}, \ldots, \boldsymbol{W}_K^{(i-1)}\right), \tag{3.13}$$

where $\boldsymbol{W}_k^{(i-1)}$ denotes the covariance matrix of user k after the ith iteration and the gradients $\boldsymbol{\Phi}_k \left(\boldsymbol{W}_1^{(i-1)}, \ldots, \boldsymbol{W}_K^{(i-1)}\right)$ compute according to

$$\begin{aligned}\boldsymbol{\Phi}_k \left(\boldsymbol{W}_1^{(i-1)}, \ldots, \boldsymbol{W}_K^{(i-1)}\right) = \left.\frac{\partial R_{\mathrm{wsr}}(\boldsymbol{W}_1, \ldots \boldsymbol{W}_K)}{\partial \boldsymbol{W}_k^{\mathsf{T}}}\right|_{\boldsymbol{W}_1^{(i-1)}, \ldots, \boldsymbol{W}_K^{(i-1)}} = \\ = \sum_{j=\hat{\pi}^{-1}(k)}^{K} \Delta\mu_j \frac{C}{\ln 2\sigma_{\mathrm{n}}^2} \boldsymbol{H}_k \left(\boldsymbol{I}_{CN_{\mathrm{T}}} + \frac{C}{\sigma_{\mathrm{n}}^2} \left[\sum_{m=1}^{j} \boldsymbol{H}_{\hat{\pi}(m)}^{\mathrm{H}} \boldsymbol{W}_{\hat{\pi}(m)}^{(i-1)} \boldsymbol{H}_{\hat{\pi}(m)} \right] \right)^{-1} \boldsymbol{H}_k^{\mathrm{H}}, \end{aligned} \tag{3.14}$$

where $\hat{\pi}^{-1}(k)$ denotes the inverse of the encoding function in the broadcast channel, i.e., $\hat{\pi}^{-1}(k)$ is the encoding position of user k. The scaling with P_{Tx} divided by the sum of the gradients in (3.13) increases the speed of convergence [30]. The "inverse step size" $d^{(i)}$ is initialized with $d^{(0)} = 1$ and possibly updated after the projections, which might be necessary as described later. As the gradients $\boldsymbol{\Phi}_k\left(\boldsymbol{W}_1^{(i-1)},\ldots,\boldsymbol{W}_K^{(i-1)}\right)$ are positive semidefinite and $d^{(i-1)} > 0$, the matrices $\hat{\boldsymbol{W}}_k^{(i)}$ are also positive semidefinite but do generally not fulfill the sum power constraint. For this reason a projection back onto the feasible domain is necessary in a second step. For this purpose an orthogonal projection is used which minimizes the error between the matrices $\hat{\boldsymbol{W}}_k^{(i)}$ and the matrices $\boldsymbol{W}_k^{(i)}$, which fulfill the sum power constraint. The error is thereby measured in terms of the Frobenius norm so that

$$\left\{\boldsymbol{W}_k^{(i)}\right\}_{k=1,\ldots,K} = \underset{\left\{\tilde{\boldsymbol{W}}_k^{(i)}\right\}_{k=1,\ldots,K}}{\operatorname{argmin}} \sum_{k=1}^{K} \left\|\hat{\boldsymbol{W}}_k^{(i)} - \tilde{\boldsymbol{W}}_k^{(i)}\right\|_{\text{F}}^2,$$

$$\text{s.t.} \sum_{k=1}^{K} \operatorname{tr}\left(\tilde{\boldsymbol{W}}_k^{(i)}\right) \leq P_{\text{Tx}}, \quad \tilde{\boldsymbol{W}}_k^{(i)} \succeq 0, \forall k. \tag{3.15}$$

Solving the Karush-Kuhn-Tucker (KKT) [33] conditions of Problem (3.15) yields

$$\boldsymbol{W}_k^{(i)} = \boldsymbol{U}_k^{(i)} \left[\boldsymbol{\Sigma}_k^{(i)} - \lambda \boldsymbol{I}_{CN_{\text{T}}}\right]^+ \boldsymbol{U}_k^{(i),\text{H}}, \tag{3.16}$$

where $[\boldsymbol{A}]^+$ sets all negative elements in the matrix $\boldsymbol{A}$ to zero and $\boldsymbol{U}_k^{(i)}$ and $\boldsymbol{\Sigma}_k^{(i)}$ stem from the Eigenvalue Decomposition (EVD) of the matrix

$$\hat{\boldsymbol{W}}_k^{(i)} = \boldsymbol{U}_k^{(i)} \boldsymbol{\Sigma}_k^{(i)} \boldsymbol{U}_k^{(i),\text{H}}. \tag{3.17}$$

The Lagrange multiplier λ is determined iteratively so that the transmit power constraint is fulfilled with equality, i.e.,

$$\sum_{k=1}^{K} \operatorname{tr}\left(\boldsymbol{W}_k^{(i)}\right) = \sum_{k=1}^{K} \operatorname{tr}\left(\left[\boldsymbol{\Sigma}_k^{(i)} - \lambda \boldsymbol{I}_{CN_{\text{T}}}\right]^+\right) = P_{\text{Tx}},$$

which can be done as described in [30, Corollary 1]. It can however happen that with the new covariance matrices $\boldsymbol{W}_k^{(i)}$ the weighted sum rate decreases compared to the previous step, i.e., $R_{\text{wsr}}\left(\boldsymbol{W}_1^{(i)},\ldots,\boldsymbol{W}_K^{(i)}\right) < R_{\text{wsr}}\left(\boldsymbol{W}_1^{(i-1)},\ldots,\boldsymbol{W}_K^{(i-1)}\right)$. In that case the step size with which the gradients $\boldsymbol{\Phi}_k\left(\boldsymbol{W}_1^{(i-1)},\ldots,\boldsymbol{W}_K^{(i-1)}\right)$ in (3.13) are multiplied has been chosen to large and must therefore be reduced. For this purpose $d_i^{(i-1)}$ is increased by one and the steps (3.13) and (3.15) must be repeated. In case the weighted sum rate is still lower than in the previous step, further repetitions are required until an increase in the objective function is obtained, where in each repetition $d_i^{(i-1)}$ is increased by one. The value for $d_i^{(i-1)}$ which finally leads to an increase in the weighted sum rate is then retained for the next iteration, i.e., $d_i^{(i)} = d_i^{(i-1)}$. The steps required to solve (3.12) are summarized in Algorithm 3.1.

Algorithm 3.1 has been proposed for single-carrier systems where $C = 1$. By applying the algo-

Algorithm 3.1 Projected Gradient Algorithm for Weighted Sum Rate Maximization in the Dual MAC using Orthogonal Projections

$\boldsymbol{W}_k^{(0)} = \frac{P_{\text{Tx}}}{C\sum_{k=1}^{K} r_k} \boldsymbol{I}_{Cr_k}\ \forall k$

$d^{(0)} = 1$

$i = 1$

repeat

  $\boldsymbol{\Phi}_k\left(\boldsymbol{W}_1^{(i-1)}, \ldots, \boldsymbol{W}_K^{(i-1)}\right) = \frac{\partial R_{\text{wsr}}(\boldsymbol{W}_1, \ldots \boldsymbol{W}_K)}{\partial \boldsymbol{W}_k^{\mathsf{T}}}\Big|_{\boldsymbol{W}_1^{(i-1)}, \ldots, \boldsymbol{W}_K^{(i-1)}},\ \forall k$

  repeat

    $\hat{\boldsymbol{W}}_k^{(i)} = \boldsymbol{W}_k^{(i-1)} + \frac{1}{d^{(i-1)}} \frac{P_{\text{Tx}}}{\sum_{k=1}^{K} \operatorname{tr}\left(\boldsymbol{\Phi}_k\left(\boldsymbol{W}_1^{(i-1)}, \ldots, \boldsymbol{W}_K^{(i-1)}\right)\right)} \boldsymbol{\Phi}_k\left(\boldsymbol{W}_1^{(i-1)}, \ldots, \boldsymbol{W}_K^{(i-1)}\right),\ \forall k$

    Compute EVDs $\hat{\boldsymbol{W}}_k^{(i)} = \boldsymbol{U}_k^{(i)} \boldsymbol{\Sigma}_k^{(i)} \boldsymbol{U}_k^{(i),\mathrm{H}},\ \forall k.$

    $\boldsymbol{W}_k^{(i)} = \boldsymbol{U}_k^{(i)} \left[\boldsymbol{\Sigma}_k^{(i)} - \lambda \boldsymbol{I}_{CN_{\mathrm{T}}}\right]^{+} \boldsymbol{U}_k^{(i),\mathrm{H}}\ \forall k$, where λ is determined via implicit equation $\sum_{k=1}^{K} \operatorname{tr}\left(\left[\boldsymbol{\Sigma}_k^{(i)} - \lambda \boldsymbol{I}_{CN_{\mathrm{T}}}\right]^{+}\right) = P_{\text{Tx}}$

    if $R_{\text{wsr}}\left(\boldsymbol{W}_1^{(i)}, \ldots, \boldsymbol{W}_K^{(i)}\right) < R_{\text{wsr}}\left(\boldsymbol{W}_1^{(i-1)}, \ldots, \boldsymbol{W}_K^{(i-1)}\right)$ **then**

      $d^{(i-1)} = d^{(i-1)} + 1$

    end if

  until $R_{\text{wsr}}\left(\boldsymbol{W}_1^{(i)}, \ldots, \boldsymbol{W}_K^{(i)}\right) > R_{\text{wsr}}\left(\boldsymbol{W}_1^{(i-1)}, \ldots, \boldsymbol{W}_K^{(i-1)}\right)$

  $d^{(i)} = d^{(i-1)}$

  $i = i + 1$

until $R_{\text{wsr}}\left(\boldsymbol{W}_1^{(i)}, \ldots, \boldsymbol{W}_K^{(i)}\right) - R_{\text{wsr}}\left(\boldsymbol{W}_1^{(i-1)}, \ldots, \boldsymbol{W}_K^{(i-1)}\right) \leq \varepsilon$

rithm to multicarrier systems one would expect the complexity to be cubic in the number of carriers, as matrix inversions and EVDs of $CN_\mathrm{T} \times CN_\mathrm{T}$ matrices are required in (3.14) and (3.17), respectively (see [34] for the complexity of matrix inversions and [35] for the complexity of EVDs). By exploiting the block-diagonal structure of the channel matrices $\boldsymbol{H}_k$ [c.f. (2.1)] however, the complexity grows only linearly in the number of carriers. That is because the matrices $\boldsymbol{W}_k^{(i)}$ are block-diagonal as long as all matrices $\boldsymbol{W}_k^{(i-1)}$ are block-diagonal, since in this case Equations (3.13), (3.14), and (3.16) preserve the block-diagonal structure. With block-diagonal $\boldsymbol{W}_k^{(0)}$ and the fact that Algorithm 3.1 always converges [30], the optimum covariance matrices $\boldsymbol{W}_k$ will therefore also exhibit a block-diagonal structure. A more rigorous proof for that fact can be found in [36, Section 3.1.2.]. Thus, Equations (3.14) and (3.17) can be decomposed into C independent equations with $N_\mathrm{T} \times N_\mathrm{T}$ matrices and therefore be solved more efficiently.
For the special case of equal weights, i.e., $\mu_1 = \ldots = \mu_K$, Problem (3.12) reduces to the pure sum rate maximization, which reads as

$$\max_{\{\boldsymbol{W}_k\}_{k=1,\ldots,K}} \log_2 \left| \boldsymbol{I}_{CN_\mathrm{T}} + \frac{C}{\sigma_\mathrm{n}^2} \left[\sum_{m=1}^{K} \boldsymbol{H}_{\hat{\pi}(m)}^\mathrm{H} \boldsymbol{W}_{\hat{\pi}(m)} \boldsymbol{H}_{\hat{\pi}(m)} \right] \right| =$$
$$\text{s.t.} \sum_{k=1}^{K} \mathrm{tr}(\boldsymbol{W}_k) \leq P_\mathrm{Tx}, \quad \boldsymbol{W}_k \succeq 0, \forall k. \tag{3.18}$$

Although Problem (3.18) can also be solved with Algorithm 3.1, the simplified objective function compared to (3.12) has induced the development of different solution methods. During one iteration of the iterative water-filling algorithm proposed in [6], first each user's noise+interference covariance matrix is computed assuming that the transmit covariance matrices of the other users remain the same as in the previous step. This assumption enables a closed-form solution for transmit covariance matrices maximizing an estimated sum rate. Those matrices are then used to update the covariance matrices in the iterative algorithm. The computation of the update requires one matrix inversion and a complete EVD for each user. To avoid the difficulties related to the convergence properties of iterative-waterfilling, a modification of the algorithm is presented in [37] and in [38] an improvement for the update rule of the covariance matrices is proposed to achieve faster convergence. In [39] Yu solves Problem (3.18) via the minimization of the Lagrange dual function with two nested loops. During the inner loop the optimum covariance matrices for a fixed Lagrange multiplier are determined in an iterative manner, where each iteration again requires K matrix inversions and K EVDs, while in the outer loop this multiplier is adjusted until convergence. An improvement of the convergence behavior can be achieved by the modification from [40]. Nevertheless, those algorithms are not able to outperform the method from [41] in terms of speed of convergence and computational complexity. The numerically complex EVDs are avoided by using a scaled projected gradient algorithm that optimizes the precoders in the dual MAC instead of the covariance matrices. Thus, complexity is further reduced, although K matrix inversions are still required in each iteration for the computation of gradients.

3.2 Spatial Zero-Forcing and Successive Resource Allocation with Dirty Paper Coding

Despite the advancements in reducing the computational complexity of finding the optimum covariance matrices for weighted sum rate maximization, that have been described in the previous section, achieving the optimum solution still remains a computationally complex problem. Algorithm 3.1 features an iterative nature and each iteration requires CK inversions of $N_\mathrm{T} \times N_\mathrm{T}$ matrices to compute the gradients of the objective function in Line 7 and CK complete eigenvalue decompositions of $N_\mathrm{T} \times N_\mathrm{T}$ matrices in Line 8. Additionally, the complexity of the transformation of the optimum covariance matrices in the dual MAC to the broadcast channel is approximately as high as the complexity of one iteration of Algorithm 3.1. For this reason an efficient non-iterative algorithm for weighted sum rate maximization will be presented in this section that is able to achieve the optimum solution closely at drastically reduced computational complexity.

3.2.1 State-of-the-Art Near Optimum Approaches

To attain an efficient near optimum method for weighted sum rate maximization two complementary simplifications will be made in the following, namely the introduction of zero-forcing constraints and a greedy resource allocation. Zero-forcing in combination with DPC has been proposed in [42] for MISO systems. The optimum encoding oder of users and the selection of users that receive non-zero powers is however only solved by an exhaustive search in [42], which becomes infeasible with increasing number of users. Tu and Blum therefore propose in [43] to determine an encoding order in a greedy manner, i.e., to encode the user with the strongest norm of its channel vector first and then to encode in each step that user that leads to the strongest increase in sum rate provided that the encoding order determined in previous steps is kept fixed. A method for complexity reduction of this algorithm based on Householder transformations is presented in [44]. For MIMO systems, zero-forcing and DPC has been combined in [45], where each user receives as many data streams as it has receive antennas and all data streams of the same user are encoded simultaneously. It is shown in [45] that this scheme is asymptotically optimum for infinite SNR, where any encoding order is optimum. For finite SNR, finding the optimum set of active users and the optimum encoding order for DPC is solved by an exhaustive search in [45], which is why a greedy determination of the encoding order is described in [46]. In each step the user to be encoded next is the user that leads to the strongest increase in sum rate. Furthermore simplified selection rules are proposed in [46] avoiding the explicit computations of sum rates for each candidate user in each step. Alternative simplified selection rules are presented in [47] and [48]. While the greedy approach can be extended straightforwardly to the problem of weighted sum rate maximization by choosing in each step the user that leads to the strongest increase in weighted sum rate instead of sum rate, the aforementioned simplifications are only applicable to the problem of sum rate maximization. Furthermore the mentioned approaches assign either r_k or zero data streams to user k, which can be sub-optimum, especially with zero-forcing constraints. For these reasons an efficient algorithm for weighted sum rate maximization exploiting spatial zero-forcing and a successive resource allocation will be presented in the following, where the number of data streams allocated to a certain user k can vary from 0 to r_k and which exhibits the potential for further complexity reductions, which will be explicated in Section 3.4. An algorithm for weighted sum rate maximization, that is based to some extend based on the same principles and to which the complexity reductions derived in Section 3.4 can also be applied, is presented in [49, Chapter 6]. However, this algorithm is much more complex as it is of iterative nature and one iteration exhibits the same complexity as the method presented in the next sections.

3.2.2 Spatial Zero-Forcing

As in Section 3.1, Problem (3.1) is the starting point for the derivation of a near-optimum efficient algorithm. To end up with such a method, first zero-forcing constraints are introduced to that problem, so that after receive filtering each user experiences no interference from other users. Mathematically, the zero-forcing constraints can be written as

$$\boldsymbol{P}_k^{\frac{1}{2}} \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{T}_m \boldsymbol{P}_m^{\frac{1}{2}} = \mathbf{0}_{d_k, d_m}, \quad \forall m \in \mathcal{I}_k, \forall k, \tag{3.19}$$

where the multiplication with $\boldsymbol{P}_k^{\frac{1}{2}}$ assures that data streams with $[\boldsymbol{P}_k]_{j,j} = 0$ do not impose zero-forcing constraints. The sets $\mathcal{I}_k$ are given as defined in (3.3) and thus Problem (3.1) reads as

$$\max_{\{\boldsymbol{T}_k, \boldsymbol{P}_k, \boldsymbol{G}_k, \hat{\pi}(k)\}_{k=1,\ldots,K}} \sum_{k=1}^{K} \mu_k \log_2 \left| \boldsymbol{I}_{Cr_k} + \frac{C}{\sigma_{\mathrm{n}}^2} \left(\boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{G}_k \right)^{-1} \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{T}_k \boldsymbol{P}_k \boldsymbol{T}_k^{\mathrm{H}} \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{G}_k \right| =$$

$$\text{s.t.} \sum_{k=1}^{K} \mathrm{tr}(\boldsymbol{P}_k) \leq P_{\mathrm{Tx}}, \quad \boldsymbol{e}_j^T \boldsymbol{T}_k^{\mathrm{H}} \boldsymbol{T}_k \boldsymbol{e}_j = 1, \forall j = 1, \ldots, d_k, \forall k, \quad \boldsymbol{P}_k \succeq 0, \boldsymbol{P}_k \text{ diagonal }, \forall k,$$
$$\boldsymbol{P}_k^{\frac{1}{2}} \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{T}_m \boldsymbol{P}_m^{\frac{1}{2}} = \mathbf{0}_{d_k, d_m}, \quad \forall m \in \mathcal{I}_k, \forall k. \tag{3.20}$$

In contrast to the optimum, the receive filters cannot be removed a priori from the optimization problem as in (3.6), because the constraints in (3.20) are no longer independent of the receive filters. Additionally, the optimization with zero-forcing constraints cannot be conducted with respect to the transmit covariance matrices. Nevertheless, for practical systems one is interested in the transmit filters anyway, for which reason an algorithm directly delivering the transmit filters and thus avoiding the final decomposition of the covariance matrices is desirable. Obviously, with the additional constraints the optimum weighted sum rate resulting from (3.20) will be smaller or equal to the optimum value of the objective function in (3.1). For high SNR, i.e., for $P_{\mathrm{Tx}} \to \infty$, the zero-forcing constraints are fulfilled at the optimum anyway and the two optimization problems lead to the same solution. Additionally to these inter-user zero-forcing constraints, the matrices $\boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{T}_k \boldsymbol{P}_k^{\frac{1}{2}}$ will be constrained to be diagonal in the following, i.e., all data streams of the same user must not interfere with each other. In contrast to the inter-user zero-forcing constraints, those intra-user zero-forcing constraints do not influence the optimum, as the matrices $\boldsymbol{T}_k \boldsymbol{P}_k^{\frac{1}{2}}$ can be multiplied from the right-hand side by any $d_k \times d_k$ orthonormal matrix without changing the objective function nor the constraints in Problem (3.20)[1] and these degrees of freedom can be used to make the matrices $\boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{T}_k \boldsymbol{P}_k^{\frac{1}{2}}$ or equivalently the matrices $\boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{T}_k$ diagonal. Additionally, the matrices $\boldsymbol{G}_k$ can be multiplied from the right-hand side by any invertible $d_k \times d_k$ matrix also without changing the objective function nor the constraints. Thus, constraining $\boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{G}_k$ to be equal to the $d_k \times d_k$ identity matrix does also not influence the optimum solution. Adding those constraints

[1]As the matrices $\boldsymbol{P}_k$ are diagonal and the matrices $\boldsymbol{T}_k$ contain columns with norm one, the power constraint can be also written as $\sum_{k=1}^{K} \mathrm{tr}(\boldsymbol{P}_k) = \sum_{k=1}^{K} \mathrm{tr}\left(\boldsymbol{T}_k \boldsymbol{P}_k^{\frac{1}{2}} \boldsymbol{P}_k^{\frac{1}{2}} \boldsymbol{T}_k^{\mathrm{H}} \right) = \sum_{k=1}^{K} \mathrm{tr}\left(\boldsymbol{T}_k \boldsymbol{P}_k^{\frac{1}{2}} \boldsymbol{U}_k \boldsymbol{U}_k^{\mathrm{H}} \boldsymbol{P}_k^{\frac{1}{2}} \boldsymbol{T}_k^{\mathrm{H}} \right) \leq P_{\mathrm{Tx}}$, where $\boldsymbol{U}_k \in \mathbb{C}^{d_k \times d_k}$ is an arbitrary orthonormal matrix with $\boldsymbol{U}_k \boldsymbol{U}_k^{\mathrm{H}} = \boldsymbol{I}_{d_k}$.

to Problem (3.20) leads to the following optimization problem

$$\begin{aligned}
&\max_{\{\boldsymbol{T}_k,\boldsymbol{P}_k,\boldsymbol{G}_k,\hat{\pi}(k)\}_{k=1,\ldots,K}} \sum_{k=1}^{K} \mu_k \log_2 \left| \boldsymbol{I}_{Cr_k} + \frac{C}{\sigma_{\mathrm{n}}^2} \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{T}_k \boldsymbol{P}_k \boldsymbol{T}_k^{\mathrm{H}} \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{G}_k \right|, \\
&\text{s.t.} \sum_{k=1}^{K} \operatorname{tr}(\boldsymbol{P}_k) \leq P_{\mathrm{Tx}}, \quad \boldsymbol{e}_j^{\mathrm{T}} \boldsymbol{T}_k^{\mathrm{H}} \boldsymbol{T}_k \boldsymbol{e}_j = 1, \forall j = 1,\ldots,d_k, \forall k, \quad \boldsymbol{P}_k \succeq 0, \boldsymbol{P}_k \text{ diagonal}, \forall k, \\
&\boldsymbol{P}_k^{\frac{1}{2}} \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{T}_m \boldsymbol{P}_m^{\frac{1}{2}} = \boldsymbol{0}_{d_k,d_m}, \quad \forall m \in \mathcal{I}_k, \forall k, \quad \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{T}_k = \text{ diagonal}, \forall k, \quad \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{G}_k = \boldsymbol{I}_{d_k}, \forall k.
\end{aligned} \tag{3.21}$$

Introducing the vectors

$$\boldsymbol{t}_{k,j} := \boldsymbol{T}_k \boldsymbol{e}_j, \qquad \boldsymbol{g}_{k,j} := \boldsymbol{G}_k \boldsymbol{e}_j$$

and the scalars

$$p_{k,j} := \boldsymbol{e}_j^{\mathrm{T}} \boldsymbol{P}_k \boldsymbol{e}_j,$$

so that $\boldsymbol{T}_k \boldsymbol{P}_k^{\frac{1}{2}} \boldsymbol{e}_j = \sqrt{p_{k,j}} \boldsymbol{t}_{k,j}$, the inter-user zero-forcing constraints $\boldsymbol{P}_k^{\frac{1}{2}} \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{T}_m \boldsymbol{P}_m^{\frac{1}{2}} = \boldsymbol{0}_{d_k,d_m}$ become

$$\sqrt{p_{k,j} p_{m,\ell}} \boldsymbol{g}_{k,j}^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{t}_{m,\ell} = 0, \quad \forall j = 1,\ldots,d_k, \forall \ell = 1,\ldots,d_m, \forall m \in \mathcal{I}_k, \forall k,$$

the intra-user zero-forcing constraints read as

$$\boldsymbol{g}_{k,j}^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{t}_{k,\ell} = 0, \forall j = 1,\ldots,d_k, \forall \ell = 1,\ldots,d_k, \ell \neq j, \forall k,$$

and the orthogonality constraints on the receive filters of different data streams allocated to the same user, i.e., $\boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{G}_k = \boldsymbol{I}_{d_k}$, can be re formulated as

$$\boldsymbol{g}_{k,m}^{\mathrm{H}} \boldsymbol{g}_{k,j} = 0, \forall j = 1,\ldots,d_k, m = 1,\ldots,d_k, j \neq m, \forall k.$$

Thus, Problem (3.21) can be rewritten as

$$\begin{aligned}
&\max_{\{\boldsymbol{t}_{k,j},\boldsymbol{g}_{k,j},p_{k,j}\}_{j=1,\ldots,d_k,k=1,\ldots K},\{\hat{\pi}(k)\}_{k=1,\ldots,K}} \sum_{k=1}^{K} \mu_k \sum_{j=1}^{d_k} \log_2 \left(1 + \frac{C}{\sigma_{\mathrm{n}}^2} p_{k,j} \boldsymbol{g}_{k,j}^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{t}_{k,j} \boldsymbol{t}_{k,j}^{\mathrm{H}} \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{g}_{k,j}\right), \\
&\text{s.t.} \sum_{k=1}^{K} \sum_{j=1}^{d_k} p_{k,j} \leq P_{\mathrm{Tx}}, \quad \boldsymbol{t}_{k,j}^{\mathrm{H}} \boldsymbol{t}_{k,j} = 1, \quad \boldsymbol{g}_{k,j}^{\mathrm{H}} \boldsymbol{g}_{k,j} = 1, \quad p_{k,j} \geq 0, \forall j = 1,\ldots,d_k, \forall k, \\
&\boldsymbol{g}_{k,m}^{\mathrm{H}} \boldsymbol{g}_{k,j} = 0, \forall j = 1,\ldots,d_k, m = 1,\ldots,d_k, j \neq m, \forall k, \\
&\sqrt{p_{k,j} p_{m,\ell}} \boldsymbol{g}_{k,j}^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{t}_{m,\ell} = 0, \quad \forall j = 1,\ldots,d_k, \forall \ell = 1,\ldots,d_m, \forall m \in \mathcal{I}_k, \forall k, \\
&\sqrt{p_{k,j} p_{k,\ell}} \boldsymbol{g}_{k,j}^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{t}_{k,\ell} = 0, \forall j = 1,\ldots,d_k, \forall \ell = 1,\ldots,d_k, \ell \neq j, \forall k.
\end{aligned} \tag{3.22}$$

Hence, the MIMO broadcast channel is decomposed into a system of scalar subchannels free of interference. The jth data stream of user k can be transmitted error-free at a rate of $\log_2\left(1 + \frac{C}{\sigma_{\mathrm{n}}^2} p_{k,j} \boldsymbol{g}_{k,j}^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{t}_{k,j} \boldsymbol{t}_{k,j}^{\mathrm{H}} \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{g}_{k,j}\right)$ and the optimum power allocation boils down to a water-filling alike algorithm so that the optimum $p_{k,j}$ are given by

$$p_{k,j} = \left[\eta \mu_k - \frac{\sigma_{\mathrm{n}}^2}{C \boldsymbol{g}_{k,j}^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{t}_{k,j} \boldsymbol{t}_{k,j}^{\mathrm{H}} \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{g}_{k,j}}\right]^+, \tag{3.23}$$

where η is determined so that the transmit power constraint is fulfilled with equality, as shown in Appendix A2. Despite the fact that the power allocation has become relatively easy through the zero-forcing constraints, Problem (3.22) is non-convex. Additionally, neither transforming (3.22) to the dual MAC nor introducing zero-forcing constraints in the dual MAC leads to a convex problem. On the other hand, directly working in the broadcast channel, as it will be done in the following, offers the advantage of avoiding the numerically expensive transformations from the dual MAC to the broadcast channel required for the optimum solution.

In multicarrier systems, where $C > 1$, the optimum broadcast covariance matrices have a block-diagonal structure, which follows from the fact that the optimum covariance matrices in the dual MAC are block-diagonal [36, Section 3.1.2] and the duality transforms from [5] preserve this structure. This implies that the optimum transmit filters $\boldsymbol{T}_k$ and the optimum receive filters $\boldsymbol{G}_k$ also exhibit a block diagonal structure. This is not necessarily the case with zero-forcing constraints at finite SNRs. Nevertheless, the $\boldsymbol{T}_k$ and the $\boldsymbol{G}_k$ will be enforced to be block-diagonal in the following. Besides the fact that the optimum filters are block-diagonal and it is desired to achieve this optimum as close as possible with a reduced complexity algorithm, the symbols can be encoded independently on all carriers in this case, while a joint encoding necessary without block-diagonal transmit filters would lead to long encoding delays and increased computational complexity. Adding these constraints to Problem (3.22) implies for the vectors $\boldsymbol{t}_{k,j}$ and $\boldsymbol{g}_{k,j}$ that all its elements must be zero except in the rows that correspond to the carrier $\gamma(k,j)$, where

$$\gamma(k,j) : \{1,\ldots,K\}, \left\{1,\ldots,\max_k d_k\right\} \mapsto \{1,\ldots,C\} : (k,j) \mapsto \gamma(k,j)$$

is the carrier over which the jth data stream of user k will be transmitted. In the following the condition of block-diagonal filters will be formulated as

$$\boldsymbol{t} = \boldsymbol{S}_{c,\mathrm{T}}^{\mathsf{T}} \boldsymbol{S}_{c,\mathrm{T}} \boldsymbol{t}, \qquad \boldsymbol{g} = \boldsymbol{S}_{c,k}^{\mathsf{T}} \boldsymbol{S}_{c,k} \boldsymbol{g} \tag{3.24}$$

with the selection matrices $\boldsymbol{S}_{c,\mathrm{T}}$ and $\boldsymbol{S}_{c,k}$ defined in (2.2). Note that with this restriction inter-carrier interference is not present, i.e., the zero-forcing constraints $\sqrt{p_{k,j}p_{m,\ell}}\boldsymbol{g}_{k,j}^{\mathrm{H}}\boldsymbol{H}_k\boldsymbol{t}_{m,\ell} = 0$ are fulfilled by (3.24), if $\boldsymbol{g}_{k,j}$ and $\boldsymbol{t}_{m,\ell}$ are the filters for data streams on different carriers, which is the case for $\gamma(k,j) \neq \gamma(m,\ell)$.

The maximum weighted sum rate in the broadcast channel can be achieved by a user-wise encoding [14], i.e., all d_k data streams of user k are encoded at the same time. Additionally, the same encoding order is applied on all carriers. With zero-forcing constraints, however, the encoding function $\hat{\pi}(i)$ as defined in (3.2) may no longer be optimum. For this reason the restriction on a user-wise encoding in (3.22) is abolished and an arbitrary encoding order of the data streams is allowed. This also implies that data streams of the same user may be encoded at two non-consecutive encoding positions, for example at first and fourth place. For sum rate maximization, i.e., equal weights $\mu_1 = \cdots = \mu_K$, this relaxation of the encoding order is one of the key properties of the Successive Encoding Successive Allocation Method (SESAM) presented in [10], which distinguishes SESAM from other successive zero-forcing approaches such as [46]. Consequently, the zero-forcing constraints in (3.23) need to be modified as follows. Due to the block-diagonal structure of the precoding vectors the ℓth data stream of user m experiences only interference from data streams allocated to the same carrier, i.e., data streams (k,j) for which $\gamma(m,\ell) = \gamma(k,j)$ holds. Interference from data streams encoded before the ℓth data stream of user m is cancelled by DPC. Those are the data streams on the carrier $\gamma(m,\ell)$ with $\pi(k,j) < \pi(m,\ell)$, where the encoding function

$$\pi(k,j) : \{1,\ldots,K\}, \left\{1,\ldots,\max_k(d_k)\right\} \mapsto \{1,\ldots,N_{\mathrm{T}}\}$$

returns the encoding position of the jth data stream of user k on carrier $\gamma(k,j)$. Thus the remaining interference must be suppressed by the precoders and the zero-forcing constraints therefore imply that

$$\sqrt{p_{m,\ell}p_{k,j}}\boldsymbol{g}_{m,\ell}^{\mathrm{H}}\boldsymbol{H}_m\boldsymbol{t}_{k,j} = 0, \forall (k,j) \text{ with } \gamma(m,\ell) = \gamma(k,j) \text{ and } \pi(k,j) > \pi(m,\ell) \tag{3.25}$$

for all possible tuples (m,ℓ). Note that the inter- as well as the intra-user zero-forcing constraints are considered with (3.25), as the case $m = k$ is not excluded from (3.25). With the modified encoding order, the optimum power allocation (3.23), and the requirement on block diagonal filters, Problem (3.22) becomes

$$\begin{aligned}
&\max_{\{\boldsymbol{t}_{k,j},\boldsymbol{g}_{k,j},\gamma(k,j),\pi(k,j)\}_{j,=1,\ldots,d_k,k=1,\ldots,K}} \sum_{k=1}^{K}\mu_k\sum_{j=1}^{d_k}\log_2\left(1+\frac{C}{\sigma_{\mathrm{n}}^2}p_{k,j}\boldsymbol{g}_{k,j}^{\mathrm{H}}\boldsymbol{H}_k\boldsymbol{t}_{k,j}\boldsymbol{t}_{k,j}^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{g}_{k,j}\right),\\
&\text{s.t. } \sum_{k=1}^{K}\sum_{j=1}^{d_k}p_{k,j} = P_{\mathrm{Tx}}, \quad p_{k,j} = \left[\eta\mu_k - \frac{\sigma_{\mathrm{n}}^2}{C\boldsymbol{g}_{k,j}^{\mathrm{H}}\boldsymbol{H}_k\boldsymbol{t}_{k,j}\boldsymbol{t}_{k,j}^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{g}_{k,j}}\right]^+,\\
&\boldsymbol{t}_{k,j}^{\mathrm{H}}\boldsymbol{t}_{k,j} = 1, \quad \boldsymbol{g}_{k,j}^{\mathrm{H}}\boldsymbol{g}_{k,j} = 1, \forall j = 1,\ldots,d_k, \forall k,\\
&\boldsymbol{t}_{k,j} = \boldsymbol{S}_{\gamma(k,j),\mathrm{T}}^{\mathsf{T}}\boldsymbol{S}_{\gamma(k,j),\mathrm{T}}\boldsymbol{t}_{k,j}, \quad \boldsymbol{g}_{k,j} = \boldsymbol{S}_{\gamma(k,j),k}^{\mathsf{T}}\boldsymbol{S}_{\gamma(k,j),k}\boldsymbol{g}_{k,j}, \forall j = 1,\ldots,d_k, \forall k,\\
&\boldsymbol{g}_{k,m}^{\mathrm{H}}\boldsymbol{g}_{k,j} = 0, \forall j = 1,\ldots,d_k, m = 1,\ldots,d_k, j \neq m, \forall k,\\
&\sqrt{p_{k,j}p_{m,\ell}}\boldsymbol{g}_{m,\ell}^{\mathrm{H}}\boldsymbol{H}_m\boldsymbol{t}_{k,j} = 0, \forall (m,\ell) \text{ with } \gamma(m,\ell) = \gamma(k,j) \text{ and } \pi(m,\ell) < \pi(k,j), \forall (k,j).
\end{aligned} \tag{3.26}$$

3.2.3 Successive Resource Allocation

Problem (3.26) is still non-convex and furthermore combinatorial. That is because for its solution one has to test all possible encodings $\pi(k,j)$ and all possible carrier allocations $\gamma(k,j)$ for the maximum weighted sum rate, which becomes infeasible already with a moderate number of users and carriers. For this reason a successive approach will be pursued in the following. This implies that for initialization all users have zero data streams, i.e., $d_k = 0, \forall k$. At first, a data stream is allocated to that user that can achieve the largest weighted single data stream rate. Consequently, the user $k(1)$ to which the first data stream is allocated, the corresponding transmit and receive filters $\boldsymbol{t}_{k(1),1}$ and $\boldsymbol{g}_{k(1),1}$, and the carrier $\gamma(k(1),1)$ on which this data stream is transmitted are determined according to

$$\begin{aligned}
\{\boldsymbol{t}_{k(1),1}, \boldsymbol{g}_{k(1),1}, \gamma(k(1),1), k(1)\} &= \operatorname*{argmax}_{\boldsymbol{t},\boldsymbol{g},c,k}\mu_k\log_2(1+P_{\mathrm{Tx}}\boldsymbol{g}^{\mathrm{H}}\boldsymbol{H}_k\boldsymbol{t}\boldsymbol{t}^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{g}),\\
\text{s.t. } \boldsymbol{t}^{\mathrm{H}}\boldsymbol{t} = 1, \quad \boldsymbol{g}^{\mathrm{H}}\boldsymbol{g} = 1, &\quad \boldsymbol{t} = \boldsymbol{S}_{c,\mathrm{T}}^{\mathsf{T}}\boldsymbol{S}_{c,\mathrm{T}}\boldsymbol{t}, \quad \boldsymbol{g} = \boldsymbol{S}_{c,k}^{\mathsf{T}}\boldsymbol{S}_{c,k}\boldsymbol{g}.
\end{aligned} \tag{3.27}$$

The maximum is achieved by aligning $\boldsymbol{g}_{k(1),1}$ and $\boldsymbol{H}_{k(1)}^{\mathrm{H}}\boldsymbol{t}_{k(1),1}$ so that

$$\boldsymbol{g}_{k(1),1} = \boldsymbol{H}_{k(1)}\boldsymbol{t}_{k(1),1}\frac{1}{\sqrt{\boldsymbol{t}_{k(1),1}^{\mathrm{H}}\boldsymbol{H}_{k(1)}^{\mathrm{H}}\boldsymbol{H}_{k(1)}\boldsymbol{t}_{k(1),1}}}$$

and Problem (3.27) reads as

$$\{\boldsymbol{t}_{k(1),1}, \gamma(k(1),1), k(1)\} = \operatorname*{argmax}_{\boldsymbol{t},c,k}\mu_k\log_2(1+P_{\mathrm{Tx}}\boldsymbol{t}^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{H}_k\boldsymbol{t}), \quad \text{s.t. } \boldsymbol{t}^{\mathrm{H}}\boldsymbol{t} = 1, \quad \boldsymbol{t} = \boldsymbol{S}_{c,\mathrm{T}}^{\mathsf{T}}\boldsymbol{S}_{c,\mathrm{T}}\boldsymbol{t}.$$

Thus, $\boldsymbol{t}_{k(1),1}$ is the unit norm eigenvector corresponding to the principal eigenvalue of the matrix $\boldsymbol{H}_{k(1)}^{\mathrm{H}}\boldsymbol{H}_{k(1)}$, which implicitly fulfills the carrier separation constraint (3.24) due to the block-diagonal structure of the channel matrices. The carrier $\gamma(k(1),1)$ is therefore determined implicitly by the block matrix within $\boldsymbol{H}_{k(1)}^{\mathrm{H}}\boldsymbol{H}_{k(1)}$, which exhibits the maximum eigenvalue amongst all block matrices. After solving the optimization, the variable d_k needs to be updated to $d_k = 1$ and $\pi(k(1),1)$ is given by $\pi(k(1),1) = 1$. Additionally, for the algorithm auxiliary variables n_c, $c = 1,\ldots,C$ will be required in the following, where n_c denotes the number of data streams allocated to carrier c and which are all initialized with zero except $n_{\gamma(k(1),1)}$, which is given by $n_{\gamma(k(1),1)} = 1$. For the future allocation steps, the carrier allocation $\gamma(k(1),1)$ the encoding position $\pi(k(1),1)$ and the filters $\boldsymbol{t}_{k(1),1}$ and $\boldsymbol{g}_{k(1),1}$ are kept fixed. The user which receives a data stream in the second step is then determined so that the weighted sum rate becomes maximum given $\gamma(k(1),1)$, $\pi(k(1),1)$, $\boldsymbol{t}_{k(1),1}$ and $\boldsymbol{g}_{k(1),1}$ of the first step. Continuing this way and keeping the carrier allocation, encoding position, transmit and receive filters fixed after each step, the optimization problem in the ith step reads as

$$
\begin{aligned}
&\{\boldsymbol{t}_{k(i),d_{k(i)}+1},\boldsymbol{g}_{k(i),d_{k(i)}+1},\gamma(k(i),d_{k(i)}+1),k(i)\} = \underset{\boldsymbol{t},\boldsymbol{g},c,k}{\mathrm{argmax}} \\
&\sum_{k'=1}^{K}\mu_{k'}\sum_{j=1}^{d_{k'}}\log_2\left(1+\frac{C}{\sigma_{\mathrm{n}}^2}p_{k',j}\left|\boldsymbol{g}_{k',j}^{\mathrm{H}}\boldsymbol{H}_{k'}\boldsymbol{t}_{k',j}\right|_2^2\right)+\mu_k\log_2\left(1+\frac{C}{\sigma_{\mathrm{n}}^2}p_{k,d_k+1}\left|\boldsymbol{g}^{\mathrm{H}}\boldsymbol{H}_k\boldsymbol{t}\right|_2^2\right), \\
&\text{s.t. } \boldsymbol{t}^{\mathrm{H}}\boldsymbol{t}=1,\quad \boldsymbol{t}=\boldsymbol{S}_{c,\mathrm{T}}^{\mathsf{T}}\boldsymbol{S}_{c,\mathrm{T}}\boldsymbol{t},\quad \boldsymbol{g}^{\mathrm{H}}\boldsymbol{g}=1,\quad \boldsymbol{g}=\boldsymbol{S}_{c,k}^{\mathsf{T}}\boldsymbol{S}_{c,k}\boldsymbol{g}, \\
&p_{k',j}=\left[\eta\mu_{k'}-\frac{\sigma_{\mathrm{n}}^2}{C\left|\boldsymbol{g}_{k',j}^{\mathrm{H}}\boldsymbol{H}_{k'}\boldsymbol{t}_{k',j}\right|_2^2}\right]^+,\quad p_{k,d_k+1}=\left[\eta\mu_k-\frac{\sigma_{\mathrm{n}}^2}{C\left|\boldsymbol{g}^{\mathrm{H}}\boldsymbol{H}_k\boldsymbol{t}\right|_2^2}\right]^+, \\
&\sum_{k'=1}^{K}\sum_{j=1}^{d_{k'}}p_{k',j}+p_{k,d_k+1}=P_{\mathrm{Tx}},\quad \sqrt{p_{k,d_k+1}p_{m,\ell}}\boldsymbol{g}_{m,\ell}^{\mathrm{H}}\boldsymbol{H}_m\boldsymbol{t}=0,\quad \forall(m,\ell)\text{ with }\gamma(m,\ell)=c,
\end{aligned}
\tag{3.28}
$$

Note all other zero-forcing constraints in (3.26) are automatically fulfilled by fixing the transmit precoders and receive filters from previous steps. That is because the vector $\boldsymbol{t}_{k,j}$ must be orthogonal to all vectors $\boldsymbol{H}_m^{\mathrm{H}}\boldsymbol{g}_{m,\ell}$ with $\pi(m,\ell) < \pi(k,j)$ and $\gamma(m,\ell) = \gamma(k,j)$, when $p_{k,d_k+1}p_{m,\ell} \neq 0$. As in each step of the successive allocation the new data stream is always encoded last on the chosen carrier, the vectors $\boldsymbol{g}_{m,\ell}$ with $\pi(m,\ell) < \pi(k,j)$ and $\gamma(m,\ell) = \gamma(k,j)$ have been determined in steps previous to the step in which the vector $\boldsymbol{t}_{k,j}$ is computed. This implies that zero-forcing constraints already considered in previous steps remain valid and are not affected by the current allocation.

As for the first allocated data stream, it is optimum in Problem (3.28) to align the receive filter of the new data stream to the corresponding transmit filter multiplied with the channel matrix so that

$$
\boldsymbol{g}_{k(i),d_{k(i)}+1}=\boldsymbol{H}_{k(i)}\boldsymbol{t}_{k(i),d_{k(i)}+1}\frac{1}{\sqrt{\boldsymbol{t}_{k(i),d_{k(i)}+1}^{\mathrm{H}}\boldsymbol{H}_{k(i)}^{\mathrm{H}}\boldsymbol{H}_{k(i)}\boldsymbol{t}_{k(i),d_{k(i)}+1}}}. \tag{3.29}
$$

That is because such an alignment maximizes the channel gain $\left|\boldsymbol{g}_{k(i),d_{k(i)}+1}\boldsymbol{H}_{k(i)}\boldsymbol{t}_{k(i),d_{k(i)}+1}\right|_2^2$ and the constraints $\boldsymbol{g}_{k(i),\ell}^{\mathrm{H}}\boldsymbol{g}_{k(i),d_{k(i)}+1} = 0$ are already considered by the zero-forcing constraints on $\boldsymbol{t}_{k(i),d_{k(i)}+1}$, as

$$
\boldsymbol{g}_{k,\ell}^{\mathrm{H}}\boldsymbol{H}_k\boldsymbol{t}_{k(i),d_{k(i)}+1}=\boldsymbol{g}_{k,\ell}^{\mathrm{H}}\boldsymbol{g}_{k(i),d_{k(i)}+1}\sqrt{\boldsymbol{t}_{k(i),d_{k(i)}+1}^{\mathrm{H}}\boldsymbol{H}_{k(i)}^{\mathrm{H}}\boldsymbol{H}_{k(i)}\boldsymbol{t}_{k(i),d_{k(i)}+1}}=0.
$$

Additionally, the carrier separation constraints on the receive filters are fulfilled, as long as the transmit filters obey to the carrier separation constraints. Let $\boldsymbol{t}_k^{(i)}$ and $\boldsymbol{g}_k^{(i)}$ be the transmit and receive vector that maximize (3.28) assuming that the next data stream is allocated to user k. For a given user allocation the optimum weighted sum rate is found by maximizing the channel gain

$$\left|\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{H}_k\boldsymbol{t}_k^{(i)}\right|_2^2 = \boldsymbol{t}_k^{(i),\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{H}_k\boldsymbol{t}_k^{(i)}$$

so that

$$\begin{aligned}\left\{\boldsymbol{t}_k^{(i)}, c_k^{(i)}\right\} = \underset{\boldsymbol{t},c}{\operatorname{argmax}}\, \boldsymbol{t}^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{H}_k\boldsymbol{t}, \quad \text{s.t. } \boldsymbol{t}^{\mathrm{H}}\boldsymbol{t} = 1, \quad \boldsymbol{t} = \boldsymbol{S}_{c,\mathrm{T}}^{\mathsf{T}}\boldsymbol{S}_{c,\mathrm{T}}\boldsymbol{t},\\ \boldsymbol{t}_{m,\ell}^{\mathrm{H}}\boldsymbol{H}_m^{\mathrm{H}}\boldsymbol{H}_m\boldsymbol{t} = 0, \quad \forall(m,\ell) \text{ with } \gamma(m,\ell) = c,\end{aligned} \tag{3.30}$$

where $c_k^{(i)}$ is the carrier to which a data stream will be allocated if the next data stream is allocated to user k and it is initially assumed that $p_{m,\ell}p_{k,d_k+1} \neq 0$. Defining the projection matrix $\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}$ that projects into

$$\operatorname{null}\left\{\boldsymbol{t}_{1,1}^{\mathrm{H}}\boldsymbol{H}_1^{\mathrm{H}}\boldsymbol{H}_1, \ldots, \boldsymbol{t}_{1,d_1}^{\mathrm{H}}\boldsymbol{H}_1^{\mathrm{H}}\boldsymbol{H}_1, \ldots, \boldsymbol{t}_{K,1}^{\mathrm{H}}\boldsymbol{H}_K^{\mathrm{H}}\boldsymbol{H}_K, \ldots, \boldsymbol{t}_{K,d_K}^{\mathrm{H}}\boldsymbol{H}_K^{\mathrm{H}}\boldsymbol{H}_K\right\}$$

the zero-forcing constraints can be inserted into the objective function of Problem (3.30) so that

$$\left\{\boldsymbol{t}_k^{(i)}, c_k^{(i)}\right\} = \underset{\boldsymbol{t},c}{\operatorname{argmax}}\, \boldsymbol{t}^{\mathrm{H}}\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}\boldsymbol{t}, \quad \text{s.t. } \boldsymbol{t}^{\mathrm{H}}\boldsymbol{t} = 1, \quad \boldsymbol{t} = \boldsymbol{S}_{c,\mathrm{T}}^{\mathsf{T}}\boldsymbol{S}_{c,\mathrm{T}}\boldsymbol{t}. \tag{3.31}$$

The maximum is achieved by choosing $\boldsymbol{t}_k^{(i)}$ to be the unit-norm eigenvector belonging to the principal eigenvalue of the matrix $\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}$. The projector $\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}$ exhibits like the Gramian channel matrices $\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{H}_k$ a block-diagonal structure, as all vectors $\boldsymbol{t}_{m,\ell}$ obey to the carrier separation constraint (3.24). Therefore all eigenvectors of the matrices $\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}$ also fulfill the carrier separation constraint in (3.24). The optimum carrier $c_k^{(i)}$ is, as in the first step, given implicitly by the position of this block of the matrix $\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}$ which exhibits the strongest principal eigenvalue. Note that the determination of the transmit filters according to (3.31) facilitates the computation of the projection matrices $\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i+1)}$ for the next step, which can be computed according to

$$\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i+1)} = \hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)} - \frac{\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}\boldsymbol{H}_{k(i)}^{\mathrm{H}}\boldsymbol{H}_{k(i)}\boldsymbol{t}_{k(i)}^{(i)}\boldsymbol{t}_{k(i)}^{(i),\mathrm{H}}\boldsymbol{H}_{k(i)}^{\mathrm{H}}\boldsymbol{H}_{k(i)}\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}}{\boldsymbol{t}_{k(i)}^{(i),\mathrm{H}}\boldsymbol{H}_{k(i)}^{\mathrm{H}}\boldsymbol{H}_{k(i)}\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}\boldsymbol{H}_{k(i)}^{\mathrm{H}}\boldsymbol{H}_{k(i)}\boldsymbol{t}_{k(i)}^{(i)}}. \tag{3.32}$$

As $\boldsymbol{t}_{k(i),d_{k(i)}+1} = \boldsymbol{t}_{k(i)}^{(i)}$ is the eigenvector corresponding to the principal eigenvalue of the matrix $\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}\boldsymbol{H}_{k(i)}^{\mathrm{H}}\boldsymbol{H}_{k(i)}\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}$, it does certainly not lie in the nullspace of the projector $\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}$, i.e.,

$$\boldsymbol{t}_{k(i)}^{(i)} = \hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}\boldsymbol{t}_{k(i)}^{(i)}, \tag{3.33}$$

and

$$\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}\boldsymbol{H}_{k(i)}^{\mathrm{H}}\boldsymbol{H}_{k(i)}\boldsymbol{t}_{k(i)}^{(i)} = \hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}\boldsymbol{H}_{k(i)}^{\mathrm{H}}\boldsymbol{H}_{k(i)}\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}\boldsymbol{t}_{k(i)}^{(i)} = \rho_1\left(\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}\boldsymbol{H}_{k(i)}^{\mathrm{H}}\boldsymbol{H}_{k(i)}\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}\right)\boldsymbol{t}_{k(i)}^{(i)}.$$

Using these properties in (3.32), the projector $\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i+1)}$ simply computes according to

$$\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i+1)} = \hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)} - \boldsymbol{t}_{k(i)}^{(i)}\boldsymbol{t}_{k(i)}^{(i),\mathrm{H}}. \tag{3.34}$$

Consequently, all transmit vectors $\boldsymbol{t}_{k,j}$ are orthogonal to each other. Another consequence of (3.33) is that the vectors $\boldsymbol{t}_k^{(i)}$ do not have to computed explicitly and it suffices to determine the principal eigenvalues $\rho_1\left(\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\right)$. That is because the weighted sum rate only depends on the terms

$$\left|\boldsymbol{g}_k^{(i),\text{H}}\boldsymbol{H}_k\boldsymbol{t}_k^{(i)}\right|^2 = \boldsymbol{t}_k^{(i),\text{H}}\boldsymbol{H}_k^{\text{H}}\boldsymbol{H}_k\boldsymbol{t}_k^{(i)} = \boldsymbol{t}_k^{(i),\text{H}}\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{t}_k^{(i)}$$

and, as $\boldsymbol{t}_k^{(i)}$ is the eigenvector corresponding to the principal eigenvalue of the matrix $\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}$,

$$\boldsymbol{t}_k^{(i),\text{H}}\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{t}_k^{(i)} = \rho_1\left(\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\right). \tag{3.35}$$

The user $k(i)$ can then be chosen according to

$$k(i) = \underset{k}{\operatorname{argmax}}\, R_{\text{WSR}}^{(i)}(k) \tag{3.36}$$

where

$$R_{\text{WSR}}^{(i)}(k) = \sum_{k'=1}^{K}\mu_{k'}\sum_{j=1}^{d_{k'}}\log_2\left(1+\frac{C}{\sigma_{\text{n}}^2}p_{k',j}\boldsymbol{t}_{k',j}^{\text{H}}\boldsymbol{H}_{k'}^{\text{H}}\boldsymbol{H}_{k'}\boldsymbol{t}_{k',j}\right)+$$
$$+\mu_k\log_2\left(1+\frac{C}{\sigma_{\text{n}}^2}p_{k,d_k+1}\rho_1\left(\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\right)\right),$$
$$p_{k',j} = \left[\eta\mu_{k'} - \frac{\sigma_{\text{n}}^2}{C\boldsymbol{t}_{k',j}^{\text{H}}\boldsymbol{H}_{k'}^{\text{H}}\boldsymbol{H}_{k'}\boldsymbol{t}_{k',j}}\right]^+, \quad p_{k,d_k+1} = \left[\eta\mu_k - \frac{\sigma_{\text{n}}^2}{C\rho_1\left(\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\right)}\right]^+,$$
$$\sum_{k'=1}^{K}\sum_{j=1}^{d_{k'}}p_{k',j} + p_{k,d_k+1} = P_{\text{Tx}}. \tag{3.37}$$

Although very unlikely, it can however happen that a previously assigned data stream receives zero power in the current allocation step, despite the fact that it has received a non-zero power in a previous step. Assume that in the ith step one obtains from (3.28) that $p_{k',j} = 0$ for some tuple (k', j). Due to the properties of the successive allocation such a case can only occur for users k' with $\mu_{k'} \neq \mu_{k(i)}$. Then all transmit filters $\boldsymbol{t}_{m,\ell}$ for data streams allocated to the same carrier $\gamma(k', c)$ after the jth data stream of user k' have been chosen in previous steps to fulfill the zero-forcing constraints

$$\boldsymbol{g}_{k',j}^{\text{H}}\boldsymbol{H}_{k'}\boldsymbol{t}_{m,\ell}\forall(m,\ell) \text{ with } \gamma(m,\ell) = \gamma(k',j) \text{ and } \pi(m,\ell) > \pi(k',j), \tag{3.38}$$

as in those steps $p_{k',j}p_{m,\ell} \neq 0$. In step i however, it is no longer necessary to fulfill (3.38), as now $p_{k',j} = 0$. A performance improvement can therefore be achieved by recomputing all transmit filters $\boldsymbol{t}_{m,\ell}$ and receive filters $\boldsymbol{g}_{m,\ell}$ with $\gamma(m,\ell) = \gamma(k',j)$ and $\pi(m,\ell) > \pi(k',j)$ not considering the zero-forcing constraints (3.38).

At the end of the ith step $\boldsymbol{t}_{k(i),d_{k(i)}+1}$ is given by $\boldsymbol{t}_{k(i),d_{k(i)}+1} = \boldsymbol{t}_{k(i)}^{(i)}$, the variables $\gamma(k(i), d_{k(i)}+1) = c_{k(i)}^{(i)}$ and $\pi(k(i), d_{k(i)}+1) = n_{c_{k(i)}^{(i)}} + 1$ are stored for future steps and $d_{k(i)}$ and $n_{\gamma(k(i),d_{k(i)}+1)}$ are incremented by one. For equal weights it is shown in [50] that if the last data stream receives zero power during testing, the corresponding user will not be selected in future steps and can therefore be excluded from the selection process without performance reductions. The algorithm terminates, if the data stream allocated last receives zero power, i.e., no improvements in weighted sum rate can be achieved any more, or at the latest, if no degrees of freedom are left to fulfill the zero-forcing constraints, which happens after $C\min(N_\mathrm{T}, \sum_{k=1}^{K} r_k)$ steps. The algorithm is summarized in Algorithm 3.2. For equal weights, i.e., $\mu_1 =, \ldots, = \mu_K$ the algorithm is identical to the Successive Encoding Successive Allocation Method (SESAM) proposed in [10]. In this case the user selection in (3.36) boils down to selecting the user, which matrix $\hat{\boldsymbol{P}}_\mathrm{DPC}^{(i)} \boldsymbol{H}_k^\mathrm{H} \boldsymbol{H}_k \hat{\boldsymbol{P}}_\mathrm{DPC}^{(i)}$ exhibits the strongest principal eigenvalue.

The method proposed in [49, Chapter 8] also works with zero-forcing, but uses a different successive allocation scheme. It considers the dual problem of the weighted sum rate maximization, where the sum power constraint is dualized. The algorithm iteratively searches for the optimum Lagrange multiplier corresponding to the power constraint by bisection, where in each step the user allocation, transmit and receive filters are determined in a successive manner so that the dual function becomes optimum. That leads to the same rules for the transmit and receive filters as (3.31) and (3.29), respectively. Thus, one iteration of this algorithm is as complex as Algorithm 3.2 in total.

3.3 Spatial Zero-Forcing and Successive Resource Allocation without Dirty Paper Coding

The algorithm presented in the previous section still relies on DPC and numerically complex methods such as vector precoding [16] or the coding scheme from [17] must be used as practical implementation for DPC. The latter scheme also exhibits long encoding delays. Tomlinson-Harashima Precoding (THP) [51, 52] is a less complex implementation of DPC, but associated with performance losses. The reasons for those degradations are explained in [53]. Nevertheless, THP still exhibits practical challenges, such as the implementation of modulo operators at all receivers due to the dynamics of the received signals. Therefore, an algorithm for weighted sum rate maximization that solely relies on linear transmit and signal processing will be proposed in this section. Besides the fact that the algorithm does not rely on DPC, it is non-iterative and therefore exhibits a low computational complexity.

3.3.1 State-of-the-Art Near Optimum Approaches

When part of the multiuser interference cannot be cancelled by DPC, the problem of maximizing the weighted sum rate is non-convex. So far, there exists no algorithm that solves this problem optimally. Only for the special case of two users and single-antenna receivers a method to achieve points on the boundary of the rate region achievable with linear precoding is presented in [54]. An algorithm that converges to a local optimum in the vicinity of the initial starting point is proposed in [55]. As the algorithm requires in each iteration uplink downlink conversions of the filters based on the duality from [56] and a geometric program must be solved in each step, it is extremely complex. The method in [57] is also iterative and relies on the solution of geometric programs. Besides aiming at maximizing the weighted sum rate, [57] considers the problem of feedback reduction for this optimization. Projected gradient methods are used in [58] and [59] for near optimum solutions to the weighted sum rate maximization with linear precoding. For the problem sum rate maxi-

Algorithm 3.2 Reduced Complexity Algorithm for Weighted Sum Rate Maximization with DPC

Initialization: $d_k = 0, \boldsymbol{T}_k = [], \boldsymbol{G}_k = [], \forall k = 1, \ldots, K, \quad n_c = 0, \forall c = 1, \ldots, C$

$\hat{\boldsymbol{P}}_{\text{DPC}}^{(1)} = \boldsymbol{I}_{CN_{\text{T}}}$

$i = 1$

while $i \leq C \min(\sum_{k=1}^{K} r_k, N_{\text{T}})$ **do**

  for $k = 1$ to K **do**

$$R_{\text{WSR}}^{(i)}(k) = \sum_{k'=1}^{K} \mu_{k'} \sum_{j=1}^{d_{k'}} \log_2 \left(1 + \frac{C}{\sigma_{\text{n}}^2} p_{k',j} \boldsymbol{t}_{k',j}^{\text{H}} \boldsymbol{H}_{k'}^{\text{H}} \boldsymbol{H}_{k'} \boldsymbol{t}_{k',j}\right) + \\ + \mu_k \log_2 \left(1 + \frac{C}{\sigma_{\text{n}}^2} p_{k,d_k+1} \rho_1 \left(\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)} \boldsymbol{H}_k^{\text{H}} \boldsymbol{H}_k \hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\right)\right),$$

$$p_{k',j} = \left[\eta \mu_{k'} - \frac{\sigma_{\text{n}}^2}{C \boldsymbol{t}_{k',j}^{\text{H}} \boldsymbol{H}_{k'}^{\text{H}} \boldsymbol{H}_{k'} \boldsymbol{t}_{k',j}}\right]^+, \quad p_{k,d_k+1} = \left[\eta \mu_k - \frac{\sigma_{\text{n}}^2}{C \rho_1 \left(\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)} \boldsymbol{H}_k^{\text{H}} \boldsymbol{H}_k \hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\right)}\right]^+,$$

$$\sum_{k'=1}^{K} \sum_{j=1}^{d_{k'}} p_{k',j} + p_{k,d_k+1} = P_{\text{Tx}}$$

    $c_k^{(i)}$: index of block within $\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)} \boldsymbol{H}_k^{\text{H}} \boldsymbol{H}_k \hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}$ that contains principal eigenvalue

  end for

  $k(i) = \operatorname*{argmax}_k R_{\text{WSR}}^{(i)}(k)$

  Remove data streams that have received zero power and recompute $R_{\text{WSR}}^{(i)}$, $\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}$ transmit and receive filters if necessary

  $\boldsymbol{t}_{k(i),d_{k(i)}+1} = \operatorname*{argmax}_{\boldsymbol{t}} \boldsymbol{t}^{\text{H}} \hat{\boldsymbol{P}}_{\text{DPC}}^{(i)} \boldsymbol{H}_{k(i)}^{\text{H}} \boldsymbol{H}_{k(i)} \hat{\boldsymbol{P}}_{\text{DPC}}^{(i)} \boldsymbol{t}, \quad$ s.t. $\boldsymbol{t}^{\text{H}} \boldsymbol{t} = 1, \quad \boldsymbol{t} = \boldsymbol{S}_{c_{k(i)}^{(i)},\text{T}}^{\text{T}} \boldsymbol{S}_{c_{k(i)}^{(i)},\text{T}} \boldsymbol{t}$

  $\boldsymbol{g}_{k(i),d_{k(i)}+1} = \boldsymbol{H}_{k(i)} \boldsymbol{t}_{k(i),d_{k(i)}+1} \frac{1}{\sqrt{\boldsymbol{t}_{k(i),d_{k(i)}+1}^{\text{H}} \boldsymbol{H}_{k(i)}^{\text{H}} \boldsymbol{H}_{k(i)} \boldsymbol{t}_{k(i),d_{k(i)}+1}}}$

  $\boldsymbol{T}_{k(i)} = [\boldsymbol{T}_{k(i)}, \boldsymbol{t}_{k(i),d_{k(i)}+1}]$, $\boldsymbol{G}_{k(i)} = [\boldsymbol{G}_{k(i)}, \boldsymbol{g}_{k(i),d_{k(i)}+1}]$

  $d_{k(i)} = d_{k(i)} + 1$, $\gamma(k(i), d_{k(i)}) = c_{k(i)}^{(i)}$, $\pi(k(i), d_{k(i)}) = n_{c_{k(i)}^{(i)}} + 1$, $n_{c_{k(i)}^{(i)}} = n_{c_{k(i)}^{(i)}} + 1$

  $\hat{\boldsymbol{P}}_{\text{DPC}}^{(i+1)} = \hat{\boldsymbol{P}}_{\text{DPC}}^{(i)} - \boldsymbol{t}_{k(i),d_{k(i)}} \boldsymbol{t}_{k(i),d_{k(i)}}^{\text{H}}$

  $i = i + 1$

end while

$p_{1,1}, \ldots, p_{K,d_K} \leftarrow$ waterfilling with channel gains $\boldsymbol{t}_{1,1}^{\text{H}} \boldsymbol{H}_1^{\text{H}} \boldsymbol{H}_1 \boldsymbol{t}_{1,1}, \ldots, \boldsymbol{t}_{K,d_K}^{\text{H}} \boldsymbol{H}_K^{\text{H}} \boldsymbol{H}_K \boldsymbol{t}_{K,d_K}$

for $k = 1$ to K **do**

  $\boldsymbol{P}_k = \operatorname{diag}(p_{k,1}, \ldots, p_{k,d_k})$

end for

mization it is proposed in [45] to determine first the covariance matrices according to the iterative water-filling algorithm optimum for the DPC case from [6] and to project these matrices afterwards so the part of the interference occurring through the absence of DPC is cancelled. Algorithms for selecting the appropriate subgroup of users for this method are proposed in [60]. In order to ease the objective function and the power allocation, zero-forcing constraints can be introduced. In case the total number of receive antennas in the system is lower than the number of transmit antennas, Block-Diagonalization (BD) as proposed in [61] can be used to decompose the MIMO broadcast channel into a system of scalar interference-free subchannels, over which the transmit power can be distributed by a weighted water-filling algorithm. However, in practical systems it is more likely that the total number of receive antennas exceeds the number of transmit antennas and there are not enough degrees of freedom to serve all users simultaneously. Therefore user selection algorithms are proposed in [62] and [63] when the weights are equal for all users.

By introducing zero-forcing constraints optimization problems in the MIMO broadcast channel become combinatorial, as one has to test all possible combinations of data stream allocations to users, where the total number of data streams must not exceed the number of transmit antennas. Even in case the total number of receive antennas is smaller than the number of transmit antennas it is not necessarily optimum to allocate as many data streams to each user as it has receive antennas, as proposed with BD in [61]. To avoid an exhaustive search, which becomes infeasible already with a moderate number of users, greedy approaches can be used, where in each step a data stream is allocated to a user so that the increase in the objective function becomes maximum. For MISO systems and sum rate as objective function such an approach has been proposed in [12], for which a low complexity implementation is presented in [64]. The complexity can be further diminished by excluding users from the allocation process which channel vectors are aligned closely to those of the already selected users as proposed in [65]. The decision whether two channel vectors are aligned closely is thereby based on an a priori fixed threshold value. Selecting channel vectors as little aligned as possible by a method from graph theory is proposed in [66]. In [67] a greedy maximization of weighted sum rate in MISO systems is considered where a lower bound for the weighted sum rate is utilized during user selection. One possibility to apply these algorithms to systems with multiple antennas also at the receivers is to perform SVDs at each receiver, apply the left singular vectors as receive filters and treat every product of right singular vector and the corresponding singular value as virtual user in a MISO system as proposed in [65] for the algorithm presented therein. An analysis for this algorithm in MIMO systems with a large number of users extending the results from [65] can be found in [68]. The concept of SVD receivers is also used in [69] and [70], where additionally the aspect of feedback reduction is considered so that each user only has to make known a singular vector to the transmitter if the corresponding singular value is above a certain threshold. In [71] the greedy approach from [12] is extended to MIMO systems by applying SVD receivers and modifying these receivers in case more than one data stream is allocated to the same user. Another possibility for applying the greedy algorithms proposed for MISO systems to MIMO is antenna selection, i.e., every receive antenna is treated as different virtual user, as considered in [72] for sum rate maximization with equal weights. The general concept of an algorithm of this kind for an arbitrary objective function and an efficient computation of the precoding vectors is described in [73]. A more sophisticated antenna selection for a preselected group of users, which have in sum less receive antennas than number of transmit antennas, is presented [74].

Thus, most of the prior works dealing with greedy zero-forcing approaches assume that the receive filters are given a priori by singular or canonical unit vectors and are, apart from [71], not changed during the execution of the algorithms. In [75] the sum rate is approximated in each step so that the receive filters can be included into the successive optimization. However, besides the fact that this approximation becomes more and more inaccurate with increasing SNR, the receive filter determination is reduced to a selection between the left singular vectors for complexity reasons at the end. The greedy algorithm from [23] chooses in each step the receive filters to maximize the SNR of the newly allocated subchannel ignoring the effect of this filter on the SNRs of previously allocated subchannels. Furthermore the algorithm is proposed for sum rate maximization only and a fixed power allocation is assumed complicating its straight-forward application to weighted sum rate maximization. For this reason, in this section an algorithm will be presented that is based on greedy approach and zero-forcing but includes the receive filters into the successive weighted sum rate maximization.

3.3.2 Spatial Zero-Forcing

When no DPC is used, the sets $\mathcal{I}_k$ read as

$$\mathcal{I}_k = \{j \in \{1, \dots, K\} | j \neq k\},$$

i.e., all other users interfere with user k. With those sets of interfering users, the same derivations can be conducted as in Section 3.2.2 for the DPC case, i.e., inter- and intra-user zero-forcing constraints are introduced, where the latter do not influence optimality, leading to the stream-wise formulation of the optimization problem in (3.22), the optimum power allocation is given by (3.23), and the carrier separation constraints are established as in (3.24). Through the abstinence of DPC, all signals can be encoded simultaneously, i.e., there is no optimum encoding order, and interference from all data streams on the same carrier has to be considered with the zero-forcing constraints. Considering these changes compared to Problem (3.26), the optimization for spatial zero-forcing without DPC is given by

$$\begin{aligned}
&\max_{\{\boldsymbol{t}_{k,j},\boldsymbol{g}_{k,j},\gamma(k,j)\}_{j=1,\dots,d_k,k=1,\dots,K}} \sum_{k=1}^{K} \mu_k \sum_{j=1}^{d_k} \log_2 \left(1 + \frac{C}{\sigma_{\mathrm{n}}^2} p_{k,j} \boldsymbol{g}_{k,j}^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{t}_{k,j} \boldsymbol{t}_{k,j}^{\mathrm{H}} \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{g}_{k,j}\right), \\
&\text{s.t.} \sum_{k=1}^{K}\sum_{j=1}^{d_k} p_{k,j} = P_{\mathrm{Tx}}, \quad p_{k,j} = \left[\eta\mu_k - \frac{\sigma_{\mathrm{n}}^2}{C\boldsymbol{g}_{k,j}^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{t}_{k,j} \boldsymbol{t}_{k,j}^{\mathrm{H}} \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{g}_{k,j}}\right]^+, \\
&\boldsymbol{t}_{k,j}^{\mathrm{H}}\boldsymbol{t}_{k,j} = 1, \quad \boldsymbol{g}_{k,j}^{\mathrm{H}}\boldsymbol{g}_{k,j} = 1, \quad \boldsymbol{t}_{k,j} = \boldsymbol{S}_{c,\mathrm{T}}^{\mathsf{T}}\boldsymbol{S}_{c,\mathrm{T}}\boldsymbol{t}_{k,j}, \quad \boldsymbol{g}_{k,j} = \boldsymbol{S}_{c,k}^{\mathsf{T}}\boldsymbol{S}_{c,k}\boldsymbol{g}_{k,j}, \quad \forall j = 1, \dots, d_k, \forall k, \\
&\boldsymbol{g}_{k,n}^{\mathrm{H}}\boldsymbol{g}_{k,j} = 0, \forall n = 1, \dots, d_k, j = 1, \dots, d_k, j \neq n, \forall k, \\
&\sqrt{p_{k,j}p_{m,\ell}}\,\boldsymbol{g}_{m,\ell}^{\mathrm{H}}\boldsymbol{H}_m\boldsymbol{t}_{k,j} = 0, \forall (m,\ell) \text{ with } \gamma(m,\ell) = \gamma(k,j), \forall (k,j) \neq (m,\ell).
\end{aligned} \tag{3.39}$$

For notational convenience it will be assumed in the following that all data streams receive non-zero powers, i.e., $p_{k,j} > 0, \forall j = 1, \dots, d_k, \forall k$. The case $p_{k,j} = 0$ for some tuple (k, j) will be revisited later and should be avoided anyway by an intelligent allocation. Then, the zero-forcing constraints read as

$$\boldsymbol{g}_{m,\ell}^{\mathrm{H}}\boldsymbol{H}_m\boldsymbol{t}_{k,j} = 0, \forall (m,\ell) \text{ with } \gamma(m,\ell) = \gamma(k,j), \forall (k,j) \neq (m,\ell)$$

and the water-level η is given by

$$\eta = \frac{P_{\mathrm{Tx}} + \sum_{m=1}^{K}\sum_{\ell=1}^{d_m} \frac{\sigma_{\mathrm{n}}^2}{C\boldsymbol{g}_{m,\ell}^{\mathrm{H}}\boldsymbol{H}_m\boldsymbol{t}_{m,\ell}\boldsymbol{t}_{m,\ell}^{\mathrm{H}}\boldsymbol{H}_m^{\mathrm{H}}\boldsymbol{g}_{m,\ell}}}{\sum_{k'=1}^{K}\mu_{k'}d_{k'}}.$$

By defining the composite channel matrix $\boldsymbol{H}_{\text{comp}}$ as

$$\boldsymbol{H}_{\text{comp}} = \begin{bmatrix} \boldsymbol{g}_{1,1}^{\text{H}} \boldsymbol{H}_1 \\ \vdots \\ \boldsymbol{g}_{1,d_1}^{\text{H}} \boldsymbol{H}_1 \\ \vdots \\ \boldsymbol{g}_{K,1}^{\text{H}} \boldsymbol{H}_K \\ \vdots \\ \boldsymbol{g}_{K,d_K}^{\text{H}} \boldsymbol{H}_K^{\text{H}} \end{bmatrix}$$

and introducing the channel gains

$$\lambda_{k,j} = \boldsymbol{g}_{k,j}^{\text{H}} \boldsymbol{H}_k \boldsymbol{t}_{k,j} \boldsymbol{t}_{k,j}^{\text{H}} \boldsymbol{H}_k^{\text{H}} \boldsymbol{g}_{k,j}$$

as new variables, the zero-forcing constraints in (3.39) can be written compactly as

$$\boldsymbol{H}_{\text{comp}} \left[\boldsymbol{t}_{1,1}, \ldots, \boldsymbol{t}_{K,d_K} \right] = \operatorname{diag} \left(\sqrt{\lambda_{1,1}}, \ldots, \sqrt{\lambda_{K,d_K}} \right)$$

so that Problem (3.39) reads as

$$\begin{aligned}
&\max_{\{\boldsymbol{g}_{k,j}, \gamma(k,j)\}_{j,=1,\ldots,d_k,k=1,\ldots,K}} \; \max_{\{\boldsymbol{t}_{k,j}, \lambda_{k,j}\}_{j,=1,\ldots,d_k,k=1,\ldots,K}} \sum_{k=1}^{K} \mu_k \sum_{j=1}^{d_k} \log_2 \left(\lambda_{k,j} \mu_k \frac{P_{\text{Tx}} \frac{C}{\sigma_{\text{n}}^2} + \sum\limits_{m=1}^{K} \sum\limits_{\ell=1}^{d_m} \frac{1}{\lambda_{m,\ell}}}{\sum_{k'=1}^{K} \mu_{k'} d_{k'}} \right), \\
&\text{s.t. } \boldsymbol{H}_{\text{comp}} \left[\boldsymbol{t}_{1,1}, \ldots, \boldsymbol{t}_{K,d_K} \right] = \operatorname{diag} \left(\sqrt{\lambda_{1,1}}, \ldots, \sqrt{\lambda_{K,d_K}} \right), \quad \lambda_{k,j} \geq 0, \quad \forall j = 1, \ldots, d_k, \forall k, \\
&\boldsymbol{t}_{k,j}^{\text{H}} \boldsymbol{t}_{k,j} = 1, \quad \boldsymbol{g}_{k,j}^{\text{H}} \boldsymbol{g}_{k,j} = 1, \quad \boldsymbol{t}_{k,j} = \boldsymbol{S}_{c,\text{T}}^{\mathsf{T}} \boldsymbol{S}_{c,\text{T}} \boldsymbol{t}_{k,j}, \quad \boldsymbol{g}_{k,j} = \boldsymbol{S}_{c,k}^{\mathsf{T}} \boldsymbol{S}_{c,k} \boldsymbol{g}_{k,j}, \quad \forall j = 1, \ldots, d_k, \forall k, \\
&\boldsymbol{g}_{k,m}^{\text{H}} \boldsymbol{g}_{k,j} = 0, \quad \forall j = 1, \ldots, d_k, m = 1, \ldots, d_k, j \neq m, \forall k, \\
&\boldsymbol{H}_{\text{comp}} = \left[\boldsymbol{H}_1^{\text{H}} \boldsymbol{g}_{1,1}, \ldots, \boldsymbol{H}_1^{\text{H}} \boldsymbol{g}_{1,d_1}, \ldots, \boldsymbol{H}_K^{\text{H}} \boldsymbol{g}_{K,1}, \ldots, \boldsymbol{H}_K^{\text{H}} \boldsymbol{g}_{K,d_K} \right]^{\text{H}}.
\end{aligned} \tag{3.40}$$

The maximization with respect to $\boldsymbol{t}_{k,j}$ and $\lambda_{k,j}$ is solved by

$$\lambda_{k,j} = \frac{1}{\boldsymbol{e}_{n_{k,j}}^{T} \left(\boldsymbol{H}_{\text{comp}} \boldsymbol{H}_{\text{comp}}^{\text{H}} \right)^{-1} \boldsymbol{e}_{n_{k,j}}}, \tag{3.41}$$

where

$$n_{k,j} = \sum_{k'=1}^{k-1} d_{k'} + j \tag{3.42}$$

denotes the row, in which the j th data stream of user k is placed in the composite channel matrix $\boldsymbol{H}_{\text{comp}}$, and

$$\begin{aligned}
\left[\boldsymbol{t}_{1,1}, \ldots, \boldsymbol{t}_{K,d_K} \right] =: \boldsymbol{T} &= \boldsymbol{H}_{\text{comp}}^{+} \operatorname{diag} \left(\sqrt{\lambda_{1,1}}, \ldots, \sqrt{\lambda_{K,d_K}} \right) \\
&= \boldsymbol{H}_{\text{comp}}^{\text{H}} \left(\boldsymbol{H}_{\text{comp}} \boldsymbol{H}_{\text{comp}}^{\text{H}} \right)^{-1} \operatorname{diag} \left(\sqrt{\lambda_{1,1}}, \ldots, \sqrt{\lambda_{K,d_K}} \right).
\end{aligned}$$

The proof can be found in Appendix A3. Due to (3.41) the sum $\sum_{m=1}^{K}\sum_{\ell=1}^{d_k}\frac{1}{\lambda_{m,\ell}}$ needed to compute the water-level η can be written compactly as

$$\sum_{m=1}^{K}\sum_{\ell=1}^{d_m}\frac{1}{\lambda_{m,\ell}} = \sum_{m=1}^{K}\sum_{\ell=1}^{d_m}\boldsymbol{e}_{n_{m,\ell}}^{T}\left(\boldsymbol{H}_{\text{comp}}\boldsymbol{H}_{\text{comp}}^{\text{H}}\right)^{-1}\boldsymbol{e}_{n_{m,\ell}} = \operatorname{tr}\left[\left(\boldsymbol{H}_{\text{comp}}\boldsymbol{H}_{\text{comp}}^{\text{H}}\right)^{-1}\right] =$$
$$= \operatorname{tr}\left[\left(\boldsymbol{H}_{\text{comp}}\boldsymbol{H}_{\text{comp}}^{\text{H}}\right)^{-1}\boldsymbol{H}_{\text{comp}}\boldsymbol{H}_{\text{comp}}^{\text{H}}\left(\boldsymbol{H}_{\text{comp}}\boldsymbol{H}_{\text{comp}}^{\text{H}}\right)^{-1}\right] = \left\|\boldsymbol{H}_{\text{comp}}^{+}\right\|_{\text{F}}^{2} \qquad (3.43)$$

and Problem (3.40) reads as

$$\max_{\{\boldsymbol{g}_{k,j},\gamma(k,j)\}_{j,=1,\ldots,d_k,k=1,\ldots,K}} \sum_{k=1}^{K}\mu_k\sum_{j=1}^{d_k}\log_2\left(\frac{\mu_k}{\boldsymbol{e}_{n_{k,j}}^{T}\left(\boldsymbol{H}_{\text{comp}}\boldsymbol{H}_{\text{comp}}^{\text{H}}\right)^{-1}\boldsymbol{e}_{n_{k,j}}}\frac{P_{\text{Tx}}\frac{C}{\sigma_{\text{n}}^{2}}+\left\|\boldsymbol{H}_{\text{comp}}^{+}\right\|_{\text{F}}^{2}}{\sum_{k'=1}^{K}\mu_{k'}d_{k'}}\right),$$
$$\text{s.t. } \boldsymbol{g}_{k,j} = \boldsymbol{S}_{c,k}^{\mathsf{T}}\boldsymbol{S}_{c,k}\boldsymbol{g}_{k,j}, \forall j = 1,\ldots,d_k, \forall k,$$
$$\boldsymbol{g}_{k,j}^{\text{H}}\boldsymbol{g}_{k,j} = 1, \quad \boldsymbol{g}_{k,m}^{\text{H}}\boldsymbol{g}_{k,j} = 0, \forall j = 1,\ldots,d_k, m = 1,\ldots,d_k, j \neq m, \forall k,$$
$$\boldsymbol{H}_{\text{comp}} = \left[\boldsymbol{H}_1^{\text{H}}\boldsymbol{g}_{1,1},\ldots,\boldsymbol{H}_1^{\text{H}}\boldsymbol{g}_{1,d_1},\ldots,\boldsymbol{H}_K^{\text{H}}\boldsymbol{g}_{K,1},\ldots,\boldsymbol{H}_K^{\text{H}}\boldsymbol{g}_{K,d_K}\right]^{\text{H}}. \qquad (3.44)$$

3.3.3 Successive Resource Allocation

Although with the zero-forcing constraints the transmit filters could be eliminated from the optimization, Problem (3.44) is still combinatorial and non-convex. For this purpose a successive resource allocation is proposed as in the DPC case. Therefore, each user receives zero data streams at the beginning, i.e., $d_k = 0$ for all users k. The user $k(1)$ to which the first data stream is allocated to, the corresponding receive filter $\boldsymbol{g}_{k(1),1}$ and the carrier $\gamma(k(1),1)$ are determined so that the weighted rate a user can achieve with single-stream transmission becomes maximum, i.e.,

$$\{\boldsymbol{g}_{k(1),1},\gamma(k(1),1),k(1)\} = \operatorname*{argmax}_{\boldsymbol{g},c,k}\mu_k\log_2\left(1+P_{\text{Tx}}\boldsymbol{g}^{\text{H}}\boldsymbol{H}_k\boldsymbol{H}_k^{\text{H}}\boldsymbol{g}\right)$$
$$\text{s.t. } \boldsymbol{g}^{\text{H}}\boldsymbol{g} = 1, \quad \boldsymbol{g} = \boldsymbol{S}_{c,k}^{\mathsf{T}}\boldsymbol{S}_{c,k}\boldsymbol{g}. \qquad (3.45)$$

Problem (3.45) is solved by choosing $\boldsymbol{g}_{k(1),1}$ as the unit norm eigenvector corresponding to the principal eigenvalue of the matrix $\boldsymbol{H}_{k(1)}\boldsymbol{H}_{k(1)}^{\text{H}}$. As with DPC, due to the block-diagonal structure of the matrices $\boldsymbol{H}_{k(1)}\boldsymbol{H}_{k(1)}^{\text{H}}$, this solution fulfills the carrier separation constraint and the carrier $\gamma(k(1),1)$ is implicitly given by the index of that block matrix within $\boldsymbol{H}_{k(1)}\boldsymbol{H}_{k(1)}^{\text{H}}$ that exhibits the strongest principal eigenvalue. The user $k(1)$ has to be found be evaluating the maximum single-stream rates for all users with the corresponding optimum receive filters. Note that for equal weights, $k(1)$ is the user with the maximum principal eigenvalue of the matrices $\boldsymbol{H}_k\boldsymbol{H}_k^{\text{H}}$. At the end of the first allocation, the variable $d_{k(1)}$ is updated to $d_{k(1)} = 1$. For the future allocation steps, the receive filter $\boldsymbol{g}_{k(1),1}$ and the carrier allocation $\gamma(k(1),1)$ are kept fixed. Proceeding similarly with the next allocation steps, the successive allocation problem in the ith step can be written as

$$\{\boldsymbol{g}_{k(i),d_{k(i)}+1}, \gamma(k(i), d_{k(i)}+1), k(i)\} =$$
$$= \underset{\boldsymbol{g},c,k}{\operatorname{argmax}} \sum_{k'=1}^{K} \mu_{k'} \sum_{j=1}^{d_{k'}} \log_2 \left(\frac{\mu_{k'} \left(P_{\text{Tx}} \frac{C}{\sigma_{\text{n}}^2} + \left\| \boldsymbol{H}_{\text{comp}}^{+} \right\|_{\text{F}}^2 \right)}{\boldsymbol{e}_{n_{k',j}}^T \left(\boldsymbol{H}_{\text{comp}} \boldsymbol{H}_{\text{comp}}^{\text{H}} \right)^{-1} \boldsymbol{e}_{n_{k',j}} \left(\sum_{m=1}^{K} \mu_m d_m + \mu_k \right)} \right) +$$
$$+ \mu_k \log_2 \left(\frac{\mu_k \left(P_{\text{Tx}} \frac{C}{\sigma_{\text{n}}^2} + \left\| \boldsymbol{H}_{\text{comp}}^{+} \right\|_{\text{F}}^2 \right)}{\boldsymbol{e}_i^T \left(\boldsymbol{H}_{\text{comp}} \boldsymbol{H}_{\text{comp}}^{\text{H}} \right)^{-1} \boldsymbol{e}_i \left(\sum_{m=1}^{K} \mu_m d_m + \mu_k \right)} \right)$$
$$\text{s.t. } \boldsymbol{g} = \boldsymbol{S}_{c,k}^{\mathsf{T}} \boldsymbol{S}_{c,k} \boldsymbol{g}, \quad \boldsymbol{g}^{\text{H}} \boldsymbol{g} = 1, \quad \boldsymbol{g}_{k,m}^{\text{H}} \boldsymbol{g} = 0, \forall m \text{ with } \gamma(k,m) = c, \quad \boldsymbol{H}_{\text{comp}} = \left[\boldsymbol{H}_{\text{comp}}^{(i-1),\text{H}} \boldsymbol{H}_k^{\text{H}} \boldsymbol{g} \right]^{\text{H}}, \tag{3.46}$$

where

$$\boldsymbol{H}_{\text{comp}}^{(i-1),\text{H}} = \left[\boldsymbol{H}_1^{\text{H}} \boldsymbol{g}_{1,1}, \ldots, \boldsymbol{H}_1^{\text{H}} \boldsymbol{g}_{1,d_1}, \ldots, \boldsymbol{H}_K^{\text{H}} \boldsymbol{g}_{K,1}, \ldots, \boldsymbol{H}_K^{\text{H}} \boldsymbol{g}_{K,d_K} \right] \tag{3.47}$$

denotes the composite channel matrix after step $i-1$. The objective function in (3.46) depends on the diagonal elements of the matrix $\left(\boldsymbol{H}_{\text{comp}} \boldsymbol{H}_{\text{comp}}^{\text{H}} \right)^{-1}$ [c.f. (3.43)] which are given by

$$\boldsymbol{e}_j^{\mathsf{T}} \left(\boldsymbol{H}_{\text{comp}} \boldsymbol{H}_{\text{comp}}^{\text{H}} \right)^{-1} \boldsymbol{e}_j = \begin{cases} \left(1 + \frac{\boldsymbol{g}^{\text{H}} \boldsymbol{H}_k \hat{\boldsymbol{P}}_j^{(i)} \boldsymbol{H}_k^{\text{H}} \boldsymbol{g}}{\boldsymbol{g}^{\text{H}} \boldsymbol{H}_k \hat{\boldsymbol{P}}_{\text{lin}}^{(i)} \boldsymbol{H}_k^{\text{H}} \boldsymbol{g}} \right) \alpha_j^{(i-1)} & j < i \\ \frac{1}{\boldsymbol{g}^{\text{H}} \boldsymbol{H}_k \hat{\boldsymbol{P}}_{\text{lin}}^{(i)} \boldsymbol{H}_k^{\text{H}} \boldsymbol{g}}, & j = i \end{cases} \tag{3.48}$$

as shown in Appendix A4.

$$\hat{\boldsymbol{P}}_j^{(i)} = \boldsymbol{T}^{(i-1)} \boldsymbol{e}_j \boldsymbol{e}_j^{\mathsf{T}} \boldsymbol{T}^{(i-1),\text{H}}$$

is a rank one projector, where

$$\boldsymbol{T}^{(i-1)} = \left[\boldsymbol{t}_{1,1}^{(i-1)}, \ldots, \boldsymbol{t}_{K,d_K}^{(i-1)} \right] = \left(\boldsymbol{H}_{\text{comp}}^{(i-1)} \right)^{+} \operatorname{diag} \left(\sqrt{\lambda_{1,1}^{(i-1)}}, \ldots, \sqrt{\lambda_{K,d_K}^{(i-1)}} \right).$$

contains the precoding vectors, if the allocation is stopped after step $i-1$. The matrix $\hat{\boldsymbol{P}}_{\text{lin}}^{(i)}$ is a projection matrix that projects into the nullspace of $\boldsymbol{H}_{\text{comp}}^{(i-1)}$ so that

$$\hat{\boldsymbol{P}}_{\text{lin}}^{(i)} = \boldsymbol{I}_{CN_{\text{T}}} - \boldsymbol{H}_{\text{comp}}^{(i-1),\text{H}} \left(\boldsymbol{H}_{\text{comp}}^{(i-1)} \boldsymbol{H}_{\text{comp}}^{(i-1),\text{H}} \right)^{-1} \boldsymbol{H}_{\text{comp}}^{(i-1)}. \tag{3.49}$$

$\hat{\boldsymbol{P}}_{\text{lin}}^{(i)}$ is block-diagonal and $\hat{\boldsymbol{P}}_{\text{lin}}^{(1)} = \boldsymbol{I}_{CN_{\text{T}}}$.

$$\alpha_j^{(i-1)} = \boldsymbol{e}_j^{\mathsf{T}} \left(\boldsymbol{H}_{\text{comp}}^{(i-1)} \boldsymbol{H}_{\text{comp}}^{(i-1),\text{H}} \right)^{-1} \boldsymbol{e}_j \tag{3.50}$$

is the inverse channel gain of the jth data stream in step $i-1$ so that

$$\lambda_{k,j}^{(i-1)} = \frac{1}{\alpha_{n_{k,j}}^{(i-1)}} = \frac{1}{\boldsymbol{e}_{n_{k,j}}^{\mathsf{T}} \left(\boldsymbol{H}_{\text{comp}}^{(i-1)} \boldsymbol{H}_{\text{comp}}^{(i-1),\text{H}} \right)^{-1} \boldsymbol{e}_{n_{k,j}}}.$$

Thus, it can be seen from (3.48) that with the additional zero-forcing constraints the channel gains of the subchannels allocated in previous steps diminish or stay equal, which occurs for every column in $\boldsymbol{T}^{(i)}$ which is orthogonal to $\boldsymbol{g}^{\mathrm{H}}\boldsymbol{H}_k$. With (3.48) the Frobenius norm of the pseudo-inverse of the composite channel matrix as defined in (3.46) computes according to

$$\left\|\boldsymbol{H}_{\mathrm{comp}}^{+}\right\|_{\mathrm{F}}^{2} = \operatorname{tr}\left[\left(\boldsymbol{H}_{\mathrm{comp}}\boldsymbol{H}_{\mathrm{comp}}^{\mathrm{H}}\right)^{-1}\right] = \frac{\sum_{j=1}^{i-1}\boldsymbol{g}^{\mathrm{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_j^{(i)}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{g}\alpha_j^{(i-1)}+1}{\boldsymbol{g}^{\mathrm{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\mathrm{lin}}^{(i)}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{g}} + \sum_{j=1}^{i-1}\alpha_j^{(i-1)} = \\ = \frac{\boldsymbol{g}^{\mathrm{H}}\boldsymbol{B}_k^{(i)}\boldsymbol{g}}{\boldsymbol{g}^{\mathrm{H}}\boldsymbol{A}_k^{(i)}\boldsymbol{g}} + \hat{\alpha}^{(i-1)}, \quad (3.51)$$

where the identity

$$\sum_{k'=1}^{K}\sum_{j=1}^{d_{k'}}\boldsymbol{g}^{\mathrm{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{n_{k',j}}^{(i)}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{g}\alpha_{n_{k,j}}^{(i-1)} = \sum_{j=1}^{i-1}\boldsymbol{g}^{\mathrm{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_j^{(i)}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{g}\alpha_j^{(i-1)}$$

and the definitions

$$\boldsymbol{A}_k^{(i)} := \boldsymbol{H}_k\hat{\boldsymbol{P}}_{\mathrm{lin}}^{(i)}\boldsymbol{H}_k^{\mathrm{H}}, \qquad \boldsymbol{B}_k^{(i)} := \sum_{j=1}^{i-1}\boldsymbol{H}_k\hat{\boldsymbol{P}}_j^{(i)}\boldsymbol{H}_k^{\mathrm{H}}\alpha_j^{(i-1)} + \boldsymbol{I}_{Cr_k}, \quad (3.52)$$

and

$$\hat{\alpha}^{(i-1)} = \sum_{j=1}^{i-1}\alpha_j^{(i-1)}.$$

have been used. Inserting these results into the objective function from (3.46) yields

$$\left\{\boldsymbol{g}_{k(i),d_{k(i)}+1},\gamma(k(i),d_{k(i)}+1),k(i)\right\} = \\ \underset{\boldsymbol{g},c,k}{\operatorname{argmax}}\log_2\left(\frac{\left(P_{\mathrm{Tx}}\frac{C}{\sigma_{\mathrm{n}}^2}+\hat{\alpha}^{(i-1)}\right)\boldsymbol{g}^{\mathrm{H}}\boldsymbol{A}_k^{(i)}\boldsymbol{g}+\boldsymbol{g}^{\mathrm{H}}\boldsymbol{B}_k^{(i)}\boldsymbol{g}}{\sum_{m=1}^{K}\mu_m d_m+\mu_k}\right)^{\sum_{m=1}^{K}\mu_m d_m+\mu_k} \\ -\sum_{k'=1}^{K}\sum_{j=1}^{d_{k'}}\log_2\left(\boldsymbol{g}^{\mathrm{H}}\left(\boldsymbol{A}_k^{(i)}+\boldsymbol{H}_k\hat{\boldsymbol{P}}_{n_{k',j}}^{(i)}\boldsymbol{H}_k^{\mathrm{H}}\right)\boldsymbol{g}\frac{\alpha_{n_{k',j}}^{(i-1)}}{\mu_{k'}}\right)^{\mu_{k'}} - \log_2\left(\frac{1}{\mu_k}\right)^{\mu_k} \\ \text{s.t. } \boldsymbol{g}=\boldsymbol{S}_{c,k}^{\mathsf{T}}\boldsymbol{S}_{c,k}\boldsymbol{g},\quad \boldsymbol{g}^{\mathrm{H}}\boldsymbol{g}=1,\quad \boldsymbol{g}_{k,m}^{\mathrm{H}}\boldsymbol{g}=0,\forall m \text{ with } \gamma(k,m)=c. \quad (3.53)$$

As with DPC, Problem (3.53) is solved in two steps. First for each user k the receive filter $\boldsymbol{g}_k^{(i)}$ and the carrier $c_k^{(i)}$ are determined to maximize the weighted sum rate assuming that the next data stream is allocated to user k. Rewriting the sums of logarithms as the product of its arguments and considering the fact that the logarithm is a monotonically increasing function, i.e., maximizing the argument of the logarithm is identical to maximizing the logarithm, the optimum $\boldsymbol{g}_k^{(i)}$ and $c_k^{(i)}$ can be found from

$$\left\{\boldsymbol{g}_k^{(i)}, c_k^{(i)}\right\} = \underset{\boldsymbol{g},c}{\operatorname{argmax}} \frac{\left(\left(P_{\text{Tx}}\frac{C}{\sigma_{\text{n}}^2} + \hat{\alpha}^{(i-1)}\right)\boldsymbol{g}^{\text{H}}\boldsymbol{A}_k^{(i)}\boldsymbol{g} + \boldsymbol{g}^{\text{H}}\boldsymbol{B}_k^{(i)}\boldsymbol{g}\right)^{\sum_{m=1}^{K}\mu_m d_m + \mu_k}}{\left(\sum_{m=1}^{K}\mu_m d_m + \mu_k\right)^{\sum_{m=1}^{K}\mu_m d_m+\mu_k} \prod_{k'=1}^{K}\prod_{j=1}^{d_{k'}}\left(\boldsymbol{g}^{\text{H}}\left(\boldsymbol{A}_k^{(i)} + \boldsymbol{H}_k\hat{\boldsymbol{P}}_{n_{k'},j}^{(i)}\boldsymbol{H}_k^{\text{H}}\right)\boldsymbol{g}\frac{\alpha_{n_{k'},j}^{(i-1)}}{\mu_{k'}}\right)^{\mu_{k'}}\left(\frac{1}{\mu_k}\right)^{\mu_k}}$$

$$\text{s.t. } \boldsymbol{g} = \boldsymbol{S}_{c,k}^{\mathsf{T}}\boldsymbol{S}_{c,k}\boldsymbol{g}, \quad \boldsymbol{g}^{\text{H}}\boldsymbol{g} = 1, \quad \boldsymbol{g}_{k,m}^{\text{H}}\boldsymbol{g} = 0, \forall m \text{ with } \gamma(k,m) = c, \tag{3.54}$$

Problem (3.54) is still non-convex and is therefore in general not easy to solve. For the special case $i = 2$, Problem (3.54) is of the kind

$$\max_{\boldsymbol{g}} \frac{\left(\boldsymbol{g}^{\text{H}}\boldsymbol{C}\boldsymbol{g}\right)^{\beta}}{\boldsymbol{g}^{\text{H}}\boldsymbol{D}\boldsymbol{g}}, \quad \text{s.t. } \boldsymbol{g}^{\text{H}}\boldsymbol{g} = 1$$

with $\beta > 1$ and the optimum receive filters $\boldsymbol{g}_k^{(i)}$ could therefore be found by the algorithm from [76], which returns receive filters that comply with the carrier separation constraint when applied to block-diagonal matrices. As this algorithm is iterative and requires a matrix inversion in each step, this option will not pursued in the following. Instead it is proposed to maximize a lower bound of the objective function, which is obtained by applying the inequality between weighted geometric and weighted arithmetic mean (e.g. [67, Lemma 1]) to the denominator in (3.54) so that

$$\begin{aligned}
&\left[\prod_{k'=1}^{K}\prod_{j=1}^{d_{k'}}\left(\boldsymbol{g}^{\text{H}}\left(\boldsymbol{A}_k^{(i)} + \boldsymbol{H}_k\hat{\boldsymbol{P}}_{n_{k'},j}^{(i)}\boldsymbol{H}_k^{\text{H}}\right)\boldsymbol{g}\frac{\alpha_{n_{k'},j}^{(i-1)}}{\mu_{k'}}\right)^{\mu_{k'}}\left(\frac{1}{\mu_k}\right)^{\mu_k}\right]^{\frac{1}{\sum_{m=1}^{K}\mu_m d_m+\mu_k}} \\
&\leq \left(\sum_{k'=1}^{K}\sum_{j=1}^{d_{k'}}\frac{\mu_{k'}}{\mu_{k'}}\alpha_{n_{k'},j}^{(i-1)}\boldsymbol{g}^{\text{H}}\left(\boldsymbol{A}_k^{(i)} + \boldsymbol{H}_k\hat{\boldsymbol{P}}_{n_{k'},j}^{(i)}\boldsymbol{H}_k^{\text{H}}\right)\boldsymbol{g} + \frac{\mu_k}{\mu_k}\right)\frac{1}{\sum_{m=1}^{K}\mu_m d_m + \mu_k} \\
&= \boldsymbol{g}^{\text{H}}\left(\boldsymbol{A}_k^{(i)}\hat{\alpha}^{(i-1)} + \boldsymbol{B}_k^{(i)}\right)\boldsymbol{g}\frac{1}{\sum_{m=1}^{K}\mu_m d_m + \mu_k}.
\end{aligned}$$

Inserting this lower bound instead of the true objective function in (3.54), the problem of finding the receive filter and carrier in step i reads as

$$\left\{\boldsymbol{g}_k^{(i)}, c_k^{(i)}\right\} = \underset{\boldsymbol{g},c}{\operatorname{argmax}}\left(\frac{P_{\text{Tx}}\frac{C}{\sigma_{\text{n}}^2}\boldsymbol{g}^{\text{H}}\boldsymbol{A}_k^{(i)}\boldsymbol{g} + \boldsymbol{g}^{\text{H}}\left(\hat{\alpha}^{(i-1)}\boldsymbol{A}_k^{(i)} + \boldsymbol{B}_k^{(i)}\right)\boldsymbol{g}}{\boldsymbol{g}^{\text{H}}\left(\hat{\alpha}^{(i-1)}\boldsymbol{A}_k^{(i)} + \boldsymbol{B}_k^{(i)}\right)\boldsymbol{g}}\right)^{\sum_{m=1}^{K}\mu_m d_m+\mu_k}$$

$$\text{s.t. } \boldsymbol{g} = \boldsymbol{S}_{c,k}^{\mathsf{T}}\boldsymbol{S}_{c,k}\boldsymbol{g}, \quad \boldsymbol{g}^{\text{H}}\boldsymbol{g} = 1, \quad \boldsymbol{g}_{k,m}^{\text{H}}\boldsymbol{g} = 0, \forall m \text{ with } \gamma(k,m) = c. \tag{3.55}$$

Note that in (3.55) the weighted sum rate $R_{\text{WSR}}^{(i)}(k)$ obtainable if user k receives a data stream in step i is lower bounded by

$$R_{\text{WSR}}^{(i)}(k) \geq \left(\sum_{m=1}^{K}\mu_m d_m + \mu_k\right)\log_2\left(1 + \frac{P_{\text{Tx}}\frac{C}{\sigma_{\text{n}}^2}}{\left\|\boldsymbol{H}_{\text{comp}}^{+}\right\|_{\text{F}}^2}\right) \tag{3.56}$$

which has been obtained with (3.51). This lower bound coincides with the bound in [77], where a MISO system is considered and a different derivation is used. It can be achieved by choosing the unit-norm beamformers $\boldsymbol{t}_{k,j}$ multiplied by the powers $p_{k,j}$ as

$$\left[\sqrt{p_{1,1}}\boldsymbol{t}_{1,1},\ldots,\sqrt{p_{1,d_1 1}}\boldsymbol{t}_{1,d_1},\ldots,\sqrt{p_{K,1}}\boldsymbol{t}_{K,1},\ldots,\sqrt{p_{K,d_K}}\boldsymbol{t}_{K,d_K},\sqrt{p_{k,d_k+1}}\boldsymbol{t}_{k,d_k+1}\right]= \\ =\boldsymbol{H}_{\text{comp}}^{+}\frac{1}{\left\|\boldsymbol{H}_{\text{comp}}^{+}\right\|_{\text{F}}},$$

i.e., the optimum zero-forcing unit-norm beamformers but a sub-optimum power allocation are used. Problem (3.55) is equivalent to

$$\left\{\boldsymbol{g}_k^{(i)},c_k^{(i)}\right\}=\underset{\boldsymbol{g},c}{\operatorname{argmax}}\frac{\boldsymbol{g}^{\text{H}}\boldsymbol{A}_k^{(i)}\boldsymbol{g}}{\boldsymbol{g}^{\text{H}}\left(\hat{\alpha}^{(i-1)}\boldsymbol{A}_k^{(i)}+\boldsymbol{B}_k^{(i)}\right)\boldsymbol{g}}$$
$$\text{s.t. }\boldsymbol{g}=\boldsymbol{S}_{c,k}^{\mathsf{T}}\boldsymbol{S}_{c,k}\boldsymbol{g},\quad \boldsymbol{g}^{\text{H}}\boldsymbol{g}=1,\quad \boldsymbol{g}_{k,m}^{\text{H}}\boldsymbol{g}=0,\forall m \text{ with } \gamma(k,m)=c. \tag{3.57}$$

Dividing the objective function in (3.57) by $\boldsymbol{g}^{\text{H}}\boldsymbol{A}_k^{(i)}\boldsymbol{g}$, Problem (3.57) is maximized by the same arguments as

$$\left\{\boldsymbol{g}_k^{(i)},c_k^{(i)}\right\}=\underset{\boldsymbol{g},c}{\operatorname{argmax}}\frac{\boldsymbol{g}^{\text{H}}\boldsymbol{A}_k^{(i)}\boldsymbol{g}}{\boldsymbol{g}^{\text{H}}\boldsymbol{B}_k^{(i)}\boldsymbol{g}}$$
$$\text{s.t. }\boldsymbol{g}=\boldsymbol{S}_{c,k}^{\mathsf{T}}\boldsymbol{S}_{c,k}\boldsymbol{g},\quad \boldsymbol{g}^{\text{H}}\boldsymbol{g}=1,\quad \boldsymbol{g}_{k,m}^{\text{H}}\boldsymbol{g}=0,\forall m \text{ with } \gamma(k,m)=c. \tag{3.58}$$

Ignoring the constraints in (3.58) for the moment and setting the derivative of the objective function to zero yields

$$\boldsymbol{A}_k^{(i)}\boldsymbol{g}=\frac{\boldsymbol{g}^{\text{H}}\boldsymbol{A}_k^{(i)}\boldsymbol{g}}{\boldsymbol{g}^{\text{H}}\boldsymbol{B}_k^{(i)}\boldsymbol{g}}\boldsymbol{B}_k^{(i)}\boldsymbol{g},$$

which is a generalized eigenvalue problem. Thus, $\boldsymbol{g}$ must be chosen to be a generalized eigenvector of the matrix pair $\boldsymbol{A}_k^{(i)}$ and $\boldsymbol{B}_k^{(i)}$. The objective function, which is equal to the generalized eigenvalue of this matrix pair at the places, where its derivative is zero, is maximized by choosing $\boldsymbol{g}_k^{(i)}$ to be the eigenvector belonging to the principal generalized eigenvalue of this matrix pair. In the following it will shown that this solution of the unconstrained maximization fulfills all constraints in (3.58). As the objective function is independent of the norm of $\boldsymbol{g}$, the norm one constraint can be easily fulfilled by taking the unit-norm eigenvector. The carrier separation constraint is fulfilled, as the matrices $\boldsymbol{A}_k^{(i)}$ and $\boldsymbol{B}_k^{(i)}$ are block-diagonal. That is due to the fact that $\hat{\boldsymbol{P}}_{\text{lin}}^{(i)}$ and all $\hat{\boldsymbol{P}}_j^{(i)}=\boldsymbol{T}^{(i-1)}\boldsymbol{e}_j\boldsymbol{e}_j^{\mathsf{T}}\boldsymbol{T}^{(i-1),\text{H}}$ are block-diagonal, as the transmit filters in $\boldsymbol{T}^{(i-1)}$ obey to the carrier separation constraint. $c_k^{(i)}$ is implicitly given by the index of this block within the matrix pair $\boldsymbol{A}_k^{(i)}$ and $\boldsymbol{B}_k^{(i)}$, that exhibits the strongest principal generalized eigenvalue. It remains to proof that $\boldsymbol{g}_k^{(i)}$ also fulfills the orthogonality constraints $\boldsymbol{g}_{k,m}\boldsymbol{g}_k^{(i)}=0$ for all tuples (k,m) with $\gamma(k,m)=c_k^{(i)}$. As the matrix $\boldsymbol{B}_k^{(i)}$ is always invertible due to the positive semidefinitenes of the matrices $\boldsymbol{H}_k\hat{\boldsymbol{P}}_j^{(i)}\boldsymbol{H}_k^{\text{H}}\alpha_j^{(i-1)}$, $\boldsymbol{g}_k^{(i)}$ is also the eigenvector corresponding to the principal eigenvalue of the matrix $\boldsymbol{B}_k^{(i),-1}\boldsymbol{A}_k^{(i)}$ and will therefore not lie in $\operatorname{null}\left\{\boldsymbol{A}_k^{(i)}\right\}=\operatorname{null}\left\{\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{lin}}^{(i)}\boldsymbol{H}_k^{\text{H}}\right\}$, i.e.,

$$\boldsymbol{g}_k^{(i)}\in\left(\operatorname{null}\left\{\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{lin}}^{(i)}\boldsymbol{H}_k^{\text{H}}\right\}\right)^{\perp}.$$

As $\boldsymbol{g}_{k,m}^{\mathrm{H}}\boldsymbol{H}_k$ is one row of $\boldsymbol{H}_{\text{comp}}^{(i-1)}$ and $\hat{\boldsymbol{P}}_{\text{lin}}^{(i)}$ projects into the nullspace of $\boldsymbol{H}_{\text{comp}}^{(i-1)}$ [c.f. (3.49)], $\hat{\boldsymbol{P}}_{\text{lin}}^{(i)}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{g}_{k,m} = \boldsymbol{0}_{CN_{\mathrm{T}},1}$ and thus

$$\boldsymbol{g}_{k,m} \in \operatorname{null}\left\{\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{lin}}^{(i)}\boldsymbol{H}_k^{\mathrm{H}}\right\}.$$

Due to the fact that $\boldsymbol{g}_k^{(i)} \in \left(\operatorname{null}\left\{\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{lin}}^{(i)}\boldsymbol{H}_k^{\mathrm{H}}\right\}\right)^{\perp}$, $\boldsymbol{g}_k^{(i)}$ is therefore orthogonal to all $\boldsymbol{g}_{k,m}$.

Note that in case no data stream has been allocated to carrier $c_k^{(i)}$ during the previous steps, i.e., $n_{c_k^{(i)}} = 0$, the receive filter $\boldsymbol{g}_k^{(i)}$ also maximizes the exact objective function in (3.54). That is because the projection matrix $\hat{\boldsymbol{P}}_{\text{lin}}^{(i)}$ has a block-diagonal structure and each block is a projector affecting only the channel matrices of one carrier in the products $\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{lin}}^{(i)}\boldsymbol{H}_k^{\mathrm{H}}$. In case no data stream has been allocated to carrier $c_k^{(i)}$ in a previous step, the blocks affecting this carrier in the projection matrices are identity matrices. As the receive filters must obey the carrier separation constraint, the products $\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{lin}}^{(i)}\boldsymbol{H}_k\boldsymbol{g}_k^{(i)}$ are in this case given by

$$\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{lin}}^{(i)}\boldsymbol{H}_k\boldsymbol{g}_k^{(i)} = \boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{A}_k^{(i)}\boldsymbol{g}_k^{(i)} = \boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{H}_k\boldsymbol{H}_k\boldsymbol{g}_k^{(i)}.$$

Additionally, $\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_j^{(i)}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{g}_k^{(i)} = 0$ and consequently $\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{B}_k^{(i)}\boldsymbol{g}_k^{(i)} = 1$, as the blocks within $\hat{\boldsymbol{P}}_j^{(i)} = \boldsymbol{T}^{(i-1)}\boldsymbol{e}_j\boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{T}^{(i-1),\mathrm{H}}$ that correspond to carrier $c_k^{(i)}$ are zero matrices, as no data stream has been allocated to this carrier in a previous step. Ignoring the terms independent of $\boldsymbol{g}$, the objective function in (3.54) for the case $n_{c_k^{(i)}} = 0$ reads as

$$\frac{\left(\left(\hat{\alpha}^{(i-1)} + P_{\text{Tx}}\frac{C}{\sigma_{\mathrm{n}}^2}\right)\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{H}_k\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{g}_k^{(i)} + 1\right)^{1+\frac{\mu_k}{\sum_{m=1}^{K}\mu_m d_m + \mu_k}}}{\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{H}_k\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{g}_k^{(i)}\hat{\alpha}^{(i-1)} + 1}. \tag{3.59}$$

This objective function is maximized under the constraints from (3.54) by choosing $\boldsymbol{g}_k^{(i)}$ to be the unit-norm eigenvector belonging to the principal eigenvalue of the block matrix with $\boldsymbol{H}_k\boldsymbol{H}_k^{\mathrm{H}}$ that corresponds to carrier $c_k^{(i)}$. This can be verified by solving (3.54) with the algorithm from [76], which can be used for objective functions as in (3.59) and which converges in this case within one iteration.

Once the filters $\boldsymbol{g}_k^{(i)}$ have been computed for all users k, the user $k(i)$ is determined so that the weighted sum rate becomes maximum, i.e.,

$$\begin{aligned} k(i) = \underset{k}{\operatorname{argmax}} \sum_{k'=1}^{K}\mu_{k'}\sum_{j=1}^{d_{k'}}\log_2\left[\min\left(1, \frac{\eta^{(i)}C\mu_{k'}}{\sigma_{\mathrm{n}}^2\left(1 + \frac{\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{n_{k',j}}^{(i)}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{g}_k^{(i)}}{\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{A}_k^{(i)}\boldsymbol{g}_k^{(i)}}\right)\alpha_{n_{k',j}}^{(i-1)}}\right)\right] + \\ +\mu_k\log_2\left[\min\left(1, \frac{C}{\sigma_{\mathrm{n}}^2}\eta^{(i)}\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{A}_k^{(i)}\boldsymbol{g}^{(i)}\mu_k\right)\right]. \end{aligned} \tag{3.60}$$

$\eta^{(i)}(k)$ is the water-level for the weighted water-filling power allocation, if the next data stream is allocated to user k. In case all data streams receive non-zero powers, $\eta^{(i)}(k)$ is given by

$$\eta^{(i)}(k) = \frac{P_{\text{Tx}} + \frac{\sigma_{\mathrm{n}}^2\hat{\alpha}^{(i-1)}}{C} + \frac{\sigma_{\mathrm{n}}^2\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{B}_k^{(i)}\boldsymbol{g}_k^{(i)}}{C\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{A}_k^{(i)}\boldsymbol{g}_k^{(i)}}}{\sum_{k'=1}^{K}\mu_{k'}d_{k'} + \mu_k} = \frac{P_{\text{Tx}} + \frac{\sigma_{\mathrm{n}}^2\hat{\alpha}^{(i-1)}}{C} + \frac{\sigma_{\mathrm{n}}^2}{C\rho_1\left(\boldsymbol{B}_k^{(i),-1}\boldsymbol{A}_k^{(i)}\right)}}{\sum_{k'=1}^{K}\mu_{k'}d_{k'} + \mu_k}, \tag{3.61}$$

which follows from the fact that

$$\rho_1\left(\boldsymbol{B}_k^{(i),-1}\boldsymbol{A}_k^{(i)}\right)\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{B}_k^{(i)}\boldsymbol{g}_k^{(i)}=\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{A}_k^{(i)}\boldsymbol{g}_k^{(i)}. \tag{3.62}$$

In case it is observed that the water-level is too small for a data stream, i.e.,

$$\eta^{(i)}(k)<\frac{C\mu_{\hat{k}}}{\sigma_\mathrm{n}^2\left(1+\frac{\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{n_{\hat{k},\hat{j}}}^{(i)}\boldsymbol{H}_k^\mathrm{H}\boldsymbol{g}_k^{(i)}}{\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{A}_k^{(i)}\boldsymbol{g}_k^{(i)}}\right)\alpha_{n_{\hat{k},\hat{j}}}^{(i-1)}},$$

for some tuple $\left(\hat{k},\hat{j}\right)$ not only the water-level and the weighted sum rate but also all channel gains on the affected carrier should be recomputed, as otherwise that data stream imposes zero-forcing constraints on the transmit filters of the other data streams and therefore reduces their channel gains, although it receives zero power. Thus, the corresponding data stream has to be removed from the pseudo-inverse $\boldsymbol{H}_\mathrm{comp}^{(i)}$ and the variable $d_{\hat{k}}$ needs to be adjusted to this removal.

In contrast to DPC, where the receive filters for the current data stream do not affect the channel gains of the previously allocated data streams, a new data stream allocation possibly deteriorates the channel gains of all subchannels on the same carrier [c.f. (3.48)]. It can therefore happen that the weighted sum rate is lower than in the previous step. In this case the last allocation is not put into effect and the algorithm is terminated. Otherwise $d_{k(i)}$ is incremented by one, $\gamma(k(i),d_{k(i)})$ and $\boldsymbol{g}_{k(i),d_{k(i)}}$ are given by $\gamma(k(i),d_{k(i)})=c_{k(i)}^{(i)}$ and $\boldsymbol{g}_{k(i),d_{k(i)}}=\boldsymbol{g}_{k(i)}^{(i)}$. The projector $\hat{\boldsymbol{P}}_\mathrm{lin}^{(i+1)}$ for the next allocation step can then be computed according to

$$\hat{\boldsymbol{P}}_\mathrm{lin}^{(i+1)}=\hat{\boldsymbol{P}}_\mathrm{lin}^{(i)}-\frac{\hat{\boldsymbol{P}}_\mathrm{lin}^{(i)}\boldsymbol{H}_{k(i)}^\mathrm{H}\boldsymbol{g}_{k(i),d_{k(i)}}\boldsymbol{g}_{k(i),d_{k(i)}}^\mathrm{H}\boldsymbol{H}_{k(i)}\hat{\boldsymbol{P}}_\mathrm{lin}^{(i)}}{\boldsymbol{g}_{k(i),d_{k(i)}}^\mathrm{H}\boldsymbol{H}_{k(i)}\hat{\boldsymbol{P}}_\mathrm{lin}^{(i)}\boldsymbol{H}_{k(i)}^\mathrm{H}\boldsymbol{g}_{k(i),d_{k(i)}}}=\hat{\boldsymbol{P}}_\mathrm{lin}^{(i)}-\boldsymbol{t}_{k(i),d_{k(i)}}^{(i)}\boldsymbol{t}_{k(i),d_{k(i)}}^{(i),\mathrm{H}}. \tag{3.63}$$

The normalized transmit vector $\boldsymbol{t}_{k(i),d_{k(i)}}^{(i)}$ for the data stream allocated in the ith step is given by

$$\boldsymbol{t}_{k(i),d_{k(i)}}^{(i)}=\frac{\hat{\boldsymbol{P}}_\mathrm{lin}^{(i)}\boldsymbol{H}_{k(i)}^\mathrm{H}\boldsymbol{g}_{k(i),d_{k(i)}}}{\sqrt{\boldsymbol{g}_{k(i),d_{k(i)}}^\mathrm{H}\boldsymbol{H}_{k(i)}\hat{\boldsymbol{P}}_\mathrm{lin}^{(i)}\boldsymbol{H}_{k(i)}^\mathrm{H}\boldsymbol{g}_{k(i),d_{k(i)}}}} \tag{3.64}$$

because $\hat{\boldsymbol{P}}_\mathrm{lin}^{(i)}\boldsymbol{H}_{k(i)}^\mathrm{H}\boldsymbol{g}_{k(i),d_{k(i)}}$ is orthogonal to the rows of $\boldsymbol{H}_\mathrm{comp}^{(i-1)}$ [c.f. (3.49)] and therefore orthogonal to the effective subchannels of all other data streams and contains no component in the nullspace of $\boldsymbol{H}_\mathrm{comp}^{(i)}$. As the row corresponding to the data stream allocated at ith place in the pseudoinverse of $\boldsymbol{H}_\mathrm{comp}^{(i)}$ fulfills the same properties (see Appendix A3), $\hat{\boldsymbol{P}}_\mathrm{lin}^{(i)}\boldsymbol{H}_{k(i)}^\mathrm{H}\boldsymbol{g}_{k(i),d_{k(i)}}$ and $\boldsymbol{t}_{k(i),d_{k(i)}}^{(i)}$ are collinear and $\boldsymbol{t}_{k(i),d_{k(i)}}^{(i)}$ can be obtained by normalizing $\hat{\boldsymbol{P}}_\mathrm{lin}^{(i)}\boldsymbol{H}_{k(i)}^\mathrm{H}\boldsymbol{g}_{k(i),d_{k(i)}}$ to one. The relationship (3.63) also enables a computationally efficient update of the matrices $\boldsymbol{A}_k^{(i)}$ so that

$$\boldsymbol{A}_k^{(i+1)}=\boldsymbol{A}_k^{(i)}-\boldsymbol{H}_k\boldsymbol{t}_{k(i),d_{k(i)}}^{(i)}\boldsymbol{t}_{k(i),d_{k(i)}}^{(i),\mathrm{H}}\boldsymbol{H}_k^\mathrm{H}. \tag{3.65}$$

The coefficients $\alpha_j^{(i)}$ can also be efficiently computed according to [c.f. (3.48)]

$$\alpha_{n_{k,j}}^{(i)}=\begin{cases}\frac{1}{\boldsymbol{g}_{k(i),d_{k(i)}}^\mathrm{H}\boldsymbol{H}_{k(i)}\hat{\boldsymbol{P}}_\mathrm{lin}^{(i)}\boldsymbol{H}_{k(i)}^\mathrm{H}\boldsymbol{g}_{k(i),d_{k(i)}}}, & j=d_{k(i)},k=k(i)\\ \left(1+\alpha_{m(i)}^{(i)}\gamma_{k,j}^{(i)}\gamma_{k,j}^{(i)*}\right)\alpha_{\hat{n}_{k,j}}^{(i-1)} & \text{else}\end{cases}, \tag{3.66}$$

where $n_{k,j}$ is computed as stated in (3.42), $\hat{n}_{k,j}$ considers the adjustment of the variable $d_{k(i)}$ from step $i-1$ to step i so that

$$\hat{n}_{k,j} = \begin{cases} n_{k,j}, & k \leq k(i) \\ n_{k,j}+1, & \text{else} \end{cases},$$

$m(i) = \sum_{k'=1}^{k(i)+1} d_{k'}$ and

$$\gamma_{k,j}^{(i)} = \boldsymbol{g}_{k(i),d_{k(i)}}^{\mathrm{H}} \boldsymbol{H}_{k(i)} \boldsymbol{t}_{k,j}^{(i-1)} \tag{3.67}$$

can be obtained via an inner vector product. Note that $\gamma_{k,j}^{(i)}$ is only different from zero for transmit vectors $\boldsymbol{t}_{k,j}^{(i-1)}$ corresponding to data streams allocated to the carrier $c_{k(i)}^{(i)}$. The transmit filters $\boldsymbol{t}_{k,j}^{(i)}$ for the other data streams with $(k,j) \neq (k(i), d_{k(i)})$ are determined according to

$$\boldsymbol{t}_{k,j}^{(i)} = \beta_{k,j}^{(i)} \left(\boldsymbol{t}_{k,j}^{(i-1)} - \boldsymbol{t}_{k(i),d_{k(i)}}^{(i)} \gamma_{k,j}^{(i)} \sqrt{\alpha_{m(i)}^{(i)}} \right), \tag{3.68}$$

where $\beta_{k,j}^{(i)}$ is a scaling factor so that $\boldsymbol{t}_{k,j}^{(i)}$ has unit norm and which is shown in Appendix A5. Each update to obtain $\boldsymbol{t}_{k,j}^{(i)}$ additionally requires only one vector subtraction with N_{T} nonzero entries and one norm computation of a vector with length N_{T}, as the complex numbers $\alpha_{m(i)}^{(i)}$ and $\gamma_{k,j}^{(i)}$ have already been determined during the computation of the $\alpha_{n_{k,j}}^{(i)}$. This update rule is more efficient than the update of the pseudo-inverse via the LQ decomposition of $\boldsymbol{H}_{\mathrm{comp}}^{(i)}$ proposed in [64], as the latter requires in each step a multiplication of a matrix with another lower-triangular matrix to compute all precoding vectors. Unfortunately, no efficient update rule to obtain the matrices $\boldsymbol{B}_k^{(i+1)}$ from the matrices $\boldsymbol{B}_k^{(i)}$ exists. That is why those matrices have to be computed from scratch with the vectors $\boldsymbol{t}_{k,j}^{(i)}$. Note that in multicarrier systems due to the carrier separation constraint the update rules described above only affect those parts of the matrices or vectors that correspond to carrier $c_{k(i)}^{(i)}$, all other parts remain unchanged. An overview of the spatial resource allocation with zero-forcing is given with Algorithm 3.3, a modification of algorithm for the MAC with individual power constraints can be found in [78]. Additional little performance improvements of Algorithm 3.3 can be achieved by the following measures. Instead of running the algorithm in the broadcast channel it can be run in the dual MAC. The filters obtained this way are then transformed via the general duality from [79] into the broadcast channel. Performance improvements are possible by replacing the receive filters computed in this manner by MMSE filters as proposed in [80]. Furthermore one can recompute the transmit and receive filters in case a user can transmit more than one data stream on the same carrier similar to [71]. An SVD of this user's channel matrix multiplied by projection matrices that consider the other users' zero-forcing constraints is performed so that this user's channel gains become maximum while the other users' channel gains are not affected by this measure as described in [81].

3.4 Further Complexity Reductions

Although compared to the optimum algorithms the computational complexity could be drastically reduced by the methods proposed in the previous two sections, those still require the computation of eigenvalues and eigenvectors for each user in each step. For this reason a user preselection

Algorithm 3.3 Reduced Complexity Algorithm for Weighted Sum Rate Maximization without DPC

Initialization: $d_k = 0, \boldsymbol{G}_k = [], \boldsymbol{A}_k^{(i)} = \boldsymbol{I}_{CN_\mathrm{T}}, \forall k = 1, \ldots, K$, $i = 1$, $\hat{\boldsymbol{P}}_\mathrm{lin}^{(1)} = \boldsymbol{I}_{CN_\mathrm{T}}$

$k(1) = \underset{k}{\operatorname{argmax}}\, \mu_k \log_2\left(1 + \frac{P_\mathrm{Tx}C}{\sigma_\mathrm{n}^2}\rho_1\left(\boldsymbol{H}_k\boldsymbol{H}_k^\mathrm{H}\right)\right)$,
$R_\mathrm{WSR}^{(1)} = \mu_{k(1)} \log_2\left(1 + \frac{P_\mathrm{Tx}C}{\sigma_\mathrm{n}^2}\rho_1\left(\boldsymbol{H}_{k(1)}\boldsymbol{H}_{k(1)}^\mathrm{H}\right)\right)$

$\boldsymbol{g}_{k(1)}^{(1)}$ eigenvector corresponding to principal eigenvalue of matrix $\boldsymbol{H}_{k(1)}\boldsymbol{H}_{k(1)}^\mathrm{H}$

repeat

  $d_{k(i)} = d_{k(i)} + 1$

  $\boldsymbol{g}_{k(i),d_{k(i)}} = \boldsymbol{g}_{k(i)}^{(i)}$, $\boldsymbol{G}_{k(i)} = \left[\boldsymbol{G}_{k(i)}, \boldsymbol{g}_{k(i),d_{k(i)}}\right]$

  $\alpha_{m(i)}^{(i)} = \frac{1}{\boldsymbol{g}_{k(i),d_{k(i)}}^\mathrm{H}\boldsymbol{A}_{k(i)}^{(i)}\boldsymbol{g}_{k(i),d_{k(i)}}}$, $\boldsymbol{t}_{k(i),d_{k(i)}}^{(i)} = \frac{\hat{\boldsymbol{P}}_\mathrm{lin}^{(i)}\boldsymbol{H}_{k(i)}^\mathrm{H}\boldsymbol{g}_{k(i),d_{k(i)}}}{\sqrt{\alpha_{m(i)}^{(i)}}}$, $m(i) = \sum_{k'=1}^{k(i)} d_{k'}$

  $\hat{\boldsymbol{P}}_\mathrm{lin}^{(i+1)} = \hat{\boldsymbol{P}}_\mathrm{lin}^{(i)} - \boldsymbol{t}_{k(i),d_{k(i)}}^{(i)}\boldsymbol{t}_{k(i),d_{k(i)}}^{(i),\mathrm{H}}$

  for $k = 1$ to K **do**

    Compute $\boldsymbol{A}_k^{(i+1)}$ from $\boldsymbol{A}_k^{(i)}$ with (3.65)

    Compute $\alpha_{n_{k,j}}^{(i)}$ from $\alpha_{\hat{n}_{k,j}}^{(i-1)}$ with (3.66) and $\gamma_{k,j}^{(i)}$ with (3.67) $\forall j = 1, \ldots, d_k$

  end for

  $i = i + 1$

  for $k = 1$ to K **do**

    $\boldsymbol{B}_k^{(i)} = \boldsymbol{H}_k\left[\sum_{k'=1}^{K}\sum_{j=1}^{d_{k'}} \boldsymbol{t}_{k,j}^{(i-1)}\boldsymbol{t}_{k,j}^{(i-1),\mathrm{H}}\alpha_{n_{k',j}}^{(i-1)}\right]\boldsymbol{H}_k^\mathrm{H} + \boldsymbol{I}_{Cr_k}$

    $\boldsymbol{g}_k^{(i)}$: generalized unit norm eigenvector for principal generalized eigenvalue of matrix pair $\boldsymbol{A}_k^{(i)}$, $\boldsymbol{B}_k^{(i)}$

    $$R_\mathrm{WSR}^{(i)}(k) = \sum_{k'=1}^{K}\mu_{k'}\sum_{j=1}^{d_{k'}}\log_2\left[\min\left(1, \frac{\eta^{(i)}C\mu_{k'}}{\sigma_\mathrm{n}^2\left(1+\frac{\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{H}_k\boldsymbol{t}_{k,j}^{(i-1)}\boldsymbol{t}_{k,j}^{(i-1),\mathrm{H}}\boldsymbol{H}_k^\mathrm{H}\boldsymbol{g}_k^{(i)}}{\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{A}_k^{(i)}\boldsymbol{g}_k^{(i)}}\right)\alpha_{n_{k',j}}^{(i-1)}}\right)\right] + \\ + \mu_k\log_2\left[\min\left(1, \frac{C}{\sigma_\mathrm{n}^2}\eta^{(i)}\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{A}_k^{(i)}\boldsymbol{g}_k^{(i)}\mu_k\right)\right]$$

    $\eta^{(i)}$: water-level for weighted water-filling with power P_Tx and subchannel gains $\frac{C}{\sigma_\mathrm{n}^2}\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{A}_k^{(i)}\boldsymbol{g}_K^{(i)}$, $\frac{C}{\sigma_\mathrm{n}^2\left(1+\frac{\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{n_{k',j}}^{(i)}\boldsymbol{H}_k^\mathrm{H}\boldsymbol{g}_k^{(i)}}{\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{A}_k^{(i)}\boldsymbol{g}_k^{(i)}}\right)\alpha_{n_{k',j}}^{(i-1)}}, j = 1, \ldots, d_{k'}, k' = 1, \ldots, K$

  end for

  $k(i) = \underset{k}{\operatorname{argmax}}\, R_\mathrm{WSR}^{(i)}(k)$, $\quad R_\mathrm{WSR}^{(i)} = \max_k R_\mathrm{WSR}^{(i)}(k)$

  Remove data streams that have received zero power and recompute transmit filters and $R_\mathrm{WSR}^{(i)}$ if necessary

until $R_\mathrm{WSR}^{(i)} < R_\mathrm{WSR}^{(i-1)}$

$p_{1,1}, \ldots, p_{K,d_K} \leftarrow$ weighted water-filling with power P_Tx and subchannel gains $\frac{C}{\sigma_\mathrm{n}^2\alpha_1^{(i-1)}}, \ldots, \frac{C}{\sigma_\mathrm{n}^2\alpha_{i-1}^{(i-1)}}$

for $k = 1$ to K **do**

  $\boldsymbol{P}_k = \operatorname{diag}\left(p_{k,1}\ldots p_{k,d_k}\right)$, $\boldsymbol{T}_k = \left[\boldsymbol{t}_{k,1}^{(i-1)}, \ldots, \boldsymbol{t}_{k,d_k}^{(i-1)}\right]$

end for

method will be presented in Section 3.4.1, which deselects users that will certainly not lead to the maximum weighted sum rate by simple decision criteria avoiding the explicit computations of the principal eigenvalues of these users. While this method enables partially drastic complexity reductions at no performance losses, the other two techniques presented in this Section inhere potential losses in weighted sum rate. A weaker lower bound for the weighted sum rate for the algorithm without DPC will be introduced in Section 3.4.2, which enables the computation of candidate channel gains via ordinary instead of generalized eigenvalue problems. In Section 3.4.3 a simplified user selection is presented for both algorithms with and without DPC, where the user selection can be conducted without explicit eigenvalue computations. A further possibility for reducing the computational complexity inducing performance losses, that will not explained here, is to apply subspace beamforming, where the space for the transmit filters is restricted to a predefined subspace, as proposed in [82]. For algorithms relying on DPC the method from [83] can be used to identify a subgroup of users, with which the gain of those algorithms becomes maximum compared to simple time or frequency division algorithms.

3.4.1 User Preselection

The most complex part in the successive resource allocation schemes is the user selection in (3.36) with DPC, which requires the determination of principal eigenvalues for each user, and in (3.60) without DPC, where a generalized eigenvalue, the corresponding eigenvector and the quadratic forms $\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{n_{k'},j}^{(i)}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{g}_k^{(i)}$ have to be computed. Correspondingly, the complexity grows linearly with the number of users. By the method to be presented next this complexity can be reduced at no performance losses. The basic principle is to compute lower and upper bounds for the weighted sum rate, if the next data stream is allocated to user k. Obviously, the bounds, which will be denoted as $R_{\text{WSR,lb}}^{(i)}(k)$ and $R_{\text{WSR,ub}}^{(i)}(k)$ in the following, must be obtainable at very low additional effort. The maximum weighted sum rate $R_{\text{WSR}}^{(i)}$ achievable in step i is lower bounded by the maximum lower bound amongst all users, i.e.,

$$\max_k R_{\text{WSR,lb}}^{(i)}(k) \leq \max_k R_{\text{WSR}}^{(i)}(k) = R_{\text{WSR}}^{(i)}.$$

If the upper bound $R_{\text{WSR,ub}}^{(i)}(m)$ of a certain user m is smaller than the maximum lower bound amongst all users that user m will certainly not lead to the maximum weighted sum rate in step i and can therefore be excluded from the user selection process in step i. Additionally, if the upper bound of a user's rate is lower than the weighted sum rate $R_{\text{WSR}}^{(i-1)}$ achieved in the previous step, this user will certainly not be served in the next step, as even if this user leads to the strongest weighted sum rate amongst all users, the allocation is not conducted as it would lead to a decrease in weighted sum rate. Thus, it suffices to conduct the user selections in (3.36) and (3.60) with a possibly reduced user set $\mathcal{S}^{(i)}$, which is given by

$$\mathcal{S}^{(i)} = \left\{ k \in \{1,\ldots,K\} \,\middle|\, R_{\text{WSR,ub}}^{(i)}(k) \geq \max_\ell R_{\text{WSR,lb}}^{(i)}(\ell), R_{\text{WSR,ub}}^{(i)}(k) \geq R_{\text{WSR}}^{(i-1)} \right\}.$$

In the remainder of this section the bounds $R_{\text{WSR,ub}}^{(i)}(k)$ and $R_{\text{WSR,lb}}^{(i)}(k)$ will be derived for the user selection with DPC and the user selection without DPC. Although the bounds are assessed very conservative, they turn out to be very effective in practice, as it will be shown by simulation results.

- **Bounds for the Weighted Sum Rate achievable with DPC:**
  The weighted sum rate with DPC is a monotonically increasing function in the channel gains,

as shown with Lemma A3.1. Thus, $R_{\text{WSR}}^{(i)}(k)$ as given by (3.37) is a monotonically increasing function in $\rho_1\left(\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\right)$. As furthermore the channel gains $\boldsymbol{t}_{k',j}^{\text{H}}\boldsymbol{H}_{k'}^{\text{H}}\boldsymbol{H}_{k'}^{\text{H}}\boldsymbol{t}_{k',j}$ are not affected by the new data stream allocation, the bounds for the principal eigenvalue of the matrix $\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}$ can be used to limit the weighted sum rate from below and above. These bounds are given by

$$\begin{aligned}\max_c \frac{\operatorname{tr}\left(\boldsymbol{S}_{c,\text{T}}\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{S}_{c,\text{T}}^{\mathsf{T}}\right)}{m_{k,c}^{(i)}} &\leq \rho_1\left(\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\right)\\ &\leq \max_c\left(\operatorname{tr}\left(\boldsymbol{S}_{c,\text{T}}\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{S}_{c,\text{T}}^{\mathsf{T}}\right)\right), \quad (3.69)\end{aligned}$$

where $m_{k,c}^{(i)}$ denotes an upper bound for the rank of the matrix $\boldsymbol{S}_{c,\text{T}}\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{S}_{c,\text{T}}^{\mathsf{T}}$. It is equal to the number of receive antennas of user k minus the number of data streams already allocated to user k on carrier c. For single-carrier systems, (3.69) is identical to Equation (2.3.7) in [35]. The extension to multi-carrier systems results from the fact that $\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}$ is block diagonal and the multiplications with $\boldsymbol{S}_c$ and $\boldsymbol{S}_c^{\mathsf{T}}$, which are defined in (2.2), select the non-zero blocks corresponding to carrier c within these matrices and therefore

$$\rho_1\left(\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\right) = \max_c \rho_1\left(\boldsymbol{S}_{c,\text{T}}\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{S}_{c,\text{T}}^{\mathsf{T}}\right). \quad (3.70)$$

The lower bound $R_{\text{WSR,lb}}^{(i)}(k)$ can then be obtained by inserting $\max_c \frac{\operatorname{tr}\left(\boldsymbol{S}_{c,\text{T}}\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{S}_{c,\text{T}}^{\mathsf{T}}\right)}{m_{k,c}^{(i)}}$ instead of $\rho_1\left(\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\right)$ into the expression for the weighted sum rate $R_{\text{WSR}}^{(i)}(k)$ in (3.37). Correspondingly, the upper bound $R_{\text{WSR,ub}}^{(i)}(k)$ is computed by using the upper bound for $\rho_1\left(\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\right)$ in (3.37). In case all weights are equal, inserting these bounds into the formula for the weighted sum rate is not necessary, as in each step this user leads to the maximum increase in weighted sum rate that exhibits the largest principal eigenvalue of the matrix $\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}$. In this special case the set $\mathcal{S}^{(i)}$ can therefore be directly determined according to

$$\begin{aligned}\mathcal{S}^{(i)} = \Bigg\{k \in \{1,\dots,K\} \Bigg| &\max_c \operatorname{tr}\left(\boldsymbol{S}_{c,\text{T}}\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{S}_{c,\text{T}}^{\mathsf{T}}\right)\\ &\geq \max_{\ell,c} \frac{\operatorname{tr}\left(\boldsymbol{S}_{c,\text{T}}\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_\ell^{\text{H}}\boldsymbol{H}_\ell\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{S}_{c,\text{T}}^{\mathsf{T}}\right)}{m_{\ell,c}^{(i)}}\Bigg\},\end{aligned}$$

as it has been proposed in [84].
The computation of the lower and upper bounds in (3.69) is possible at little extra computational complexity, as the matrices $\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}$ are required for the user selection process anyway. Additionally, the expressions $\operatorname{tr}\left(\boldsymbol{S}_{c,\text{T}}\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_\ell^{\text{H}}\boldsymbol{H}_\ell\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{S}_{c,\text{T}}^{\mathsf{T}}\right)$ only change compared to the previous step for the carrier $c = c_{k(i-1)}^{(i-1)}$ to which a data stream has been allocated to in step $i-1$.

- **Bounds for the Weighted Sum Rate achievable without DPC:**
When no DPC is used, it has already been stated in (3.56) that $R_{\text{WSR}}^{(i)}(k)$ can be lower bounded by

$$
\begin{aligned}
R_{\text{WSR}}^{(i)}(k) &\geq \left(\sum_{m=1}^{K}\mu_m d_m + \mu_k\right)\log_2\left(1+\frac{P_{\text{Tx}}\frac{C}{\sigma_{\text{n}}^2}}{\left\|\boldsymbol{H}_{\text{comp}}^{+}\right\|_{\text{F}}^2}\right)\\
&= \left(\sum_{m=1}^{K}\mu_m d_m + \mu_k\right)\log_2\left(1+\frac{P_{\text{Tx}}\frac{C}{\sigma_{\text{n}}^2}}{\hat{\alpha}_i + \frac{1}{\rho_1\left(\boldsymbol{B}_k^{(i),-1}\boldsymbol{A}_k^{(i)}\right)}}\right),
\end{aligned}
$$

where (3.62) has been applied. As $\boldsymbol{B}_k^{(i),-1}\boldsymbol{A}_k^{(i)}$ is block-diagonal, $\rho_1\left(\boldsymbol{B}_k^{(i),-1}\boldsymbol{A}_k^{(i)}\right)$ is the maximum eigenvalue of the principal eigenvalues of the matrices $\boldsymbol{S}_{c,k}\boldsymbol{B}_k^{(i),-1}\boldsymbol{A}_k^{(i)}\boldsymbol{S}_{c,k}^{\mathsf{T}}$, where $\boldsymbol{S}_{c,k}$ is defined in (2.2). Thus, one obtains

$$
\rho_1\left(\boldsymbol{B}_k^{(i),-1}\boldsymbol{A}_k^{(i)}\right) = \max_c \rho_1\left(\boldsymbol{S}_{c,k}\boldsymbol{B}_k^{(i),-1}\boldsymbol{A}_k^{(i)}\boldsymbol{S}_{c,k}^{\mathsf{T}}\right). \tag{3.71}
$$

By using the inequality

$$
\begin{aligned}
\rho_1\left(\boldsymbol{S}_{c,k}\boldsymbol{B}_k^{(i),-1}\boldsymbol{A}_k^{(i)}\boldsymbol{S}_{c,k}^{\mathsf{T}}\right) &= \rho_1\left(\left(\boldsymbol{S}_{c,k}\boldsymbol{B}_k^{(i)}\boldsymbol{S}_{c,k}^{\mathsf{T}}\right)^{-1}\boldsymbol{S}_{c,k}\boldsymbol{A}_k^{(i)}\boldsymbol{S}_{c,k}^{\mathsf{T}}\right)\\
&\geq \frac{\operatorname{tr}\left(\boldsymbol{S}_{c,k}\boldsymbol{A}_k^{(i)}\boldsymbol{S}_{c,k}^{\mathsf{T}}\right)}{r_k\left(1+\operatorname{tr}\left(\boldsymbol{S}_{c,k}\boldsymbol{B}_k^{(i)}\boldsymbol{S}_{c,k}^{\mathsf{T}}-\boldsymbol{I}_{r_k}\right)\right)}
\end{aligned}
$$

derived in Appendix A6, a lower bound for the weighted sum rate is given by

$$
R_{\text{WSR,lb}}^{(i)}(k) = \left(\sum_{m=1}^{K}\mu_m d_m + \mu_k\right)\log_2\left(1+\frac{P_{\text{Tx}}\frac{C}{\sigma_{\text{n}}^2}}{\hat{\alpha}_i + \min_c \frac{r_k\left(1-r_k+\operatorname{tr}\left(\boldsymbol{S}_{c,k}\boldsymbol{B}_k^{(i)}\boldsymbol{S}_{c,k}^{\mathsf{T}}\right)\right)}{\operatorname{tr}\left(\boldsymbol{S}_{c,k}\boldsymbol{A}_k^{(i)}\boldsymbol{S}_{c,k}^{\mathsf{T}}\right)}}\right). \tag{3.72}
$$

For the upper bound $R_{\text{WSR,ub}}^{(i)}(k)$ again Lemma A3.1 is used. Thus, the subchannel gains can bounded from above and inserted into the weighted sum rate expression to obtain an upper bound for the weighted sum rate. It is assumed that the newly allocated subchannel has no detrimental effect on the previously allocated subchannels, which implies that a lower bound for the inverse channel gains, i.e., an upper bound for the subchannel gains, can be stated as

$$
\left(1+\frac{\boldsymbol{g}_k^{(i),\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{n_{k',j}}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{g}_k^{(i)}}{\boldsymbol{g}_k^{(i),\text{H}}\boldsymbol{A}_k^{(i)}\boldsymbol{g}_k^{(i)}}\right)\alpha_{n_{k',j}}^{(i-1)} \geq \alpha_{n_{k',j}}^{(i-1)}.
$$

The gain of the newly allocated subchannel can be upper bounded according to

$$
\boldsymbol{g}_k^{(i),\text{H}}\boldsymbol{A}_k^{(i)}\boldsymbol{g}_k^{(i)} \leq \max_c \rho_1\left(\boldsymbol{S}_{c,k}\boldsymbol{A}_k^{(i)}\boldsymbol{S}_{c,k}^{\mathsf{T}}\right) \leq \max_c \operatorname{tr}\left(\boldsymbol{S}_{c,k}\boldsymbol{A}_k^{(i)}\boldsymbol{S}_{c,k}^{\mathsf{T}}\right)
$$

which follows from the fact that $\boldsymbol{A}_k^{(i)}$ is positive semi-definite. Using these inequalities in (3.60), an upper bound for $R_{\text{WSR}}^{(i)}(k)$ is given by

$$R_{\text{WSR,ub}}^{(i)}(k) = \sum_{k'=1}^{K} \mu_{k'} \sum_{j=1}^{d_{k'}} \log_2 \left[\min \left(1, \frac{\eta_{\text{ub}}^{(i)} C \mu_{k'}}{\sigma_{\text{n}}^2 \alpha_{n_{k',j}}^{(i-1)}} \right) \right] + \\ + \mu_k \log_2 \left[\min \left(1, \frac{C}{\sigma_{\text{n}}^2} \eta_{\text{ub}}^{(i)} \max_c \operatorname{tr} \left(\boldsymbol{S}_{k,c} \boldsymbol{A}_k^{(i)} \boldsymbol{S}_{c,k}^{\mathsf{T}} \right) \mu_k \right) \right], \quad (3.73)$$

with the water-level $\eta_{\text{ub}}^{(i)}$, which is, in case all estimated subchannel gains can be activated, given by

$$\eta_{\text{ub}}^{(i)} = \frac{P_{\text{Tx}} + \frac{\sigma_{\text{n}}^2 \hat{\alpha}^{(i-1)}}{C} + \frac{\sigma_{\text{n}}^2}{C \max_c \operatorname{tr}\left(\boldsymbol{S}_{c,k} \boldsymbol{A}_k^{(i)} \boldsymbol{S}_{c,k}^{\mathsf{T}} \right)}}{\sum_{k'=1}^{K} \mu_{k'} d_{k'} + \mu_k}.$$

Note that this upper bound is also valid for any choice of receive filters with unit norm. Likewise to the DPC case, the bounds can be computed at almost no extra effort, as the matrices $\boldsymbol{A}_k^{(i)}$ and $\boldsymbol{B}_k^{(i)}$ are required for the user selection process anyway and the terms $\operatorname{tr}\left(\boldsymbol{S}_{c,k}\boldsymbol{A}_k^{(i)}\boldsymbol{S}_{c,k}^{\mathsf{T}}\right)$ and $\operatorname{tr}\left(\boldsymbol{S}_{c,k}\boldsymbol{B}_k^{(i)}\boldsymbol{S}_{c,k}^{\mathsf{T}}\right)$ only change for the carrier $c = c_{k(i-1)}^{(i)}$.

Note that the bounds for the eigenvalue $\rho_1\left(\boldsymbol{S}_{c,\text{T}}\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{S}_{c,\text{T}}^{\mathsf{T}}\right)$ stated in (3.69) and the bounds for the eigenvalue $\rho_1\left(\boldsymbol{S}_{c,k}\boldsymbol{B}_k^{(i),-1}\boldsymbol{A}_k^{(i)}\boldsymbol{S}_{c,k}^{\mathsf{T}}\right)$ derived in Appendix A6 can also be used for a carrier preselection as shown in the following by means of the DPC case. According to (3.70) for each user the carrier $\hat{c}$ must be selected which matrix $\boldsymbol{S}_{\hat{c},\text{T}}\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{S}_{\hat{c},\text{T}}^{\mathsf{T}}$ exhibits the strongest principal eigenvalue. If the upper bound for the principal eigenvalue on a certain carrier is smaller than the maximum lower bound over all carriers this carrier will certainly not exhibit the strongest principal eigenvalue amongst all carriers and can therefore be deselected without explicitly computing the principal eigenvalue of its channel matrix.

3.4.2 Maximization of a Weaker Lower Bound for Weighted Sum Rate without DPC

Solving a generalized eigenvalue problem for each user in each allocation step still exhibits a considerable amount of computational complexity due to the necessity of inverting the matrices $\boldsymbol{B}_k^{(i)}$. For this reason a weaker lower bound for the weighted sum rate will be derived in the following, which enables a determination of the receive filters via ordinary eigenvalue problems. The denominator in (3.58) can be upper bounded by

$$\boldsymbol{g}^{\text{H}}\boldsymbol{B}_k^{(i)}\boldsymbol{g} \leq \boldsymbol{g}^{\text{H}}\boldsymbol{g}\rho_1\left(\boldsymbol{B}_k^{(i)}\right) = \rho_1\left(\boldsymbol{B}_k^{(i)}\right). \quad (3.74)$$

which stems from the fact that $\boldsymbol{g}$ is constrained to have norm one. Thus, using (3.74) the receive filters $\boldsymbol{g}_{\text{lb2}}^{(i)}(k)$ maximizing a weaker lower bound than the $\boldsymbol{g}_k^{(i)}$ determined from (3.58) are given by

$$\left\{\boldsymbol{g}_{\text{lb2}}^{(i)}(k), c_{\text{lb2}}^{(i)}(k)\right\} = \operatorname*{argmax}_{\boldsymbol{g},c} \frac{\boldsymbol{g}^{\text{H}}\boldsymbol{A}_k^{(i)}\boldsymbol{g}}{\rho_1\left(\boldsymbol{B}_k^{(i)}\right)}$$

$$\text{s.t. } \boldsymbol{g} = \boldsymbol{S}_{c,k}^{\mathsf{T}}\boldsymbol{S}_{c,k}\boldsymbol{g}, \quad \boldsymbol{g}^{\text{H}}\boldsymbol{g} = 1, \quad \boldsymbol{g}_{k,m}^{\text{H}}\boldsymbol{g} = 0, \forall m \text{ with } \gamma(k,m) = c, \quad (3.75)$$

which is solved by choosing $\boldsymbol{g}_{\mathrm{lb2}}^{(i)}(k)$ to be the unit norm eigenvector corresponding to the principal eigenvalue of the matrix $\boldsymbol{A}_k^{(i)} = \boldsymbol{H}_k \hat{\boldsymbol{P}}_{\mathrm{lin}}^{(i)} \boldsymbol{H}_k^{\mathrm{H}}$. The carrier $c_{\mathrm{lb2}}^{(i)}(k)$ is given implicitly by the index of this block within the matrix $\boldsymbol{A}_k^{(i)}$, which exhibits the maximum principal eigenvalue. The fact that the optimum $\boldsymbol{g}_{\mathrm{lb2}}^{(i)}(k)$ is orthogonal to all other receive filters of the same user can be shown in the same way as with the vectors $\boldsymbol{g}_k^{(i)}$. The user $k_{\mathrm{lb2}}(i)$ is then determined as in (3.60) with the receive filters $\boldsymbol{g}_{\mathrm{lb2}}^{(i)}(k)$. Correspondingly, after the ith step the variable $d_{k_{\mathrm{lb2}}(i)}$ is incremented by one and $\boldsymbol{g}_{k_{\mathrm{lb2}}(i),d_{k_{\mathrm{lb2}}}}$ is given by $\boldsymbol{g}_{k_{\mathrm{lb2}}(i),d_{k_{\mathrm{lb2}}}} = \boldsymbol{g}_{\mathrm{lb2}}^{(i)}(k_{\mathrm{lb2}}(i))$. Note that by determining the receive filters according to (3.75) besides avoiding the complexity associated with computing generalized eigenvalues, the matrices $\boldsymbol{B}_k^{(i)}$ do not have to be computed explicitly which redundantizes line 15 in Algorithm 3.3. Additionally, the quadratic forms $\boldsymbol{g}_{\mathrm{lb2}}^{(i),\mathrm{H}}(k)\boldsymbol{A}_k^{(i)}\boldsymbol{g}_{\mathrm{lb2}}^{(i)}(k)$ needed in (3.60) are equal to the principal eigenvalue of the matrices $\boldsymbol{A}_k^{(i)}$ and do therefore not have to be computed explicitly. The user and carrier preselection explained in the previous section can be applied with these receive filters as well, as the lower bound and the upper bound are independent of the choice of receive filters.

Determining the receive filters according to (3.75) corresponds to maximizing the channel gain of the newly allocated subchannel, which is given by $\boldsymbol{g}_{\mathrm{lb2}}^{(i),\mathrm{H}}(k(i))\boldsymbol{H}_{k(i)}\hat{\boldsymbol{P}}_{\mathrm{lin}}^{(i)}\boldsymbol{H}_{k(i)}^{\mathrm{H}}\boldsymbol{g}_{\mathrm{lb2}}^{(i)}(k(i))$, ignoring the effect of its receive filter on the previously allocated subchannels [c.f. (3.48)]. This choice would be optimum, if DPC was applied in the successive scheme, as in this case the choice of the receive filters would not affect the gains of previously allocated subchannels. Applying the receive filters optimum for a successive approach with DPC to scenarios, where no DPC is used, has been proposed in [36, Section 4.1.3.1]. For sum rate maximization, i.e., equal weights, the successive algorithm, where the receive filters are determined as in (3.58), has been derived in [81] and [85]. Due to its linear and successive nature it has been named LInear Successive Allocation (LISA). In [23] the receive filters for sum rate maximization without DPC are determined as in (3.58). However, as therein a fixed allocation of power to data streams is assumed, its application to the problem of weighted sum rate maximization is not straightforward.

3.4.3 User Selection based on Upper Bounds for the Weighted Sum Rate

To reduce the computational complexity even further so that for each subchannel allocation only one principal eigenvalue and the corresponding eigenvector are required, it is proposed to select the user in the ith step to maximize the corresponding upper bound for the weighted sum rate derived in Section 3.4.1. Thus, in line 6 of Algorithm 3.2 and line 17 of Algorithm 3.3 the upper bounds $R_{\mathrm{WSR,ub}}^{(i)}(k)$ are computed instead of the actual weighted sum rates $R_{\mathrm{WSR}}^{(i)}(k)$ and correspondingly the user $k(i)$ is determined to maximize this upper bound. Although it might be more intuitive to select the user in the ith step according to the maximum lower bound, it is proposed to use the upper bound due to the reduced complexity required for its computation, especially in the linear case. While the lower bound in (3.72) requires the knowledge of the matrices $\boldsymbol{B}_k^{(i)}$, which cannot be obtained by a simple update rule from the matrices $\boldsymbol{B}_k^{(i-1)}$, the upper bound in (3.73) is independent of these matrices.

3.5 Numerical Results

In order to show the performance of the near optimum low complexity algorithms proposed in this chapter a scenario with $N_{\mathrm{T}} = 4$ antennas at the transmitter and $K = 10$ users is chosen.

Each user is equipped with $r_k = 2$ receive antennas. The transmit signal $\boldsymbol{x}(t)$ propagates to each user over 4 different paths in time, where it is assumed that each channel matrix $\tilde{\boldsymbol{H}}_{k,\ell}$, $\ell = 1, \ldots, 4$, $k = 1, \ldots, K$ contains circularly symmetric entries, which are independently drawn from a complex Gaussian distribution with zero mean and unit variance.The temporal distance between two consecutively arriving symbols is equal to 160μs, i.e., $\tau_{k,m} - \tau_{k,m-1} = 160\mu$s, $m = 2, \ldots, 4$, $k = 1, \ldots, K$. The intersymbol interference is canceled by OFDM with $C = 128$ subcarrier, i.e., a cyclic prefix of sufficient length is used, and the center frequency is given by $f_c = 5$GHz. The system's bandwidth is equal to $B = 125$MHz. Figure 3.1 exhibits the weighted sum rates per subcarrier over the SNR averaged over 1000 channel realizations in such a system, where the SNR is computed from the ratio of transmit power to noise variance. 5 users in the system have twice

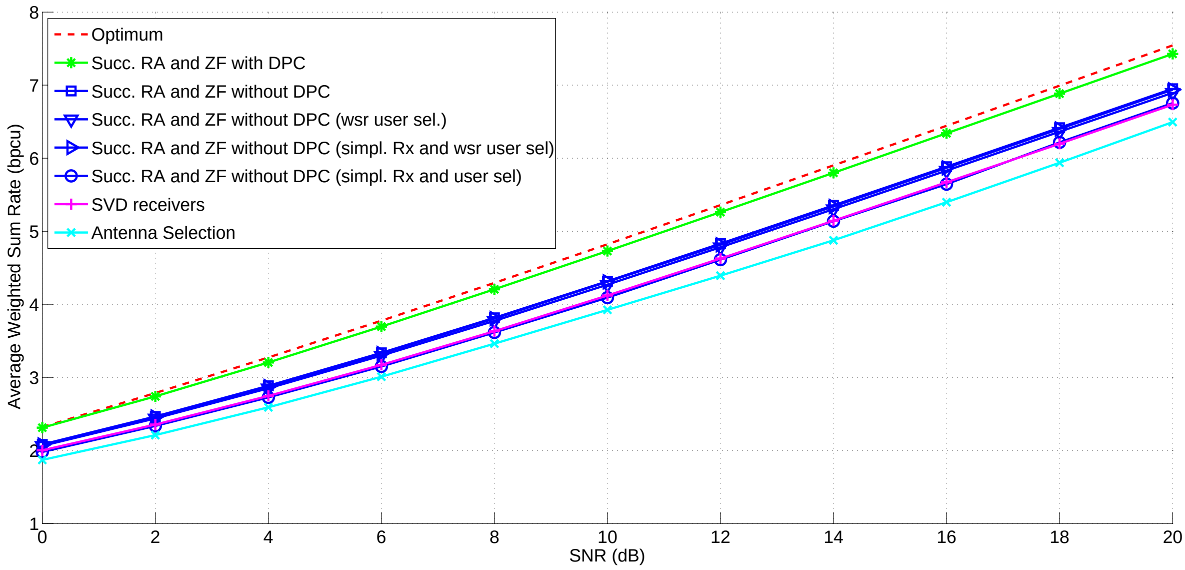

Figure 3.1: Average weighted sum rates in a system with $K = 10$ users with $r_k = 2$ receive antennas, $N_T = 4$ transmit antennas, $C = 128$ carrier and $B = 125$MHz. $\mu_1 = \ldots, \mu_5 = \frac{1}{15}$, $\mu_6 = \ldots, \mu_{10} = \frac{2}{15}$

the priority of the other users, so that

$$\mu_1 = \ldots, \mu_5 = \frac{1}{15}, \qquad \mu_6 = \ldots, \mu_{10} = \frac{2}{15},$$

where the weights are normalized so that $\sum_{k=1}^{K} \mu_k = 1$. The proposed successive resource allocation and spatial zero-forcing with DPC, denoted as "Succ. RA and ZF with DPC" in Figure 3.1, can almost achieve the same weighted sum rate as the optimum algorithm. Giving up DPC leads to further small performance losses. Less than 2dB transmit power are required to achieve the same weighted sum rate as with DPC. Furthermore it can be observed that the methods for complexity reduction proposed in Sections 3.4.2 and 3.4.3 lead to negligible performance losses[2] compared to Algorithm 3.3, which is denoted as "Succ. RA and ZF without DPC" in Figure 3.1. The adjunct "simpl. Rx" implies that the receiver filters are chosen to maximize a weaker lower bound for the sum rate according to (3.75) and "simpl. Rx and user sel" additionally includes a simplified user selection to maximize an upper bound for the weighted sum rate. For comparison the performance

[2]Recall that the method for complexity reduction from Section 3.4.1 does not lead to any performance losses at all.

of Orthogonal Frequency Division Multiple Access (OFDMA), where one carrier is exclusively occupied by one user, and for which a near-optimum algorithm is proposed in [86] is included into Figure 3.1. Additionally the average weighted sum rates achievable by using left singular vectors ("SVD receivers") and unit canonical vectors ("Antenna Selection") are shown. Thereby the receive filters are fixed a priori and the successive framework with zero-forcing of Algorithm 3.3 is used to determine the user allocation and transmit filters, where each product of receive filter and channel matrix is treated as "virtual user". However, although those methods perform only slightly inferior to the proposed algorithms, their complexity is higher, as the methods for complexity reduction from Sections 3.4.1 to 3.4.3 cannot be applied to them.

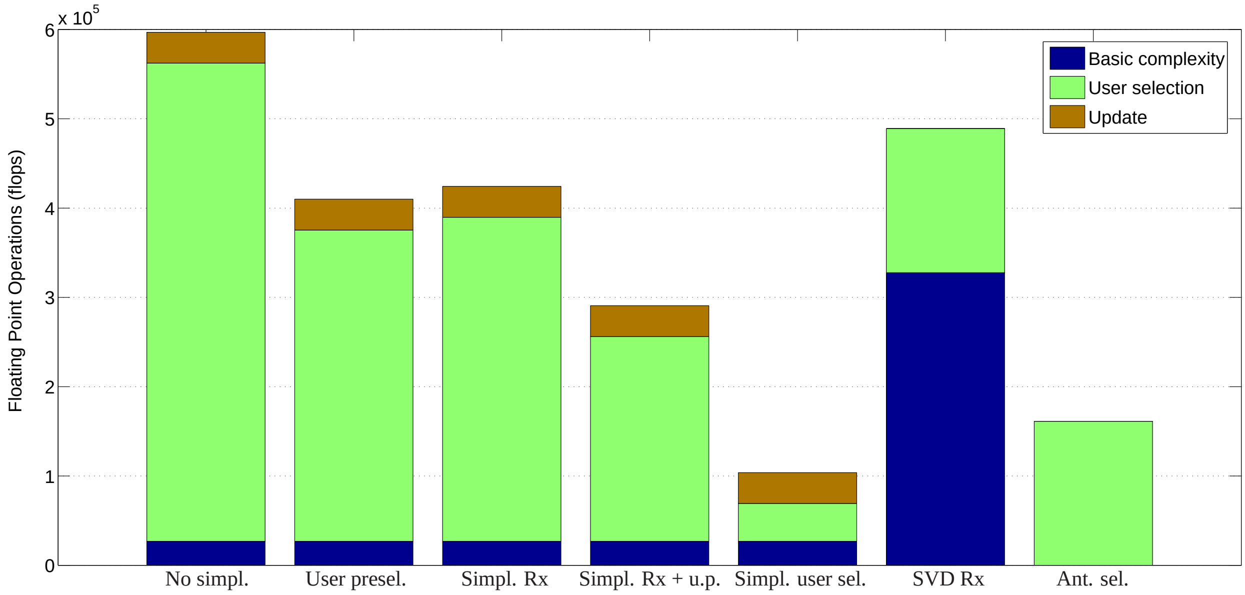

Figure 3.2: Main complexity of different algorithms in a system with $K = 10$ users with $r_k = 2$ receive antennas, $N_\mathrm{T} = 4$ transmit antennas, SNR=20dB, $C = 128$ carrier and $B = 125$MHz.$\mu_1 = \ldots, \mu_5 = \frac{1}{15}, \quad \mu_6 = \ldots, \mu_{10} = \frac{2}{15}$

Figure 3.2 compares the computational complexity of the most costly parts of the algorithms avoiding DPC with each other for the simulation parameters stated above. The computational complexity is measured in the number of Floating Point Operations (flops), where one flop corresponds to one complex addition or multiplication. The block "Basic complexity" includes the SVD's of the channel matrices $\boldsymbol{H}_k$ for the SVD receivers, which are computed according to the R-SVD method [35] (see [35, Ch. 5.4.5] for its complexity quantification). For the successive resource allocation and zero-forcing methods this block contains the complexity of the matrix-matrix multiplications $\boldsymbol{H}_k \boldsymbol{H}_k^\mathrm{H}$, from which the principal eigenvalues are computed. The number of flops required for the basic operations as matrix-matrix multiplications and additions as well as matrix inversions, is adopted from [34]. Additionally, the block-diagonal structure of the involved matrices is taken into account. The block "User selection" encompasses the complexity of all data stream allocations conducted during the algorithms. For the proposed method without DPC that includes all steps from line 14 to line 18 of Algorithm 3.3 apart from the scalar water-filling, which complexity can be neglected compared to the other parts. This block therefore contains the number of flops required for the computation and the inversion of the matrices $\boldsymbol{B}_k^{(i)}$, the matrix-matrix products $\boldsymbol{B}_k^{(i),-1}\boldsymbol{A}_k^{(i)}$, the computation of its principal eigenvalues and the corresponding eigenvectors, and the computation of the terms $\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{H}_k\boldsymbol{t}_{k,j}^{(i-1)}\boldsymbol{t}_{k,j}^{(i-1),\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{g}_k^{(i)}$. For the determination of the principal eigenvalue and the corresponding eigenvector the power method [35, Ch. 7.3.1] is used, where the number of iterations required by this method has been acquired from averaging over the number of iterations obtained from simulations with the same parameters as in Figure 3.1 at 20dB. Note that for reasons of fairness, the complexity of the iterative power method has been considered here, although for the special case of $r_k = 2$ simpler ways for determining the maximum eigenvalue and the corresponding eigenvector exist. When the simplified receivers according to (3.75) are used, the computation of the matrices $\boldsymbol{B}_k^{(i)}$, its inversions and its multiplications with the matrices $\boldsymbol{A}_k^{(i)}$ are not necessary and consequently not considered in the corresponding blocks. Regarding the proposed user preselection, which does not incur any performance losses, the average number of preselected users has been determined with the simulation setup from Figure 3.1. In case of a priori fixed receive filters, the user selection is a little bit different. As each product of receive filter and channel matrix is treated as different "virtual" user, the weighted sum rate must be computed $\sum_{k=1}^{K} r_k$ instead of K times. In turn, no eigenvalues and eigenvectors are needed for the user selection. Furthermore, for those algorithms the complexity of the update of the matrices $\boldsymbol{A}_k^{(i)}$ according to (3.65) is not necessary. That is because in this case for the sum rate computation only the products $\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{A}_k^{(i)}\boldsymbol{g}_k^{(i)}$ are required. For fixed receivers the products $\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{H}_k$ can be replaced by $\hat{\boldsymbol{h}}_{k'}^{\mathrm{H}}$, where $\hat{\boldsymbol{h}}_{\boldsymbol{k}'}^{\mathrm{H}}$ denotes the effective channel of the k'th virtual user. Consequently the quadratic forms $\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{A}_k^{(i)}\boldsymbol{g}_k^{(i)}$ can be obtained by inner vector products according to

$$\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{A}_k^{(i)}\boldsymbol{g}_k^{(i)} = \boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\mathrm{lin}}^{(i)}\boldsymbol{g}_k^{(i)}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{g}_k^{(i)} = \hat{\boldsymbol{h}}_{k'}^{\mathrm{H}}\hat{\boldsymbol{P}}_{\mathrm{lin}}^{(i-1)}\hat{\boldsymbol{h}}_{k'} - \hat{\boldsymbol{h}}_{k'}^{\mathrm{H}}\boldsymbol{t}_{k(i-1),d_{k(i-1)}}^{(i-1)}\boldsymbol{t}_{k(i-1),d_{k(i-1)}}^{(i-1),\mathrm{H}}\hat{\boldsymbol{h}}_{k'}$$

[c.f. (3.63)]. As by contrast, the algorithms proposed in this chapter require the explicit knowledge of the matrices $\boldsymbol{A}_k^{(i)}$, the complexity required for their updates is subsumed in the bar "Update".

As shown in Figure 3.2, without further measures for complexity reduction, denoted as "No simpl." in Figure 3.2, the proposed method is more complex than using SVD receivers, which complexity is visualized in the second bar from the right denoted as "SVD Rx". However, as soon as the receive filters are determined from an eigenvalue instead of a generalized eigenvalue problem, the corresponding complexity depicted in the third bar from the left ("Simpl Rx") is already lower than that of using SVD receivers, although the performance of the latter is on average worse than that of the proposed methods. Applying the user preselection described in Section 3.4.1 reduces the computational complexity shown with the bars "User presel." and "Simpl. Rx + u.p. ", of the proposed approaches by 31.5% without performance loses. This way, the computational complexity is still higher than that of antenna selection ("Ant. sel."), which exhibits the same complexity as SVD receivers without considering the number of flops required for the SVDs. Nevertheless, by conducting the user selection based on the upper bound for weighted sum rate and with simplified receive filters, which complexity bar is labelled "Simpl. user sel." in Figure 3.2, antenna selection is outperformed in terms of average weighted sum rate and computational complexity.

For $K = 2$ users the rate regions achievable with the different algorithms can be visualized, as it is done in Figure 3.3 for an SNR of 20dB. The rate regions have been obtained by varying the weights of user 1 between 0 and 1, setting the weight of user 2 to $1-\mu_1$ and running the algorithms for weighted sum rate maximization. One channel realization is used, where as before, the entries of the $L_k = 4$, $k = 1, 2$, temporal channel matrices for each user are independently drawn from a Gaussian distribution with zero mean and unit variance and the temporal difference between two consecutive paths is equal to 160μs. The OFDM system has $C = 128$ subcarrier at a center fre-

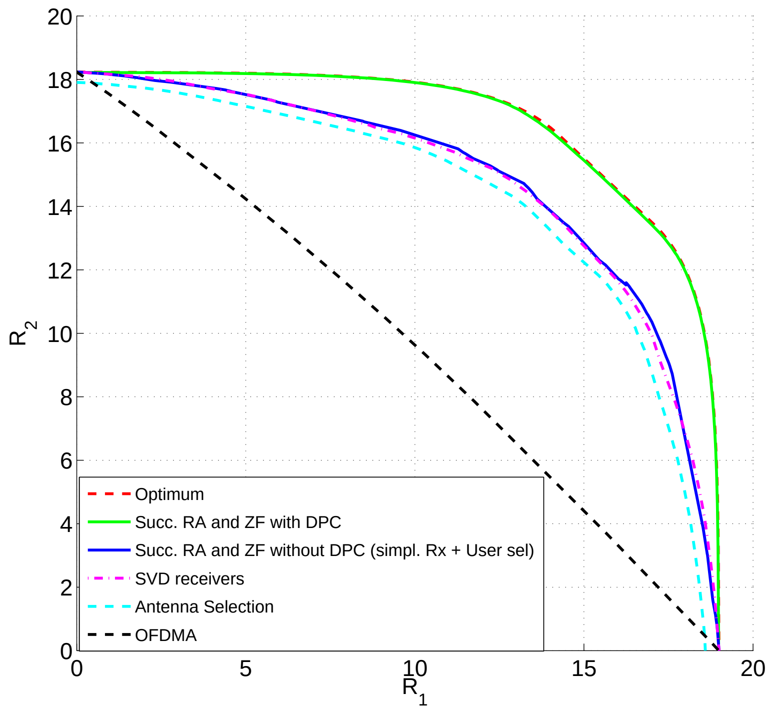

Figure 3.3: Rate regions for $K = 2$ users with $r_k = 2$ receive antennas, $N_\mathrm{T} = 4$ transmit antennas, $C = 128$, $B = 125$MHz, SNR$= 20$dB.

quency of 5 GHz and a bandwidth of $B = 125$MHz is available. The region achievable with DPC and successive resource allocation and spatial zero-forcing almost coincides with the optimum capacity region. The approaches without DPC are close to each other, where for a better visibility from the proposed methods only the variant performing worst with simplified receivers and user selection has been included. It performs negligible worse than SVD receivers only at some parts of the rate region and outperforms antenna selection anywhere on the rate region. Note that antenna selection is the only method that does not use SVD receivers in case $\mu_1 = 1$ or $\mu_2 = 1$, which is why it is worse than all other methods at the borders of the rate region.

4. Quality of Service Constrained Utility Maximization in the MIMO Broadcast Channel

Although, as mentioned in the previous chapter, weighted sum rate maximization is an important tool for considering requirements of higher layers in the communication system, in many cases a priori finding the appropriate weights fulfilling the demands in the system turns out to be a non-trivial problem. Assigning a higher weight to a certain user than to another user for example, does not automatically imply that this user attains a higher rate in the end. Additionally, some mobile applications like video streams, require guaranteed minimum transmission rates to function properly. For this reason Quality of Service (QoS) constraints are introduced in this section and considered for utility maximization. First the general problem setup will be explained and three popular QoS constrained utility maximization problems will be introduced in Section 4.1, before their optimum solution via the dual problem will be given in Section 4.2. The remainder of the chapter efficient is dedicated to near optimum approaches. After an overview of state-of-the-art near optimum algorithms in Section 4.3, the concepts of spatial zero-forcing and successive resource allocation will be applied to the QoS constrained optimization problems in Section 4.4, before the chapter is concluded with further complexity reductions in Section 4.5 and numerical results in Section 4.6.

4.1 Problem Setup

In general a QoS constrained utility maximization problem can be written as

$$\begin{aligned} &\max_{\{\boldsymbol{W}_k\}_{k=1,\dots,K},\boldsymbol{r}} u\left(\boldsymbol{W}_1,\dots,\boldsymbol{W}_K,\boldsymbol{r}\right), \\ &\text{s.t. } \boldsymbol{h}_1\left(\boldsymbol{r}\right) \leq \boldsymbol{0}_{c_1,1}, \quad \boldsymbol{h}_2\left(\boldsymbol{r}\right) = \boldsymbol{0}_{c_2,1}, \quad \{\boldsymbol{W}_1,\dots,\boldsymbol{W}_K\} \in \mathcal{C}_{\mathrm{P}}, \\ &\boldsymbol{r} \in \mathcal{P}\left(\boldsymbol{H}_1^{\mathrm{H}},\dots,\boldsymbol{H}_K^{\mathrm{H}},\boldsymbol{W}_1,\dots,\boldsymbol{W}_K\right) \end{aligned} \tag{4.1}$$

where the maximization is carried out with respect to the rates $\boldsymbol{r} = [R_1,\dots,R_K]^{\mathsf{T}}$ and the covariance matrices $\boldsymbol{W}_k$ in the dual MAC, from which the transmit covariance matrices in the broadcast channels can be obtained with the duality transformations from [5]. $u\left(\boldsymbol{W}_1,\dots,\boldsymbol{W}_K,\boldsymbol{r}\right)$, $\mathcal{C}_{\mathrm{P}}$, $\boldsymbol{h}_1\left(R_1,\dots,R_K\right) \in \mathbb{C}^{c_1}$ and $\boldsymbol{h}_2\left(R_1,\dots,R_K\right) \in \mathbb{C}^{c_2}$ denote the utility function, the constraint set on the covariance matrices $\boldsymbol{W}_1,\dots,\boldsymbol{W}_K$, the vector of QoS inequality constraint functions and the vector of QoS equality constraint functions, respectively. c_1 and c_2 are the number of inequality and equality constraints, respectively. Along the lines of [87, Problem 3.1], the relation between the rates $R_1,\dots,R_K$ and the covariance matrices is considered by the set $\mathcal{P}\left(\boldsymbol{H}_1^{\mathrm{H}},\dots,\boldsymbol{H}_K^{\mathrm{H}},\boldsymbol{W}_1,\dots,\boldsymbol{W}_K\right)$. It contains the rate region of the MAC achievable with fixed covariance matrices $\boldsymbol{W}_1,\dots,\boldsymbol{W}_K$ and is given by

$$\begin{aligned} &\mathcal{P}\left(\boldsymbol{H}_1^{\mathrm{H}},\dots,\boldsymbol{H}_K^{\mathrm{H}},\boldsymbol{W}_1,\dots,\boldsymbol{W}_K\right) = \\ &\quad = \left\{\boldsymbol{r} \,\middle|\, \sum_{i\in\mathcal{S}} R_i \leq \log_2\left|\boldsymbol{I}_{CN_{\mathrm{T}}} + \sum_{i\in\mathcal{S}} \boldsymbol{H}_i^{\mathrm{H}}\boldsymbol{W}_i\boldsymbol{H}_i\right|, \forall\mathcal{S}\subseteq\{1,\dots,K\}, R_k \geq 0, \forall k\right\}, \end{aligned}$$

which can be derived from [13, Theorem 15.3.6] and the fact that the input symbols $\boldsymbol{x}_k[n]$ and the noise are multivariate Gaussian distributed. $\mathcal{P}\left(\boldsymbol{H}_1^{\mathrm{H}}, \ldots, \boldsymbol{H}_K^{\mathrm{H}}, \boldsymbol{W}_1, \ldots, \boldsymbol{W}_K\right)$ is a polyhedron with $2^K + 1$ vertices and it can be shown that it fulfills the properties of a polymatroid as defined in [27, Definition 3.1]. The union of all polymatroids, whose covariance matrices fulfill the sum power constraint P_{Tx} leads to the capacity region of the broadcast channel [88, Chapter 2.4], i.e.

$$\mathcal{C}\left(\boldsymbol{H}_1^{\mathrm{H}}, \ldots, \boldsymbol{H}_K^{\mathrm{H}}, P_{\mathrm{Tx}}\right) = \bigcup_{\sum_{k=1}^{K} \mathrm{tr}(\boldsymbol{W}_k) \leq P_{\mathrm{Tx}}} \mathcal{P}\left(\boldsymbol{H}_1^{\mathrm{H}}, \ldots, \boldsymbol{H}_K^{\mathrm{H}}, \boldsymbol{W}_1, \ldots, \boldsymbol{W}_K\right).$$

Here, three popular QoS constrained utility maximization problems will be considered, although the presented framework can also be applied to other problems with a concave utility function and a convex constraint set. In the following these problems will be introduced by explaining the utility functions, the power constraint set and the constraint functions in (4.1).

- **Weighted Sum Rate Maximization with Minimum Rate Requirements**
  The utility function is the weighted sum of the user's rates and each user's rate must not be lower than a predefined minimum rate, where $R_{k,\min}$ denotes the minimum required rate for user k. The average sum transmit power must not exceed P_{Tx} and all transmit covariance must be positive semidefinite. Thus for the weighted sum rate maximization under minimum rate requirements the functions and sets in (4.1) read as

$$u\left(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_K, \boldsymbol{r}\right) = \boldsymbol{\mu}^{\mathsf{T}} \boldsymbol{r} \quad \text{with } \boldsymbol{\mu} = [\mu_1, \ldots, \mu_K]^{\mathsf{T}}$$

$$\boldsymbol{h}_1\left(\boldsymbol{r}\right) = \begin{bmatrix} R_{1,\min} \\ \vdots \\ R_{K,\min} \end{bmatrix} - \begin{bmatrix} R_1 \\ \vdots \\ R_K \end{bmatrix} = \boldsymbol{r}_{\min} - \boldsymbol{r}$$

$$\boldsymbol{h}_2\left(\boldsymbol{r}\right) = \boldsymbol{0}_{c_2,1}$$

$$\mathcal{C}_{\mathrm{P}} = \left\{ \boldsymbol{W}_1 \in \mathbb{C}^{Cr_1 \times Cr_1}, \ldots, \boldsymbol{W}_K \in \mathbb{C}^{Cr_K \times Cr_K} \,\middle|\, \boldsymbol{W}_k \succeq 0, \forall k = 1, \ldots, K, \sum_{k=1}^{K} \mathrm{tr}(\boldsymbol{W}_k) \leq P_{\mathrm{Tx}} \right\}.$$

  An algorithm to solve this specific QoS constrained utility maximization optimally has been proposed in [8]. It is based on an iterative search for appropriate weight vectors with which a weighted sum rate maximization without minimum rate constraints leads to the same user rates as with those constraints. In each iteration the weights of all except one user are kept fixed and for that user its weight is determined by bisection so that its minimum rate requirement is fulfilled with equality. Obviously, each iteration requires the repeated solution of a weighted sum rate maximization by Algorithm 3.1. Using the ellipsoid method [89] as presented in the following, has been proposed by the same authors in [90]. Clearly, the problem at hand can be infeasible, as the transmit power constraint can be to low to fulfill all minimum rate requirements. Nevertheless, the algorithm presented in the next section is able to detect such a case.

- **Sum Rate Maximization with Relative Rate Constraints**
  With this problem the resulting rates of each pair of users must fulfill predefined ratios. Choosing without loss of generality user 1 to be the reference user, these ratios are a priori given by the values $\rho_k = \frac{R_k}{R_1}$ for $k = 2, \ldots, K$. The utility function is the sum rate, which is equivalent to using the rate of user 1 as utility due to the fixed ratios ρ_k. Additionally, a sum power

constraint is imposed. Hence, the utility function and the constraints can be stated as

$$
\begin{aligned}
&u\left(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_K, \boldsymbol{r}\right) = R_1 \\
&\boldsymbol{h}_1\left(\boldsymbol{r}\right) = \boldsymbol{0}_{c_1,1} \\
&\boldsymbol{h}_2\left(\boldsymbol{r}\right) = \begin{bmatrix} R_1 \\ \vdots \\ R_1 \end{bmatrix} - \begin{bmatrix} \frac{R_2}{\rho_2} \\ \vdots \\ \frac{R_K}{\rho_K} \end{bmatrix} \\
&\mathcal{C}_{\mathrm{P}} = \left\{ \boldsymbol{W}_1 \in \mathbb{C}^{Cr_1 \times Cr_1}, \ldots, \boldsymbol{W}_K \in \mathbb{C}^{Cr_K \times Cr_K} \,\middle|\, \boldsymbol{W}_k \succeq 0, \forall k = 1, \ldots, K, \sum_{k=1}^{K} \mathrm{tr}(\boldsymbol{W}_k) \leq P_{\mathrm{Tx}} \right\}.
\end{aligned}
$$

With such a problem the intersection point of a line through the origin and the point $[1, \rho_2, \ldots, \rho_K]$ and the boundary of the capacity region achievable with the transmit power P_{Tx} is determined. Such a problem occurs for instance in [91, Section 3.4] and is also known as rate balancing. The optimum solution has been proposed by Lee in Jindal in [9] and fits perfectly into the framework presented in the next section.

- **Sum Power Minimization with Minimum Rate Requirements**
  In this case the users' minimum rate requirements should be achieved with the minimum possible transmit power. Thus, the utility is the negative sum of the traces of the transmit covariance matrices, and $\mathcal{C}_p$ is the set of positive semidefinite matrices, so that

$$
\begin{aligned}
&u\left(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_K, \boldsymbol{r}\right) = -\sum_{k=1}^{K} \mathrm{tr}\left(\boldsymbol{W}_k\right) \\
&\boldsymbol{h}_1\left(\boldsymbol{r}\right) = \begin{bmatrix} R_{1,\mathrm{min}} \\ \vdots \\ R_{K,\mathrm{min}} \end{bmatrix} - \begin{bmatrix} R_1 \\ \vdots \\ R_K \end{bmatrix} = \boldsymbol{r}_{\mathrm{min}} - \boldsymbol{r} \\
&\boldsymbol{h}_2\left(\boldsymbol{r}\right) = \boldsymbol{0}_{c_2,1} \\
&\mathcal{C}_{\mathrm{P}} = \left\{ \boldsymbol{W}_1 \in \mathbb{C}^{Cr_1 \times Cr_1}, \ldots, \boldsymbol{W}_K \in \mathbb{C}^{Cr_K \times Cr_K} \,|\, \boldsymbol{W}_k \succeq 0, \forall k = 1, \ldots, K \right\}.
\end{aligned}
$$

  An algorithm attempting to solve the sum power minimization for MIMO systems has first been presented in [92], which is however not optimum. In [93] an optimum algorithm for scenarios with $\sum_{k=1}^{K} r_k \leq N_{\mathrm{T}}$ is described. For general antenna configurations it is proposed in [9] to iteratively solve the rate balancing problem with $\rho_k = \frac{R_{k,\mathrm{min}}}{R_{1,\mathrm{min}}}$ until a power constraint is found so that the minimum rate constraints are fulfilled with equality. In [8] Wunder and Michel solve the power minimization problem in a way to the one they propose to solve the weighted sum rate maximization under minimum rate constraints by alternating adjustments of the users' weights. Solving the power minimization problem with the framework to be presented next corresponds to using the algorithm from [87].

4.2 Optimum Algorithm

In this section a general framework for the optimum solution to the QoS constrained utility maximization problems presented in the last section will be given. In Section 4.2.1 the solution via the dual problem will be described, where some algorithmic details are given in Section 4.2.2.

4.2.1 Solution of the Dual Problem

For all instances of the QoS constrained utility maximization mentioned in the previous section the constraint set on the variables $\boldsymbol{W}_1, \ldots, \boldsymbol{W}_K, \boldsymbol{r}$ is convex. That can be proofed by showing that any convex combination of two feasible variable sets $\boldsymbol{W}_1^{(1)}, \ldots, \boldsymbol{W}_K^{(1)}, \boldsymbol{r}^{(1)}$ and $\boldsymbol{W}_1^{(2)}, \ldots, \boldsymbol{W}_K^{(2)}, \boldsymbol{r}^{(2)}$ is again feasible. This is very simple for the inequality and equality constraint functions and the set $\mathcal{C}_\mathrm{P}$, which are all linear in the rates and the covariance matrices, respectively for the presented optimization problems. The fact that

$$\lambda \boldsymbol{r}^{(1)} + (1-\lambda)\boldsymbol{r}^{(2)} \in \mathcal{P}\left(\boldsymbol{H}_1^\mathrm{H}, \ldots, \boldsymbol{H}_K^\mathrm{H}, \lambda \boldsymbol{W}_1^{(1)} + (1-\lambda)\boldsymbol{W}_1^{(2)}, \ldots, \boldsymbol{W}_K^{(1)} + (1-\lambda)\boldsymbol{W}_K^{(2)}\right)$$

$\forall \lambda, 0 \leq \lambda \leq 1$ can be shown via the concavity of the logarithm in the determinant of a positive definite Hermitian matrix [94, Theorem 7.6.7]. As additionally all utility functions introduced above are concave in the rates and covariance matrices, Problem (4.1) becomes a convex optimization problem for the considered QoS maximizations. To obtain an algorithm for its solution the concept of Lagrangian duality [33, Chapter 6] will be applied. The dual problem of (4.1) with respect to the rate equality and inequality constraints is given by

$$\min_{\boldsymbol{\theta}_1 \geq \boldsymbol{0}, \boldsymbol{\theta}_2} \sup_{\substack{\boldsymbol{r} \in \mathcal{P}\left(\boldsymbol{H}_1^\mathrm{H}, \ldots, \boldsymbol{H}_K^\mathrm{H}, \boldsymbol{W}_1, \ldots, \boldsymbol{W}_K\right) \\ \boldsymbol{W}_1, \ldots, \boldsymbol{W}_K \in \mathcal{C}_\mathrm{P}}} \left\{ u\left(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_K, \boldsymbol{r}\right) - \boldsymbol{\theta}_1^\mathsf{T} \boldsymbol{h}_1(\boldsymbol{r}) - \boldsymbol{\theta}_2^\mathsf{T} \boldsymbol{h}_2(\boldsymbol{r}) \right\}. \tag{4.2}$$

The sets $\mathcal{C}_\mathrm{P}$ and $\mathcal{P}\left(\boldsymbol{H}_1^\mathrm{H}, \ldots, \boldsymbol{H}_K^\mathrm{H}, \boldsymbol{W}_1, \ldots, \boldsymbol{W}_K\right)$ are convex, there exists a set of rates $\hat{\boldsymbol{r}}$, for which $\boldsymbol{h}_1(\hat{\boldsymbol{r}}) < \boldsymbol{0}_{c_1,1}$ and $\boldsymbol{h}_2(\hat{\boldsymbol{r}}) = \boldsymbol{0}_{c_2,1}$ provided that the problem is feasible and $\boldsymbol{0}_{c_2,1} \in \mathrm{int}\{\boldsymbol{h}_2(\boldsymbol{r})\}$. Thus, the conditions for strong duality from [33, Theorem 6.2.4] are satisfied and the duality gap is zero, i.e., the optimum in (4.2) is identical to the optimum in the primal problem (4.1). For this reason the solution of (4.2) will be pursued now. For notational convenience Problem (4.2) is reformulated as

$$\min_{\boldsymbol{\theta}} \max_{\substack{\boldsymbol{r} \in \mathcal{P}\left(\boldsymbol{H}_1^\mathrm{H}, \ldots, \boldsymbol{H}_K^\mathrm{H}, \boldsymbol{W}_1, \ldots, \boldsymbol{W}_K\right) \\ \boldsymbol{W}_1, \ldots, \boldsymbol{W}_K \in \mathcal{C}_\mathrm{P}}} \left\{ u\left(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_K, \boldsymbol{r}\right) - \boldsymbol{\theta}^\mathsf{T} \boldsymbol{h}(\boldsymbol{r}) \right\} = \min_{\boldsymbol{\theta}} g(\boldsymbol{\theta}), \quad \text{s.t. } \boldsymbol{C}_\mathrm{c} \boldsymbol{\theta} \leq \boldsymbol{d}_\mathrm{c}, \tag{4.3}$$

with

$$\boldsymbol{\theta} = \left[\boldsymbol{\theta}_1^\mathsf{T} \boldsymbol{\theta}_2^\mathsf{T}\right]^\mathsf{T}, \qquad \boldsymbol{h}(\boldsymbol{r}) = \left[\boldsymbol{h}_1^\mathsf{T}(\boldsymbol{r}) \boldsymbol{h}_2^\mathsf{T}(\boldsymbol{r})\right]^\mathsf{T}$$

and $g(\boldsymbol{\theta})$ denotes the dual function. $\boldsymbol{C}_\mathrm{c}$ and $\boldsymbol{d}_\mathrm{c}$ resemble the linear constraints on the dual variables in $\boldsymbol{\theta}$ and are given by

$$\boldsymbol{C}_\mathrm{c} = \begin{cases} -\boldsymbol{I}_K, & \text{minimum rate constraints} \\ \begin{bmatrix} -\boldsymbol{I}_{K-1} \\ \boldsymbol{1}_{K-1}^\mathsf{T} \end{bmatrix}, & \text{rate balancing} \end{cases}, \boldsymbol{d}_\mathrm{c} = \begin{cases} \boldsymbol{0}_{K,1}, & \text{minimum rate constraints} \\ \begin{bmatrix} \boldsymbol{0}_{K-1,1} \\ 1 \end{bmatrix}, & \text{rate balancing} \end{cases}. \tag{4.4}$$

The supremum operator in (4.2) has been replaced by the minimum operator as the constraint set is compact, i.e., closed and bounded. The constraints on $\boldsymbol{\theta}$ introduced above for the rate balancing problem do not change the result of the optimization problem, but reduce the search space for the optimum value of $\boldsymbol{\theta}$, a property that will be helpful for the algorithm to be described next. That is because for the rate balancing problem the dual function reads as

$$g(\boldsymbol{\theta}) = \max_{\substack{\boldsymbol{r} \in \mathcal{P}\left(\boldsymbol{H}_1^\mathrm{H}, \ldots, \boldsymbol{H}_K^\mathrm{H}, \boldsymbol{W}_1, \ldots, \boldsymbol{W}_K\right) \\ \boldsymbol{W}_1, \ldots, \boldsymbol{W}_K \in \mathcal{C}_\mathrm{P}}} \left\{ \left(1 - \sum_{k=2}^{K} [\boldsymbol{\theta}_2]_{k-1}\right) R_1 + \sum_{k=2}^{K} [\boldsymbol{\theta}_2]_{k-1} \frac{R_k}{\rho_k} \right\} \tag{4.5}$$

In case $[\boldsymbol{\theta}_2]_{k+1} < 0$ the maximum operator sets the corresponding rate R_k to zero, as any other feasible value for R_k leads to a lower value of the objective function. Thus, $[\boldsymbol{\theta}_2]_{k+1} < 0$ has the same effect as $[\boldsymbol{\theta}_2]_{k+1} = 0$, which is considered by the constraint function in (4.4). The same argument holds for $1 - \sum_{j=1}^{K-1} [\boldsymbol{\theta}_2]_j < 0$ and the corresponding rate R_1. The same result has also been stated in [9, Theorem 1].

As every dual problem, (4.3) constitutes a convex optimization problem and the ellipsoid method [89], [95] will be used in the following for its solution. In each step the range of possible solutions is confined to an ellipsoid, where the volume covered by the ellipsoid shrinks with every iteration until convergence is achieved. The ellipsoid $\mathcal{E}^{(i)}$ after the ith iteration is characterized by a center $\boldsymbol{\theta}^{(i)}$ and a matrix $\boldsymbol{E}^{(i)}$ so that

$$\mathcal{E}^{(i)} = \left\{ \boldsymbol{\theta} \,\middle|\, \left(\boldsymbol{\theta} - \boldsymbol{\theta}^{(i)}\right)^{\mathsf{T}} \boldsymbol{E}^{(i)^{-1}} \left(\boldsymbol{\theta} - \boldsymbol{\theta}^{(i)}\right) \leq 1 \right\},$$

where the determination of the initial ellipsoid $\mathcal{E}^{(0)}$ is described in the next section. In each iteration first the halfspace

$$\mathcal{H}^{(i)} = \left\{ \boldsymbol{\theta} \,\middle|\, \left(\boldsymbol{\theta}^{\mathsf{T}} - \boldsymbol{\theta}^{(i),\mathsf{T}}\right) \boldsymbol{a}^{(i)} \leq b^{(i)} \right\}, \tag{4.6}$$

is determined. It is characterized by the vector $\boldsymbol{a}^{(i)}$ and the scalar $b^{(i)}$ and will be used to shrink the size of the ellipsoid. For that purpose it should not contain the ellipsoid $\mathcal{E}^{(i)}$ completely but still the optimum solution. In case $\boldsymbol{\theta}^{(i)}$ is feasible, i.e., $\boldsymbol{C}_{\mathrm{c}}\boldsymbol{\theta}^{(i)} \leq \boldsymbol{d}_{\mathrm{c}}$, the fact that $-\boldsymbol{h}\left(\hat{\boldsymbol{r}}^{(i)}\right)$ is a subgradient of $g(\boldsymbol{\theta})$ at $\boldsymbol{\theta} = \boldsymbol{\theta}^{(i)}$ is exploited , where

$$\left\{ \hat{\boldsymbol{r}}^{(i)}, \hat{\boldsymbol{W}}_1^{(i)}, \ldots, \hat{\boldsymbol{W}}_K^{(i)} \right\} = \underset{\substack{\boldsymbol{r} \in \mathcal{P}\left(\boldsymbol{H}_1^{\mathrm{H}}, \ldots, \boldsymbol{H}_K^{\mathrm{H}}, \boldsymbol{W}_1, \ldots, \boldsymbol{W}_K\right) \\ \boldsymbol{W}_1, \ldots, \boldsymbol{W}_K \in \mathcal{C}_{\mathrm{P}}}}{\operatorname{argmax}} \left\{ u\left(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_K, \boldsymbol{r}\right) - \boldsymbol{\theta}^{(i),\mathsf{T}} \boldsymbol{h}(\boldsymbol{r}) \right\}. \tag{4.7}$$

That is because

$$\begin{aligned} g(\boldsymbol{\theta}) &\geq u\left(\hat{\boldsymbol{W}}_1^{(i)}, \ldots, \hat{\boldsymbol{W}}_K^{(i)}, \hat{\boldsymbol{r}}^{(i)}\right) - \boldsymbol{\theta}^{\mathsf{T}} \boldsymbol{h}\left(\hat{\boldsymbol{r}}^{(i)}\right) \\ &\geq g\left(\boldsymbol{\theta}^{(i)}\right) - \left(\boldsymbol{\theta}^{\mathsf{T}} - \boldsymbol{\theta}^{(i),\mathsf{T}}\right) \boldsymbol{h}\left(\hat{\boldsymbol{r}}^{(i)}\right), \quad \forall \boldsymbol{\theta} \text{ with } \boldsymbol{C}_{\mathrm{c}}\boldsymbol{\theta} \leq \boldsymbol{d}_{\mathrm{c}}, \end{aligned} \tag{4.8}$$

which corresponds to the inequality from the definition of a subgradient of a convex function at $\boldsymbol{\theta}^{(i)}$ [33, Definition 3.2.3]. Another consequence from this definition is that for all feasible $\boldsymbol{\theta}$ with $\left(\boldsymbol{\theta}^{\mathsf{T}} - \boldsymbol{\theta}^{(i),\mathsf{T}}\right) \boldsymbol{h}\left(\hat{\boldsymbol{r}}^{(i)}\right) < 0$ the dual function $g(\boldsymbol{\theta})$ will be greater than $g\left(\boldsymbol{\theta}^{(i)}\right)$, i.e., the optimum solution to will not lie in the half-space defined this way. Thus for feasible $\boldsymbol{\theta}^{(i)}$ the parameters of the halfspace $\mathcal{H}^{(i)}$ in (4.6) read as

$$\boldsymbol{a}^{(i)} = -\boldsymbol{h}\left(\hat{\boldsymbol{r}}^{(i)}\right), \qquad b^{(i)} = 0.$$

In case $\boldsymbol{\theta}^{(i)}$ violates the jth constraint, i.e., $\left[\boldsymbol{C}_{\mathrm{c}}\boldsymbol{\theta}^{(i)}\right]_j = \boldsymbol{e}_j^{\mathsf{T}} \boldsymbol{C}_{\mathrm{c}} \boldsymbol{\theta}^{(i)} =: \boldsymbol{c}_{j,\mathrm{c}}^{\mathsf{T}} \boldsymbol{\theta}^{(i)} > [\boldsymbol{d}_{\mathrm{c}}]_j$, the subgradient $\boldsymbol{h}\left((\hat{\boldsymbol{r}}^{(i)}\right)$ does not exist. Instead the halfspace $\mathcal{H}^{(i)} = \left\{ \boldsymbol{\theta} \,\middle|\, \boldsymbol{c}_{j,\mathrm{c}}^{\mathsf{T}} \boldsymbol{\theta} - [\boldsymbol{d}_{\mathrm{c}}]_j \leq 0 \right\}$ is taken, which parameters are given by

$$\boldsymbol{a}^{(i)} = \boldsymbol{c}_{j,\mathrm{c}}, \qquad b^{(i)} = [\boldsymbol{d}_{\mathrm{c}}]_j - \boldsymbol{c}_{j,\mathrm{c}}^{\mathsf{T}} \boldsymbol{\theta}^{(i)}.$$

The ellipsoid $\mathcal{E}^{(i+1)}$ is then determined to be the ellipsoid with minimum volume covering the space

$$\mathcal{E}^{(i)} \cap \mathcal{H}^{(i)}. \tag{4.9}$$

As proven in [95], this design criterion leads to the following update rules for the parameters of the ellipsoid, which are given by

$$\boldsymbol{\theta}^{(i+1)} = \boldsymbol{\theta}^{(i)} - \frac{1+\alpha n}{(n+1)\sqrt{\boldsymbol{a}^{(i),\mathsf{T}}\boldsymbol{E}^{(i)}\boldsymbol{a}^{(i)}}}\boldsymbol{E}^{(i)}\boldsymbol{a}^{(i)}$$
$$\boldsymbol{E}^{(i+1)} = \frac{n^2(1-\alpha^2)}{n^2-1}\left(\boldsymbol{E}^{(i)} - \frac{2(1+\alpha n)}{(n+1)(1+\alpha)\boldsymbol{a}^{(i),\mathsf{T}}\boldsymbol{E}^{(i)}\boldsymbol{a}^{(i)}}\boldsymbol{E}^{(i)}\boldsymbol{a}^{(i)}\boldsymbol{a}^{(i),\mathsf{T}}\boldsymbol{E}^{(i)}\right),$$

where $\alpha = 0$ in case $\boldsymbol{\theta}^{(i)}$ is feasible and

$$\alpha = \frac{-b^{(i)}}{\sqrt{\boldsymbol{a}^{(i),\mathsf{T}}\boldsymbol{E}^{(i)}\boldsymbol{a}^{(i)}}} = \frac{\boldsymbol{c}_{j,\mathrm{c}}^{\mathsf{T}}\boldsymbol{\theta}^{(i)} - [\boldsymbol{d}_{\mathrm{c}}]_j}{\sqrt{\boldsymbol{a}^{(i),\mathsf{T}}\boldsymbol{E}^{(i)}\boldsymbol{a}^{(i)}}}$$

otherwise. n denotes the number of variables contained in $\boldsymbol{\theta}$, i.e., $n = K$ for minimum rate constraints and $n = K-1$ for the rate balancing problem.
In [89] it is shown that the difference between the optimum value of the dual function and $g\left(\boldsymbol{\theta}^{(i)}\right)$ can be upper bounded by $\sqrt{\boldsymbol{h}^{\mathsf{T}}\left((\hat{\boldsymbol{r}}^{(i)}\right)\boldsymbol{E}^{(i)}\boldsymbol{h}\left((\hat{\boldsymbol{r}}^{(i)}\right)}$, i.e., the algorithm can be terminated, if this bound falls below a predefined threshold ε leading to the stopping criterion

$$\sqrt{\boldsymbol{h}^{\mathsf{T}}\left((\hat{\boldsymbol{r}}^{(i)}\right)\boldsymbol{E}^{(i)}\boldsymbol{h}\left((\hat{\boldsymbol{r}}^{(i)}\right)} \leq \varepsilon.$$

If on the other hand it is observed that $\boldsymbol{\theta}^{(i)}$ is infeasible and

$$\boldsymbol{c}_{j,\mathrm{c}}^{\mathsf{T}}\boldsymbol{\theta}^{(i)} - [\boldsymbol{d}_{\mathrm{c}}]_j - \sqrt{\boldsymbol{a}^{(i),T}\boldsymbol{E}^{(i)}\boldsymbol{a}^{(i)}} > 0$$

the problem is infeasible, which can happen, if the minimum rate requirements for the weighted sum rate maximization cannot be fulfilled with the given transmit power constraint.

4.2.2 Algorithmic Details

In this section the open issues of the algorithm to solve QoS constrained utility maximization problems presented above will be covered, namely the initialization of the ellipsoid, the computation of the subgradients $\boldsymbol{h}\left((\hat{\boldsymbol{r}}^{(i)}\right)$ and the reconstruction of the primal solution.

Initialization of the Ellipsoid

Obviously, the initial ellipsoid $\mathcal{E}^{(0)}$ has to contain the optimum Lagrange multiplier minimizing (4.3). Furthermore its parameters should be easily computable, and the volume of the ellipsoid should be small to achieve a fast convergence. For rate balancing, the parameters of the initial ellipsoid are given by [9]

$$\boldsymbol{\theta}^{(0)} = \frac{1}{K}\mathbf{1}_{K-1}$$
$$\boldsymbol{E}^{(0)} = \left(1 - \frac{1}{K}\right)\boldsymbol{I}_{K-1}.$$

For the remaining QoS constrained utility maximization problems introduced in this chapter, the optimum vector of Lagrangian multipliers $\boldsymbol{\theta}^*$ can be constrained to lie in the cuboid

$$\mathcal{C}_{\mathrm{u}} = \left\{\boldsymbol{\theta}\,|\mathbf{0}_{n,1} \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}_{\max}\right\}.$$

$\mathcal{E}^{(0)}$ is then chosen to be the ellipsoid with minimum volume completely covering $\mathcal{C}_\mathrm{u}$. Thus, the parameters of the initial ellipsoid are given by

$$\begin{aligned}\boldsymbol{\theta}^{(0)} &= \frac{1}{2}\boldsymbol{\theta}_\mathrm{max}\\ \boldsymbol{E}^{(0)} &= \mathrm{diag}\left(\left[\boldsymbol{\theta}^2_\mathrm{max}\right]_1, \ldots, \left[\boldsymbol{\theta}^2_\mathrm{max}\right]_n\right) K/4\end{aligned}$$

For the weighted sum rate maximization under minimum rate requirements, $\boldsymbol{\theta}_\mathrm{max}$ has been derived in [90] to be equal to

$$\boldsymbol{\theta}_\mathrm{max} = \sum_{j=1}^{K} \mu_j \left(R_{j,\mathrm{su}}(P_\mathrm{Tx}) - \bar{R}_j\right) \begin{bmatrix} \frac{1}{\left(\bar{R}_1 - R_{1,\mathrm{min}}\right)} \\ \vdots \\ \frac{1}{\left(\bar{R}_K - R_{K,\mathrm{min}}\right)} \end{bmatrix},$$

where the single-user rate

$$R_{j,\mathrm{su}}(P_\mathrm{Tx}) = \max_{\boldsymbol{W}_j} \log_2 \left|\boldsymbol{I}_{CN_\mathrm{T}} + \boldsymbol{H}_j^\mathrm{H}\boldsymbol{W}_j\boldsymbol{H}_j\right|, \quad \text{s.t. } \mathrm{tr}\left(\boldsymbol{W}_j\right) \leq P_\mathrm{Tx}$$

corresponds to the rate user j can achieve if it receives all available transmit power. The rates $\bar{R}_j$ must all be greater than the corresponding minimum rates $R_{j,\mathrm{min}}$ and achievable with a sum power constraint less than P_Tx. Such a rate vector can for example be achieved by solving the rate balancing problem with $\rho_k = \frac{R_{k,\mathrm{min}}}{R_{1,\mathrm{min}}}$, for $k = 2, \ldots, K$ under a sum power constraint of $P_\mathrm{Tx} - \epsilon$, where $\epsilon > 0$. For this purpose also the near optimum algorithm to be presented in Section 4.4 can be used. For the power minimization, in [87] it is proposed to determine for each user k a set of sub-optimum covariance matrices $\boldsymbol{W}_1^{(k)}, \ldots, \boldsymbol{W}_K^{(k)}$ in the following way. For each user $j \neq k$ the minimum rate requirement $R_{j,\mathrm{min}}$ has to be fulfilled with equality and the kth user's rate must be equal to $R_{k,\mathrm{min}} + 1$. A fixed decoding order of the users in the MAC is assumed and the matrices $\boldsymbol{W}_j^{(k)}$ are determined successively. First the covariance matrix of the user decoded last is computed so that its minimum rate requirement is fulfilled with equality. This matrix is kept fixed and considered in the interference terms of the other users' rates. With this simplification the covariance matrix of the user decoded second last can be computed so that its minimum rate requirement is fulfilled with equality. This process is continued until the covariance matrix of the user decoded first is computed and repeated until all possible scenarios have been considered where one user's minimum rate requirement has been increased by one. Finally, $\boldsymbol{\theta}_\mathrm{max}$ is given by

$$\boldsymbol{\theta}_\mathrm{max} = \begin{bmatrix} \sum_{j=1}^K \mathrm{tr}\left(\boldsymbol{W}_j^{(1)}\right) \\ \vdots \\ \sum_{j=1}^K \mathrm{tr}\left(\boldsymbol{W}_j^{(K)}\right) \end{bmatrix}.$$

Computation of Subgradients

The computation of the subgradients $\boldsymbol{h}\left((\hat{\boldsymbol{r}}^{(i)}\right)$ requires the solution of (4.7). When a sum power constraint is contained in $\mathcal{C}_\mathrm{P}$, Problem (4.7) is a weighted sum rate maximization, that can be solved with Algorithm 3.1. That is because for the weighted sum rate maximization under minimum rate requirements

$$g(\boldsymbol{\theta}) = \max_{\substack{\boldsymbol{r}\in\mathcal{P}\left(\boldsymbol{H}_1^{\mathrm{H}},\ldots,\boldsymbol{H}_K^{\mathrm{H}},\boldsymbol{W}_1,\ldots,\boldsymbol{W}_K\right)\\ \boldsymbol{W}_1,\ldots,\boldsymbol{W}_K\in\mathcal{C}_{\mathrm{P}}}} \left\{u\left(\boldsymbol{W}_1,\ldots,\boldsymbol{W}_K,\boldsymbol{r}\right) - \boldsymbol{\theta}^{(i),\mathsf{T}}\boldsymbol{h}(\boldsymbol{r})\right\} =$$
$$= \max_{\substack{\boldsymbol{r}\in\mathcal{P}\left(\boldsymbol{H}_1^{\mathrm{H}},\ldots,\boldsymbol{H}_K^{\mathrm{H}},\boldsymbol{W}_1,\ldots,\boldsymbol{W}_K\right)\\ \boldsymbol{W}_1,\ldots,\boldsymbol{W}_K\in\mathcal{C}_{\mathrm{P}}}} \left\{\left(\boldsymbol{\mu}^{\mathsf{T}} + \boldsymbol{\theta}_1^{(i),\mathsf{T}}\right)\boldsymbol{r}\right\} - \boldsymbol{\theta}_1^{(i),\mathsf{T}}\boldsymbol{r}_{\text{min}} \tag{4.10}$$

holds and for rate balancing this can be seen from (4.5). For the sum power minimization, Problem (4.7) reads as

$$\left\{(\hat{\boldsymbol{r}}^{(i)},\hat{\boldsymbol{W}}_1^{(i)},\ldots,\hat{\boldsymbol{W}}_K^{(i)}\right\} = \operatorname*{argmax}_{\substack{\boldsymbol{r}\in\mathcal{P}\left(\boldsymbol{H}_1^{\mathrm{H}},\ldots,\boldsymbol{H}_K^{\mathrm{H}},\boldsymbol{W}_1,\ldots,\boldsymbol{W}_K\right)\\ \boldsymbol{W}_1\succeq\mathbf{0},\ldots,\boldsymbol{W}_K\succeq\mathbf{0}}} \left\{\boldsymbol{\theta}_1^{(i),\mathsf{T}}\boldsymbol{r} - \sum_{k=1}^{K}\operatorname{tr}\left(\boldsymbol{W}_k\right)\right\}. \tag{4.11}$$

Denoting the used transmit power as $P_{\text{aux}} = \sum_{k=1}^{K}\operatorname{tr}\left(\boldsymbol{W}_k\right)$, the maximization in (4.11) can be split according to

$$\max_{P_{\text{aux}}}\left\{\max_{\substack{\boldsymbol{r}\in\mathcal{P}\left(\boldsymbol{H}_1^{\mathrm{H}},\ldots,\boldsymbol{H}_K^{\mathrm{H}},\boldsymbol{W}_1,\ldots,\boldsymbol{W}_K\right)\\ \boldsymbol{W}_1\succeq\mathbf{0},\ldots,\boldsymbol{W}_K\succeq\mathbf{0},\sum_{k=1}^{K}\operatorname{tr}(\boldsymbol{W}_k)\leq P_{\text{aux}}}} \boldsymbol{\theta}_1^{(i),\mathsf{T}}\boldsymbol{r} - P_{\text{aux}}\right\} = \max_{P_{\text{aux}}}\left\{R_{\text{WSRmax}}(P_{\text{aux}}) - P_{\text{aux}}\right\}$$

The inner maximization corresponds to a weighted sum rate maximization, which can be solved efficiently using Algorithm 3.1 and the outer maximization is conducted over the function $R_{\text{WSRmax}}(P_{\text{aux}}) - P_{\text{aux}}$, which is concave in the scalar variable P_{aux} and can therefore be solved by the bisection method [33, Chapter 8.2]. That is because $R_{\text{WSRmax}}(P_{\text{aux}})$ is concave in P_{aux} and the sum of concave functions is also concave. The first fact can be shown via

$$R_{\text{WSRmax}}\left(\lambda P_{\text{aux}}^{(1)} + (1-\lambda)P_{\text{aux}}^{(2)}\right) \geq R_{\text{wsr}}\left(\lambda\boldsymbol{W}_1^{(1)} + (1-\lambda)\boldsymbol{W}_1^{(2)},\ldots,\lambda\boldsymbol{W}_K^{(1)} + (1-\lambda)\boldsymbol{W}_K^{(2)}\right)$$
$$\geq \lambda R_{\text{wsr}}\left(\boldsymbol{W}_1^{(1)},\ldots,\boldsymbol{W}_K^{(1)}\right) + (1-\lambda)R_{\text{wsr}}\left(\boldsymbol{W}_1^{(2)},\ldots,\boldsymbol{W}_K^{(2)}\right) \tag{4.12}$$

where $R_{\text{wsr}}(\boldsymbol{W}_1,\ldots,\boldsymbol{W}_K)$ is given by (3.12) and the $\boldsymbol{W}_k^{(i)}$ are the optimum covariance matrices for the power constraint $P_{\text{aux}}^{(i)}$,i.e.,

$$\left\{\boldsymbol{W}_1^{(i)},\ldots,\boldsymbol{W}_K^{(i)}\right\} = \operatorname*{argmax}_{\boldsymbol{W}_1\succeq\mathbf{0},\ldots,\boldsymbol{W}_K\succeq\mathbf{0}} R_{\text{wsr}}(\boldsymbol{W}_1,\ldots,\boldsymbol{W}_K), \quad \text{s.t.} \sum_{k=1}^{K}\operatorname{tr}(\boldsymbol{W}_k) \leq P_{\text{aux}}^{(i)}.$$

The first inequality in (4.12) stems from the fact that the covariance matrices $\lambda\boldsymbol{W}_k^{(1)} + (1-\lambda)\boldsymbol{W}_k^{(2)}$, $k = 1,\ldots,K$ fulfill the sum power constraint $\lambda P_{\text{aux}}^{(1)} + (1-\lambda)P_{\text{aux}}^{(2)}$ but are not necessarily optimum for that sum power constraint. The second inequality directly follows from the concavity of the weighted sum rate in the covariance matrices of the dual MAC.
The bisection method requires the evaluation of the gradients

$$\frac{\partial(R_{\text{WSRmax}}(P_{\text{aux}}) - P_{\text{aux}})}{\partial P_{\text{aux}}} = \frac{\partial R_{\text{WSRmax}}(P_{\text{aux}})}{\partial P_{\text{aux}}} - 1, \tag{4.13}$$

where the power P_{opt} has to be found, so that the gradient in (4.13) becomes zero leading to the equality

$$\left.\frac{\partial R_{\text{WSRmax}}(P_{\text{aux}})}{\partial P_{\text{aux}}}\right|_{P_{\text{opt}}} = 1.$$

The derivative

$$\left.\frac{\partial R_{\text{WSRmax}}(P_{\text{aux}})}{\partial P_{\text{aux}}}\right|_{P} = \nu\left(P\right)$$

at a certain power P is given by the Lagrange multiplier from the weighted sum rate maximization corresponding to the sum power constraint P, i.e.,

$$\nu\left(P\right) = \underset{\nu}{\operatorname{argmin}} \max_{\substack{\boldsymbol{r}\in\mathcal{P}\left(\boldsymbol{H}_1^{\text{H}},\ldots,\boldsymbol{H}_K^{\text{H}},\boldsymbol{W}_1,\ldots,\boldsymbol{W}_K\right)\\ \boldsymbol{W}_1\succeq\boldsymbol{0},\ldots,\boldsymbol{W}_K\succeq\boldsymbol{0}}} \left\{\boldsymbol{\theta}_1^{(i),\mathsf{T}}\boldsymbol{r} - \nu\left(\sum_{k=1}^{K}\operatorname{tr}\left(\boldsymbol{W}_k\right) - P\right)\right\},$$

which follows from the local sensitivity analysis of a perturbed convex optimization problem [29, Section 5.6.3]. Thus, given an interval $[0; P_{\text{max}}]$ containing P_{opt}, first a weighted sum rate maximization has to be solved with sum power constraint $\frac{P_{\text{max}}}{2}$ using for example Algorithm 3.1 and the corresponding Lagrange multiplier $\nu\left(\frac{P_{\text{max}}}{2}\right)$ has to be determined. If $\nu\left(\frac{P_{\text{max}}}{2}\right)$ is greater than one, the bisection method is continued in the interval $[0; P_{\text{max}}/2]$, otherwise in the interval $[P_{\text{max}}/2; P_{\text{max}}]$. Note that in comparison with the algorithm for power minimization from [9], where this problem is solved via an iterative application of the rate balancing problem, the presented approach requires only one execution of the ellipsoid method, which shows a slow convergence behavior, whereas the method from [9] relies on an iterative application of the ellipsoid method.

Reconstruction of the Primal Solution

After the convergence of the ellipsoid method, it is still necessary to compute the rates $\tilde{\boldsymbol{r}}$ and the covariance matrices $\tilde{\boldsymbol{W}}_1,\ldots,\tilde{\boldsymbol{W}}_K$ optimum for (4.1) from the solution of the dual problem. Let

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta}}{\operatorname{argmin}}\, g(\boldsymbol{\theta}), \quad \text{s.t. } \boldsymbol{C}_{\text{c}}\boldsymbol{\theta} \leq \boldsymbol{d}_{\text{c}}$$

be the optimum solution of the dual problem and

$$\begin{aligned}\{\boldsymbol{r}^*, \boldsymbol{W}_1^*, \ldots, \boldsymbol{W}_K^*\} &= \underset{\substack{\boldsymbol{r}\in\mathcal{P}\left(\boldsymbol{H}_1^{\text{H}},\ldots,\boldsymbol{H}_K^{\text{H}},\boldsymbol{W}_1,\ldots,\boldsymbol{W}_K\right)\\ \boldsymbol{W}_1,\ldots,\boldsymbol{W}_K\in\mathcal{C}_{\text{P}}}}{\operatorname{argmax}} \left\{u\left(\boldsymbol{W}_1,\ldots,\boldsymbol{W}_K,\boldsymbol{r}\right) - \boldsymbol{\theta}^{*,\mathsf{T}}\boldsymbol{h}(\boldsymbol{r})\right\} = \\ &= \underset{\substack{\boldsymbol{r}\in\mathcal{P}\left(\boldsymbol{H}_1^{\text{H}},\ldots,\boldsymbol{H}_K^{\text{H}},\boldsymbol{W}_1,\ldots,\boldsymbol{W}_K\right)\\ \boldsymbol{W}_1,\ldots,\boldsymbol{W}_K\in\mathcal{C}_{\text{P}}}}{\operatorname{argmax}} \left\{\hat{\boldsymbol{\mu}}^{\mathsf{T}}(\boldsymbol{\theta}^*)\boldsymbol{r} - \eta\sum_{k=1}^{K}\operatorname{tr}\left(\boldsymbol{W}_k\right)\right\},\end{aligned} \tag{4.14}$$

be the rates and covariance matrices leading to the maximum of the dual function, where

$$\hat{\boldsymbol{\mu}}(\boldsymbol{\theta}^*) = \boldsymbol{\mu} + \boldsymbol{\theta}^*, \qquad \eta = 0$$

for the weighted sum rate maximization with minimum rate requirements [c.f. (4.10)],

$$\hat{\boldsymbol{\mu}}(\boldsymbol{\theta}^*) = \begin{bmatrix} 1 - \sum_{k=1}^{K-1}\left[\boldsymbol{\lambda}^*\right]_j \\ \boldsymbol{\theta}^* \end{bmatrix}, \qquad \eta = 0$$

for rate balancing [c.f. (4.5)] and

$$\hat{\boldsymbol{\mu}}(\boldsymbol{\theta}^*) = \boldsymbol{\theta}^*, \qquad \eta = 1$$

for the power minimization with minimum rate requirements [c.f. (4.11)]. In case $\{\boldsymbol{r}^*, \boldsymbol{W}_1^*, \ldots, \boldsymbol{W}_K^*\}$ is a unique solution to (4.14), the maximizing variables to the primal problem are identical to the optimum variables of the dual problem, i.e., $\tilde{\boldsymbol{r}} = \boldsymbol{r}^*$ and $\tilde{\boldsymbol{W}}_k = \boldsymbol{W}_k^*$, $\forall k = 1, \ldots, K$. As it has been shown in the previous section, the rates $\boldsymbol{r}^*$ and the covariance matrices $\boldsymbol{W}_1^*, \ldots, \boldsymbol{W}_K^*$ result from a weighted sum rate maximization. It can therefore easily be detected whether the solution of the dual problem is unique. In case all entries in the vector $\hat{\boldsymbol{\mu}}\left(\boldsymbol{\theta}^*\right)$ are different[1], there is only one single optimum decoding order in the dual MAC, which is given by the order of the weights [see (3.10)] and thus only one set of optimum rates and covariance matrices. This can be shown as in the proof of Lemma 3.3 in [87]. Otherwise, no unique decoding order results from (3.10). Any decoding order within the users having equal weights and using the same covariance matrices $\boldsymbol{W}_1^*, \ldots, \boldsymbol{W}_K^*$ for all decoding orders leads to different rate vectors but the same value of the dual function. Additionally, any convex combination of the rate vectors obtained with different decoding orders is also optimum with respect to the dual function. This implies that the dual function is maximized on a time-sharing region of the capacity region. Those regions correspond to these faces of the polytope $\mathcal{P}(\boldsymbol{H}_1^{\mathrm{H}}, \ldots, \boldsymbol{H}_K^{\mathrm{H}}, \boldsymbol{W}_1^*, \ldots, \boldsymbol{W}_K^*)$, where the constraints $\sum_{i\in\mathcal{S}_j} R_i \leq \log_2 \left| \boldsymbol{I}_{CN_{\mathrm{T}}} + \sum_{i\in\mathcal{S}_j} \boldsymbol{H}_i^{\mathrm{H}} \boldsymbol{W}_i^* \boldsymbol{H}_i \right|$, $j = 1, \ldots, L$ are fulfilled with equality. $\mathcal{S}_j$ denotes the jth set containing at least two users having equal weights and L is the total number of sets containing at least two users with equal weights. Points on the time-sharing region are not directly achievable by successive decoding [36, Chapter 3.3.3], but only by time-sharing between the corner points. While the dual function is maximized on every point of the time-sharing region, the primal problem may be infeasible on parts of the time-sharing region or may even be solved by one point on the time-sharing region only. For this reason, additional measures have to be taken to reconstruct the primal solution from the dual in case some users have equal weights.

A naive approach would be to compute the rates for all possible decoding orders and determine the convex combination of the rate vectors obtained this way solving the primal problem (4.1). With increasing number of sets $\mathcal{S}_j$ and increasing cardinality of these sets, such an approach becomes more and more complex, as $\left(\sum_{j=1}^{L} |\mathcal{S}_j|\right)!$ different decoding orders are possible. For this reason a successive algorithm for the computation of the time-sharing solution has been proposed in [96], which is based on the method from [97]. The basic idea is to iteratively approximate the polytope $\mathcal{P}\left(\boldsymbol{H}_1^{\mathrm{H}}, \ldots, \boldsymbol{H}_K^{\mathrm{H}}, \boldsymbol{W}_1^*, \ldots, \boldsymbol{W}_K^*\right)$ by another polytope with less extreme points, which is contained completely in $\mathcal{P}\left(\boldsymbol{H}_1^{\mathrm{H}}, \ldots, \boldsymbol{H}_K^{\mathrm{H}}, \boldsymbol{W}_1^*, \ldots, \boldsymbol{W}_K^*\right)$. Denoting the polytope of the ith iteration as $\hat{\mathcal{P}}^{(i)}\left(\boldsymbol{H}_1^{\mathrm{H}}, \ldots, \boldsymbol{H}_K^{\mathrm{H}}, \boldsymbol{W}_1^*, \ldots, \boldsymbol{W}_K^*\right)$, the primal problem in the ith iteration reads as

$$\begin{aligned} &\max_{\boldsymbol{r}} u\left(\boldsymbol{W}_1^*, \ldots, \boldsymbol{W}_K^*, \boldsymbol{r}\right), \\ &\text{s.t. } \boldsymbol{h}_1\left(\boldsymbol{r}\right) \leq \boldsymbol{0}_{c_1,1}, \quad \boldsymbol{h}_2\left(\boldsymbol{r}\right) = \boldsymbol{0}_{c_2,1}, \quad \boldsymbol{r} \in \hat{\mathcal{P}}^{(i)}\left(\boldsymbol{H}_1^{\mathrm{H}}, \ldots, \boldsymbol{H}_K^{\mathrm{H}}, \boldsymbol{W}_1^*, \ldots, \boldsymbol{W}_K^*\right), \end{aligned} \tag{4.15}$$

where the maximization with respect to the covariance matrices can be skipped due to the fact that the $\boldsymbol{W}_1^*, \ldots, \boldsymbol{W}_K^*$ are optimum on the whole time sharing region. The condition $\boldsymbol{r} \in$

[1]Due to the finite accuracy of the ellipsoid method, the case that some users have equal weights will occur very rarely in practice. Thus, two weights λ_k^* and λ_j^* should only be considered as different, if $|\lambda_j^* - \lambda_k^*| \geq \alpha$, where α denotes an appropriate threshold.

$\hat{\mathcal{P}}^{(i)}\left(\boldsymbol{H}_1^{\mathrm{H}},\ldots,\boldsymbol{H}_K^{\mathrm{H}},\boldsymbol{W}_1^*,\ldots,\boldsymbol{W}_K^*\right)$ can be rewritten as

$$\boldsymbol{r}\leq\hat{\boldsymbol{R}}^{(i)}\boldsymbol{\alpha},\quad \boldsymbol{\alpha}\geq\mathbf{0}_{K+i,1},\quad \sum_{j=1}^{K+i}[\boldsymbol{\alpha}]_j\leq 1,$$

where $\hat{\boldsymbol{R}}^{(i)}=\left[\hat{\boldsymbol{R}}^{(1)},\hat{\boldsymbol{r}}^{(1)},\ldots,\hat{\boldsymbol{r}}^{(i-1)}\right]$ contains the extreme points of the polytope $\hat{\mathcal{P}}^{(i)}\left(\boldsymbol{H}_1^{\mathrm{H}},\ldots,\boldsymbol{H}_K^{\mathrm{H}},\boldsymbol{W}_1^*,\ldots,\boldsymbol{W}_K^*\right)$ and $\hat{\boldsymbol{r}}^{(j)}$ denotes the extreme point added after step j. The matrix $\hat{\boldsymbol{R}}^{(1)}$ for the first iteration is given by

$$\hat{\boldsymbol{R}}^{(1)}=\begin{bmatrix}\log_2\left|\boldsymbol{I}_{CN_{\mathrm{T}}}+\boldsymbol{H}_1^{\mathrm{H}}\boldsymbol{W}_1^*\boldsymbol{H}_1\right| \\ \mathbf{0}_{K-1,1}\end{bmatrix},\ldots,\begin{matrix}\mathbf{0}_{K-1,1}\\ \log_2\left|\boldsymbol{I}_{CN_{\mathrm{T}}}+\boldsymbol{H}_K^{\mathrm{H}}\boldsymbol{W}_K^*\boldsymbol{H}_K\right|\end{matrix},\boldsymbol{r}_{\mathrm{min}}+\epsilon\mathbf{1}_K\Bigg],$$

where $\epsilon>0$ should be chosen so that $\boldsymbol{r}_{\mathrm{min}}$ lies in the interior of the initial polytope, i.e., the initial problem should be feasible. Clearly for the rate balancing problem, which is always feasible, the last entry in $\hat{\boldsymbol{R}}^{(1)}$ can be skipped. To save computational complexity by using an initial polytope with less extreme points, for the power minimization one can use the utility and constraint functions of the rate balancing problem for the determination of the time-sharing solution, where the ratios $\rho_k=\frac{R_{k,\mathrm{min}}}{R_{1,\mathrm{min}}}$ are given by the ratios of the minimum rates. That is because in the optimum all rate inequalities are fulfilled with equality, as otherwise the transmit power could be reduced without violating the minimum rate constraints. Furthermore the solution of such a rate balancing solution with fixed covariance matrices will lie in the same time-sharing region. Thus, Problem (4.15) reads as

$$\max_{\boldsymbol{r},\boldsymbol{\alpha}} u\left(\boldsymbol{W}_1^*,\ldots,\boldsymbol{W}_K^*,\boldsymbol{r}\right),$$

$$\text{s.t. } \boldsymbol{h}_1(\boldsymbol{r})\leq\mathbf{0}_{c_1,1},\quad \boldsymbol{h}_2(\boldsymbol{r})=\mathbf{0}_{c_2,1},\quad \boldsymbol{r}-\hat{\boldsymbol{R}}^{(i)}\boldsymbol{\alpha}\leq\mathbf{0},\quad \boldsymbol{\alpha}\geq\mathbf{0}_{K+i,1},\quad \sum_{j=1}^{K+i}[\boldsymbol{\alpha}]_j\leq 1, \tag{4.16}$$

which is a linear program for the QoS constrained problems introduced in this chapter, that can be efficiently solved (e.g. [33, Chapter 2.7]). Once the solution $\bar{\boldsymbol{r}}^{(i)}$ and $\bar{\boldsymbol{\alpha}}^{(i)}$ for Problem (4.16) has been obtained, the Lagrange multiplier $\bar{\boldsymbol{\lambda}}^{(i)}$ corresponding to the third constraint in (4.16) is computed according to

$$\bar{\boldsymbol{\lambda}}^{(i)}=\underset{\boldsymbol{\theta}\geq\mathbf{0}}{\operatorname{argmin}}\,\bar{g}(\boldsymbol{\theta})=\underset{\boldsymbol{\theta}\geq\mathbf{0}}{\operatorname{argmin}}\left\{u\left(\boldsymbol{W}_1^*,\ldots,\boldsymbol{W}_K^*,\bar{\boldsymbol{r}}^{(i)}\right)-\boldsymbol{\theta}^{\mathsf{T}}\left(\bar{\boldsymbol{r}}^{(i)}-\hat{\boldsymbol{R}}^{(i)}\bar{\boldsymbol{\alpha}}^{(i)}\right)\right\}.$$

In case the maximum of the dual function $\bar{g}\left(\bar{\boldsymbol{\lambda}}^{(i)}\right)$ of Problem (4.16) is identical to the maximum of the dual function $g(\boldsymbol{\theta}^*)$ of the original problem in (4.1), the optimum rates for the primal problem have been found with the rates $\bar{\boldsymbol{r}}^{(i)}$ and $\tilde{\boldsymbol{r}}=\bar{\boldsymbol{r}}^{(i)}$. That is because the optimum utility from (4.16) gives a lower bound for the optimum utility in (4.1), as the constraint set $\hat{\mathcal{P}}^{(i)}\left(\boldsymbol{H}_1^{\mathrm{H}},\ldots,\boldsymbol{H}_K^{\mathrm{H}},\boldsymbol{W}_1^*,\ldots,\boldsymbol{W}_K^*\right)$ is completely contained in the polytope from (4.1), i.e., the constraint set in (4.15) is stricter than in the original Problem (4.1), and for both problems the duality gap is zero. Correspondingly, the dual function $\bar{g}\left(\bar{\boldsymbol{\lambda}}^{(i)}\right)$ is a lower bound for $g(\boldsymbol{\theta}^*)$. Thus, in case the optimum dual functions are equal, both problems lead to the same optimum solution. Otherwise, i.e., if $\bar{g}\left(\bar{\boldsymbol{\lambda}}^{(i)}\right)<g(\boldsymbol{\theta}^*)$, a new extreme point has to be added to the polytope $\hat{\mathcal{P}}^{(i)}\left(\boldsymbol{H}_1^{\mathrm{H}},\ldots,\boldsymbol{H}_K^{\mathrm{H}},\boldsymbol{W}_1^*,\ldots,\boldsymbol{W}_K^*\right)$. This point $\hat{\boldsymbol{r}}^{(i)}$ solves the problem

$$\hat{\boldsymbol{r}}^{(i)}=\underset{\boldsymbol{r}}{\operatorname{argmax}}\,\bar{\boldsymbol{\lambda}}^{(i),\mathsf{T}}\boldsymbol{r},\quad \text{s.t. } \boldsymbol{r}\in\mathcal{P}\left(\boldsymbol{H}_1^{\mathrm{H}},\ldots,\boldsymbol{H}_K^{\mathrm{H}},\boldsymbol{W}_1^*,\ldots,\boldsymbol{W}_K^*\right), \tag{4.17}$$

which is a linear program that is solved by one of the extreme points of $\mathcal{P}\left(\boldsymbol{H}_1^{\mathrm{H}},\ldots,\boldsymbol{H}_K^{\mathrm{H}},\boldsymbol{W}_1^*,\ldots,\boldsymbol{W}_K^*\right)$. As pointed out in [97], $\bar{\boldsymbol{\lambda}}^{(i)}$ is a subgradient of the optimum utility in (4.16) with respect to changes in the constraint vector $\hat{\boldsymbol{R}}^{(i)}\boldsymbol{\alpha}$. Hence, by (4.17) this extreme point of $\mathcal{P}\left(\boldsymbol{H}_1^{\mathrm{H}},\ldots,\boldsymbol{H}_K^{\mathrm{H}},\boldsymbol{W}_1^*,\ldots,\boldsymbol{W}_K^*\right)$ is added to the set $\hat{\mathcal{P}}^{(i)}\left(\boldsymbol{H}_1^{\mathrm{H}},\ldots,\boldsymbol{H}_K^{\mathrm{H}},\boldsymbol{W}_1^*,\ldots,\boldsymbol{W}_K^*\right)$ that is aligned closest to the direction of strongest increase of the utility function. The extreme points of $\mathcal{P}\left(\boldsymbol{H}_1^{\mathrm{H}},\ldots,\boldsymbol{H}_K^{\mathrm{H}},\boldsymbol{W}_1^*,\ldots,\boldsymbol{W}_K^*\right)$ only differ in the encoding order of users having equal weights $\hat{\boldsymbol{\mu}}(\boldsymbol{\theta}^*)$. The optimum extreme point with respect to (4.17) is then achieved by decoding the users within each group $\mathcal{S}_j$ in increasing order of the corresponding weights in $\bar{\boldsymbol{\lambda}}^{(i)}$, i.e., the user with $\mathrm{argmax}_{k\in\mathcal{S}_j}\left[\bar{\boldsymbol{\lambda}}^{(i)}\right]_k$ is decoded last amongst the users in group $\mathcal{S}_j$, whereas the decoding order amongst the different groups and the users having unique weights $\hat{\boldsymbol{\mu}}(\boldsymbol{\theta}^*)$ remains the same as in the original problem. The proof for optimality of this decoding order can be found in [96]. The whole algorithm for QoS constrained utility maximization is summarized in Algorithm 4.1.

As it has been shown in this section, QoS constrained optimization problems can be solved by an iterative solution of weighted sum rate maximizations. One approach for efficient algorithms would therefore be to apply the methods presented in the previous chapter for weighted sum rate maximization and determine the weights optimum for the original QoS constrained problem with the ellipsoid method. As it exhibits a considerable amount of computational complexity and slow convergence behavior, more efficient algorithms for near optimum solutions to QoS constrained optimization problems will be presented next.

4.3 State-of-the-Art Near Optimum Approaches

Zero-forcing is also a popular concept for solving QoS constrained utility maximization problems. The simplest zero-forcing method is to allow only one user per carrier and separate the users by the carriers, which leads to Orthogonal Frequency Division Multiple Access (OFDMA). The problem of power minimization under minimum rate requirements is solved almost optimally under the OFDMA constraint in [86] via the dual problem. Although the algorithm is presented for Single-Input Single-Output (SISO) systems in [86], its extension to MIMO is straightforward by applying left and right singular vectors as receive and transmit filters, respectively, as it is done for example in [98]. In this paper one special case of the rate balancing problem, the so-called Kalai-Smorodinsky bargaining solution is considered. In [99] the rate balancing problem for SISO OFDMA systems is considered, where an equal power allocation is assumed and the carriers are successively allocated to the users. Thereby that user receives a carrier in each step that rate achievable with an equal power allocation is furthest away from the desired proportion of the sum rate. By serving only one user per carrier, the spatial degrees can however not fully exploited, which is why OFDMA approaches inhere considerable performance losses. One way to overcome this drawback but remain with zero-forcing is to apply Block-Diagonalization (BD) [61] on each carrier and perform QoS constrained power allocation over the resulting scalar subchannels. In [100] the performance of such a scheme is improved for the power minimization problem by recalculating the receive and transmit filters. BD is combined with DPC in [101] for that problem, where the main focus is put on determining the optimum encoding order. As already mentioned in the last chapter, BD requires the total number of receive antennas in the system to be smaller than the number of transmit antennas, an unlikely setup in practical systems. For that reason an algorithm to allocate data streams to users for the power minimization problem is presented in [102], where zero-forcing transmit filters are applied at the transmitter and the left singular vectors at the receivers. For the same problem, a successive allocation of data streams to users in a MISO system

Algorithm 4.1 Algorithm for QoS Constrained Utility Maximization Using the Ellipsoid Method

Initialize ellipsoid:
Rate Balancing: $\boldsymbol{\theta}^{(0)} = \frac{1}{K}\mathbf{1}_{K-1}, \quad \boldsymbol{E}^{(0)} = \left(1 - \frac{1}{K}\right)\boldsymbol{I}_{K-1}$
Power minimization and weights sum rate maximization with minimum rate requirements: $\boldsymbol{\theta}^{(0)} = \frac{1}{2}\boldsymbol{\theta}_{\text{max}}$, $\boldsymbol{E}^{(0)} = \text{diag}\left([\boldsymbol{\theta}^2_{\text{max}}]_1, \ldots, [\boldsymbol{\theta}^2_{\text{max}}]_n\right) K/4$
$i = 1$, convergence =FALSE
repeat
  if $\boldsymbol{\theta}^{(i-1)}$ feasible **then**
    Compute subgradient $-\boldsymbol{h}\left((\hat{\boldsymbol{r}}^{(i)}\right)$
    $\boldsymbol{a}^{(i)} = -\boldsymbol{h}\left((\hat{\boldsymbol{r}}^{(i)}\right)$, $b^{(i)} = 0$, $\alpha = 0$
    if $\left(\sqrt{\boldsymbol{h}^{\mathsf{T}}\left((\hat{\boldsymbol{r}}^{(i)}\right)\boldsymbol{E}^{(i)}\boldsymbol{h}\left((\hat{\boldsymbol{r}}^{(i)}\right)} \leq \varepsilon\right)$ **then**
      convergence=TRUE
    end if
  else
    Determine violated constraint $\boldsymbol{c}^{\mathsf{T}}_{j,\text{c}}\boldsymbol{\theta}^{(i)} > [\boldsymbol{d}_{\text{c}}]_j$
    $\boldsymbol{a}^{(i)} = \boldsymbol{c}_{j,\text{c}}$, $b^{(i)} = [\boldsymbol{d}_{\text{c}}]_j - \boldsymbol{c}^{\mathsf{T}}_{j,\text{c}}\boldsymbol{\theta}^{(i)}$, $\alpha = \frac{-b^{(i)}}{\sqrt{\boldsymbol{a}^{(i),\mathsf{T}}\boldsymbol{E}^{(i)}\boldsymbol{a}^{(i)}}}$
  end if
  Update ellipsoid:
  $\boldsymbol{\theta}^{(i+1)} = \boldsymbol{\theta}^{(i)} - \frac{1+\alpha n}{(n+1)\sqrt{\boldsymbol{a}^{(i),\mathsf{T}}\boldsymbol{E}^{(i)}\boldsymbol{a}^{(i)}}}\boldsymbol{E}^{(i)}\boldsymbol{a}^{(i)}$
  $\boldsymbol{E}^{(i+1)} = \frac{n^2(1-\alpha^2)}{n^2-1}\left(\boldsymbol{E}^{(i)} - \frac{2(1+\alpha n)}{(n+1)(1+\alpha)\boldsymbol{a}^{(i),\mathsf{T}}\boldsymbol{E}^{(i)}\boldsymbol{a}^{(i)}}\boldsymbol{E}^{(i)}\boldsymbol{a}^{(i)}\boldsymbol{a}^{(i),\mathsf{T}}\boldsymbol{E}^{(i)}\right)$
  $i = i + 1$
until convergence=TRUE or detection of infeasibilty $\left(\boldsymbol{c}^{\mathsf{T}}_{j,\text{c}}\boldsymbol{\theta}^{(i)} - [\boldsymbol{d}_{\text{c}}]_j - \sqrt{\boldsymbol{a}^{(i-1),T}\boldsymbol{E}^{(i)}\boldsymbol{a}^{(i-1)}} > 0\right)$
Find users having equal weights $\hat{\boldsymbol{\mu}}\left(\boldsymbol{\theta}^{(i-1)}\right)$
if group of at least two users having equal weights exists **then**
  $\left\{\boldsymbol{r}^*, \boldsymbol{W}^*_1, \ldots, \boldsymbol{W}^*_K\right\} = \text{argmax}_{\substack{\boldsymbol{r}\in\mathcal{P}\left(\boldsymbol{H}^{\text{H}}_1,\ldots,\boldsymbol{H}^{\text{H}}_K,\boldsymbol{W}_1,\ldots,\boldsymbol{W}_K\right)\\ \boldsymbol{W}_1,\ldots,\boldsymbol{W}_K\in\mathcal{C}_{\text{P}}}}\left\{u\left(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_K, \boldsymbol{r}\right) - \boldsymbol{\theta}^{(i),\mathsf{T}}\boldsymbol{h}(\boldsymbol{r})\right\}$
  $\hat{\boldsymbol{R}}^{(1)} = \begin{bmatrix} \log_2\left|\boldsymbol{I}_{CN_{\text{T}}} + \boldsymbol{H}^{\text{H}}_1\boldsymbol{W}^*_1\boldsymbol{H}_1\right| \\ \mathbf{0}_{K-1,1} \end{bmatrix}, \ldots, \begin{matrix} \mathbf{0}_{K-1,1} \\ \log_2\left|\boldsymbol{I}_{CN_{\text{T}}} + \boldsymbol{H}^{\text{H}}_K\boldsymbol{W}^*_K\boldsymbol{H}_K\right| \end{matrix}, \boldsymbol{R}_{\text{min}} + \epsilon\mathbf{1}_K \Bigg]$
  $j = 1$
  repeat
    Obtain $\tilde{\boldsymbol{r}}^{(j)}$ and $\tilde{\alpha}^{(j)}$ as maximizers in (4.16)
    $\tilde{\boldsymbol{\lambda}}^{(j)} = \text{argmin}_{\boldsymbol{\theta}\geq\mathbf{0}}\,\tilde{g}(\boldsymbol{\theta}) = \text{argmin}_{\boldsymbol{\theta}\geq\mathbf{0}}\left\{u\left(\boldsymbol{W}^*_1, \ldots, \boldsymbol{W}^*_K, \tilde{\boldsymbol{r}}^{(j)}\right) - \boldsymbol{\theta}^{\mathsf{T}}\left(\tilde{\boldsymbol{r}}^{(j)} - \hat{\boldsymbol{R}}^{(j)}\tilde{\boldsymbol{\alpha}}^{(j)}\right)\right\}$
    $\hat{\boldsymbol{r}}^{(j)} = \text{argmax}_{\boldsymbol{r}}\,\tilde{\boldsymbol{\lambda}}^{(j),\mathsf{T}}\boldsymbol{r}$, s.t. $\boldsymbol{r} \in \mathcal{P}\left(\boldsymbol{H}^{\text{H}}_1, \ldots, \boldsymbol{H}^{\text{H}}_K, \boldsymbol{W}^*_1, \ldots, \boldsymbol{W}^*_K\right)$
    $\hat{\boldsymbol{R}}^{(j+1)} = \left[\hat{\boldsymbol{R}}^{(j)}, \hat{\boldsymbol{r}}^{(j)}\right]$
    $j = j + 1$
  until $g\left(\boldsymbol{\theta}^{(i)}\right) - \tilde{g}\left(\tilde{\boldsymbol{\lambda}}^{(j-1)}\right) < \varepsilon$
end if

with zero-forcing at the transmitter is proposed in [103]. QoS constrained utility maximization problems without zero-forcing constraints and linear transmit and receive filters are considered in [104], [55] and [57], where in the latter per-stream and not per-user SINR requirements are treated. Per-antenna power constraints are introduced in [105] to QoS constrained problems with linear transmit and receive processing. All these algorithms have in common, that they are very complex, as they work iteratively and in each iteration uplink-downlink transformations are required, and that they are not guaranteed to converge to the global optimum solution. In [106] an optimum iterative algorithm for linear transmit and receive filters is presented for the problem of sum power minimization under maximum requirements for each user's Mean Square Error (MSE), which are related to rate constraints by an inequality between MSE and rate. A reduced complexity but still iterative algorithm for the sum rate maximization with relative rate constraints without DPC is presented in [107], a non-iterative approach for the same problem but with DPC at the transmitter is proposed in [11].

4.4 Successive Resource Allocation and Spatial Zero-Forcing

An efficient method to solve the general QoS constrained Problem (4.1) will be presented in this section. For this purpose its dual problem

$$\min_{\boldsymbol{\theta}_1 \geq \mathbf{0}_{c_1,1}, \boldsymbol{\theta}_2} \sup_{\substack{\boldsymbol{r} \in \mathcal{P}\left(\boldsymbol{H}_1^{\mathrm{H}}, \ldots, \boldsymbol{H}_K^{\mathrm{H}}, \boldsymbol{W}_1, \ldots, \boldsymbol{W}_K\right) \\ \boldsymbol{W}_1, \ldots, \boldsymbol{W}_K \in \mathcal{C}_{\mathrm{P}}}} \left\{ u\left(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_K, \boldsymbol{r}\right) - \boldsymbol{\theta}_1^{\mathsf{T}} \boldsymbol{h}_1(\boldsymbol{r}) - \boldsymbol{\theta}_2^{\mathsf{T}} \boldsymbol{h}_2(\boldsymbol{r}) \right\}, \tag{4.18}$$

which is identical to Problem (4.2), will be used. The optimum value $\min\limits_{\boldsymbol{\theta}_1 \geq \mathbf{0}_{c_1,1}, \boldsymbol{\theta}_2} g(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ of the dual problem is equal to the optimum value of the utility function in the primal problem (4.1), due to the strong duality shown in Section 4.2.1, which is why the dual problem will be used as a starting point for simplifications in this section. The concept of successive resource allocation and spatial zero-forcing, which proofed to be effective for weighted sum rate maximization in the previous chapter, is applied to the dual function for QoS constrained utility maximization problems in this section. This implies that in each step the data stream allocations and receive filters obtained in the previous steps are kept fixed and the filters for the newly allocated data stream and the corresponding user allocation are determined so that the increase in the optimum value of the dual function becomes maximum. Thus, in the ith step a solution to

$$\max_{k} \min_{\boldsymbol{\theta}_1 \geq \mathbf{0}_{c_1,1}, \boldsymbol{\theta}_2} \max_{\boldsymbol{p}^{(i)}(k) \geq \mathbf{0}_{i,1}, \mathbf{1}_i^{\mathsf{T}} \boldsymbol{p}^{(i)}(k) \leq P_{\mathrm{Tx}}} \max_{\boldsymbol{r} \in \mathcal{C}_{\mathrm{R}}\left(\boldsymbol{p}^{(i)}(k)\right)} \left\{ u\left(\boldsymbol{p}^{(i)}(k), \boldsymbol{r}\right) - \boldsymbol{\theta}_1^{\mathsf{T}} \boldsymbol{h}_1(\boldsymbol{r}) - \boldsymbol{\theta}_2^{\mathsf{T}} \boldsymbol{h}_2(\boldsymbol{r}) \right\} \tag{4.19}$$

has to be found. The maximization over k is done last as it is desired to find that user with the maximum value of the dual function. For the power minimization problem it is necessary to set $P_{\mathrm{Tx}} = \infty$ and

$$\boldsymbol{p}^{(i)}(k) = \left[p_{1,1}^{(i)}, \ldots, p_{1,d_k}^{(i)}, \ldots, p_{k-1,d_{k-1}}^{(i)}, p_{k,1}^{(i)}, \ldots, p_{k,d_k+1}^{(i)}, p_{k+1,1}^{(i)}, \ldots, p_{K,d_K}^{(i)} \right]^{\mathsf{T}}$$

contains the powers allocated to the scalar subchannels when the ith subchannel is allocated to user k. As described with the weighted sum rate maximization problem in the previous chapter, these subchannels result from the decomposition of the MIMO broadcast channel The variables d_k are again initialized with zero and incremented by one each time a subchannel is allocated to user

k. With the successive resource allocation and zero-forcing, the dependency of the utility from the uplink covariance matrices $\boldsymbol{W}_k$ can be replaced by the power vector $\boldsymbol{p}^{(i)}(k)$. The set $\mathcal{C}_{\mathrm{R}}\left(\boldsymbol{p}^{(i)}(k)\right)$ contains all rates achievable if a new data stream is allocated to user k, the receive filters and user allocations from previous steps are kept fixed, the power allocation $\boldsymbol{p}^{(i)}(k)$ is used and the receive filter for the new data stream and the transmit filters obey to the carrier separation and zero-forcing constraints. The rates in the vector $\boldsymbol{r}$ are computed from the filters and powers used in the broadcast channel. Due to the duality from [5] these rates are identical to the rates in the dual MAC used in Problem (4.18). While such a rate computation destroys the convexity of the problem in case the optimum solution has to be found as in Section 4.2, the successive problem can be easier solved this way as shown in the following. As already pointed out in Section 4.2.1, due to the fact that $u\left(\boldsymbol{p}^{(i)}(k), \boldsymbol{r}\right)$, $\boldsymbol{h}_1(\boldsymbol{r})$ and $\boldsymbol{h}_2(\boldsymbol{r})$ are linear functions in $\boldsymbol{r}$, the maximization over $\boldsymbol{p}^{(i)}(k)$ and $\boldsymbol{r}$ in (4.18) corresponds to a weighted sum rate maximization for fixed Lagrange multipliers $\boldsymbol{\theta}$ or, in case of the power minimization problem, can be solved via an iterative application of a weighted sum rate maximization. Thus, the optimum receive filter for the newly allocated data stream and the optimum transmit filters result from a weighted sum rate maximization. Therefore, the optimum transmit filter for the newly allocated data stream for DPC can be obtained from from (3.31) and the corresponding receive filter from (3.29). Accordingly, in case no DPC is applied the receive filter maximizing a lower bound for the weighted sum rate is given by (3.58) and (3.64) and the update rules (3.68) can be used to obtain the transmit filters. As all these filters are independent of the weights, the maximization over $\boldsymbol{r} \in \mathcal{C}_{\mathrm{R}}\left(\boldsymbol{p}^{(i)}(k)\right)$ can be solved by using the rates

$$\boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right) = \left[R_1\left(\boldsymbol{p}^{(i)}(k)\right), \ldots, R_K\left(\boldsymbol{p}^{(i)}(k)\right)\right]^{\mathsf{T}},$$

where using the results from Sections 3.2 and 3.3 the rates $R_m\left(\boldsymbol{p}^{(i)}(k)\right)$ can be stated as

$$R_m\left(\boldsymbol{p}^{(i)}(k)\right) = \sum_{j=1}^{\hat{d}_m} \log_2\left(1 + p_{m,j}^{(i)}\lambda_{m,j}^{(i)}(k)\right), \tag{4.20}$$

where

$$\hat{d}_m = \begin{cases} d_m, & m \neq k \\ d_m + 1, & m = k \end{cases}. \tag{4.21}$$

The channel gains $\lambda_{m,j}^{(i)}(k)$ for $j = 1, \ldots, d_m$, $m = 1, \ldots, K$ are given by

$$\lambda_{m,j}^{(i)}(k) = \frac{C}{\sigma_{\mathrm{n}}^2}\boldsymbol{t}_{m,j}^{\mathrm{H}}\boldsymbol{H}_m^{\mathrm{H}}\boldsymbol{H}_m\boldsymbol{t}_{m,j}$$

for the DPC case [c.f. (3.37)] and by

$$\lambda_{m,j}^{(i)}(k) = \frac{C}{\sigma_{\mathrm{n}}^2\left(1 + \frac{\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{n_{m,j}}^{(i)}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{g}_k^{(i)}}{\boldsymbol{g}_k^{(i),\mathrm{H}}\boldsymbol{A}_k^{(i)}\boldsymbol{g}_k^{(i)}}\right)\alpha_{n_{m,j}}^{(i-1)}}$$

when no DPC can be used [c.f. (3.60)]. The gain of the newly allocated subchannel $\lambda_{k,d_k+1}^{(i)}(k)$ computes according to

$$\lambda_{k,d_k+1}^{(i)}(k) = \frac{C}{\sigma_{\mathrm{n}}^2}\rho_1\left(\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}\right)$$

in case DPC is applied at the transmitter and

$$\lambda^{(i)}_{k,d_k+1}(k) = \frac{C}{\sigma_{\mathrm{n}}^2} \boldsymbol{g}_k^{(i),\mathrm{H}} \boldsymbol{A}_k^{(i)} \boldsymbol{g}_k^{(i)}$$

otherwise. As in the previous chapter, the carrier $c_k^{(i)}$ on which the newly allocated subchannel is transmitted if this subchannel is allocated to user k is given by that block within the matrix $\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)} \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{H}_k \hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}$ in the DPC case and the matrix $\boldsymbol{B}_k^{(i),-1} \boldsymbol{A}_k^{(i)}$ otherwise that exhibits the largest principal eigenvalue. Thus, the carrier preselection proposed at the end of Section 3.4.1 can be applied to the QoS constrained utility maximization problems as well. Problem (4.19) then reads as

$$\max_{k} \min_{\boldsymbol{\theta}_1 \geq \mathbf{0}_{c_1,1}, \boldsymbol{\theta}_2} \max_{\boldsymbol{p}^{(i)}(k) \geq \mathbf{0}_{i,1}, \mathbf{1}_i^{\mathsf{T}} \boldsymbol{p}^{(i)}(k) \leq P_{\mathrm{Tx}}} \left\{ u\left(\boldsymbol{p}^{(i)}(k), \boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right)\right) - \right.$$
$$\left. - \boldsymbol{\theta}_1^{\mathsf{T}} \boldsymbol{h}_1\left(\boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right)\right) - \boldsymbol{\theta}_2^{\mathsf{T}} \boldsymbol{h}_2\left(\boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right)\right) \right\}. \quad (4.22)$$

The optimum value of the dual function for a fixed user k is equal to $-\infty$, if the constraints of the primal problem are violated, i.e., there exists an index j with $\boldsymbol{e}_j^{\mathsf{T}} \boldsymbol{h}_1\left(\boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right)\right) > 0$ or $\boldsymbol{e}_j^{\mathsf{T}} \boldsymbol{h}_2\left(\boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right)\right) \neq 0$. That is definitely the case if the rate of a user m is equal to zero, which happens if no subchannel has been allocated to that user so far. As afterwards the maximum is taken with respect to the user index, that user would never be served. In order to work properly, the algorithm therefore needs an initialization phase, in which each user receives one subchannel. For this purpose a set $\mathcal{S}^{(i)}$ will be defined for the first K allocation steps, which contains all users that have been served in steps 1 to $i-1$. Those users are then excluded from the user selection process in the steps i to K. Furthermore in the dual function only the constraints of those users and the user k are considered so that the case that the optimum value of the dual function becomes $-\infty$ is avoided. During the initialization phase instead of (4.22) the following optimization problem has to be solved

$$\max_{k \in \{1,\ldots,K\} \setminus \mathcal{S}^{(i)}} \min_{\boldsymbol{\theta}_1 \geq \mathbf{0}_{c_1,1}, \boldsymbol{\theta}_2} \max_{\boldsymbol{p}^{(i)}(k) \geq \mathbf{0}_{i,1}, \mathbf{1}_i^{\mathsf{T}} \boldsymbol{p}^{(i)}(k) \leq P_{\mathrm{Tx}}}$$
$$\left\{ u\left(\boldsymbol{p}^{(i)}(k), \boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right)\right) - \boldsymbol{\theta}_1^{\mathsf{T}} \hat{\boldsymbol{h}}_1^{(i)}\left(k, \boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right)\right) - \boldsymbol{\theta}_2^{\mathsf{T}} \hat{\boldsymbol{h}}_2^{(i)}\left(k, \boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right)\right) \right\}. \quad (4.23)$$

$\hat{\boldsymbol{h}}_1^{(i)}\left(k, \boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right)\right)$ contains a zero in every row, that corresponds to a minimum rate constraint of a user that has not been served so far except user k, i.e.,

$$\boldsymbol{e}_j^{\mathsf{T}} \hat{\boldsymbol{h}}_1^{(i)}\left(k, \boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right)\right) = \begin{cases} \boldsymbol{e}_j^{\mathsf{T}} \boldsymbol{h}_1\left(\boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right)\right), & j \in \mathcal{S}^{(i)}, j = k \\ 0 & \text{else} \end{cases}.$$

For the sum rate maximization with relative rate constraints, the first subchannel is allocated to user 1 and from the second step onwards $\hat{\boldsymbol{h}}_2^{(i)}\left(k, \boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right)\right)$ contains a zero in every row that corresponds to a relative rate constraint of a user that has not been served so far except user k, which implies that

$$\boldsymbol{e}_j^{\mathsf{T}} \hat{\boldsymbol{h}}_2^{(i)}\left(k, \boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right)\right) = \begin{cases} \boldsymbol{e}_j^{\mathsf{T}} \boldsymbol{h}_2\left(\boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right)\right), & j+1 \in \mathcal{S}^{(i)}, j = k-1 \\ 0 & \text{else} \end{cases}.$$

Clearly, the QoS constrained utility maximization problems cannot be solved with the proposed method, if there are too few degrees of freedom to allocate at least one subchannel to each user, i.e., CN_{T} must be at least equal to the number of users with non-zero rate requirements. A more complicated initialization algorithm is proposed in [103], where in the ith step that user is chosen that exhibits the largest difference in the objective function between allocating the next data stream to the optimum carrier and the second best carrier. However, as this algorithm requires twice as many solutions of the dual problem as the initialization described above, it will be not considered here.

When the problem of weighted sum rate maximization under a sum power constraint and minimum rate requirements is considered, it can happen that the problem remains infeasible, i.e., the optimum value of the dual function goes to $-\infty$, even if at least one subchannel has been allocated to each user. That is because the transmit power constraint P_{Tx} can be too low to fulfill the minimum rate requirements with a small number of subchannels. Thus, for this problem at first the power constraint is omitted in (4.22) and (4.23), i.e., $P_{\mathrm{Tx}} \to \infty$, which corresponds to minimizing the transmit power required to fulfill the minimum rates. If in a certain step j the used transmit power falls below P_{Tx}, i.e., $\mathbf{1}_j^{\mathsf{T}} \boldsymbol{p}^{(j)}(k(j)) \leq P_{\mathrm{Tx}}$, the actual problem of weighted sum rate maximization with minimum rate constraints is considered, as it has been proposed in [57]. If all degrees of freedom are exploited, i.e., $C \min\left(N_{\mathrm{T}}, \sum_{k=1}^{K} r_k\right)$ subchannels have been allocated and the power required to fulfill the minimum rate requirements is still higher than P_{Tx}, the problem is infeasible with the proposed successive allocation and spatial zero-forcing method. In general, the algorithm terminates when it is observed that the optimum value of the dual function would decrease with a new data stream allocation or all degrees of freedom are used, i.e., after $C \min\left(N_{\mathrm{T}}, \sum_{k=1}^{K} r_k\right)$ steps.

In the remainder it will shown how Problem (4.22) can be solved for the QoS constrained utility maximization problems introduced in Section 4.1. The solution of Problem (4.23) during initialization is similar and will be skipped for notational convenience. The maximization with respect to k has to be solved via an exhaustive search, which is why in the following Problem (4.22) will be solved for fixed k. All three QoS constrained problems have in common that the maximization over $\boldsymbol{p}^{(i)}(k)$ leads to the following optimum powers

$$p_{m,j}^{(i)} = \left[\frac{\nu_m}{\ln(2)} - \frac{1}{\lambda_{m,j}^{(i)}(k)}\right]^{+}, \tag{4.24}$$

which follows from evaluating the Karush-Kuhn-Tucker (KKT) conditions of such a problem. ν_m is a function of the weight μ_m and the Lagrangian multipliers $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$. Due to the zero-forcing constraints, the minimization with respect to the Lagrangian multipliers does not have to be solved with the ellipsoid method, but with much simpler methods as explained in the following for the three problems from Section 4.1.

- **Weighted Sum Rate Maximization with Minimum Rate Requirements**
  Inserting the utility and constraint functions of this QoS constrained optimization problem into (4.22) leads to the dual problem

$$\min_{\boldsymbol{\theta}_1 \geq \mathbf{0}_{K,1}} \max_{\boldsymbol{p}^{(i)}(k) \geq \mathbf{0}_{i,1}, \mathbf{1}_i^{\mathsf{T}} \boldsymbol{p}^{(i)}(k) \leq P_{\mathrm{Tx}}} \left\{ \left(\boldsymbol{\mu}^{\mathsf{T}} + \boldsymbol{\theta}_1^{\mathsf{T}}\right) \boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right) \right\} - \boldsymbol{\theta}_1^{\mathsf{T}} \boldsymbol{r}_{\min} = \min_{\boldsymbol{\theta}_1 \geq \mathbf{0}_{K,1}} g(\boldsymbol{\theta}_1)$$

  Thus, the inner maximization corresponds to a weighted sum rate maximization for fixed $\boldsymbol{\theta}_1$ and $\nu_m = \boldsymbol{e}_m^{\mathsf{T}} (\boldsymbol{\mu} + \boldsymbol{\theta}_1)\, \eta$ for the optimum powers in (4.24), where η is chosen such that the sum power constraint is fulfilled with equality (see Appendix A2). The optimum Lagrange multipliers $\boldsymbol{\theta}_1^*$ are then found in an iterative manner. $\boldsymbol{\theta}_1^*$ minimizes the dual function, if

$$g(\boldsymbol{\theta}_1) - g(\boldsymbol{\theta}_1^*) \geq 0, \quad \forall \boldsymbol{\theta}_1 \geq \mathbf{0}_{K,1}.$$

$g(\boldsymbol{\theta}_1)$ can be lower bounded by using other powers than the optimum ones from (4.24) with the Lagrange multipliers $\boldsymbol{\theta}_1$ so that

$$\begin{aligned} g(\boldsymbol{\theta}_1) &= \max_{\boldsymbol{p}^{(i)}(k) \geq \mathbf{0}_{i,1}, \mathbf{1}_i^\mathsf{T} \boldsymbol{p}^{(i)}(k) \leq P_{\text{Tx}}} \left\{ \left(\boldsymbol{\mu}^\mathsf{T} + \boldsymbol{\theta}_1^\mathsf{T}\right) \boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right) \right\} - \boldsymbol{\theta}_1^\mathsf{T} \boldsymbol{r}_{\text{min}} \\ &= \left(\boldsymbol{\mu}^\mathsf{T} + \boldsymbol{\theta}_1^\mathsf{T}\right) \boldsymbol{r}(\boldsymbol{\theta}_1) - \boldsymbol{\theta}_1^\mathsf{T} \boldsymbol{r}_{\text{min}} \\ &\geq \left(\boldsymbol{\mu}^\mathsf{T} + \boldsymbol{\theta}_1^\mathsf{T}\right) \boldsymbol{r}(\boldsymbol{\theta}_1^*) - \boldsymbol{\theta}_1^\mathsf{T} \boldsymbol{r}_{\text{min}} \\ &= g(\boldsymbol{\theta}_1^*) - \left(\boldsymbol{\theta}_1^\mathsf{T} - \boldsymbol{\theta}_1^{*,\mathsf{T}}\right) \left(\boldsymbol{r}_{\text{min}} - \boldsymbol{r}(\boldsymbol{\theta}_1^*)\right) \end{aligned}$$

where $\boldsymbol{r}(\boldsymbol{\theta}_1)$ results from inserting the optimum powers from (4.24) into the rate expressions (4.20) and the inequality results from the fact that the powers optimum for $\boldsymbol{\theta}_1$ are not necessarily optimum for $\boldsymbol{\theta}_1^*$. Consequently,

$$g(\boldsymbol{\theta}_1) - g(\boldsymbol{\theta}_1^*) \geq \left(\boldsymbol{\theta}_1^{*,\mathsf{T}} - \boldsymbol{\theta}_1^\mathsf{T}\right) \left(\boldsymbol{r}_{\text{min}} - \boldsymbol{r}(\boldsymbol{\theta}_1^*)\right), \tag{4.25}$$

Thus, $\boldsymbol{\theta}_1^*$ minimizes the dual function, if

$$\left(\boldsymbol{\theta}_1^{*,\mathsf{T}} - \boldsymbol{\theta}_1^\mathsf{T}\right) \left(\boldsymbol{r}_{\text{min}} - \boldsymbol{r}(\boldsymbol{\theta}_1^*)\right) \geq 0, \quad \forall \boldsymbol{\theta}_1 \geq \mathbf{0}_{K,1}.$$

This inequality is fulfilled in case

$$\boldsymbol{\theta}_1^{*,\mathsf{T}} \left(\boldsymbol{r}_{\text{min}} - \boldsymbol{r}(\boldsymbol{\theta}_1^*)\right) = 0, \quad \text{and} \quad \boldsymbol{r}_{\text{min}} - \boldsymbol{r}(\boldsymbol{\theta}_1^*) \leq \mathbf{0}_{K,1}, \tag{4.26}$$

which represent the primal feasibility and complementary slackness constraints of the weighted sum rate maximization with minimum rate requirements, when the subchannel gains are fixed and the optimum subchannel powers should be found. Of course $\boldsymbol{\theta}_1^*$ could be found be an exhaustive search, i.e., all possible combinations of zero entries in $\boldsymbol{\theta}_1^*$ are tested until (4.26) is fulfilled. In the following a faster approach will be presented. Initially $\boldsymbol{\theta}_1^{(0)} = \mathbf{0}_{K,1}$ is taken and the corresponding rates $\boldsymbol{r}\left(\boldsymbol{\theta}_1^{(0)}\right)$ are computed. In case all minimum rate requirements are fulfilled, the optimum Lagrange multiplier has already been found. Otherwise, a new Lagrange multiplier will be determined as described for the jth step in the following. For all users m, which rates are below the minimum rates, i.e., $\boldsymbol{e}_m^\mathsf{T} \boldsymbol{r}\left(\boldsymbol{\theta}_1^{(j-1)}\right) < R_{\text{min},m}$, the corresponding entries in $\boldsymbol{\theta}_1^{(j)}$ will be non-zero and chosen such that the rates resulting from the optimum power allocation (4.24) are equal to the corresponding minimum rates, all other entries in $\boldsymbol{\theta}_1^{(j)}$ are set to zero. Thus, the non-zero entries in $\boldsymbol{\theta}_1^{(j)}$ are determined from the following implicit equations:

$$\begin{aligned} &\sum_{j=1}^{\hat{d}_m} \log_2 \left(\min \left(1, \frac{\boldsymbol{e}_m^\mathsf{T} \boldsymbol{\theta}_1^{(j)} \eta}{\ln(2)} \lambda_{m,j}^{(i)}(k) \right) \right) = R_{m,\text{min}}, \quad \forall m \in \mathcal{M}^{(j)} \\ &\sum_{m \in \mathcal{M}^{(j)}} \left[\frac{\left(\boldsymbol{e}_m^\mathsf{T} \boldsymbol{\theta}_1^{(j)} + \mu_m\right) \eta}{\ln(2)} - \frac{1}{\lambda_{m,j}^{(i)}(k)} \right]^+ + \sum_{m \notin \mathcal{M}^{(j)}} \left[\frac{\mu_m \eta}{\ln(2)} - \frac{1}{\lambda_{m,j}^{(i)}(k)} \right]^+ = P_{\text{Tx}}, \end{aligned} \tag{4.27}$$

where $\hat{d}_m$ is defined in (4.21) and $\mathcal{M}^{(j)}$ contains all user which rates $\boldsymbol{e}_\ell^{\mathsf{T}}\boldsymbol{r}\left(\boldsymbol{\theta}_1^{(j-1)}\right)$ are smaller than or equal to the corresponding minimum rate requirement, i.e.,

$$\mathcal{M}^{(j)} = \left\{\ell = 1,\ldots,K \,\middle|\, \boldsymbol{e}_\ell^{\mathsf{T}}\boldsymbol{r}\left(\boldsymbol{\theta}_1^{(j-1)}\right) \leq R_{\text{min},\ell}\right\}.$$

This choice for $\boldsymbol{\theta}_1^{(j)}$ assures that

$$\boldsymbol{\theta}_1^{(j),\mathsf{T}}\left(\boldsymbol{r}_{\text{min}} - \boldsymbol{r}\left(\boldsymbol{\theta}_1^{(j)}\right)\right) = 0, \tag{4.28}$$

as $\boldsymbol{e}_m^{\mathsf{T}}\boldsymbol{\theta}_1^{(j)} = 0$ for $m \notin \mathcal{M}^{(j)}$ and $\boldsymbol{e}_m^{\mathsf{T}}\left(\boldsymbol{r}_{\text{min}} - \boldsymbol{r}\left(\boldsymbol{\theta}_1^{(j)}\right)\right) = 0$ for $m \in \mathcal{M}^{(j)}$. The set $\mathcal{M}^{(j-1)}$ is completely contained in $\mathcal{M}^{(j)}$, i.e., $\mathcal{M}^{(j-1)} \subset \mathcal{M}^{(j)}$, as that set contains all users whose rates in step $j-1$ are lower or equal to the minimum rate requirements and for all users in the set $\mathcal{M}^{(j-1)}$ the minimum rate requirements are fulfilled with equality [c.f. (4.28)]. Correspondingly the vector $\boldsymbol{\theta}_1^{(j-1)}$ contains zeroes at all places, where $\boldsymbol{\theta}_1^{(j)}$ has zero entries, and additionally at places which correspond to users that are contained in $\mathcal{M}^{(j)}$ but not in $\mathcal{M}^{(j-1)}$. Consequently,

$$\boldsymbol{\theta}_1^{(j-1),\mathsf{T}}\left(\boldsymbol{r}_{\text{min}} - \boldsymbol{r}\left(\boldsymbol{\theta}_1^{(j)}\right)\right) = 0. \tag{4.29}$$

Inserting (4.28) and (4.29) into (4.25) it can be concluded that

$$g\left(\boldsymbol{\theta}_1^{(j-1)}\right) - g\left(\boldsymbol{\theta}_1^{(j)}\right) \geq \left(\boldsymbol{\theta}_1^{(j),\mathsf{T}} - \boldsymbol{\theta}_1^{(j-1),\mathsf{T}}\right)\left(\boldsymbol{r}_{\text{min}} - \boldsymbol{r}\left(\boldsymbol{\theta}_1^{(j)}\right)\right) = 0.$$

Thus, with the proposed iterative method the optimum of dual function cannot get worse than in the previous step. In case all rate requirements are fulfilled, the inequality and the equality in (4.26) hold and the optimum $\boldsymbol{\theta}_1^*$ has been found. As (4.26) are the primal feasibility and complementary slackness conditions of the primal problem of allocating power to scalar subchannels leading to the maximum weighted sum rate under a sum power and minimum rate constraints, and the powers resulting from (4.24) fulfill the dual feasibility conditions of this problem, the powers obtained this way lead to a KKT point of the primal problem. As in this case the objective function is concave and the constraint functions convex in the subchannel powers, these powers are also optimum for the primal power allocation problem.

- **Sum Rate Maximization with Relative Rate Constraints**
  For this instance of a QoS constrained utility maximization problem the dual problem in (4.22) reads as

$$\min_{\boldsymbol{\theta}_2} \max_{\boldsymbol{p}^{(i)}(k)\geq \mathbf{0}_{i,1}, \mathbf{1}_i^{\mathsf{T}}\boldsymbol{p}^{(i)}(k)\leq P_{\text{Tx}}} \left\{ R_1\left(\boldsymbol{p}^{(i)}(k)\right) - \boldsymbol{\theta}_2^{\mathsf{T}}\left(\mathbf{1}_{K-1}R_1\left(\boldsymbol{p}^{(i)}(k)\right) - \begin{bmatrix} \frac{R_2\left(\boldsymbol{p}^{(i)}(k)\right)}{\rho_2} \\ \vdots \\ \frac{R_K\left(\boldsymbol{p}^{(i)}(k)\right)}{\rho_K} \end{bmatrix}\right)\right\} = \min_{\boldsymbol{\theta}_2} g(\boldsymbol{\theta}_2).$$

As before, the inner maximization corresponds to weighted sum rate maximization for fixed $\boldsymbol{\theta}_2$, so that

$$\nu_1 = \left(1 - \boldsymbol{\theta}_2^{\mathsf{T}}\mathbf{1}_{K-1}\right)\eta, \quad \text{and} \quad \nu_m = \boldsymbol{e}_{m-1}^{\mathsf{T}}\boldsymbol{\theta}_2\eta, \quad m = 2,\ldots,K$$

in (4.24), where η is chosen such that the transmit power constraint is fulfilled with equality. Similarly to (4.25), the following inequality must hold

$$g\left(\boldsymbol{\theta}_2\right) - g\left(\boldsymbol{\theta}_2^*\right) \geq \left(\boldsymbol{\theta}_2^{*,\mathsf{T}} - \boldsymbol{\theta}_2^{\mathsf{T}}\right)\left(\mathbf{1}_{K-1} R_1\left(\boldsymbol{p}^{(i)}(k)\right) - \begin{bmatrix} \frac{R_2\left(\boldsymbol{p}^{(i)}(k)\right)}{\rho_2} \\ \vdots \\ \frac{R_K\left(\boldsymbol{p}^{(i)}(k)\right)}{\rho_K} \end{bmatrix}\right) \geq 0, \quad \forall \boldsymbol{\theta}_2$$

so that $\boldsymbol{\theta}_2^*$ minimizes the dual function. Together with the sum power constraint, the following set of equalities must therefore be fulfilled

$$\begin{aligned} & R_1\left(\boldsymbol{\theta}_2^*\right)\rho_m = R_m\left(\boldsymbol{\theta}_2^*\right), \quad m = 2, \ldots, K \\ & \sum_{j=1}^{\hat{d}_1}\left[\frac{\left(1 - \boldsymbol{\theta}_2^{*,\mathsf{T}}\mathbf{1}_{K-1}\right)\eta}{\ln(2)} - \frac{1}{\lambda_{1,j}^{(i)}(k)}\right]^+ + \sum_{m=2}^{K}\sum_{j=1}^{\hat{d}_m}\left[\frac{\boldsymbol{e}_{m-1}^{\mathsf{T}}\boldsymbol{\theta}_2^*\eta}{\ln(2)} - \frac{1}{\lambda_{m,j}^{(i)}(k)}\right]^+ = P_{\mathrm{Tx}}, \end{aligned} \tag{4.30}$$

which has a unique solution but must be solved iteratively, as done in [11]. For fixed $\kappa = \frac{\left(1 - \boldsymbol{\theta}_2^{*\mathsf{T}}\mathbf{1}_{K-1}\right)}{\eta}$ the first $K-1$ equations in (4.30) can be solved relatively easy. It i therefore proposed in [11] to solve (4.30) by a bisection search for the η fulfilling all equalities in this system of equations. Together with (4.24), the equations in (4.30) are identical to the KKT conditions of the primal power allocation problem, i.e., when the optimum powers for the rate balancing problem with fixed subchannel gains should be found. As a KKT point is necessary for the optimum and (4.30) has only one solution [11], the powers resulting from (4.30) are also optimum for the primal power allocation.

- **Sum Power Minimization with Minimum Rate Requirements**
  For this problem, the optimization of the dual function in (4.22) reads as

$$\min_{\boldsymbol{\theta}_1 \geq \mathbf{0}_{K-1,1}} \max_{\boldsymbol{p}^{(i)}(k) \geq \mathbf{0}_{i,1}} \left\{\boldsymbol{\theta}_1^{\mathsf{T}}\boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right) - \boldsymbol{p}^{(i),\mathsf{T}}(k)\mathbf{1}_K\right\} - \boldsymbol{\theta}_1^{\mathsf{T}}\boldsymbol{r}_{\mathrm{min}} = \min_{\boldsymbol{\theta}_1 \geq \mathbf{0}_{K-1,1}} g(\boldsymbol{\theta}_1),$$

which leads to $\nu_m = \boldsymbol{e}_m^{\mathsf{T}}\boldsymbol{\theta}_1$ in the optimum powers in (4.24) via evaluating the KKT. By inserting the optimum powers into the rate expressions and deriving $g(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}_1$ and setting it to zero one obtains

$$\frac{\mathrm{d}\, g(\boldsymbol{\theta}_1)}{\mathrm{d}\, \boldsymbol{\theta}_1} = \boldsymbol{r}_{\mathrm{min}} - \boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right) = 0.$$

This leads to the following implicit equations for $\boldsymbol{e}_m^{\mathsf{T}}\boldsymbol{\theta}_1$

$$\sum_{j=1}^{\hat{d}_m}\log_2\left(\min\left(1, \frac{\boldsymbol{e}_m^{\mathsf{T}}\boldsymbol{\theta}_1}{\ln(2)}\lambda_{m,j}^{(i)}(k)\right)\right) = R_{m,\mathrm{min}}, \quad m = 1, \ldots, K \tag{4.31}$$

As these equations lead to positive values for the elements in $\boldsymbol{\theta}_1$, the global minimum of $g(\boldsymbol{\theta}_1)$ does not violate the constraints $\boldsymbol{\theta}_1 \geq \mathbf{0}_{K-1,1}$. From these implicit equations it can also be concluded that the powers obtained by the solution of the dual problem lead to a feasible solution of the primal problem.

To summarize, the reduced complexity algorithms for utility constrained optimization problems have the same structure as Algorithm 3.2 in case DPC can be applied at the transmitter and as Algorithm 3.3 otherwise. Instead of the weighted sum rate in line 6 of Algorithm 3.2 and line 17 of Algorithm 3.3, for each user the value of the dual function

$$d^*(k) = \min_{\boldsymbol{\theta}_1 \geq \mathbf{0}_{c_1,1}, \boldsymbol{\theta}_2} \max_{\boldsymbol{p}^{(i)}(k) \geq \mathbf{0}_{i,1}, \mathbf{1}_i^{\mathsf{T}} \boldsymbol{p}^{(i)}(k) \leq P_{\text{Tx}}} \max_{\boldsymbol{r} \in \mathcal{C}_{\text{R}}\left(\boldsymbol{p}^{(i)}(k)\right)} \left\{ u\left(\boldsymbol{p}^{(i)}(k), \boldsymbol{r}\right) - \boldsymbol{\theta}_1^{\mathsf{T}} \boldsymbol{h}_1(\boldsymbol{r}) - \boldsymbol{\theta}_2^{\mathsf{T}} \boldsymbol{h}_2(\boldsymbol{r}) \right\}$$

optimum for the successive resource allocation and spatial zero-forcing method is computed and used for the selection process. During the initialization phase the modified optimum value of the dual function

$$d^*(k) = \min_{\boldsymbol{\theta}_1 \geq \mathbf{0}_{c_1,1}, \boldsymbol{\theta}_2} \max_{\boldsymbol{p}^{(i)}(k) \geq \mathbf{0}_{i,1}, \mathbf{1}_i^{\mathsf{T}} \boldsymbol{p}^{(i)}(k) \leq P_{\text{Tx}}}$$
$$\left\{ u\left(\boldsymbol{p}^{(i)}(k), \boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right)\right) - \hat{\boldsymbol{\theta}}_1^{(i)\,\mathsf{T}}(k) \hat{\boldsymbol{h}}_1^{(i)}\left(k, \boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right)\right) - \hat{\boldsymbol{\theta}}_2^{(i)\,\mathsf{T}}(k) \hat{\boldsymbol{h}}_2^{(i)}\left(k, \boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right)\right) \right\}$$

is used, where additionally only users can be selected that have not received a data stream yet. In line 2 of Algorithm 3.3 furthermore that user has to be chosen that minimum rate requirement can be fulfilled with a single data stream with minimum power or in case of the rate balancing problem, user 1 is chosen. Additionally, the power allocations in line 18 of Algorithm 3.2 and line 22 of Algorithm 3.3 have to be replaced by the corresponding QoS constrained power allocation problems leading to the powers (4.24) with the water-levels stemming from the implicit equations (4.27), (4.30) and (4.31). As shown previously, these powers lead to feasible solutions, which redundantizes the primal reconstruction and a possible implementation of time-sharing required for the optimum algorithm.

4.5 Further Complexity Reductions

Apart from the power minimization problem, the determination of the Lagrange multipliers $\boldsymbol{\theta}_1^*$ and $\boldsymbol{\theta}_2^*$ optimum in (4.19) can only be done in an iterative manner, as shown in the previous section, although it has been possible to avoid the use of ellipsoid methods. Especially, when the number of users grows large, the complexity of the user selection can become intractable, as for each user in each step the optimum value of the dual function has to be computed. For this reason a simplified user selection will be presented next, which has been proposed by the author in [108]. Instead of looking for the minimum with respect to $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ in (4.22), the optimum Lagrange multipliers from step $i-1$, denoted as $\boldsymbol{\theta}_1^{(i-1)}$ and $\boldsymbol{\theta}_2^{(i-1)}$, are taken so that Problem (4.22) simplifies to

$$\max_{k} \max_{\boldsymbol{p}^{(i)}(k) \geq \mathbf{0}_{i,1}, \mathbf{1}_i^{\mathsf{T}} \boldsymbol{p}^{(i)}(k) \leq P_{\text{Tx}}} \left\{ u\left(\boldsymbol{p}^{(i)}(k), \boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right)\right) - \boldsymbol{\theta}_1^{(i-1),\mathsf{T}} \boldsymbol{h}_1\left(\boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right)\right) - \right.$$
$$\left. - \boldsymbol{\theta}_2^{(i-1),\mathsf{T}} \boldsymbol{h}_2\left(\boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right)\right) \right\}. \quad (4.32)$$

This choice is motivated by the fact that, with increasing number of allocation steps, the implicit equations (4.27), (4.30) and (4.31) less and less change through adding a new subchannel from step $i-1$ to step i, which implies that $\boldsymbol{\theta}_1^{(i-1)}$ and $\boldsymbol{\theta}_2^{(i-1)}$ are good estimates for the optimum Lagrange multipliers in the previous step. This way, the inner maximization in (4.32) becomes a weighted sum rate maximization for the the QoS constrained utility maximization problems with a sum power constraint. For the weighted sum rate maximization with minimum rate requirements, Problem (4.32) reads as

$$\max_k \max_{\boldsymbol{p}^{(i)}(k)\geq \boldsymbol{0}_{i,1}, \boldsymbol{1}_i^{\mathsf{T}}\boldsymbol{p}^{(i)}(k)\leq P_{\mathrm{Tx}}} \left(\boldsymbol{\mu}^{\mathsf{T}} + \boldsymbol{\theta}_1^{(i-1),\mathsf{T}}\right) \boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right)$$

and for the rate balancing problem, (4.32) can be written as

$$\max_k \max_{\boldsymbol{p}^{(i)}(k)\geq \boldsymbol{0}_{i,1}, \boldsymbol{1}_i^{\mathsf{T}}\boldsymbol{p}^{(i)}(k)\leq P_{\mathrm{Tx}}} \left(1 - \boldsymbol{1}_{K-1}^{\mathsf{T}}\boldsymbol{\theta}_2^{(i-1)}\right) R_1\left(\boldsymbol{p}^{(i)}(k)\right) + \boldsymbol{\theta}_2^{\mathsf{T},(i-1)} \begin{bmatrix} \frac{R_2\left(\boldsymbol{p}^{(i)}(k)\right)}{\rho_2} \\ \vdots \\ \frac{R_K\left(\boldsymbol{p}^{(i)}(k)\right)}{\rho_K} \end{bmatrix}.$$

The dual problem of the power minimization is transformed into a weighted sum rate maximization by introducing an artificial power constraint, which implies that the transmit power must be equal to the optimum power of the previous step given by $\boldsymbol{1}_{i-1}^{\mathsf{T}}\boldsymbol{p}^{(i-1)}(k(i-1))$ so that Problem (4.22) reads as

$$\max_k \max_{\boldsymbol{p}^{(i)}(k)\geq \boldsymbol{0}_{i,1}, \boldsymbol{1}_i^{\mathsf{T}}\boldsymbol{p}^{(i)}(k)\leq \boldsymbol{1}_{i-1}^{\mathsf{T}}\boldsymbol{p}^{(i-1)}(k(i-1))} \boldsymbol{\theta}_1^{\mathsf{T},(i-1)}\boldsymbol{r}\left(\boldsymbol{p}^{(i)}(k)\right).$$

Thus, the user selection is done as for a weighted sum rate maximization and the optimum Lagrange multipliers $\boldsymbol{\theta}_1^{(i)}$ and $\boldsymbol{\theta}_2^{(i)}$ need only be computed as described in the previous section for the user $k(i)$ selected in step i. While this measure on its own does not gain much complexity reductions, it enables the application of user preselection presented in Section 3.4.1 and the user selection based on an upper bound for weighted sum rate from Section 3.4.3. Both methods have turned out to be very effective for the weighted sum rate maximization and will also lead to drastic complexity reductions at almost no performance losses with QoS constrained utility maximization problems. Nevertheless, the proposed complexity reduction cannot be applied during initialization. However, during this phase the number of candidate users that have to be tested shrinks with each step, as a user is excluded from the selection process after a subchannel has been allocated to it.

4.6 Numerical Results

In Figure 4.1 the average weighted sum rates per subcarrier are plotted versus the SNR, which is defined as the tenfold logarithm of the ratio of transmit power P_{Tx} to noise variance σ_{n}^2. There are $K = 5$ users in the system, each equipped with $r_k = 2$ antennas and the transmitter has $N_{\mathrm{Tx}} = 4$ antennas. The weighted sum rates are averaged over 500 channel realizations, where for each channel realization there are $L_k = 4$ temporal propagation paths for each user. Each channel matrix $\boldsymbol{H}_{k,\ell}$, $k = 1, \ldots, K$, $\ell = 1, \ldots, 4$ consists of circularly symmetric Gaussian entries with zero mean and unit variance, where the temporal distance between two consecutively arriving symbols is equal to $160\mu s$, i.e., $\tau_{k,m} - \tau_{k,m-1} = 160\mu s$ for all $k = 1 \ldots, K$, $m = 2, \ldots, 4$. OFDM with $C = 16$ carrier is employed to mitigate intersymbol interference and the bandwidth is equal to $B = 150$kHz at a center frequency of $f_c = 5$GHz. As utility function the weighted sum rate is used in Figure 4.1, where the even users have twice the priority of the odd users so that

$$\mu_1 = \mu_3 = \mu_5 = \frac{1}{7}, \qquad \mu_2 = \mu_4 = \frac{2}{7},$$

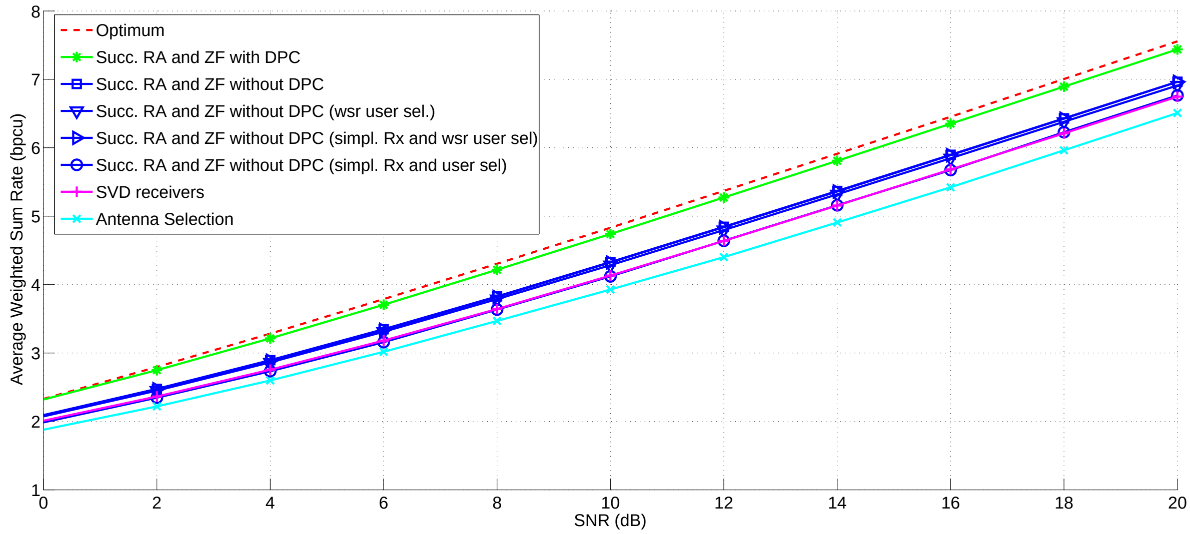

Figure 4.1: Average weighted sum rates with minimum rate constraints in a system with $K = 5$ users with $r_k = 2$ receive antennas, $N_\mathrm{T} = 4$ transmit antennas, $C = 16$ carrier and $B = 150$kHz. $\mu_1 = \mu_3 = \mu_5 = \frac{1}{7}$, $\mu_2 = \mu_4 = \frac{2}{7}$.

and minimum rate constraints are imposed for each user. Those rate constraints increase linearly with the SNR, so that

$$\frac{R_{k,\mathrm{min}}}{C} = R_{k,0} + \mathrm{SNR}/4$$

with $R_{k,0} = 1$bpcu for all users k. Thus, at 0dB each user has a minimum rate requirement of 1bpcu per subcarrier. The increase of the rate requirements with SNR is done so that the minimum rate requirements come into effect for all SNR values. This way, the minimum rate requirements are not fulfilled by simply maximizing the weighted sum rate, i.e., the optimization is not solved by the algorithms presented in Chapter 3.

As in the case of weighted sum rate maximization without minimum rate requirements, the loss of giving up DPC with the proposed method (“Succ RA and ZF without DPC”) is acceptable. The adjunct “wsr user sel.” implies that the user selection is based on a weighted sum rate maximization with the Lagrange multipliers from the previous step. As it can be seen in Figure 4.1, this measure leads to negligible performance losses compared to the method, where in each step the dual function is evaluated exactly. Furthermore with the simplified user selection the user preselection explained in Section 3.4.1 can be applied, which has reduced the complexity of the user selection by 21.4% in the simulations. Additionally using simplified receivers according to (3.75) leads to the curve labelled as “Succ RA and ZF without DPC (simpl. Rx and wsr user sel.)” in Figure 4.1. By selecting in each step the user for the next data stream according to an upper bound for the weighted sum rate obtained with the Lagrange multipliers from the previous step, which is labeled with the adjunct “(simpl. Rx and user sel.)” leads to small performance losses compared to the other methods. Nevertheless, the same average sum rate as with SVD receivers, where the left singular vectors are used as receive filters, can be achieved and antenna selection with canonical unit vectors as receive filters is outperformed for all SNR values.

5. Large System Analysis

The close-to-optimum performance of the algorithms proposed in the last two chapters has been shown by simulations results. Although impressing and similar performance can be obtained in other scenarios, these results are only valid for a special simulation setup and cannot be generalized to an arbitrary set of parameters. In this chapter some analytical results will therefore be presented for the some of the derived algorithms, when the temporal channel matrices contain Gaussian i.i.d. entries. Direct expressions for the average objective function, as the weighted sum rate, are however difficult to obtain. A common approach in the literature for the analysis of greedy zero-forcing approaches, which aim at sum rate maximization, is therefore to let the number of users grow to infinity and to show that in this case the optimum can be achieved [65], [64], [68]. Conclusions to the performance of the algorithm in a system with finite number of users can however not directly be drawn. For this reason results of large system analysis will be applied in this chapter, where at least two parameters go to infinity at a finite fixed ratio. In the large system limit many expressions of random variables, especially the eigenvalues of large random matrices, become deterministic. This facilitates to find analytical expressions for the objective functions. As it will be seen, the results obtained this way approximate the system performance quite well already with parameters of moderate finite size. After giving a short overview of literature on large system analysis in Section 5.1 and explaining some preliminaries in Section 5.2, the case of infinite number of users and transmit antennas in MISO systems will be treated in Section 5.3. MIMO systems with a finite number of users but infinite numbers of transmit and receive antennas will be considered in Section 5.4.

5.1 Related Work

In this section a short overview on application examples of large systems analysis to communication systems will be given. For an extensive overview of fundamentals of large system analysis and a some application examples the reader is referred to [109]. In Code Division Multiple Access (CDMA) systems, large system analysis is a popular tool for performance evaluation, when the number of users and the length of the spreading sequences tend towards infinity. In [110] expressions for asymptotic signal-to-interference-ratios in the uplink are given for different receive filters. This work is extended to time-varying channels and imperfect channel knowledge in [111]. SINRs in the downlink of large CDMA systems for two kinds of receive filters are derived in [112] and [113], where the results are also applicable to multi-antenna systems with the number of transmit and receive antennas going to infinity at a finite fixed ratio. Furthermore two kind of random matrices are considered in the analysis, namely those containing i.i.d. entries and those consisting of random orthonormal columns. In [114] inter-cell interference is additionally taken into account in these CDMA scenarios. For MIMO systems with channel matrices containing i.i.d. entries expressions for the ergodic capacity under statistical channel knowledge in the large system limit are given in [115], where white Gaussian noise as well as colored noise originating from multi-user interference is considered. An high SNR analysis of the scenario with white noise is conducted in [116]. Regularized beamforming in multiuser MISO systems is analyzed in [117] for an infinite

number of transmit antennas and users, where the asymptotically optimum regularization parameter is determined for those parameters. The same large system limit is used in [118] to obtain asymptotic expressions for the ergodic sum rate, when zero-forcing and DPC are used at the transmitter with quantized feedback of channel state information. Results for the large system rate in single-user MISO and MIMO systems using Random Vector Quantization (RVQ) are presented in [119], where besides the number of antennas also the number of feedback bits grows towards infinity.
Recently, the results from large system analysis have been applied to multi-cell MIMO systems. While the examples for large system analysis in single-cell systems cited above are based on random matrices which contain random i.i.d. entries with zero mean or random orthonormal entries as in [112] and [113], for multi-cell systems the random matrices contain entries with different variances, as channel matrices modeling the interference from neighboring cells in practice contain on average smaller entries than the other channel matrices. In [120] large system methods are used to simplify the problem of finding the optimum power and user allocation for weighted ergodic sum rate maximization with zero-forcing beamforming in multi-cell MISO systems, where the parameters growing to infinity are the number of antennas at the base stations and the number of users. Thus, numerically complex Monte-Carlo simulations can be avoided. The same aim is pursued in [121], where a similar problem as in [120] is considered but without zero-forcing beamformers. This algorithm is used iteratively in [122] to take fairness criteria into account. Different cooperation schemes in a two cell MISO scenario are compared with each other for an infinite number of transmit antennas and users in [123] and a multi-cell MIMO system, where the number of transmit and receive antennas tends towards infinity and only one user per cell is active, is analyzed in [124]. Considering correlations between the entries of the channel matrices or introducing random variables with non-zero means makes the large system analysis more complicated than with independent entries. That is why for those cases only results for the ergodic capacity in single-user MIMO systems exist. Reference [125] treats correlations between the entries of the channel matrices and in [126] Rician fading channels, where the random channel matrices have non-zero mean, are considered. An algorithm for finding the transmit covariance matrix maximizing the ergodic sum capacity in a single-user MIMO system under Rician fading is presented in [127].

5.2 Preliminaries

The analysis in this chapter is focussed on scenarios, where the temporal channel matrices $\tilde{\boldsymbol{H}}_{k,\ell}$ contain circularly symmetric Gaussian i.i.d. entries with zero mean and variance $\sigma_{k,\ell}^2$, i.e.

$$\operatorname{vec}\left(\tilde{\boldsymbol{H}}_{k,\ell}\right) \sim \mathcal{CN}\left(\mathbf{0}_{r_k N_\mathrm{T},1}, \sigma_{k,\ell}^2 \boldsymbol{I}_{r_k N_\mathrm{T}}\right).$$

Furthermore the channel matrices $\tilde{\boldsymbol{H}}_{k,\ell}$ are uncorrelated with each other. As the sum of independent circularly symmetric Gaussian i.i.d. variables is again circularly symmetric (e.g. [2, Lemma 4]), where the mean and the variance of the new random variable is given by the sum of the means and the variances, respectively, of the summands, the entries of the carrier channel matrices

$$\boldsymbol{H}_{k,c} = \boldsymbol{S}_{c,k} \boldsymbol{H}_k \boldsymbol{S}_{c,\mathrm{T}}^\mathsf{T} = \sum_{\ell=1}^{L_k} \tilde{\boldsymbol{H}}_{k,\ell} \exp\left(j 2\pi f_c \left(\tau_{k,1} - \tau_{k,\ell}\right)\right)$$

are i.i.d. and are drawn from a circularly symmetric Gaussian distribution with zero mean and variance $\sum_{\ell=1}^{L_k} \sigma_{k,\ell}^2$. In the remainder it will be assumed that all users' channel matrices in the frequency domain have the same variance

$$\sigma_{\mathrm{H}}^2 = \sum_{\ell=1}^{L_k} \sigma_{k,\ell}^2$$

so that

$$\operatorname{vec}\left(\boldsymbol{H}_{k,c}\right) \sim \mathcal{CN}\left(\boldsymbol{0}_{r_k N_{\mathrm{T}},1}, \sigma_{\mathrm{H}}^2 \boldsymbol{I}_{r_k N_{\mathrm{T}}}\right). \tag{5.1}$$

Clearly, the channel matrices of different carriers are correlated with each other, as they are computed from the same realizations of the temporal channel matrices $\tilde{\boldsymbol{H}}_{k,\ell}$. Nevertheless, as will be seen in the remainder, this property is not relevant in the large system limit.
If the matrices $\boldsymbol{H}_{k,c}$ are random variables, so are the eigenvalues $\rho_1\left(\boldsymbol{H}_{k,c}^{\mathrm{H}}\boldsymbol{H}_{k,c}\right), \ldots, \rho_{N_{\mathrm{T}}}\left(\boldsymbol{H}_{k,c}^{\mathrm{H}}\boldsymbol{H}_{k,c}\right)$ of the matrices $\boldsymbol{H}_{k,c}^{\mathrm{H}}\boldsymbol{H}_{k,c}$. In general the empirical eigenvalue distribution of an arbitrary random Hermitian matrix $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ with $\boldsymbol{A} \in \mathbb{C}^{n\times m}$

$$F_{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}(x) = \frac{1}{m}\left|\left\{\rho_i\left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\right), i = 1, \ldots, m \,\middle|\, \rho_i\left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\right) \leq x\right\}\right|,$$

which states the fraction of eigenvalues that are smaller or equal than x, is different for each realization of the random matrix $\boldsymbol{A}$. As the dimensions of the matrix $\boldsymbol{A}$ grow towards infinity at a finite fixed ratio $\beta = \frac{m}{n}$, i.e., $m \to \infty$, $n \to \infty$, the empirical eigenvalue distribution converges to an asymptotic limit $F_{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}^{(\infty)}(x)$ for many random matrices, which is the same for all realizations of the random matrix and only depends on β. The derivation of $F_{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}^{(\infty)}(x)$ is called the asymptotic eigenvalue distribution (aed) $f_{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}^{(\infty)}(x)$ so that

$$f_{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}^{(\infty)}(x) = \frac{\mathrm{d}\, F_{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}^{(\infty)}(x)}{\mathrm{d}\, x}.$$

For the case of Gaussian i.i.d. entries with zero-mean and variance $\frac{1}{n}$ in the matrix $\boldsymbol{A}$, the asymptotic eigenvalue distribution $f_{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}^{(\infty)}(x)$ is given by the Marčenko-Pastur distribution [128]

$$f_{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}^{(\infty)}(x) = f_{\mathrm{MP}}^{(\infty)}\left(x, \frac{m}{n}\right),$$

where

$$f_{\mathrm{MP}}^{(\infty)}(x, \hat{\beta}) = \left[1 - \frac{1}{\hat{\beta}}\right]^{+} \delta(x) + \frac{\sqrt{[x-a]^{+}\,[b-x]^{+}}}{2\pi\hat{\beta}x} \tag{5.2}$$

with

$$a = \left(1 - \sqrt{\hat{\beta}}\right)^2, \qquad b = \left(1 + \sqrt{\hat{\beta}}\right)^2$$

and $\hat{\beta}$ is equal to the ratio of columns to rows in the matrix $\boldsymbol{A}$. An important application of the asymptotic eigenvalue distributions, that will be used in the remainder, is the replacement of the sum of a function of the eigenvalues by an integral in the large system limit, where the sum is taken over all eigenvalues of a matrix. That implies that

$$\frac{1}{m}\sum_{i=1}^{m} g\left(\rho_i\left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\right)\right) \underset{\substack{m,n\to\infty \\ \frac{m}{n}=\beta}}{\longrightarrow} \int_0^{\infty} g(x) f_{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}^{(\infty)}(x)\,\mathrm{d}\, x \tag{5.3}$$

can be computed explicitly and is independent of the realization of the random matrix $\boldsymbol{A}$ in the large system limit. In the following some important examples for $g\left(\rho_i\left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\right)\right)$ are introduced that will be useful in the remainder of this chapter. Choosing

$$g\left(\rho_i\left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\right)\right) = \frac{1}{\rho_i\left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\right) - z},$$

leads to the Stieltjes transform $m_{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}(z)$ of the matrix $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$, which is defined as

$$m_{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}(z) = \frac{1}{m}\sum_{i=1}^{m}\frac{1}{\rho_i\left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\right) - z} = \frac{1}{m}\operatorname{tr}\left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A} - z\boldsymbol{I}_m\right)^{-1} = \int_0^{\infty}\frac{1}{x-z}f^{(\infty)}_{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}(x)\,\mathrm{d}\,x$$

(e.g. [109, Chapter 2.1.1]). Conversely, given the Stieltjes transform of the matrix $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ its asymptotic eigenvalue distribution can be obtained from the inversion formula

$$f^{(\infty)}_{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}(x) = \lim_{\omega\to 0}\frac{1}{\pi}\operatorname{Im}\left\{m_{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}(x + j\omega)\right\}. \tag{5.4}$$

The η-transform $\eta_{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}(\gamma)$ is defined for $\gamma \geq 0$ as

$$\eta_{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}(\gamma) = \int_0^{\infty}\frac{1}{1+\gamma x}f^{(\infty)}_{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}(x)\,\mathrm{d}\,x$$

and is related to the Stieltjes-transform via

$$\eta_{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}(\gamma) = \frac{1}{\gamma}m_{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}\left(-\frac{1}{\gamma}\right). \tag{5.5}$$

Finally, the Shannon transform $\mathcal{V}_{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}(\gamma)$ of the matrix $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ is given as

$$\mathcal{V}_{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}(\gamma) = \int_0^{\infty}\log_2\left|\boldsymbol{I}_m + \gamma\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\right| f^{(\infty)}_{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}(x)\,\mathrm{d}\,x, \tag{5.6}$$

where the notation from [109] has been used.

5.3 Large System Analysis in MISO Systems

In this section large system analysis is carried out in MISO systems, where the number of transmit antennas and the number of users go to infinity at a finite fixed ratio α, i.e.,

$$N_{\mathrm{T}} \to \infty, \qquad K \to \infty, \qquad \alpha = \frac{N_{\mathrm{T}}}{K}.$$

All other parameters such as transmit power P_{Tx} and number of carriers C, remain finite, the number of receive antennas is given by $r_k = 1$ in MISO systems anyway. The analysis is carried out for weighted sum rate as objective function under a sum power constraint. It is assumed that there are N disjoint subgroups of users, where the nth subgroup contains $\beta_n K$ users and all users in this group have the same weighting factor μ_n. The β_n are finite and describe the percentage of users having priority μ_n, so that $\sum_{n=1}^{N}\beta_n = 1$. Correspondingly, each group contains infinitely many users. For notational convenience it is assumed that the indices n are chosen such that

$$\mu_1 > \cdots > \mu_N.$$

When the same QoS constraints are imposed on each subgroup of infinitely many users, the achievable rates in the large system limit can be derived in a similar manner than for weighted sum rate maximization. However, as the expressions would get more involved than with weighted sum rate, the following analysis is restricted to the weighted sum rate maximization problem.
Clearly the results are directly applicable to MIMO systems, when antenna selection is applied at the receivers. In that case α is given by the ratio of transmit antennas to sum of receive antennas. Thus, the large system analysis with finite α holds for an infinite number of transmit antennas and either a finite number of users and an infinite number of receive antennas at each user or an infinite number of users with finite number of transmit antennas. In the following, expressions for the optimum weighted sum and the weighted sum rates achievable with successive resource allocation and zero-forcing with and without DPC will be given in the large system limit.

- **Optimum Weighted Sum Rate**:
  As in Section 3.1, the problem of weighted sum rate maximization is solved in the dual MAC in the large system limit as well, because the rates achievable in the dual MAC are the same is in the broadcast channel due to the duality from [5]. Considering the case of subgroups of users having equal weights the problem of weighted sum rate maximization from (3.12) reads as

$$\max_{\{\boldsymbol{W}_k\}_{k=1,\ldots,K}} \sum_{n=1}^{N} \Delta\mu_n \log_2 \left| \boldsymbol{I}_{CN_T} + \frac{C}{\sigma_\mathrm{n}^2} \left[\sum_{m=1}^{\sum_{n'=1}^{n} \beta_{n'} K} \boldsymbol{H}_{\hat{\pi}(m)}^\mathrm{H} \boldsymbol{W}_k \boldsymbol{H}_{\hat{\pi}(m)} \right] \right| =$$

$$\text{s.t.} \sum_{k=1}^{K} \mathrm{tr}(\boldsymbol{W}_k) \leq P_\mathrm{Tx}, \quad \boldsymbol{W}_k \succeq 0, \forall k, \tag{5.7}$$

where $\hat{\pi}(m)$ denotes the user encoded at mth place in the broadcast channel and

$$\Delta\mu_n = \begin{cases} \mu_N, & n = N, \\ \mu_n - \mu_{n+1}, & n < N \end{cases}.$$

When the number of users and the number of transmit antennas tend towards infinity, Algorithm 3.18 can still be applied to obtain the optimum solution. Nevertheless, it is possible in the large system limit to derive an analytical expression for the weighted sum rate, which is given by

$$\frac{R_\text{WSR,opt}}{CN_\mathrm{T}} \underset{\substack{N_\mathrm{T},K\to\infty,\\ \frac{N_\mathrm{T}}{K}=\alpha}}{\longrightarrow} \sum_{n=1}^{N} \Delta\mu_n \left[\frac{1}{\alpha} \sum_{j=1}^{n} \beta_j \log_2 \left(1 + \frac{\alpha}{\sigma_\mathrm{n}^2 \beta_j} w_j m_n \right) + \right.$$
$$\left. + \log_2 \left(\frac{\sigma_\mathrm{H}^2}{m_n} \right) + \left(\frac{m_n}{\sigma_\mathrm{H}^2} - 1 \right) \log_2(\mathrm{e}) \right], \tag{5.8}$$

where e is Euler's number. The variables m_n and w_n stem from the system of implicit equations

$$\begin{aligned}\sum_{j=n}^{N} \Delta\mu_j \frac{m_j}{1+w_n m_j \frac{\alpha}{\sigma_{\mathrm{n}}^2 \beta_n}} w_n &= \lambda w_n, \quad n=1,\ldots,N \\ \lambda\left(\sum_{n=1}^{N} w_n - P_{\mathrm{Tx}}\right) &= 0, \quad \lambda \geq 0, \quad w_n \geq 0, n=1,\ldots,N \\ \sum_{n'=1}^{n} \beta_{n'} \frac{1}{\alpha} &= \frac{1-\frac{m_n}{\sigma_{\mathrm{H}}^2}}{1-\left(\sum_{m=1}^{n} \frac{\beta_m}{\sum_{n'=1}^{n}\beta_{n'}} \frac{1}{1+m_n w_m \frac{\alpha}{\sigma_{\mathrm{n}}^2 \beta_m}}\right)} \quad n=1,\ldots,N,\end{aligned} \tag{5.9}$$

which is derived in Appendix A7, where it is also shown that the system of implicit equations (5.9) always has a valid solution. In case all users' weights are equal, there is only one subgroup, i.e., $N=1$. Then, the system of equations (5.9) can be solved explicitly so that

$$w_1 = P_{\mathrm{Tx}}$$

and

$$m_1 = \frac{\sigma_{\mathrm{H}}^2}{2}\left(1-\frac{1}{\alpha}\right) - \frac{\sigma_{\mathrm{n}}^2}{2P_{\mathrm{Tx}}\alpha} + \sqrt{\left(\frac{\sigma_{\mathrm{H}}^2}{2}\left(1-\frac{1}{\alpha}\right) - \frac{\sigma_{\mathrm{n}}^2}{2P_{\mathrm{Tx}}\alpha}\right)^2 + \frac{\sigma_{\mathrm{H}}^2\sigma_{\mathrm{n}}^2}{P_{\mathrm{Tx}}\alpha}},$$

which follows from taking the positive solution resulting from the quadratic equation of the last line in (5.9). Thus, the large system sum capacity is given by

$$\frac{C_{\mathrm{sum}}}{CN_{\mathrm{T}}} \underset{\substack{N_{\mathrm{T}},K\to\infty,\\ \frac{N_{\mathrm{T}}}{K}=\alpha}}{\longrightarrow} \left[\frac{1}{\alpha}\log_2\left(1+\frac{\alpha P_{\mathrm{Tx}}}{\sigma_{\mathrm{n}}^2} m_1\right) + \log_2\left(\frac{\sigma_{\mathrm{H}}^2}{m_1}\right) + \left(\frac{m_1}{\sigma_{\mathrm{H}}^2}-1\right)\log_2(\mathrm{e})\right].$$

- **Successive Resource Allocation and Spatial Zero-Forcing with DPC**:
  In MISO systems, there is no optimization of receive filters. For equal weights, Algorithm 3.2 is therefore identical to the algorithm proposed in [43]. While this paper only considers an analysis, where the number of users goes to infinity, in this section a large system analysis is carried out for an infinite number of users as well as an infinite number of transmit antennas. As derived in Appendix A8, the weighted sum rate in this large system limit can be lower bounded by

$$\begin{aligned}\frac{R_{\mathrm{WSR},DPC}}{CN_{\mathrm{T}}} \underset{\substack{K,N_{\mathrm{T}}\to\infty\\ \frac{N_{\mathrm{T}}}{K}=\alpha}}{\longrightarrow} &\frac{1}{\alpha}\sum_{n=1}^{n_{\max}-1} \beta_n \mu_n \log_2(\eta\mu_n) + \left(\frac{j_{\max}}{N_{\mathrm{T}}} - \frac{1}{\alpha}\sum_{n=1}^{n_{\max}-1}\beta_n\right)\mu_{n_{\max}}\log_2(\eta\mu_{n_{\max}}) \\ &+ \sum_{n=1}^{n_{\max}-1} \mu_n\left[\hat{\beta}_{n-1}\log_2\left(\hat{\beta}_{n-1}\right) - \hat{\beta}_n\log_2\left(\hat{\beta}_n\right) + \frac{1}{\ln 2}\left(\hat{\beta}_n - \hat{\beta}_{n-1}\right)\right] + \\ + \mu_{n_{\max}}&\left\{\left(1-\frac{j_{\max}}{N_{\mathrm{T}}}\right)\left[\frac{1}{\ln 2} - \log_2\left(1-\frac{j_{\max}}{N_{\mathrm{T}}}\right)\right] + \left(\hat{\beta}_{n_{\max}-1}\right)\left[\log_2\left(\hat{\beta}_{n_{\max}-1}\right) - \frac{1}{\ln 2}\right]\right\},\end{aligned} \tag{5.10}$$

where

$$\hat{\beta}_n = 1 - \sum_{n'=1}^{n}\frac{\beta_{n'}}{\alpha}, \quad \text{and} \quad \eta = \sigma_{\mathrm{H}}^2\alpha \frac{\frac{P_{\mathrm{Tx}}}{\sigma_{\mathrm{n}}^2} - \frac{1}{\sigma_{\mathrm{H}}^2}\ln\left(1-\frac{j_{\max}}{N_{\mathrm{T}}}\right)}{\sum_{n=1}^{n_{\max}-1}\beta_n\mu_n + \left(\frac{j_{\max}}{K} - \sum_{n=1}^{n_{\max}-1}\beta_n\right)\mu_{n_{\max}}}.$$

$j_{\max}$ is determined so that

$$\eta\mu_{n_{\max}} \geq \frac{1}{1-\frac{j_{\max}}{N_{\mathrm{T}}}} \qquad \text{and} \qquad \frac{j_{\max}}{N_{\mathrm{T}}} \leq \min\left(1, \frac{1}{\alpha}\right), \tag{5.11}$$

where due to water-filling one of these inequalities is always fulfilled with equality. Equation (5.10) requires a user allocation in decreasing order of weights, i.e., first all users in the subgroup with the strongest weight are served, then the users in the subgroup with the second largest weight and so forth. The index of this subgroup with the lowest weight containing users that are served is denoted by the index $n_{\max}$ in (5.10). Although (5.10) states a lower bound for the weighted sum rate, which is tight for an arbitrary user selection within the same group, this lower bound is valuable for analysis purposes, as it can be used to state an upper bound for the loss compared to the optimum and a lower bound for the gain achievable with DPC compared to algorithms without DPC.

For the special case of sum rate maximization, i.e., all users' weights are equal, Equation (5.10) can be simplified and the sum rate can be lower bounded in the large system limit by

$$\frac{R_{\mathrm{SR},DPC}}{CN_{\mathrm{T}}} \underset{\substack{K,N_{\mathrm{T}}\to\infty \\ \frac{N_{\mathrm{T}}}{K}=\alpha}}{\longrightarrow} \frac{j_{\max}}{N_{\mathrm{T}}}\log_2(\eta) - \left(1-\frac{j_{\max}}{N_{\mathrm{T}}}\right)\log_2\left(1-\frac{j_{\max}}{N_{\mathrm{T}}}\right) - \frac{j_{\max}}{\ln 2 N_{\mathrm{T}}}.$$

The water-level η is given by

$$\eta = \sigma_{\mathrm{H}}^2 \frac{\frac{P_{\mathrm{Tx}}}{\sigma_{\mathrm{N}}^2} - \frac{1}{\sigma_{\mathrm{H}}^2}\ln\left(1-\frac{j_{\max}}{N_{\mathrm{T}}}\right)}{\frac{j_{\max}}{N_{\mathrm{T}}}}$$

and $j_{\max}$ is determined as described above with $\mu_{n_{\max}} = 1$.

- **Successive Resource Allocation and Spatial Zero-Forcing without DPC**:
Similarly to the DPC case, in MISO systems Algorithm 3.3 with equal weights is identical to an algorithm from literature, in this case it corresponds to the method presented in [12]. Again, this paper considers only the case of infinite number of users, whereas in the following both the number of users and number of transmit antennas tend towards infinity at a finite fixed ratio α. The asymptotic weighted sum rate in this limit can be lower bounded by

$$\frac{R_{\mathrm{WSR,lin}}}{CN_{\mathrm{T}}} \underset{\substack{K,N_{\mathrm{T}}\to\infty \\ \frac{N_{\mathrm{T}}}{K}=\alpha}}{\longrightarrow} R_{\mathrm{WSR,lin}}^{(\infty)}(\alpha) \tag{5.12}$$

where

$$R_{\mathrm{WSR,lin}}^{(\infty)}(\alpha) = \max_{\rho} \frac{1}{\alpha}\left[\sum_{n=1}^{n(\rho)-1}\beta_n\mu_n\log_2(\mu_n) + \left(\rho - \sum_{n=1}^{n(\rho)-1}\beta_n\right)\mu_{n(\rho)}\log_2(\mu_{n(\rho)})\right.$$
$$\left. + \left(\sum_{n=1}^{n(\rho)-1}\mu_n\beta_n + \left(\rho-\sum_{n=1}^{n(\rho)-1}\beta_n\right)\mu_{n(\rho)}\right)\log_2\left(\frac{\frac{P_{\mathrm{Tx}}}{\sigma_{\mathrm{n}}^2}\sigma_{\mathrm{H}}^2(\alpha-\rho)+\rho}{\sum_{n=1}^{n(\rho)-1}\beta_n\mu_n + \left(\rho-\sum_{n=1}^{n(\rho)-1}\beta_n\right)\mu_{n(\rho)}}\right)\right],$$
$$\text{s.t. } \rho \leq \min(\alpha, 1), \tag{5.13}$$

and $n(\rho)$ is the index of the subgroup, which user $j_{\max} = \rho K$ belongs to. The derivation of (5.12) is based on the fact that the gains of the scalar subchannels $\lambda_{k,j}^{(j_{\max})}$ after $j_{\max}$ allocation steps all converge to the same asymptotic limit

$$\lambda_{k,j}^{(j_{\max})} = \frac{1}{\boldsymbol{e}_{n_{k,j}}^T \left(\boldsymbol{H}_{\text{comp}}^{(j_{\max})} \boldsymbol{H}_{\text{comp}}^{(j_{\max}),\text{H}} \right)^{-1} \boldsymbol{e}_{n_{k,j}}} \underset{\substack{j_{\max}, N_\text{T}, K \to \infty \\ \frac{j_{\max}}{K} = \rho, \frac{N_\text{T}}{K} = \alpha}}{\longrightarrow} \sigma_\text{H}^2 j_{\max} \left(\frac{\alpha}{\rho} - 1 \right) \tag{5.14}$$

where the computation of the finite $\lambda_{k,j}^{(j_{\max})}$ is given by (3.50), $n_{k,j}$ is defined in (3.42), and ρ denotes the fraction of users that are served. The large system limit in (5.14) is also used in [120] and can obtained by the asymptotic limit

$$\begin{aligned} \boldsymbol{e}_{n_{k,j}}^T \left(\boldsymbol{H}_{\text{comp}}^{(j_{\max})} \boldsymbol{H}_{\text{comp}}^{(j_{\max}),\text{H}} \right)^{-1} \boldsymbol{e}_{n_{k,j}} &= \operatorname{tr} \left[\left(\boldsymbol{H}_{\text{comp}}^{(j_{\max})} \boldsymbol{H}_{\text{comp}}^{(j_{\max}),\text{H}} \right)^{-1} \boldsymbol{e}_{n_{k,j}} \boldsymbol{e}_{n_{k,j}}^{\mathsf{T}} \right] \\ &\underset{j_{\max}, N_\text{T} \to \infty}{\longrightarrow} \frac{1}{j_{\max}} \operatorname{tr} \left[\left(\boldsymbol{H}_{\text{comp}}^{(j_{\max})} \boldsymbol{H}_{\text{comp}}^{(j_{\max}),\text{H}} \right)^{-1} \right] \operatorname{tr} \left(\boldsymbol{e}_{n_{k,j}} \boldsymbol{e}_{n_{k,j}}^{\mathsf{T}} \right) = \frac{1}{\sigma_\text{H}^2 j_{\max}} \frac{1}{\frac{N_\text{T}}{j_{\max}} - 1} \end{aligned} \tag{5.15}$$

As the matrices $\left(\boldsymbol{H}_{\text{comp}}^{(j_{\max})} \boldsymbol{H}_{\text{comp}}^{(j_{\max}),\text{H}} \right)^{-1}$ and $\boldsymbol{e}_{n_{k,j}} \boldsymbol{e}_{n_{k,j}}^{\mathsf{T}}$ are asymptotically free (see [109, Example 2.45]), the trace of the product of these matrices normalized by the number of rows can be written as a product of the normalized traces of the asymptotically free matrices in the large system limit (e.g. [109, Equation (2.185)]), where $\boldsymbol{H}_{\text{comp}}^{(j_{\max})}$ and $\boldsymbol{e}_{n_{k,j}} \boldsymbol{e}_{n_{k,j}}^{\mathsf{T}}$ have $j_{\max}$ rows. Considering that the matrix $\boldsymbol{H}_{\text{comp}}^{(j_{\max})}$ contains circularly symmetric i.i.d. entries with variance σ_H^2 and N_T columns, and applying [109, Equation (2.104)] leads to the last equality in (5.15). As $j_{\max} = N_\text{T}$ would imply that the smallest eigenvalue of the matrix $\boldsymbol{H}_{\text{comp}}^{(N_\text{T})} \boldsymbol{H}_{\text{comp}}^{(N_\text{T}),\text{H}}$ becomes zero in the large system limit, which is a consequence of the quarter circle law (e.g. [109, Equation (1.21)]), and $\boldsymbol{H}_{\text{comp}}^{(N_\text{T})} \boldsymbol{H}_{\text{comp}}^{(N_\text{T}),\text{H}}$ would therefore not be invertible, $j_{\max}$ is for sure smaller than N_T in the large system limit leading to the constraint $\rho \leq \min(\alpha, 1)$ in (5.13). Computing the weighted sum rate with the asymptotic channel gains (5.14) and the optimum power allocation, which can be done similar to the DPC case shown in Appendix A8, leads to the asymptotic limit in (5.13). Because all channel gains are equal in the large system limit, the water-level is sufficiently high to serve the selected users. In (5.13) the optimum fraction of active users ρ leading to the maximum asymptotic weighted sum rate has to be determined, which can be done as shown in Appendix A9. The assumption of $\boldsymbol{H}_{\text{comp}}^{(j_{\max})}$ containing i.i.d. entries is only valid for an arbitrary, i.e., not optimized selection of active users. That is because an optimum user selection implies that the active users are chosen so that the rows in $\boldsymbol{H}_{\text{comp}}^{(j_{\max})}$ are as orthogonal to each other as possible and therefore not independent. Thus, Equations (5.12) and (5.13) state a lower bound for the achievable weighted sum rate in the large system limit. The bound becomes tight, if it is asymptotically optimum to serve all users in the system or a random user selection is applied instead of a greedy one. The results of (5.15) can also be used to derive asymptotic expressions for Quality of Service constrained utility maximization problems by adjusting the asymptotic user and power allocation accordingly.

For sum rate maximization, Equation (5.13) simplifies to

$$R_{\text{WSR,lin}}^{(\infty)}(\alpha) = \max_{\rho} \frac{1}{\alpha} \rho \log_2 \left(1 + \frac{P_\text{Tx}}{\sigma_\text{n}^2} \sigma_\text{H}^2 \left(\frac{\alpha}{\rho} - 1 \right) \right), \quad \text{s.t. } \rho \leq \min(\alpha, 1),$$

which is a concave optimization problem and can therefore be solved for example by bisection [33, Chapter 8.2].

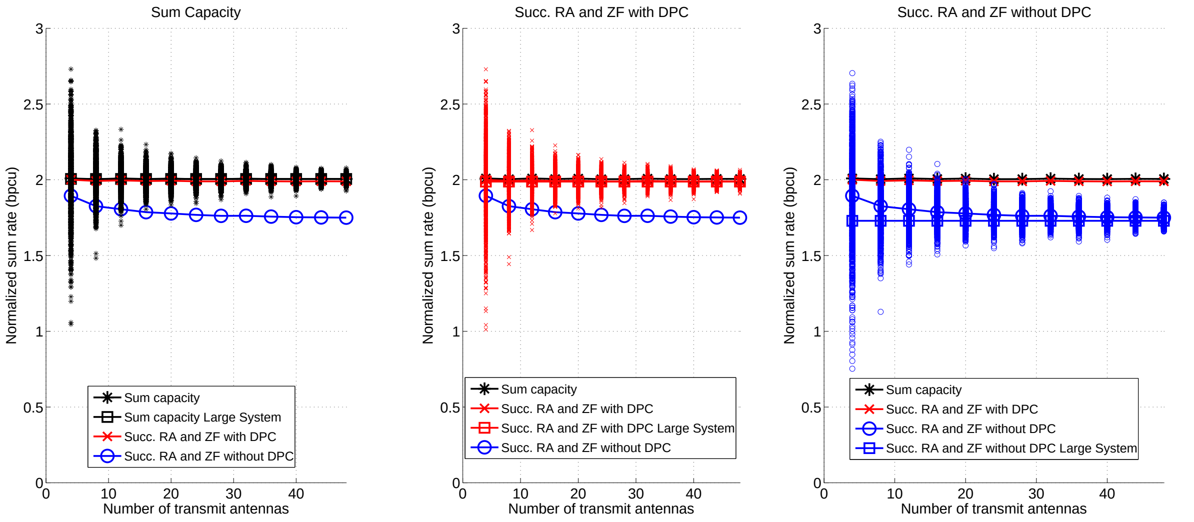

Figure 5.1: Comparison of normalized average sum rates to large system sum rates in MISO systems with $\alpha = 2$, $C = 1$, $\sigma_{\mathrm{H}}^2 = 1$, SNR$= 10$dB.

Figure 5.1 exhibits the sum rates normalized to the number of transmit antennas and averaged over 1000 circularly symmetric Gaussian channel matrices with $\sigma_{\mathrm{H}}^2 = 1$ in a single-carrier system with $C = 1$. Correspondingly, there is no multipath propagation, i.e., $L_k = 1, \forall k = 1, \ldots, K$. The ratio α of transmit antennas to users is set to $\alpha = 2$ and SNR is equal to 10dB, i.e., $\frac{P_{\mathrm{Tx}}}{\sigma_{\mathrm{n}}^2} = 10$. The weights of all users are equal so that there is one subgroup of users, i.e., $n = 1$ and $\mu_1 = 1$. In each subfigure the large system sum rates of the corresponding algorithms are compared to the average sum rates obtained by simulations. Additionally, the sum rates of each channel realization are plotted in the corresponding subfigures by markers. For the sum capacity and the successive resource allocation with zero-forcing and DPC (“Succ. RA and ZF with DPC”), the large system sum rates serve as very good approximations for the average sum rates already with only a few number of transmit antennas and users. When DPC is not applied at the transmitter (“Succ. RA and ZF without DPC”), the large system sum rate is approached slowlier by the average sum rate, so that it serves only for $N_{\mathrm{T}} \geq 20$ as a very good approximation of the average sum rate. For all three algorithms under consideration, the variances of the sum rates shrink with increasing number of transmit antennas, which is why the large system sum rates also become good estimates of the instantaneously achievable sum rates with an arbitrary channel realization for larger number of transmit antennas.

When the number of users exceeds the number of transmit antennas, as in Figure 5.2, where the same parameters as in the previous plot are used except for α, which is now set to $\alpha = \frac{1}{2}$, the large system sum rates for the successive resource allocation schemes only serve as lower bounds for the average sum rates. As already mentioned, that is because the large system analysis does not take into account the effect that with the successive resource allocation in each step a user is taken that channel exhibits the strongest energy in the subspace orthogonal to the previously selected channels, which is why the projection matrices can no longer be assumed to be independent of the channel matrices of the selected users. When the users are selected randomly in each step, however, this assumption is correct. The large system sum rates therefore serve as a very good approximation for such an allocation scheme. The average rates achievable with that scheme are

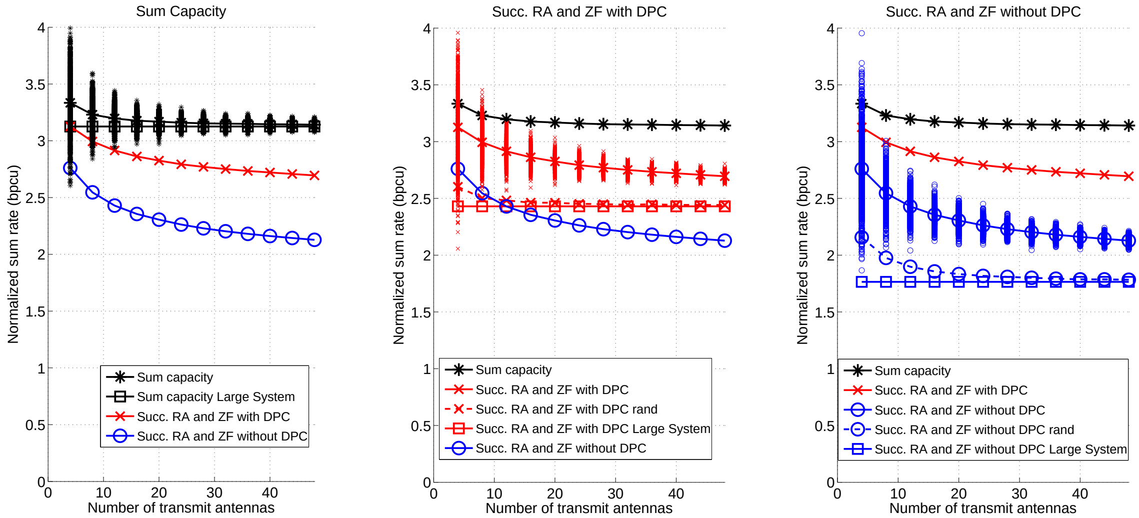

Figure 5.2: Comparison of normalized average sum rates compared to large system sum rates in MISO systems with $\alpha = \frac{1}{2}$, $C = 1$, $\sigma_{\mathrm{H}}^2 = 1$, SNR= 10dB.

plotted as dashed lines and denoted by the adjunct “rand” in Figure 5.2. For the sum capacity, the analytical result is still valid, which is why the asymptotic rate still approximates the average sum capacity very well also in such a parameter setup, although convergence to the large system limit is slower than with $\alpha = 2$.

5.4 Large System Analysis of Successive Resource Allocation and Spatial Zero-Forcing with Dirty Paper Coding in MIMO Systems

By going from MISO to MIMO systems the large system analysis gets more involved. For this reason the problem of sum rate maximization, i.e., equal priorities of the users, is considered and the analysis is carried out with the successive resource allocation and spatial zero-forcing with DPC explained in Section 3.2 only. Thus, the Successive Encoding Successive Allocation Method (SESAM) from [10] will be analyzed in the following. An approximation of the sum rate achievable with SESAM and a random user allocation has been presented by the author and others in [129] together with a large system analysis of competing algorithms as Block Diagonalization with DPC [130] and without DPC [61]. The parameters growing towards infinity are the number of transmit and receive antennas, where the ratio β of both is fixed and finite. Furthermore, for simplicity it is assumed that all users have the same number r of receive antennas, i.e., $r_k = r, \forall k = 1, \ldots, K$, so that

$$N_{\mathrm{T}} \to \infty, \quad r \to \infty, \quad \beta = \frac{N_{\mathrm{T}}}{r}.$$

The number of users K, the number of carriers C and the transmit power P_{Tx} remain finite. To obtain a large system expression for the sum rate achievable with SESAM, in the following a numerical method will be presented to compute the asymptotic distribution $f_{\mathrm{SESAM}}^{(\infty)}(x)$ of the sub-channel gains $\frac{\lambda_{k,j}}{r\sigma_{\mathrm{H}}^2}$, which are normalized by the variance of the random channel matrices and the number of receive antennas for reasons that will become clear soon. Similarly to the definition of the asymptotic eigenvalue distribution, the integral

$$\int_0^\lambda f_{\text{SESAM}}^{(\infty)}(x)\,\mathrm{d}\,x$$

states the percentage of normalized subchannel gains that are smaller or equal to λ. Analogously to (5.3), sums over functions of the subchannel gains can be replaced by integrals over $f_{\text{SESAM}}^{(\infty)}(x)$. This implies that the sum rate achievable with SESAM in the large system limit can be computed according to

$$\frac{R_{\text{sum,SESAM}}}{Cm} = \frac{1}{m}\sum_{i=1}^{m}\log_2\left(\max[1,\eta\lambda_i]\right) \underset{m\to\infty}{\longrightarrow} \int_{\lambda_{\min}}^{\infty}\log_2\left(\eta x\right) f_{\text{SESAM}}^{(\infty)}(x)\,\mathrm{d}\,x, \tag{5.16}$$

where $m = \min(N_{\text{T}}, Kr)$ denotes the maximum number of subchannels that can be allocated on one carrier and

$$\lambda_i = \frac{1}{r\sigma_{\text{H}}^2}\rho_1\left(\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_{k(i)}^{\text{H}}\boldsymbol{H}_{k(i)}\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\right)$$

is the normalized gain of the ith subchannel [c.f. (3.35)]. Note that through this normalization, $\frac{1}{r\sigma_{\text{H}}^2}\boldsymbol{H}_{k(i)}^{\text{H}}\boldsymbol{H}_{k(i)}$ are the Gramian matrcies of $r \times N_{\text{T}}$ matrices with Gaussian i.i.d. entries with variance $\frac{1}{r}$. The asymptotic eigenvalue distribution of these matrices is therefore given by the Marčenko-Pastur distribution $f_{\text{MP}}(x,\beta)$ [c.f. (5.2)] with parameter $\hat{\beta} = \beta$. As the matrices considered have the same statistical properties on all carriers and therefore the same asymptotic sum rate can be achieved on all carriers like in the MISO case, the derivations in this section are presented for single-carrier systems, i.e., the matrices $\boldsymbol{H}_k$ have N_{T} rows. The extension to multicarrier systems can then be done by multiplying the asymptotic sum rate with the number of carriers C, dividing the noise variance by C and allocation the power $\frac{P_{\text{Tx}}}{C}$ on each subcarrier, which has been considered in (5.16) and (5.17) to obtain general results. The variable η computes as

$$\eta = r\sigma_{\text{H}}^2\frac{\frac{P_{\text{Tx}}}{\sigma_{\text{N}}^2}+\sum\limits_{i=1}^{i_{\max}}\frac{1}{\lambda_i r\sigma_{\text{H}}^2}}{i_{\max}} \underset{m\to\infty}{\longrightarrow} r\sigma_{\text{H}}^2\frac{\frac{P_{\text{Tx}}}{\sigma_{\text{N}}^2}+\frac{m}{r\sigma_{\text{H}}^2}\int\limits_{\lambda_{\min}}^{\infty}\frac{1}{x}f_{\text{SESAM}}^{(\infty)}(x)\,\mathrm{d}\,x}{m\int\limits_{\lambda_{\min}}^{\infty}f_{\text{SESAM}}^{(\infty)}(x)\,\mathrm{d}\,x} = \frac{\frac{P_{\text{Tx}}}{\sigma_{\text{N}}^2}\frac{\sigma_{\text{H}}^2}{\min(\beta,K)}+\int\limits_{\lambda_{\min}}^{\infty}\frac{1}{x}f_{\text{SESAM}}^{(\infty)}(x)\,\mathrm{d}\,x}{\int\limits_{\lambda_{\min}}^{\infty}f_{\text{SESAM}}^{(\infty)}(x)\,\mathrm{d}\,x}, \tag{5.17}$$

which corresponds to the water-level multiplied by $r\sigma_{\text{H}}^2$. $\lambda_{\min} = \lambda_{i_{\max}}$ is the minimum subchannel gain that receives non-zero power, i.e., it is given by the implicit equation

$$\frac{P_{\text{Tx}}\sigma_{\text{n}}^2\frac{\sigma_{\text{H}}^2}{\min(\beta,K)}+\int\limits_{\lambda_{\min}}^{\infty}\frac{1}{x}f_{\text{SESAM}}^{(\infty)}(x)\,\mathrm{d}\,x}{\int\limits_{\lambda_{\min}}^{\infty}f_{\text{SESAM}}^{(\infty)}(x)\,\mathrm{d}\,x} = \frac{1}{\lambda_{\min}}.$$

In order to obtain $f_{\text{SESAM}}^{(\infty)}(x)$, the asymptotic distribution of the maximum eigenvalues of the matrices $\frac{1}{r\sigma_{\text{H}}^2}\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\boldsymbol{H}_{k(i)}^{\text{H}}\boldsymbol{H}_{k(i)}\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}$, $i = 1,\ldots,m$ has to be found. For $i = 1$ there is $\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)} = \boldsymbol{I}_{N_{\text{T}}}$ and as the eigenvalues of $\frac{1}{r\sigma_{\text{H}}^2}\boldsymbol{H}_k^{\text{H}}\boldsymbol{H}_k$ asymptotically follow the Marčenko-Pastur distribution (5.2) with parameter $\hat{\beta} = \beta$, the normalized gain of the first subchannel is given by $\lambda_1 = \left(1+\sqrt{\beta}\right)^2$.

Because the channel matrices of all users are assumed to have the same statistical properties, all maximum eigenvalues converge to the same asymptotic limit and the algorithm will select an arbitrary user for the first subchannel. In the next step the largest eigenvalue of the matrix $\frac{1}{r\sigma_{\mathrm{H}}^2}\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(2)}\boldsymbol{H}_{k(1)}^{\mathrm{H}}\boldsymbol{H}_{k(1)}\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(2)}$ is given by the second largest eigenvalue of the Marčenko-Pastur distribution with parameter $\hat{\beta} = \beta$, as the projector $\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(2)}$ projects into the null-space of the eigenvector corresponding to the principal eigenvector of the matrix $\boldsymbol{H}_{k(1)}^{\mathrm{H}}\boldsymbol{H}_{k(1)}$. When the dimensions of the matrix $\boldsymbol{H}_{k(1)}^{\mathrm{H}}\boldsymbol{H}_{k(1)}$ grow infinitely large, however, the difference of this value to $\lambda_1 = \left(1+\sqrt{\beta}\right)^2$ can be hardly measured with finite precision. For the other users $k \neq k(1)$, the maximum eigenvalues of the matrices $\frac{1}{r\sigma_{\mathrm{H}}^2}\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(2)}\boldsymbol{H}_{k}^{\mathrm{H}}\boldsymbol{H}_{k}\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(2)}$ are all given by $\left(1+\sqrt{\beta\frac{N_{\mathrm{T}}-1}{N_{\mathrm{T}}}}\right)^2$. That is because

$$\rho_1\left(\frac{1}{r\sigma_{\mathrm{H}}^2}\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(2)}\boldsymbol{H}_{k}^{\mathrm{H}}\boldsymbol{H}_{k}\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(2)}\right) = \rho_1\left(\frac{1}{r\sigma_{\mathrm{H}}^2}\boldsymbol{V}_{\mathrm{DPC}}^{(2)}\boldsymbol{V}_{\mathrm{DPC}}^{(2),\mathrm{H}}\boldsymbol{H}_{k}\boldsymbol{H}_{k}^{\mathrm{H}}\boldsymbol{V}_{\mathrm{DPC}}^{(2)}\boldsymbol{V}_{\mathrm{DPC}}^{(2),\mathrm{H}}\right) = \\ = \rho_1\left(\frac{1}{r\sigma_{\mathrm{H}}^2}\boldsymbol{V}_{\mathrm{DPC}}^{(2),\mathrm{H}}\boldsymbol{H}_{k}^{\mathrm{H}}\boldsymbol{H}_{k}\boldsymbol{V}_{\mathrm{DPC}}^{(2)}\right),$$

where $\boldsymbol{V}_{\mathrm{DPC}}^{(2)} \in \mathbb{C}^{N_{\mathrm{T}}\times N_{\mathrm{T}}-1}$ is an orthonormal basis of $\mathrm{span}\left\{\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(2)}\right\}$ independent of $\boldsymbol{H}_k$. The matrices $\frac{1}{\sqrt{r}\sigma_{\mathrm{H}}}\boldsymbol{H}_k\boldsymbol{V}_{\mathrm{DPC}}^{(2)} \in \mathbb{C}^{r\times N_{\mathrm{T}}-1}$ therefore contain Gaussian i.i.d. entries with variance $\frac{1}{r}$, which is why the asymptotic eigenvalue distribution can be obtained via the Marčenko-Pastur distribution with parameter $\hat{\beta} = \beta\frac{N_{\mathrm{T}}-1}{N_{\mathrm{T}}}$. Similarly to the case of user $k(1)$, this effect is hardly measurable on a computer with finite precision, as the factor $\frac{N_{\mathrm{T}}-1}{N_{\mathrm{T}}}$ is almost equal to one for infinite N_{T}, i.e., the strongest eigenvalue of the matrix $\frac{1}{r\sigma_{\mathrm{H}}^2}\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(2)}\boldsymbol{H}_{k(1)}^{\mathrm{H}}\boldsymbol{H}_{k(1)}\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(2)}$ and the matrices $\frac{1}{r\sigma_{\mathrm{H}}^2}\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(2)}\boldsymbol{H}_{k}^{\mathrm{H}}\boldsymbol{H}_{k}\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(2)}$, $k \neq k(1)$ are hardly distinguishable. For this reason in the following a user allocation will be considered, where an infinite amount of subchannels is consecutively allocated to the same user. That implies that in each step δm subchannels are all allocated to the same user, where $0 < \delta \leq 1$ denotes a predefined constant depending on the desired numerical accuracy. Although the algorithm as proposed in [10] probably leads to a different user allocation in the large system limit and this allocation scheme might not be optimum in the successive scheme, it leads to a lower bound for the sum rate. This bound becomes tight for small δ, as the error introduced this way is hardly measurable. With this allocation scheme, the asymptotic distributions of the gains of the subchannels consecutively allocated to the same user can be computed. The asymptotic distribution of all subchannel gains is then given by the normalized sum of these distributions. Let $f_j^{(\infty)}(x)$ denote the asymptotic distribution of channel gains in the jth group of subchannels allocated to the same user. Then $f_{\mathrm{SESAM}}^{(\infty)}(x)$ is given by

$$f_{\mathrm{SESAM}}^{(\infty)}(x) = \sum_{j=1}^{\frac{1}{\delta}} f_j^{(\infty)}(x).$$

For simplicity it is assumed in the following that δ is chosen so that $\frac{1}{\delta} \in \mathbb{N}$, although the following results can be easily extended to the general case. The $f_j^{(\infty)}(x)$ are the appropriately normalized tails of the asymptotic eigenvalue distributions containing the δm strongest eigenvalues of the matrices $\frac{1}{r\sigma_{\mathrm{H}}^2}\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{((j-1)\delta m+1)}\boldsymbol{H}_{\hat{k}(j)}^{\mathrm{H}}\boldsymbol{H}_{\hat{k}(j)}\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{((j-1)\delta m+1)}$, where $\hat{k}(j) = k((j-1)\delta m+1)$ is the user to which the jth subgroup of subchannels is allocated to. Those eigenvalues correspond to the subchannel gains allocated at once to the same user from step $(j-1)\delta m+1$ to step $j\delta m$. These eigenvalues are also the strongest eigenvalues of the matrices

$$\boldsymbol{C}_j := \frac{1}{r\sigma_{\mathrm{H}}^2}\boldsymbol{V}_{\mathrm{DPC}}^{((j-1)\delta m+1),\mathrm{H}}\boldsymbol{H}_{\hat{k}(j)}^{\mathrm{H}}\boldsymbol{H}_{\hat{k}(j)}\boldsymbol{V}_{\mathrm{DPC}}^{((j-1)\delta m+1)} \in \mathbb{C}^{N_{\mathrm{T}}-(j-1)\delta m\times N_{\mathrm{T}}-(j-1)\delta m}, \tag{5.18}$$

where $\boldsymbol{V}_{\mathrm{DPC}}^{((j-1)\delta m+1)} \in \mathbb{C}^{N_{\mathrm{T}}\times N_{\mathrm{T}}-(j-1)\delta m}$ denotes an orthonormal basis of $\operatorname{span}\left\{\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{((j-1)\delta m+1)}\right\}$ so that

$$\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{((j-1)\delta m+1)} = \boldsymbol{V}_{\mathrm{DPC}}^{((j-1)\delta m+1)}\boldsymbol{V}_{\mathrm{DPC}}^{((j-1)\delta m+1),\mathrm{H}}.$$

With these definitions the $f_j^{(\infty)}(x)$ are given by

$$f_j(x) = \begin{cases} \frac{1-(j-1)\delta\xi}{\xi} f_{\boldsymbol{C}_j}^{(\infty)}(x) & x \geq \hat{\lambda}_j \\ 0, & \text{else} \end{cases}$$

and

$$\int_{\hat{\lambda}_j}^{\infty} f_{\boldsymbol{C}_j}^{(\infty)}(x) = \frac{\delta m}{N_{\mathrm{T}}-(j-1)\delta m} = \frac{\delta\xi}{1-(j-1)\delta\xi}, \tag{5.19}$$

where

$$\xi = \frac{m}{N_{\mathrm{T}}} = \min\left(1, \frac{K}{\beta}\right).$$

The factor $\frac{1-(j-1)\delta\xi}{\xi}$ normalizes the tails of $f_{\boldsymbol{C}_j}^{(\infty)}(x)$ so that each subgroup contributes δ to the integral $\int_0^\infty f_{\mathrm{SESAM}}^{(\infty)}(x)\,\mathrm{d}\,x$ and the whole integral is equal to one, i.e.,

$$\int_0^\infty f_{\mathrm{SESAM}}^{(\infty)}(x)\,\mathrm{d}\,x = \sum_{j=1}^{\frac{1}{\delta}}\int_0^\infty f_j^{(\infty)}(x)\,\mathrm{d}\,x = \sum_{j=1}^{\frac{1}{\delta}}\delta = 1.$$

For $j=1$, $\boldsymbol{C}_1 = \frac{1}{r\sigma_{\mathrm{H}}^2}\boldsymbol{H}_{k(1)}^{\mathrm{H}}\boldsymbol{H}_{k(1)}$ and $f_{\boldsymbol{C}_1}^{(\infty)}(x)$ is therefore given by the tail of the Marčenko-Pastur distribution with $\hat{\beta}=\beta$, i.e.,

$$f_{\boldsymbol{C}_1}^{(\infty)}(x) = f_{\mathrm{MP}}(x,\beta).$$

For $j=2$, the next subgroup of users can either be allocated to the same user $k(1)$. In this case $f_{\boldsymbol{C}_2}^{(\infty)}(x)$ would be given by a truncated Marčenko-Pastur distribution, so that $f_{\boldsymbol{C}_2}(x) = \frac{1}{1-\delta\xi}f_{\mathrm{MP}}(x,\beta)$ for $x \leq \hat{\lambda}_1$ and $f_{\boldsymbol{C}_2}^{(\infty)}(x) = 0$ otherwise. For the other users, $f_{\boldsymbol{C}_2}^{(\infty)}(x)$ would be given by the appropriate tails of the asymptotic eigenvalue distributions of the matrices $\frac{1}{r\sigma_{\mathrm{H}}^2}\boldsymbol{V}_{\mathrm{DPC}}^{(\delta m+1)}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{H}_k\boldsymbol{V}_{\mathrm{DPC}}^{(\delta m+1)}$. Assuming that the orthonormal basis $\boldsymbol{V}_{\mathrm{DPC}}^{(\delta m+1)}$ of $\operatorname{span}\left\{\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(\delta m+1)}\right\}$ is independent of $\boldsymbol{H}_k$ for $k \neq k(1)$, the the a.e.d. of the matrices $\frac{1}{r\sigma_{\mathrm{H}}^2}\boldsymbol{V}_{\mathrm{DPC}}^{(\delta m+1)}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{H}_k\boldsymbol{V}_{\mathrm{DPC}}^{(\delta m+1)}$ is given by the Marčenko-Pastur distribution with $\hat{\beta} = \beta\frac{N_{\mathrm{T}}-\delta m}{N_{\mathrm{T}}} = \beta(1-\delta\xi)$. Thus, in case the second group of subchannels is allocated to another user than user $k(1)$, $f_{\boldsymbol{C}_2}^{(\infty)}(x)$ is given by $f_{\boldsymbol{C}_2}^{(\infty)}(x) = f_{\mathrm{MP}}(x,\beta(1-\delta\xi))$. That is because multiplying a matrix with Gaussian i.i.d. entries like $\boldsymbol{H}_k$ with an independent orthonormal matrix like $\boldsymbol{V}_{\mathrm{DPC}}^{(\delta m+1)}$ leads again to a matrix with Gaussian i.i.d. entries having the same mean and variance as in the original Gaussian matrix but different dimensions. This way, $\frac{1}{\sqrt{r}\sigma_{\mathrm{H}}}\boldsymbol{H}_k\boldsymbol{V}_{\mathrm{DPC}}^{(\delta m+1)}$ is a $r\times N_{\mathrm{T}}-\delta m$ matrix with Gaussian i.i.d. entries with zero mean and variance $\frac{1}{r}$. Likewise to the MISO case in the previous section, the assumption of $\boldsymbol{H}_k$ and $\boldsymbol{V}_{\text{DPC}}^{(\delta m+1)}$ being independent leads to a lower bound for the weighted sum rate, as the user selected for the second group of subchannels would be optimally chosen so that projecting $\boldsymbol{H}_k$ with $\hat{\boldsymbol{P}}_{\text{DPC}}^{(\delta m+1)}$ almost preserves the strongest eigenvalues of $\boldsymbol{H}_k$ and mostly affects the small eigenvalues. For the large system analysis the user for the second group of subchannels is randomly chosen from the set of users that have not been served so far. Although, this choice might not lead to the maximum increase in the asymptotic sum rate, it definitely leads to a higher sum rate than allocating the second group of subchannels to the same user than the first group. Proceeding this way, it can be shown by evaluating the large system sum rates that the first K subgroups of subchannels are all allocated to distinct users so that

$$f_{\boldsymbol{C}_j}^{(\infty)}(x) = f_{\text{MP}}(x, \beta(1-(j-1)\delta\xi)), \quad j = 1, \ldots, K.$$

In the following it will be assumed that the first subgroup of subchannels is allocated to user 1, the second to user 2 and so on. As shown in Appendix A10, the user scheduling in the large system limit continues this way also for $j > K$, so that user k receives a group of subchannels for $j = k$, $j = k + K$ and so on. For $j > K$, $f_{\boldsymbol{C}_j}^{(\infty)}(x)$ can however not be stated explicitly anymore. Instead the Stieltjes-transform $m_{\boldsymbol{C}_j}(z)$ of the matrix $\boldsymbol{C}_j$ is given by the implicit equation

$$\int_0^{\hat{\lambda}_{j-K}} \frac{f_{\boldsymbol{C}_{j-K}}(x)}{1-\tilde{\beta}_j + (x-z)m_{\boldsymbol{C}_j}(z)} \,\mathrm{d}\,x = \frac{1-(j-K)\delta\xi}{1-(j-K-1)\delta\xi}, \tag{5.20}$$

where

$$\tilde{\beta}_j = \frac{1-(j-1)\delta\xi}{1-(j-K)\delta\xi},$$

and which is derived in Appendix A10. Unfortunately, there is no explicit solution neither for $m_{\boldsymbol{C}_j}(z)$ nor for $f_{\boldsymbol{C}_j}^{(\infty)}(x)$ from (5.20). For this reason $f_{\boldsymbol{C}_j}^{(\infty)}(x)$ has to be sampled as described in the following. First Equation (5.20) is solved for $m_{\boldsymbol{C}_j}(z)$ with $z = \hat{\lambda}_{j-K}$. The imaginary part of $m_{\boldsymbol{C}_j}(\hat{\lambda}_{j-K})$ divided by π is then equal to $f_{\boldsymbol{C}_j}^{(\infty)}(\hat{\lambda}_{j-K})$, as given by (5.4). Due to the projections from step $(j-K)\delta m+1$ to step $(j-1)\delta m$, the principal eigenvalue of the matrix $\boldsymbol{C}_j$ will certainly not be larger than $\hat{\lambda}_{j-K}$, which is the channel gain in step $(j-K)\delta m$ [c.f. (5.19)], the last step the same user has received a subchannel. Thus, $f_{\boldsymbol{C}_j}^{(\infty)}(x) = 0$ for $x > \hat{\lambda}_{j-K}$ and $\hat{\lambda}_{j-K}$ can be used as a starting point for the sampling process. After, $f_{\boldsymbol{C}_j}^{(\infty)}(\hat{\lambda}_{j-K})$ has been computed z is reduced by a constant sampling distance $\varDelta$ and Equation (5.20) is solved for $m_{\boldsymbol{C}_j}(z)$ with $z = \hat{\lambda}_{j-K} - \varDelta$. This sampling is continued until $z = 0$. The integrals with $f_{\boldsymbol{C}_j}^{(\infty)}(x)$ required in (5.19) and (5.20) can then be evaluated numerically for example with the trapezoidal method (e.g. [131]), where $f_{\boldsymbol{C}_j}^{(\infty)}(x)$ is interpolated linearly between two neighboring samples.

In Figure 5.3 the ergodic sum rates normalized to the number of transmit antennas and averaged over 1000 Gaussian channel realizations with $\sigma_{\text{H}}^2 = 1$ are plotted versus the number of transmit antennas in a single carrier system with $C = 1$ at an SNR of 10 dB. The ratio β of transmit antennas to receive antennas is set to $\beta = 2$. The average sum rate of SESAM is plotted as a line marked as “SESAM”, whereas additionally each of the 1000 sum rates that has been achieved with one channel realization is marked with a star. The large system sum rate obtained as described above with $\delta = 0.05$ is plotted as a line with circles, where a sampling distance of $\varDelta = 0.001$ has been

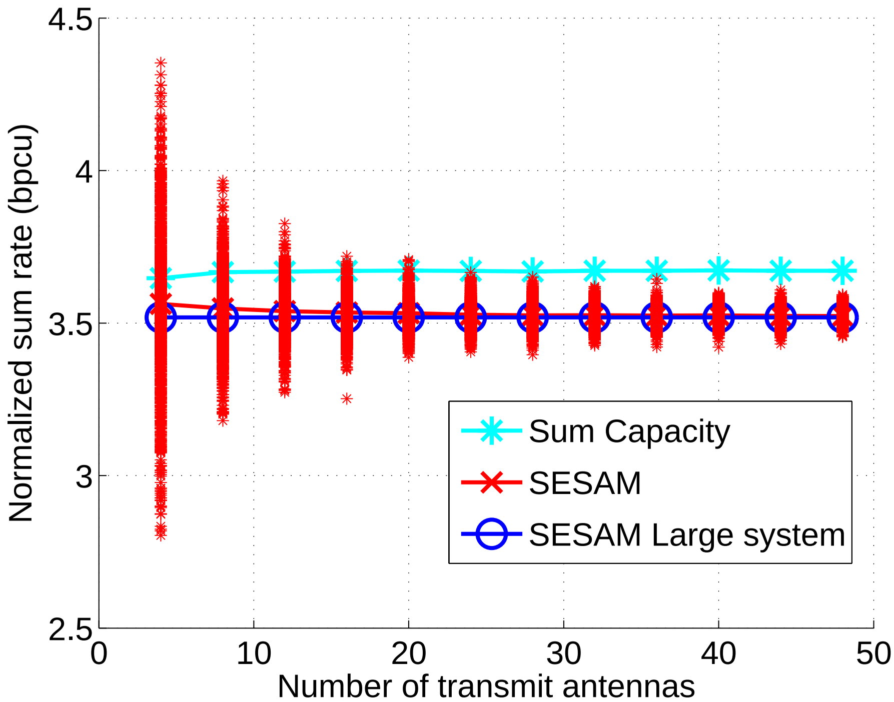

Figure 5.3: Comparison of ergodic sum rates with large system sum rate in a system with $K = 5$ users, $C = 1$ carrier, $\beta = 2$, SNR= 10dB and $\delta = 0.05$.

used to determine the asymptotic distributions $f_{\boldsymbol{C}_j}^{(\infty)}(x)$. The large system sum rate obtained this way serves as a very good approximation for the ergodic sum rate achievable with SESAM also for finite system parameters and becomes exact for systems with $N_\mathrm{T} \geq 16$. Furthermore the variance of the sum rate achievable with SESAM decreases with increasing number of transmit antennas. For comparison the average sum capacity has also been included in Figure 5.3.

6. Conclusion

Efficient low-complexity algorithms have been presented in this book for the MIMO broadcast channel. The problem of weighted sum rate maximization under a sum transmit power constraint has been tackled as well as several Quality of Service constrained utility maximization problems. Weighted sum rate and transmission power served as utility, whereas the QoS constraints have been given by minimum rates or relative rate requirements. For all optimization problems it has been shown by simulation results that the proposed algorithms are able to achieve the optimum solutions closely in multi-path Rayleigh fading scenarios, when DPC can be applied at the transmitter. Avoiding DPC by mitigating multi-user interference solely by means of linear transmit and receive signal processing leads to further small performance losses but drastic reductions in computational complexity. The efficient algorithms are based on the principles of successive resource allocation and spatial zero-forcing and work non-iteratively. Some of the presented algorithms have been analyzed in the large system limit, where at least two system parameters go to infinity at a finite fixed ratio. The analytical results obtained this way have been shown to be good approximations for the ergodic performance of the system with finite parameters having the same ratio as in the large system analysis.

The algorithms derived in this book require perfect channel knowledge of all users' channel matrices at the transmitter. This assumption can no longer be maintained in scenarios with short channel coherence times, i.e., when the channel matrices remain constant only for a short period of time. Thus, in those scenarios it is necessary to take the erroneous channel knowledge into account when designing transmit and receive signal processing filters. A possible extension to the work presented in this book would therefore be to make the presented algorithms robust against erroneous or outdated channel state information. Another aspect of future work is the consideration of inter-cell interference. In this book, isolated cells have been considered, but for an efficient use of bandwidth it might be necessary to operate neighboring cells in the same frequency band. The inter-cell interference caused this way must be taken into account during filter design. In [132] and [133], the method of successive resource allocation and zero-forcing without DPC has already been adjusted to a two-cell interference scenario.

Appendix

A1. Derivation of the OFDM Channel Model

In this appendix the relation between the OFDM channel matrices $\boldsymbol{H}_k$ and the channels $\tilde{\boldsymbol{H}}_k(t)$ in the time domain, as stated in (2.1), and the derivation of the additive Gaussian noise $\boldsymbol{\eta}_k[n]$ will be given. It is based on the model in [36, Ch. 1.3.]. At first, the modulation and cyclic prefix will be explained, a block diagram detailing the corresponding block in Figure 2.1 is given in Figure A1. The vector $\boldsymbol{x}(t)$ is convolved with rectangular pulses $g(t,T)$ of duration T and height 1, where

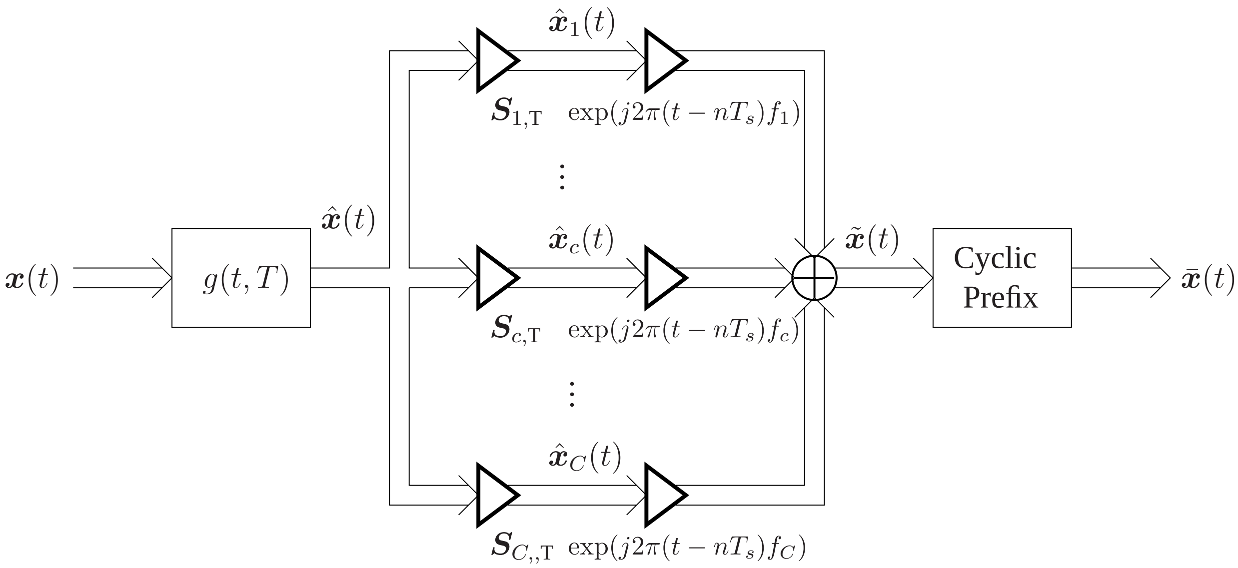

Figure A1: Modulation and Introduction of Cyclic Prefix (Mod. + CP)

$$g(t,T)=\begin{cases}1, & 0\le t\le T,\\ 0, & \text{else.}\end{cases}$$

Thus,

$$\hat{\boldsymbol{x}}(t)=\sum_{n=-\infty}^{\infty} g(t-nT_s,T)\boldsymbol{x}[n].$$

$\hat{\boldsymbol{x}}(t)\in\mathbb{C}^{CN_\mathrm{T}}$ is then split into C vectors of length N_T, where each of those vectors $\hat{\boldsymbol{x}}_c(t)\in\mathbb{C}^{N_\mathrm{T}}$ is transmitted on a different carrier and obtained by selecting N_T rows of $\hat{\boldsymbol{x}}(t)$ according to

$$\hat{\boldsymbol{x}}_c(t)=\left[\boldsymbol{0}_{N_\mathrm{T},(c-1)N_\mathrm{T}},\boldsymbol{I}_{N_\mathrm{T}},\boldsymbol{0}_{N_\mathrm{T},(C-c)N_\mathrm{T}}\right]\hat{\boldsymbol{x}}(t)=\boldsymbol{S}_{c,\mathrm{T}}\hat{\boldsymbol{x}}(t).$$

Each $\hat{\boldsymbol{x}}_c(t)$ is modulated by a different basis function $\exp(j2\pi(t-nT_s)f_c)$, where each basis function defines a carrier of the OFDM system and

$$f_c=f_0-B/2+(c-1)B/C \tag{A1}$$

is the frequency over which the signal $\hat{\boldsymbol{x}}_c(t)$ is transmitted.

$$B = \frac{C}{T} \tag{A2}$$

and f_0 denote the bandwidth and the center frequency, respectively. The resulting signals are added up, so that

$$\tilde{\boldsymbol{x}}(t) = \sum_{c=1}^{C} \hat{\boldsymbol{x}}_c(t)\exp(j2\pi(t-nT_s)f_c) = \sum_{n=-\infty}^{\infty}\sum_{c=1}^{C} \exp(j2\pi(t-nT_s)f_c)g(t-nT_s,T)\boldsymbol{S}_{c,\mathrm{T}}\boldsymbol{x}[n].$$

Finally, a cyclic prefix is added to each of the signals

$$\tilde{\boldsymbol{x}}_n(t) = \sum_{c=1}^{C} \exp(j2\pi(t-nT_s)f_c)g(t-nT_s,T)\boldsymbol{S}_{c,\mathrm{T}}\boldsymbol{x}[n] \in \mathbb{C}^{N_\mathrm{T}}.$$

That implies that the original signals $\tilde{\boldsymbol{x}}_n(t)$ are delayed by a certain time T_{cp}. As it will become clear in the remainder, T_{cp} should be chosen to be larger or equal to $\max_{k\in\mathcal{S}_k}(\tau_{k,L_k}-\tau_{k,1})$, i.e.,

$$T_{\mathrm{cp}} \geq \max_{k\in\mathcal{S}_k}(\tau_{k,L_k}-\tau_{k,1}), \tag{A3}$$

where $\mathcal{S}_k = \{k \in \{1,\ldots,K\} |\, \mathrm{tr}(\boldsymbol{P}_k) > 0\}$ contains all users that receive non-zero power. The interval between nT_s and $nT_s + T_{\mathrm{cp}}$ is filled by the last T_{cp} seconds of the original signal $\tilde{\boldsymbol{x}}_n(t)$ such that the transmit signal $\bar{\boldsymbol{x}}_n(t) \in \mathbb{C}^{N_\mathrm{T}}$ results in

$$\bar{\boldsymbol{x}}_n(t) = \begin{cases} \tilde{\boldsymbol{x}}_n(t+T-T_{\mathrm{cp}}), & nT_s \leq t \leq nT_s + T_{\mathrm{cp}} \\ \tilde{\boldsymbol{x}}_n(t-T_{\mathrm{cp}}), & nT_s + T_{\mathrm{cp}} \leq t \leq nT_s + T_{\mathrm{cp}} + T. \end{cases} \tag{A4}$$

In case all f_c are integer multiples of $1/T$, which can always be assured by an appropriate choice of f_0, the functions $\exp(j2\pi(t-nT_s)f_c)$ are periodic with a period of length T. Hence by using the enlarged windowing function $g(t, T+T_{\mathrm{cp}})$, $\bar{\boldsymbol{x}}_n(t)$ can be rewritten as

$$\bar{\boldsymbol{x}}_n(t) = \sum_{c=1}^{C} \exp(j2\pi(t-nT_s-T_{\mathrm{cp}})f_c)g(t-nT_s,T+T_{\mathrm{cp}})\boldsymbol{S}_{c,\mathrm{T}}\boldsymbol{x}[n]. \tag{A5}$$

By choosing $T_s = T + T_{\mathrm{cp}}$ the nonzero parts of the signals $\bar{\boldsymbol{x}}_m(t)$ and $\bar{\boldsymbol{x}}_p(t)$ do not overlap in time for $m \neq p$. Finally, the transmit signal $\bar{\boldsymbol{x}}(t)$ reads as $\bar{\boldsymbol{x}}(t) = \sum_{n=-\infty}^{\infty} \bar{\boldsymbol{x}}_n(t)$.
At the kth user, the signal

$$\bar{\boldsymbol{y}}_k(t) = \sum_{\ell=1}^{L_k} \tilde{\boldsymbol{H}}_{k,\ell}\bar{\boldsymbol{x}}(t-\tau_{k,\ell}) + \tilde{\boldsymbol{\eta}}_k(t) = \sum_{n=-\infty}^{\infty}\sum_{\ell=1}^{L_k} \tilde{\boldsymbol{H}}_{k,\ell}\bar{\boldsymbol{x}}_n(t-\tau_{k,\ell}) + \tilde{\boldsymbol{\eta}}_k(t) = \sum_{n=-\infty}^{\infty} \bar{\boldsymbol{y}}_{k,n}(t) \tag{A6}$$

is received. The following removal of the cyclic prefix, demodulation and sampling is depicted in Figure A2 exemplarily for user k. In case of multi-path propagation, i.e., $L_k > 1$, the non-zero parts of the signals $\bar{\boldsymbol{y}}_{k,n-1}(t)$ and $\bar{\boldsymbol{y}}_{k,n}(t)$ overlap for $nT_s + \tau_{k,1} \leq t \leq nT_s + \tau_{k,L_k}$ leading to

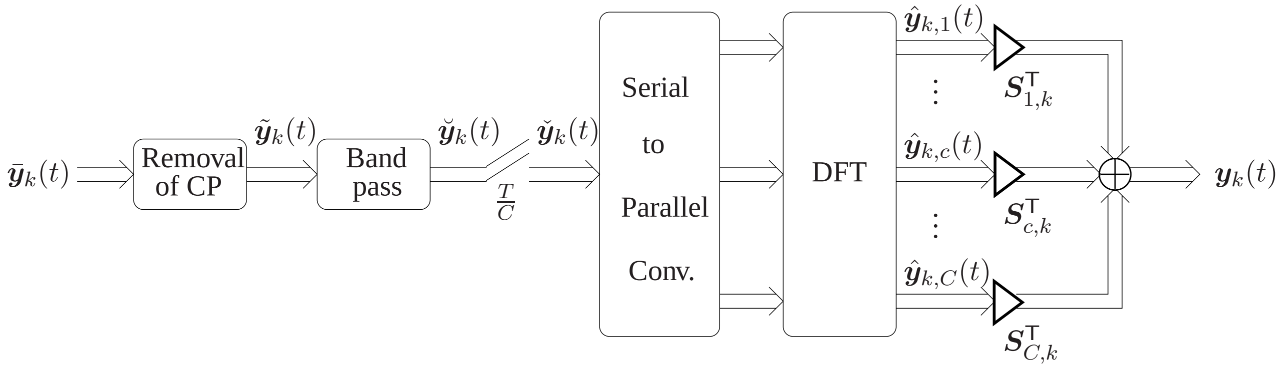

Figure A2: Removal of Cyclic Prefix and Demodulation at Receiver k

undesired intersymbol interference of the symbol $\boldsymbol{x}[n-1]$ on the symbol $\boldsymbol{x}[n]$. By setting the signal $\bar{\boldsymbol{y}}_{k,n}(t)$ to zero during this period, i.e.,

$$\tilde{\boldsymbol{y}}_{k,n}(t) = \begin{cases} 0, & nT_s + \tau_{k,1} \leq t \leq nT_s + \tau_{k,1} + T_{\text{cp}}, \\ \bar{\boldsymbol{y}}_{k,n}(t), & \text{else}, \end{cases} \tag{A7}$$

and choosing T_{cp} according to (A3), all parts of $\bar{\boldsymbol{y}}_k(t)$ containing intersymbol interference are ignored. Thus, the signals $\tilde{\boldsymbol{y}}_{k,n}(t) \in \mathbb{C}^{r_k}$ can be processed independently and only the further processing of these signals is considered in the remainder. Additionally, due to the cyclic prefix in (A4), the information useful for the detection of the signal $\boldsymbol{x}[n]$ contained in the interval $nT_s + \tau_{k,1} \leq t \leq nT_s + \tau_{k,1} + T_{\text{cp}}$ is repeated in the interval

$$\underbrace{nT_s + \tau_{k,1} + T}_{=(n+1)T_s + \tau_{k,1} - T_{\text{cp}}} \leq t \leq \underbrace{nT_s + \tau_{k,1} + T_{\text{cp}} + T}_{=(n+1)T_s + \tau_{k,1}}$$

and therefore not lost by the removal of the cyclic prefix. On the other hand this interference cancellation can only be achieved at the cost of additional delay T_{cp}. The bandpass filter suppresses all signals outside the frequency band $f_0 - B/2 \leq f \leq f_0 + B/2$. In the following it will be assumed that this filter does not affect the signal $\sum_{n=-\infty}^{\infty} \sum_{\ell=1}^{L_k} \tilde{\boldsymbol{H}}_{k,\ell} \bar{\boldsymbol{x}}_n(t - \tau_{k,\ell})$. Although that signal is not strictly band-limited, the energy outside the interval $f_0 - B/2 \leq f \leq f_0 + B/2$ is very low and will therefore be neglected in the following. The noise signal is definitely influenced and the noise $\check{\boldsymbol{\eta}}_k(t)$ after bandpass filtering remains Gaussian circularly symmetric with zero mean, but the covariance matrix becomes $B\tilde{\boldsymbol{R}}_k = \frac{C}{T}\tilde{\boldsymbol{R}}_k$ and noise samples obtained at different time instances are no longer uncorrelated, where the auto-correlation of the band-limited Gaussian noise is given by

$$r(\tau) = \mathrm{E}\left[\check{\boldsymbol{\eta}}_k(t)\check{\boldsymbol{\eta}}_k^{\mathrm{H}}(t-\tau)\right] = \tilde{\boldsymbol{R}}_k \frac{\sin(\pi B t)}{\pi t} = \tilde{\boldsymbol{R}}_k \frac{\sin\left(\pi \frac{C}{T} t\right)}{\pi t}. \tag{A8}$$

The non-zero-part of each signal $\tilde{\boldsymbol{y}}_{k,n}(t)$ is then sampled every T/C seconds leading to the signal

$$\check{\boldsymbol{y}}_{k,n}(t) = \sum_{c=1}^{C} \delta(t - t_c(n))\tilde{\boldsymbol{y}}_{k,n}(t_c(n)) = \sum_{c=1}^{C} \delta(t - t_c(n))\check{\boldsymbol{y}}_{k,c}[n], \tag{A9}$$

where

$$t_c(n) = nT_s + T_{\text{cp}} + \tau_{k,1} + \frac{c-1}{C}T.$$

These C samples are collected by the serial-to-parallel converter and made available at $t = (n+1)T_s + T_{\text{cp}} + \tau_{k,1}$ to the Discrete-Fourier Transform (DFT) block[1]. The DFT generates C vectors $\hat{\boldsymbol{y}}_{k,1}[n], \ldots, \hat{\boldsymbol{y}}_{k,C}[n] \in \mathbb{C}^{r_k}$, where

$$\hat{\boldsymbol{y}}_{k,c}[n] = \frac{1}{C}\sum_{d=1}^{C} \check{\boldsymbol{y}}_{k,d}[n] \exp\left(-j2\pi f_c \frac{d-1}{C}T\right). \tag{A10}$$

The signal $\hat{\boldsymbol{y}}_{k,c}(t)$ is therefore discrete and given by

$$\hat{\boldsymbol{y}}_{k,c}(t) = \sum_{n=-\infty}^{\infty} \delta(t - ((n+1)T_s + T_{\text{cp}} + \tau_{k,1}))\hat{\boldsymbol{y}}_{k,c}[n]. \tag{A11}$$

With equations (A5), (A6), (A7), (A9), and (A10), the signals $\hat{\boldsymbol{y}}_{k,c}[n]$ compute according to

$$\begin{aligned}\hat{\boldsymbol{y}}_{k,c}[n] = &\sum_{\ell=1}^{L_k} \tilde{\boldsymbol{H}}_{k,\ell} \sum_{b=1}^{C} \frac{1}{C} \sum_{d=1}^{C} \exp\left(j2\pi(f_b - f_c)\frac{d-1}{C}T\right) \exp(j2\pi f_b(\tau_{k,1} - \tau_{k,\ell}))\boldsymbol{S}_{c,\mathrm{T}}\boldsymbol{x}[n] + \\ &+ \frac{1}{C}\sum_{d=1}^{C} \breve{\boldsymbol{\eta}}_k(t_d(n)) \exp\left(-j2\pi f_c \frac{d-1}{C}T\right)\end{aligned} \tag{A12}$$

By using the identity $f_b - f_c = \frac{b-c}{T}$ [c.f. (A1), (A2)] and the fact that

$$\frac{1}{C}\sum_{d=1}^{C} \exp\left(j2\pi\frac{(d-1)(b-c)}{C}\right) = \begin{cases} 1, & b = c \\ 0, & \text{else} \end{cases},$$

Equation (A12) can be simplified to

$$\hat{\boldsymbol{y}}_{k,c}[n] = \sum_{\ell=1}^{L_k} \tilde{\boldsymbol{H}}_{k,\ell} \exp(j2\pi f_c(\tau_{k,1} - \tau_{k,\ell}))\boldsymbol{S}_{c,\mathrm{T}}\boldsymbol{x}[n] + \boldsymbol{\eta}_{k,c}[n],$$

where

$$\boldsymbol{\eta}_{k,c}[n] = \frac{1}{C}\sum_{d=1}^{C} \breve{\boldsymbol{\eta}}_k(t_d(n)) \exp\left(-j2\pi f_c \frac{d-1}{C}T\right).$$

By plugging the temporal distances $\tau = \frac{n}{C}T$, $n \in \{1, 2, \ldots, C-1\}$ between the noise samples $\breve{\boldsymbol{\eta}}_k(t_d(n))$ into the autocorrelation function from (A8), it can be derived that the noise samples $\breve{\boldsymbol{\eta}}_k(t_d(n))$ are mutually uncorrelated. Therefore, $\boldsymbol{\eta}_{k,c}$ is a sum of uncorrelated Gaussian variables and hence also Gaussian distributed with zero mean and covariance matrix

$$\begin{aligned}\hat{\boldsymbol{R}}_k &= \mathrm{E}\left[\boldsymbol{\eta}_{k,c}[n]\boldsymbol{\eta}_{k,c}^{\mathrm{H}}[n]\right] = \\ &= \frac{1}{C^2}\sum_{b=1}^{C}\sum_{d=1}^{C} \mathrm{E}\left[\breve{\boldsymbol{\eta}}_k(t_b(n))\breve{\boldsymbol{\eta}}_k^{\mathrm{H}}(t_d(n))\right] \exp\left(-j2\pi f_c \frac{d-b}{C}T\right) = \\ &= \frac{1}{C^2}\sum_{d=1}^{C} \frac{C}{T}\tilde{\boldsymbol{R}}_k = \frac{1}{T}\tilde{\boldsymbol{R}}_k,\end{aligned} \tag{A13}$$

[1] In case $C = 2^p$ for integer p the Fast-Fourier Transform can be used as an efficient implementation for the DFT block.

which is identical for all carriers c. Additionally, the noise vectors $\boldsymbol{\eta}_{k,c}[n]$ and $\boldsymbol{\eta}_{k,c'}[n]$ of different carriers $c \neq c'$ are uncorrelated as

$$\mathrm{E}\left[\boldsymbol{\eta}_{k,c}\boldsymbol{\eta}_{k,c'}^{\mathrm{H}}\right] = \frac{1}{C^2}\sum_{d=1}^{C}\tilde{\boldsymbol{R}}_k \exp\left(j2\pi\frac{T}{C}(d-1)f_c\right) = 0, \tag{A14}$$

which can be stated as the f_c are integer multiples of $1/T$. Finally, the multiplication with the selection matrices $\boldsymbol{S}_{c,k}^{\mathsf{T}}$ and the following addition stacks the signals $\hat{\boldsymbol{y}}_{k,c}[n]$ received on the different carriers into one vector $\boldsymbol{y}_k[n] \in \mathbb{C}^{Cr_k}$ so that

$$\begin{aligned}\boldsymbol{y}_k[n] &= \sum_{c=1}^{C}\boldsymbol{S}_{c,k}^{\mathsf{T}}\hat{\boldsymbol{y}}_{k,c}[n] = \sum_{c=1}^{C}\boldsymbol{S}_{c,k}^{\mathsf{T}}\sum_{\ell=1}^{L_k}\tilde{\boldsymbol{H}}_{k,\ell}\exp(j2\pi f_c(\tau_{k,1}-\tau_{k,\ell}))\boldsymbol{S}_{c,\mathrm{T}}\boldsymbol{x}[n] + \sum_{c=1}^{C}\boldsymbol{S}_{c,k}^{\mathsf{T}}\boldsymbol{\eta}_{k,c}[n] = \\ &= \boldsymbol{H}_k\boldsymbol{x}[n] + \boldsymbol{\eta}[n],\end{aligned}$$

which corresponds to (2.3). The covariance matrices $\boldsymbol{R}_k = \mathrm{E}\left[\boldsymbol{\eta}_k[n]\boldsymbol{\eta}_k[n]^{\mathrm{H}}\right]$ as given in (2.4) can then be derived from (A13) and (A14).

A2. Optimum Power Allocation for Weighted Sum Rate Maximization over Scalar Channels

In this section the optimum power allocation

$$\{p_{k,j}\}_{j=1,\ldots,d_k,k=1,\ldots,K} = \underset{\{\hat{p}_{k,j}\}_{j=1,\ldots,d_k,k=1,\ldots,K}}{\operatorname{argmax}} \sum_{k=1}^{K}\mu_k\sum_{j=1}^{d_k}\log_2\left(1+\hat{p}_{k,j}\frac{C}{\sigma_{\mathrm{n}}^2}\lambda_{k,j}\right),$$
$$\text{s.t.} \sum_{k=1}^{K}\sum_{j=1}^{d_k}\hat{p}_{k,j} \leq P_{\mathrm{Tx}}$$

will be derived. The Karush-Kuhn-Tucker (KKT) conditions of this problem read as

$$\begin{gathered}\frac{\mu_k\frac{C}{\sigma_{\mathrm{n}}^2}\lambda_{k,j}}{\ln 2\left(1+\hat{p}_{k,j}\frac{C}{\sigma_{\mathrm{n}}^2}\lambda_{k,j}\right)} - \hat{\eta} - \nu_{k,j} = 0, \quad \forall j=1,\ldots,d_k, k=1,\ldots,K \\ \hat{\eta}\left(\sum_{k=1}^{K}\sum_{j=1}^{d_k}\hat{p}_{k,j} - P_{\mathrm{Tx}}\right) = 0, \quad \hat{\eta} \geq 0, \quad \nu_{k,j}\hat{p}_{k,j} = 0, \quad \nu_{k,j} \geq 0, \forall j=1,\ldots,d_k, k=1,\ldots,K.\end{gathered} \tag{A15}$$

Multiplying the equations in the first line with the corresponding $\hat{p}_{k,j}$, it can be shown that

$$p_{k,j} = \left[\frac{\mu_k \ln 2}{\hat{\eta}} - \frac{\sigma_{\mathrm{n}}^2}{C\lambda_{k,j}}\right]^+ = \left[\eta\mu_k - \frac{\sigma_{\mathrm{n}}^2}{C\lambda_{k,j}}\right]^+,$$

where η is chosen so that the transmit power constraint is fulfilled with equality, fulfills all KKTs in (A15).

A3. Proof of Optimality of Using the Pseudo-inverse of the Composite Channel Matrix as Precoder under Zero-Forcing Constraints

In this section it is proofed that the optimum solution to the optimization problem

$$\max_{\boldsymbol{T},\{\lambda_{k,j}\}_{j=1,\ldots,d_k,k=1,\ldots,K}} f(\lambda_{1,1},\ldots,\lambda_{d_K,K})$$

$$\text{s.t. } \lambda_{k,j}\geq 0, \quad \boldsymbol{e}_{n_{k,j}}^{\mathsf{T}}\boldsymbol{T}^{\mathrm{H}}\boldsymbol{T}\boldsymbol{e}_{n_{k,j}}=1, \forall j=1,\ldots,d_k,\forall k, \quad \boldsymbol{H}_{\text{comp}}\boldsymbol{T}=\mathrm{diag}\left(\sqrt{\lambda_{1,1}},\ldots,\sqrt{\lambda_{d_K,K}}\right),$$

where

$$n_{k,j}=\sum_{k'=1}^{k-1}d_{k'}+j, \qquad \boldsymbol{T}=\left[\boldsymbol{t}_{1,1},\ldots,\boldsymbol{t}_{K,d_k}\right],$$

and

$$f(\lambda_{1,1},\ldots,\lambda_{d_K,K})=\log_2\left(\lambda_{k,j}\frac{\mu_k C}{\sigma_{\mathrm{n}}^2}\frac{P_{\mathrm{Tx}}+\sum_{k'=1}^{K}\sum_{j=1}^{d_{k'}}\frac{\sigma_{\mathrm{n}}^2}{C\lambda_{k',j}}}{\sum_{k'=1}^{K}\mu_{k'}d_{k'}}\right),$$

is given by

$$\lambda_{k,j}=\frac{1}{\boldsymbol{e}_{n_{k,j}}^{\mathsf{T}}\left(\boldsymbol{H}_{\text{comp}}\boldsymbol{H}_{\text{comp}}^{\mathrm{H}}\right)^{-1}\boldsymbol{e}_{n_{k,j}}} \tag{A16}$$

and

$$\boldsymbol{T}=\boldsymbol{H}_{\text{comp}}^{+}\,\mathrm{diag}\left(\sqrt{\lambda_{1,1}},\ldots,\sqrt{\lambda_{d_K,K}}\right). \tag{A17}$$

Using the definition of a generalized inverse from [134], the constraint $\boldsymbol{H}_{\text{comp}}\boldsymbol{T} = \mathrm{diag}\left(\sqrt{\lambda_{1,1}},\ldots,\sqrt{\lambda_{d_K,K}}\right)$ is fulfilled by choosing

$$\boldsymbol{T}=\left(\boldsymbol{H}_{\text{comp}}^{+}+\boldsymbol{P}_{\perp}\boldsymbol{U}\right)\mathrm{diag}\left(\sqrt{\lambda_{1,1}},\ldots,\sqrt{\lambda_{d_K,K}}\right)=\left(\boldsymbol{H}_{\text{comp}}^{+}+\boldsymbol{P}_{\perp}\boldsymbol{U}\right)\boldsymbol{\Lambda}^{\frac{1}{2}}, \tag{A18}$$

where $\boldsymbol{P}_{\perp}=\boldsymbol{I}-\boldsymbol{H}_{\text{comp}}^{+}\boldsymbol{H}_{\text{comp}}$ is a projector into the nullspace of $\boldsymbol{H}_{\text{comp}}^{+}$ and $\boldsymbol{U}$ is an arbitrary matrix. Inserting (A18) into the constraint functions $\boldsymbol{e}_{n_{k,j}}^{\mathsf{T}}\boldsymbol{T}^{\mathrm{H}}\boldsymbol{T}\boldsymbol{e}_{n_{k,j}}=1$ yields

$$\begin{aligned}
\boldsymbol{e}_{n_{k,j}}^{\mathsf{T}}\boldsymbol{T}^{\mathrm{H}}\boldsymbol{T}\boldsymbol{e}_{n_{k,j}} &= \boldsymbol{e}_{n_{k,j}}^{\mathsf{T}}\boldsymbol{\Lambda}^{\frac{1}{2}}\left(\boldsymbol{H}_{\text{comp}}^{+}+\boldsymbol{P}_{\perp}\boldsymbol{U}\right)^{\mathrm{H}}\left(\boldsymbol{H}_{\text{comp}}^{+}+\boldsymbol{P}_{\perp}\boldsymbol{U}\right)\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{e}_{n_{k,j}}\\
&= \boldsymbol{e}_{n_{k,j}}^{T}\boldsymbol{\Lambda}^{\frac{1}{2}}\left(\boldsymbol{H}_{\text{comp}}^{+}\right)^{\mathrm{H}}\boldsymbol{H}_{\text{comp}}^{+}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{e}_{n_{k,j}}+\boldsymbol{e}_{n_{k,j}}^{T}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{P}_{\perp}\boldsymbol{U}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{e}_{n_{k,j}}\\
&= \lambda_{k,j}\left(\boldsymbol{e}_{n_{k,j}}^{T}\left(\boldsymbol{H}_{\text{comp}}^{+}\right)^{\mathrm{H}}\boldsymbol{H}_{\text{comp}}^{+}\boldsymbol{e}_{n_{k,j}}+\boldsymbol{e}_{n_{k,j}}^{T}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{P}_{\perp}\boldsymbol{U}\boldsymbol{e}_{n_{k,j}}\right)=1.
\end{aligned}$$

From these equality constraints the $\lambda_{k,j}$ can be computed explicitly as

$$\lambda_{k,j}=\frac{1}{\boldsymbol{e}_{n_{k,j}}^{T}\left(\boldsymbol{H}^{+}\right)^{\mathrm{H}}\boldsymbol{H}^{+}\boldsymbol{e}_{n_{k,j}}+\boldsymbol{e}_{n_{k,j}}^{T}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{P}_{\perp}\boldsymbol{U}\boldsymbol{e}_{n_{k,j}}}.$$

As the $\boldsymbol{U}^{\mathrm{H}}\boldsymbol{P}_{\perp}\boldsymbol{U}$ are positive semidefinite, i.e., $\boldsymbol{e}_{n_{k,j}}^{T}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{P}_{\perp}\boldsymbol{U}\boldsymbol{e}_{n_{k,j}}\geq 0$, at the optimum $\boldsymbol{U}=\boldsymbol{0}_{\sum d_k,\sum d_k}$ holds, if the objective function is a monotonically increasing function in each $\lambda_{k,j}$. Together with (A18) this leads to (A17). By furthermore using the fact that

$$\left(\boldsymbol{H}_{\text{comp}}^{+}\right)^{\mathrm{H}}\boldsymbol{H}_{\text{comp}}^{+}=\left(\boldsymbol{H}_{\text{comp}}\boldsymbol{H}_{\text{comp}}^{\mathrm{H}}\right)^{-1}\boldsymbol{H}_{\text{comp}}\boldsymbol{H}_{\text{comp}}^{\mathrm{H}}\left(\boldsymbol{H}_{\text{comp}}\boldsymbol{H}_{\text{comp}}^{\mathrm{H}}\right)^{-1}=\left(\boldsymbol{H}_{\text{comp}}\boldsymbol{H}_{\text{comp}}^{\mathrm{H}}\right)^{-1},$$

the result in (A16) can be obtained. Thus, to complete the proof it has to be shown that the objective function is monotonically increasing in $\lambda_{k,j}$. This can be done by using the following lemma:

Lemma A3.1 *For fixed channel gains* $\lambda_{1,1},\ldots,\lambda_{1,d_1},\ldots,\lambda_{K,1},\ldots,\lambda_{K,d_K}$ *the weighted sum rate* $R_{WSR}(\lambda_{1,1},\ldots,\lambda_{d_K,K})$ *obtainable with an optimum power allocation is a monotonically increasing function in each channel gain* $\lambda_{k,j}$.

Proof. At places, where $R_{\mathrm{WSR}}(\lambda_{1,1},\ldots,\lambda_{d_K,K})$ is differentiable the lemma can be proofed by taking the partial derivative with respect to $\lambda_{k,j}$, which is done exemplarily for the case that all subchannels receive non-zero powers in the following.

$$\begin{aligned}\frac{\partial R_{\mathrm{WSR}}(\lambda_{1,1},\ldots,\lambda_{d_K,K})}{\partial\lambda_{k,j}} &= \frac{\mu_k}{\ln 2\lambda_{k,j}} - \frac{1}{\ln 2\left(\sum\limits_{k'=1}^{K}\mu_{k'}d_{k'}\right)\frac{P_{\mathrm{Tx}}+\sum\limits_{k'=1}^{K}\sum\limits_{j'=1}^{d_{k'}}\frac{\sigma_{\mathrm{n}}^2}{C\lambda_{k',j'}}}{\sum_{k'=1}^{K}\mu_{k'}d_{k'}}}\frac{\sigma_{\mathrm{n}}^2}{C\lambda_{k,j}}\sum_{m=1}^{K}\sum_{\ell=1}^{d_m}\mu_m = \\ &= \frac{1}{\ln 2\lambda_{k,j}}\left(\mu_k - \frac{\sigma_{\mathrm{n}}^2}{\eta C\lambda_{k,j}}\right) = \frac{p_{k,j}}{\ln 2\lambda_{k,j}\eta} > 0.\end{aligned}$$

It remains to show that at places $\lambda_{k,j}=\hat{\lambda}_{k,j}$, where $\hat{\lambda}_{k,j}$ is equal to $\frac{\sigma_{\mathrm{n}}^2}{\mu_k\eta C}$, i.e.,

$$\hat{\lambda}_{k,j} = \frac{\sigma_{\mathrm{n}}^2}{\mu_k\eta C}$$

the weighted sum rate does not increase when $\lambda_{k,j}$ is decreased. Assuming that all other channel gains remain constant this implies that

$$\eta = \frac{P_{\mathrm{Tx}} + \sum\limits_{\substack{k'=1\\k'\neq k}}^{K}\sum\limits_{j=1}^{d_{k'}}\frac{\sigma_{\mathrm{n}}^2}{C\lambda_{k',j}} + \sum\limits_{\substack{j'=1\\j'\neq j}}^{d_k}\frac{\sigma_{\mathrm{n}}^2}{C\lambda_{k,j'}} + \frac{\sigma_{\mathrm{n}}^2}{C\hat{\lambda}_{k,j}}}{\sum_{k'=1}^{K}\mu_{k'}d_{k'}} = \frac{\sigma_{\mathrm{n}}^2}{\mu_k C\hat{\lambda}_{k,j}} \tag{A19}$$

The weighted sum rate $R_{\mathrm{WSR}}(\hat{\lambda}_{k,j})$ at $\hat{\lambda}_{k,j}$ is therefore given by

$$\begin{aligned}R_{\mathrm{WSR}}\left(\hat{\lambda}_{k,j}\right) &= \left(\sum_{\substack{k'=1\\k'\neq k}}^{K}\mu_{k'}d_{k'} + \mu_k(d_k-1)\right)\log_2(\eta) - \sum_{\substack{k'=1\\k'\neq k}}^{K}\sum_{j'=1}^{d_{k'}}\log_2\left(\frac{\mu_{k'}C\lambda_{k',j'}}{\sigma_{\mathrm{n}}^2}\right) - \\ &- \sum_{\substack{j'=1\\j'\neq j}}^{d_k}\log_2\left(\frac{\mu_k C\lambda_{k,j'}}{\sigma_{\mathrm{n}}^2}\right) + \mu_k\left(\log_2(\eta) - \log_2\left(\frac{\mu_k C\hat{\lambda}_{k,j}}{\sigma_{\mathrm{n}}^2}\right)\right) = \\ &= \left(\sum_{\substack{k'=1\\k'\neq k}}^{K}\mu_{k'}d_{k'} + \mu_k(d_k-1)\right)\log_2(\eta) - \sum_{\substack{k'=1\\k'\neq k}}^{K}\sum_{j'=1}^{d_{k'}}\log_2\left(\frac{\mu_{k'}C\lambda_{k',j'}}{\sigma_{\mathrm{n}}^2}\right) - \sum_{\substack{j'=1\\j'\neq j}}^{d_k}\log_2\left(\frac{\mu_k C\lambda_{k,j'}}{\sigma_{\mathrm{n}}^2}\right)\end{aligned}$$

When $\hat{\lambda}_{k,j}$ is decreased by $\varepsilon>0$, the jth subchannel of user k is deactivated and the water-level has to be computed without considering the corresponding subchannel. That is why the weighted sum rate is non-differentiable at these places. Then a constant sum rate can be achieved which is given by

$$R_{\text{WSR}}\left(\hat{\lambda}_{k,j}-\varepsilon\right)=\left(\sum_{\substack{k'=1\\k'\neq k}}^{K}\mu_{k'}d_{k'}+\mu_k(d_k-1)\right)\log_2(\hat{\eta})-\sum_{\substack{k'=1\\k'\neq k}}^{K}\sum_{j'=1}^{d_{k'}}\log_2\left(\frac{\mu_{k'}C\lambda_{k',j'}}{\sigma_{\text{n}}^2}\right),$$

with the waterlevel

$$\hat{\eta}=\frac{P_{\text{Tx}}+\sum\limits_{\substack{k'=1\\k'\neq k}}^{K}\sum\limits_{j'=1}^{d_{k'}}\frac{\sigma_{\text{n}}^2}{C\lambda_{k',j'}}+\sum\limits_{\substack{j'=1\\j'\neq k}}^{d_k}\frac{\sigma_{\text{n}}^2}{C\lambda_{k,j'}}}{\sum\limits_{k'=1}^{K}\mu_{k'}d_{k'}-\mu_k}=\frac{\eta\left(\sum\limits_{k'=1}^{K}\mu_{k'}d_{k'}\right)-\frac{\sigma_{\text{n}}^2}{C\hat{\lambda}_{k,j}}}{\sum\limits_{k'=1}^{K}\mu_{k'}d_{k'}-\mu_k}=\frac{\sigma_{\text{n}}^2}{\mu_k C\hat{\lambda}_{k,j}}=\eta,$$

where (A19) has been used to obtain the second last equality. Thus, although $R_{\text{WSR}}(\lambda_{k,j})$ is not differentiable at $\lambda_{k,j}=\hat{\lambda}_{k,j}$, there is no discontinuity at this place, i.e., $R_{\text{WSR}}\left(\hat{\lambda}_{k,j}-\varepsilon\right)-R_{\text{WSR}}\left(\hat{\lambda}_{k,j}\right)=0$, and the sum rate does not increase for $\varepsilon>0$. □

In Problem (3.40) additionally the carrier separation constraints $\boldsymbol{t}_{k,j}=\boldsymbol{S}_{\gamma(k,j)}^{\mathsf{T}}\boldsymbol{S}_{\gamma(k,j)}\boldsymbol{S}_{\gamma(k,j)}\boldsymbol{t}_{k,j}$ have to be fulfilled in multicarrier systems with $C>1$. As long as the receive filters $\boldsymbol{g}_{k,j}$ obey the carrier separation constraints, the transmit filters from (A17) fulfill the carrier separation constraints as well. That is because

$$\begin{aligned}\boldsymbol{t}_{k,j}&=\boldsymbol{T}\boldsymbol{e}_{n_{k,j}}=\boldsymbol{H}_{\text{comp}}^{\text{H}}\left(\boldsymbol{H}_{\text{comp}}\boldsymbol{H}_{\text{comp}}^{\text{H}}\right)^{-1}\operatorname{diag}\left(\sqrt{\lambda_{1,1}},\ldots,\sqrt{\lambda_{d_K,K}}\right)\boldsymbol{e}_{n_{k,j}}\\&=\boldsymbol{H}_{\text{comp}}^{\text{H}}\boldsymbol{\Pi}\left(\boldsymbol{\Pi}\boldsymbol{H}_{\text{comp}}\boldsymbol{H}_{\text{comp}}^{\text{H}}\boldsymbol{\Pi}\right)^{-1}\boldsymbol{\Pi}\boldsymbol{e}_{n_{k,j}}\sqrt{\lambda_{k,j}},\end{aligned}$$

where $\boldsymbol{\Pi}$ is a permutation matrix with $\boldsymbol{\Pi}=\boldsymbol{\Pi}^{\mathsf{T}}=\boldsymbol{\Pi}^{-1}$ that permutes the rows of $\boldsymbol{H}_{\text{comp}}$ so that $\boldsymbol{\Pi}\boldsymbol{H}_{\text{comp}}$ becomes block-diagonal, which is possible, as longs as the $\boldsymbol{g}_{k,j}$ fulfill the carrier separation constraints. Thus, all n_1 vectors $\boldsymbol{g}_{k,j}^{\text{H}}\boldsymbol{H}_k$ with $\gamma(k,j)=1$ are accumulated in the first n_1 rows of $\boldsymbol{\Pi}\boldsymbol{H}_{\text{comp}}$, followed by all n_2 vectors $\boldsymbol{g}_{k,j}^{\text{H}}\boldsymbol{H}_k$ with $\gamma(k,j)=2$ and so on, where n_c denotes the number of data streams allocated to carrier c, and $\boldsymbol{\Pi}\boldsymbol{H}_{\text{comp}}$ can be written as

$$\boldsymbol{\Pi}\boldsymbol{H}_{\text{comp}}=\operatorname{blockdiag}\left(\boldsymbol{H}_{\text{comp},1},\ldots,\boldsymbol{H}_{\text{comp},C}\right),$$

where $\boldsymbol{H}_{\text{comp},c}\in\mathbb{C}^{n_c\times N_{\text{T}}}$ denotes the composite channel matrix on carrier c. Hence,

$$\begin{aligned}\boldsymbol{t}_{k,j}&=\operatorname{blockdiag}\left(\boldsymbol{H}_{\text{comp},1}^{\text{H}}\left(\boldsymbol{H}_{\text{comp},1}\boldsymbol{H}_{\text{comp},1}^{\text{H}}\right)^{-1},\ldots,\boldsymbol{H}_{\text{comp},C}^{\text{H}}\left(\boldsymbol{H}_{\text{comp},C}\boldsymbol{H}_{\text{comp},C}^{\text{H}}\right)^{-1}\right)\boldsymbol{\Pi}\boldsymbol{e}_{n_{k,j}}\sqrt{\lambda_{k,j}}\\&=\hat{\boldsymbol{T}}\boldsymbol{\Pi}\boldsymbol{e}_{n_{k,j}}\sqrt{\lambda_{k,j}}.\end{aligned}$$

As $\hat{\boldsymbol{T}}$ is block-diagonal, the columns in $\hat{\boldsymbol{T}}$ are arranged carrier-wise, i.e., the first n_1 columns are collinear with the precoders on carrier 1, and the permutation matrix $\boldsymbol{\Pi}$ rearranges the columns of $\hat{\boldsymbol{T}}$ so that the $n_{k,j}$th column of $\hat{\boldsymbol{T}}\boldsymbol{\Pi}$ is collinear with the precoder of the jth data stream of user k, the $\boldsymbol{t}_{k,j}$ fulfill the carrier separation constraint.

A4. Computation of Inverse Channel Gains for Zero-Forcing without DPC

In this section, Equation (3.48) for the computation of the inverse channel gains $\boldsymbol{e}_j^{\mathsf{T}} \left(\boldsymbol{H}_{\text{comp}}\boldsymbol{H}_{\text{comp}}^{\text{H}}\right)^{-1} \boldsymbol{e}_j$ will be derived. These are given by

$$\boldsymbol{e}_j^{\mathsf{T}} \left(\boldsymbol{H}_{\text{comp}}\boldsymbol{H}_{\text{comp}}^{\text{H}}\right)^{-1} \boldsymbol{e}_j = \frac{|\boldsymbol{H}_{\text{comp},-j}\boldsymbol{H}_{\text{comp},-j}^{\text{H}}|}{|\boldsymbol{H}_{\text{comp}}\boldsymbol{H}_{\text{comp}}^{\text{H}}|}, \tag{A20}$$

which follows from determining the matrix inverse via its adjoint matrix (e.g. [94, Ch. 0.8.2]) and where $\boldsymbol{H}_{\text{comp},-j}$ is built by removing the jth row from $\boldsymbol{H}_{\text{comp}}$. From (3.47) the determinant $|\boldsymbol{H}_{\text{comp}}\boldsymbol{H}_{\text{comp}}^{\text{H}}|$ can be computed according to

$$\begin{aligned}
\left|\boldsymbol{H}_{\text{comp}}\boldsymbol{H}_{\text{comp}}^{\text{H}}\right| &= \left|\begin{bmatrix} \boldsymbol{H}_{\text{comp}}^{(i-1)}\boldsymbol{H}_{\text{comp}}^{(i-1),\text{H}} & \boldsymbol{H}_{\text{comp}}^{(i-1)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{g} \\ \boldsymbol{g}^{\text{H}}\boldsymbol{H}_k\boldsymbol{H}_{\text{comp}}^{(i-1),\text{H}} & \boldsymbol{g}^{\text{H}}\boldsymbol{H}_k\boldsymbol{H}_k^{\text{H}}\boldsymbol{g} \end{bmatrix}\right| \\
&= \left|\begin{pmatrix} \boldsymbol{H}_{\text{comp}}^{(i-1)}\boldsymbol{H}_{\text{comp}}^{(i-1),\text{H}} & \boldsymbol{H}_{\text{comp}}^{(i-1)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{g} & \boldsymbol{0}_{i-1,i-2} \\ \boldsymbol{g}^{\text{H}}\boldsymbol{H}_k\boldsymbol{H}_{\text{comp}}^{(i-1),\text{H}} & \boldsymbol{g}^{\text{H}}\boldsymbol{H}_k\boldsymbol{H}_k^{\text{H}}\boldsymbol{g} & \boldsymbol{0}_{1,i-2} \\ \boldsymbol{0}_{i-2,i-1} & 0 & \boldsymbol{I}_{i-2} \end{pmatrix}\right| \\
&= \boldsymbol{g}^{\text{H}}\boldsymbol{H}_k\boldsymbol{H}_k^{\text{H}}\boldsymbol{g}\,\Big|\boldsymbol{H}_{\text{comp}}^{(i-1)}\boldsymbol{H}_{\text{comp}}^{(i-1),\text{H}} - \\
&- \begin{bmatrix} \boldsymbol{H}_{\text{comp}}^{(i-1)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{g} & \boldsymbol{0}_{i-1,i-2} \end{bmatrix} \begin{bmatrix} \frac{1}{\boldsymbol{g}^{\text{H}}\boldsymbol{H}_k\boldsymbol{H}_k^{\text{H}}\boldsymbol{g}} & \boldsymbol{0}_{1,i-2} \\ 0 & \boldsymbol{I}_{i-2} \end{bmatrix} \begin{bmatrix} \boldsymbol{g}^{\text{H}}\boldsymbol{H}_k\boldsymbol{H}_{\text{comp}}^{(i-1),\text{H}} \\ \boldsymbol{0}_{i-2,i-1} \end{bmatrix}\Big| \\
&= \left|\boldsymbol{H}_{\text{comp}}^{(i-1)}\boldsymbol{H}_{\text{comp}}^{(i-1),\text{H}}\right| \Big(\boldsymbol{g}^{\text{H}}\boldsymbol{H}_k\boldsymbol{H}_k^{\text{H}}\boldsymbol{g} - \\
&- \boldsymbol{g}^{\text{H}}\boldsymbol{H}_k\boldsymbol{H}_{\text{comp}}^{(i-1),\text{H}}\left(\boldsymbol{H}_{\text{comp}}^{(i-1)}\boldsymbol{H}_{\text{comp}}^{(i-1),\text{H}}\right)^{-1}\boldsymbol{H}_{\text{comp}}^{(i-1)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{g}\Big) = \\
&= \left|\boldsymbol{H}_{\text{comp}}^{(i-1)}\boldsymbol{H}_{\text{comp}}^{(i-1),\text{H}}\right| \boldsymbol{g}^{\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{\text{lin}}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{g},
\end{aligned} \tag{A21}$$

where the formula for the determinant of block matrices from [135] has been used in the third line and $\hat{\boldsymbol{P}}_{\text{lin}}^{(i)}$ is given by (3.49). As shown in the following, the $\hat{\boldsymbol{P}}_{\text{lin}}^{(i)}$ have a block-diagonal structure. Let $\boldsymbol{\Pi}$ be a permutation matrix with $\boldsymbol{\Pi} = \boldsymbol{\Pi}^{\mathsf{T}} = \boldsymbol{\Pi}^{-1}$ that permutes the columns of $\boldsymbol{H}_{\text{comp}}^{(i-1)}$ so that the first n_1 rows contain the effective channels $\boldsymbol{g}_{k,j}^{\text{H}}\boldsymbol{H}_k$ of the first carrier, i.e., all data streams with $\gamma(k,j) = 1$, followed by the n_2 effective channels of the second carrier and so on. As the receive filters $\boldsymbol{g}_{k,j}$ obey to the carrier separation constraint, the matrix $\boldsymbol{\Pi}\boldsymbol{H}_{\text{comp}}^{(i-1)}$ has a block-diagonal structure. Additionally, due to the properties of the permutation matrix, the projector $\hat{\boldsymbol{P}}_{\text{lin}}^{(i)}$ can be written as

$$\begin{aligned}
\hat{\boldsymbol{P}}_{\text{lin}}^{(i)} &= \boldsymbol{I}_{CN_{\text{T}}} - \boldsymbol{H}_{\text{comp}}^{(i-1),\text{H}}\left(\boldsymbol{H}_{\text{comp}}^{(i-1)}\boldsymbol{H}_{\text{comp}}^{(i-1),\text{H}}\right)^{-1}\boldsymbol{H}_{\text{comp}}^{(i-1)} = \\
&= \boldsymbol{I}_{CN_{\text{T}}} - \boldsymbol{H}_{\text{comp}}^{(i-1),\text{H}}\boldsymbol{\Pi}\left(\boldsymbol{\Pi}\boldsymbol{H}_{\text{comp}}^{(i-1)}\boldsymbol{H}_{\text{comp}}^{(i-1),\text{H}}\boldsymbol{\Pi}\right)^{-1}\boldsymbol{\Pi}\boldsymbol{H}_{\text{comp}}^{(i-1)},
\end{aligned}$$

where the second term is a product of block-diagonal matrices and therefore also block-diagonal. The determinants $\left|\boldsymbol{H}_{\text{comp},-j}\boldsymbol{H}_{\text{comp},-j}^{\text{H}}\right|$ for $j < i$ can be obtained in a similar manner so that

$$\left|\boldsymbol{H}_{\text{comp},-j}\boldsymbol{H}_{\text{comp},-j}^{\text{H}}\right| = \left|\boldsymbol{H}_{\text{comp},-j}^{(i-1)}\boldsymbol{H}_{\text{comp},-j}^{(i-1),\text{H}}\right| \boldsymbol{g}^{\text{H}}\boldsymbol{H}_k\hat{\boldsymbol{P}}_{-j}^{(i)}\boldsymbol{H}_k^{\text{H}}\boldsymbol{g},$$

where

$$\hat{\boldsymbol{P}}_{-j}^{(i)} = \boldsymbol{I}_{CN_{\text{T}}} - \boldsymbol{H}_{\text{comp},-j}^{(i-1),\text{H}}\left(\boldsymbol{H}_{\text{comp},-j}^{(i-1)}\boldsymbol{H}_{\text{comp},-j}^{(i-1),\text{H}}\right)^{-1}\boldsymbol{H}_{\text{comp},-j}^{(i-1)} \tag{A22}$$

projects into $\operatorname{null}\left\{\boldsymbol{H}_{\text{comp},-j}^{(i-1)}\right\}$ and is block-diagonal. As $\hat{\boldsymbol{P}}_{-j}^{(i)}$ has the same nullspace as $\hat{\boldsymbol{P}}_{\text{lin}}^{(i)}$ except for this component of the jth row of $\boldsymbol{H}_{\text{comp}}^{(i-1)}$, which is orthogonal to all other rows of $\boldsymbol{H}_{\text{comp}}^{(i-1)}$, it can also be written as

$$\hat{\boldsymbol{P}}_{-j}^{(i)} = \hat{\boldsymbol{P}}_{\text{lin}}^{(i)} + \frac{\hat{\boldsymbol{P}}_{-j}^{(i)} \boldsymbol{H}_{\text{comp}}^{(i-1),\text{H}} \boldsymbol{e}_j \boldsymbol{e}_j^{\mathsf{T}} \boldsymbol{H}_{\text{comp}}^{(i-1)} \hat{\boldsymbol{P}}_{-j}^{(i)}}{\boldsymbol{e}_j^{\mathsf{T}} \boldsymbol{H}_{\text{comp}}^{(i-1)} \hat{\boldsymbol{P}}_{-j}^{(i)} \boldsymbol{H}_{\text{comp}}^{(i-1),\text{H}} \boldsymbol{e}_j} = \hat{\boldsymbol{P}}_{\text{lin}}^{(i)} + \boldsymbol{T}^{(i-1)} \boldsymbol{e}_j \boldsymbol{e}_j^{\mathsf{T}} \boldsymbol{T}^{(i-1),\text{H}}. \tag{A23}$$

The vector $\hat{\boldsymbol{P}}_{-j}^{(i)} \boldsymbol{H}_{\text{comp}}^{(i-1),\text{H}} \boldsymbol{e}_j$ is collinear with $\boldsymbol{T}^{(i-1)} \boldsymbol{e}_j$, as $\boldsymbol{T}^{(i-1)} \boldsymbol{e}_j$ lies in the nullspace of $\boldsymbol{H}_{\text{comp},-j}^{(i-1)}$, i.e., $\boldsymbol{T}^{(i-1)} \boldsymbol{e}_j \in \operatorname{null}\left\{\boldsymbol{H}_{\text{comp},-j}^{(i-1)}\right\} = \operatorname{null}\left\{\hat{\boldsymbol{P}}_{-j}^{(i)}\right\}$ and has no component in the nullspace of $\boldsymbol{H}_{\text{comp}}^{(i-1)}$, which follows from the properties of the columns of the pseudo-invers of $\boldsymbol{H}_{\text{comp}}$ (c.f. Appendix A3). Thus,

$$\left|\boldsymbol{H}_{\text{comp},-j} \boldsymbol{H}_{\text{comp},-j}^{\text{H}}\right| = \left|\boldsymbol{H}_{\text{comp},-j}^{(i-1)} \boldsymbol{H}_{\text{comp},-j}^{(i-1),\text{H}}\right| \boldsymbol{g}^{\text{H}} \boldsymbol{H}_k \left(\hat{\boldsymbol{P}}_{\text{lin}}^{(i)} + \boldsymbol{T}^{(i-1)} \boldsymbol{e}_j \boldsymbol{e}_j^{\mathsf{T}} \boldsymbol{T}^{(i-1),\text{H}}\right) \boldsymbol{H}_k^{\text{H}} \boldsymbol{g}. \tag{A24}$$

For $j = i$ it can be easily seen that

$$\left|\boldsymbol{H}_{\text{comp},-i} \boldsymbol{H}_{\text{comp},-i}^{\text{H}}\right| = \left|\boldsymbol{H}_{\text{comp}}^{(i-1)} \boldsymbol{H}_{\text{comp}}^{(i-1),\text{H}}\right|, \tag{A25}$$

which is independent of $\boldsymbol{g}$. Inserting (A21), (A24) and (A25) into (A20) and introducing the variables

$$\alpha_j^{(i-1)} = \frac{\left|\boldsymbol{H}_{\text{comp},-j}^{(i-1)} \boldsymbol{H}_{\text{comp},-j}^{(i-1),\text{H}}\right|}{\left|\boldsymbol{H}_{\text{comp}}^{(i-1)} \boldsymbol{H}_{\text{comp}}^{(i-1),\text{H}}\right|} = \boldsymbol{e}_j^{\mathsf{T}} \left(\boldsymbol{H}_{\text{comp}}^{(i-1)} \boldsymbol{H}_{\text{comp}}^{(i-1),\text{H}}\right)^{-1} \boldsymbol{e}_j$$

leads to the desired result.

A5. Derivation of the Update Rule for the Transmit Filters in a Successive Algorithm without DPC

In this section the update rule for the transmit filters in a successive algorithm without DPC given in Equation (3.68), which is conform with

$$\boldsymbol{t}_{k,j}^{(i)} = \beta_{k,j}^{(i)} \left(\boldsymbol{t}_{k,j}^{(i-1)} - \boldsymbol{t}_{k(i),d_{k(i)}}^{(i)} \gamma_{k,j}^{(i)} \sqrt{\alpha_{m(i)}^{(i)}}\right),$$

will be proofed. The projection matrix $\hat{\boldsymbol{P}}_{-n_{k,j}}^{(i+1)}$ as defined in (A22) projects into

$$\operatorname{null}\left\{\boldsymbol{H}_{\text{comp},-\hat{n}_{k,j}}^{(i-1)}\right\} \cup \operatorname{null}\left\{\boldsymbol{g}_{k(i),d_{k(i)}}^{\text{H}} \boldsymbol{H}_{k(i)}\right\} = \operatorname{null}\left\{\boldsymbol{H}_{\text{comp},-n_{k,j}}^{(i)}\right\}$$

and can therefore be written as

$$\hat{\boldsymbol{P}}_{-n_{k,j}}^{(i+1)} = \hat{\boldsymbol{P}}_{-\hat{n}_{k,j}}^{(i)} - \frac{\hat{\boldsymbol{P}}_{-\hat{n}_{k,j}}^{(i)} \boldsymbol{H}_{k(i)}^{\text{H}} \boldsymbol{g}_{k(i),d_{k(i)}} \boldsymbol{g}_{k(i),d_{k(i)}}^{\text{H}} \boldsymbol{H}_{k(i)} \hat{\boldsymbol{P}}_{-\hat{n}_{k,j}}^{(i)}}{\boldsymbol{g}_{k(i),d_{k(i)}}^{\text{H}} \boldsymbol{H}_{k(i)} \hat{\boldsymbol{P}}_{-\hat{n}_{k,j}}^{(i)} \boldsymbol{H}_{k(i)}^{\text{H}} \boldsymbol{g}_{k(i),d_{k(i)}}}.$$

As explained in Appendix A4, $\boldsymbol{t}_{k,j}^{(i)}$ is collinear with $\hat{\boldsymbol{P}}_{-n_{k,j}}^{(i+1)} \boldsymbol{H}_{\text{comp}}^{(i),\text{H}} \boldsymbol{e}_{n_{k,j}}$ so that

$$\begin{aligned}
\boldsymbol{t}_{k,j}^{(i)} &= \hat{\beta}_j^{(i)} \hat{\boldsymbol{P}}_{-n_{k,j}}^{(i+1)} \boldsymbol{H}_{\text{comp}}^{(i),\text{H}} \boldsymbol{e}_{n_{k,j}} \\
&= \hat{\beta}_{k,j}^{(i)} \left(\hat{\boldsymbol{P}}_{-\hat{n}_{k,j}}^{(i)} - \frac{\hat{\boldsymbol{P}}_{-\hat{n}_{k,j}}^{(i)} \boldsymbol{H}_{k(i)}^{\text{H}} \boldsymbol{g}_{k(i),d_{k(i)}} \boldsymbol{g}_{k(i),d_{k(i)}}^{\text{H}} \boldsymbol{H}_{k(i)} \hat{\boldsymbol{P}}_{-\hat{n}_{k,j}}^{(i)}}{\boldsymbol{g}_{k(i),d_{k(i)}}^{\text{H}} \boldsymbol{H}_{k(i)} \hat{\boldsymbol{P}}_{-\hat{n}_{k,j}}^{(i)} \boldsymbol{H}_{k(i)}^{\text{H}} \boldsymbol{g}_{k(i),d_{k(i)}}} \right) \boldsymbol{H}_{\text{comp}}^{(i),\text{H}} \boldsymbol{e}_{n_{k,j}} \\
&= \tilde{\beta}_{k,j}^{(i)} \left(\boldsymbol{t}_{k,j}^{(i-1)} - \frac{\hat{\boldsymbol{P}}_{-\hat{n}_{k,j}}^{(i)} \boldsymbol{H}_{k(i)}^{\text{H}} \boldsymbol{g}_{k(i),d_{k(i)}} \boldsymbol{g}_{k(i),d_{k(i)}}^{\text{H}} \boldsymbol{H}_{k(i)} \boldsymbol{t}_{k,j}^{(i-1)}}{\boldsymbol{g}_{k(i),d_{k(i)}}^{\text{H}} \boldsymbol{H}_{k(i)} \hat{\boldsymbol{P}}_{-\hat{n}_{k,j}}^{(i)} \boldsymbol{H}_{k(i)}^{\text{H}} \boldsymbol{g}_{k(i),d_{k(i)}}} \right) \\
&= \tilde{\beta}_{k,j}^{(i)} \left(\boldsymbol{t}_{k,j}^{(i-1)} - \frac{\left(\hat{\boldsymbol{P}}_{\text{lin}}^{(i)} + \boldsymbol{t}_{k,j}^{(i-1)} \boldsymbol{t}_{k,j}^{(i-1),\text{H}} \right) \boldsymbol{H}_{k(i)}^{\text{H}} \boldsymbol{g}_{k(i),d_{k(i)}} \boldsymbol{g}_{k(i),d_{k(i)}}^{\text{H}} \boldsymbol{H}_{k(i)} \boldsymbol{t}_{k,j}^{(i-1)}}{\boldsymbol{g}_{k(i),d_{k(i)}}^{\text{H}} \boldsymbol{H}_{k(i)} \left(\hat{\boldsymbol{P}}_{\text{lin}}^{(i)} + \boldsymbol{t}_{k,j}^{(i-1)} \boldsymbol{t}_{k,j}^{(i-1),\text{H}} \right) \boldsymbol{H}_{k(i)}^{\text{H}} \boldsymbol{g}_{k(i),d_{k(i)}}} \right), \\
&= \tilde{\beta}_{k,j}^{(i)} \frac{\boldsymbol{t}_{k,j}^{(i-1)} \boldsymbol{g}_{k(i),d_{k(i)}}^{\text{H}} \boldsymbol{H}_{k(i)} \hat{\boldsymbol{P}}_{\text{lin}}^{(i)} \boldsymbol{H}_{k(i)}^{\text{H}} \boldsymbol{g}_{k(i),d_{k(i)}} - \hat{\boldsymbol{P}}_{\text{lin}}^{(i)} \boldsymbol{H}_{k(i)}^{\text{H}} \boldsymbol{g}_{k(i),d_{k(i)}} \boldsymbol{g}_{k(i),d_{k(i)}}^{\text{H}} \boldsymbol{H}_{k(i)} \boldsymbol{t}_{k,j}^{(i-1)}}{\boldsymbol{g}_{k(i),d_{k(i)}}^{\text{H}} \boldsymbol{H}_{k(i)} \left(\hat{\boldsymbol{P}}_{\text{lin}}^{(i)} + \boldsymbol{t}_{k,j}^{(i-1)} \boldsymbol{t}_{k,j}^{(i-1),\text{H}} \right) \boldsymbol{H}_{k(i)}^{\text{H}} \boldsymbol{g}_{k(i),d_{k(i)}}}
\end{aligned}$$

where (A23) has been used in the last but one line and $\hat{\beta}_{k,j}^{(i)}$ and $\tilde{\beta}_{k,j}^{(i)}$ scale the vectors on the right hand side so that $\boldsymbol{t}_{k,j}^{(i)}$ has unit norm. By using the fact that

$$\hat{\boldsymbol{P}}_{\text{lin}}^{(i)} \boldsymbol{H}_{k(i)}^{\text{H}} \boldsymbol{g}_{k(i),d_{k(i)}} = \sqrt{\boldsymbol{g}_{k(i),d_{k(i)}}^{\text{H}} \boldsymbol{H}_{k(i)} \hat{\boldsymbol{P}}_{\text{lin}}^{(i)} \boldsymbol{H}_{k(i)}^{\text{H}} \boldsymbol{g}_{k(i),d_{k(i)}}}\, \boldsymbol{t}_{k(i),d_{k(i)}}^{(i)}$$

from Equation (3.64) and using the scaling factor

$$\beta_{k,j}^{(i)} = \tilde{\beta}_{k,j}^{(i)} \frac{\boldsymbol{g}_{k(i),d_{k(i)}}^{\text{H}} \boldsymbol{H}_{k(i)} \hat{\boldsymbol{P}}_{\text{lin}}^{(i)} \boldsymbol{H}_{k(i)}^{\text{H}} \boldsymbol{g}_{k(i),d_{k(i)}}}{\boldsymbol{g}_{k(i),d_{k(i)}}^{\text{H}} \boldsymbol{H}_{k(i)} \left(\hat{\boldsymbol{P}}_{\text{lin}}^{(i)} + \boldsymbol{t}_{k,j}^{(i-1)} \boldsymbol{t}_{n_{k,j}}^{(i-1),\text{H}} \right) \boldsymbol{H}_{k(i)}^{\text{H}} \boldsymbol{g}_{k(i),d_{k(i)}}}$$

one obtains

$$\boldsymbol{t}_{k,j}^{(i)} = \beta_{k,j}^{(i)} \left(\boldsymbol{t}_{k,j}^{(i-1)} - \boldsymbol{t}_{k(i),d_{k(i)}}^{(i)} \frac{\boldsymbol{g}_{k(i),d_{k(i)}}^{\text{H}} \boldsymbol{H}_{k(i)} \boldsymbol{t}_{k,j}^{(i-1)}}{\sqrt{\boldsymbol{g}_{k(i),d_{k(i)}}^{\text{H}} \boldsymbol{H}_{k(i)} \hat{\boldsymbol{P}}_{\text{lin}}^{(i)} \boldsymbol{H}_{k(i)}^{\text{H}} \boldsymbol{g}_{k(i),d_{k(i)}}}} \right),$$

which leads to the desired result by using the definition of $\alpha_{m(i)}^{(i)}$ in (3.66) and the definition of $\gamma_{n_{k,j}}^{(i)}$ in (3.67).

A6. Lower and Upper Bounds for a Generalized Eigenvalue Problem

In this section we will derive a lower and an upper bound for the maximum eigenvalue of a matrix $(\boldsymbol{I}_r + \boldsymbol{D})^{-1} \boldsymbol{C}$, which are given by

$$\frac{\text{tr}(\boldsymbol{C})}{r\left[1 + \text{tr}(\boldsymbol{D})\right]} \leq \rho_1 \left((\boldsymbol{I}_r + \boldsymbol{D})^{-1} \boldsymbol{C} \right) \leq \text{tr}\left(\boldsymbol{C}\right)$$

in case $\boldsymbol{C} \in \mathbb{C}^{r \times r}$ and $\boldsymbol{D} \in \mathbb{C}^{r \times r}$ are positive semi-definite Hermitian matrices. For the eigenvalue from Equation (3.71) one obtains $\boldsymbol{C} = \boldsymbol{S}_{c,k} \boldsymbol{A}_k^{(i)} \boldsymbol{S}_{c,k}^{\mathsf{T}}$ and

$$\boldsymbol{D} = \boldsymbol{S}_{c,k} \boldsymbol{B}_k^{(i)} \boldsymbol{S}_{c,k}^{\mathsf{T}} - \boldsymbol{I}_{r_k} = \sum_{j=1}^{i-1} \boldsymbol{S}_{c,k} \boldsymbol{H}_k \hat{\boldsymbol{P}}_j^{(i)} \boldsymbol{H}_k^{\text{H}} \boldsymbol{S}_{c,k}^{\text{H}} \alpha_j^{(i)}.$$

This maximum eigenvalue can be upper bounded by (e.g. [136, Ch.9, Theorem H.1.a])

$$\rho_1\left((\boldsymbol{I}_r+\boldsymbol{D})^{-1}\boldsymbol{C}\right) \leq \rho_1\left((\boldsymbol{I}_r+\boldsymbol{D})^{-1}\right)\rho_1(\boldsymbol{C}).$$

As the minimum eigenvalue of the matrix $\boldsymbol{I}_r+\boldsymbol{D}$ is greater than or equal to one, when $\boldsymbol{D}$ is positive semi-definite, the maximum eigenvalue of its inverse is equal to or smaller than one. Furthermore the maximum eigenvalue of a positive semi-definite matrix $\boldsymbol{C}$ is smaller than or equal to the trace of that matrix which leads to the upper bound

$$\rho_1\left((\boldsymbol{I}_r+\boldsymbol{D})^{-1}\boldsymbol{C}\right) \leq \mathrm{tr}(\boldsymbol{C}).$$

For the derivation of the lower bound the lower bound for the maximum eigenvalue from [35, Ch. 2.3] is used, so that

$$\frac{\mathrm{tr}\left((\boldsymbol{I}_r+\boldsymbol{D})^{-1}\boldsymbol{C}\right)}{r} \leq \rho_1\left((\boldsymbol{I}_r+\boldsymbol{D})^{-1}\boldsymbol{C}\right), \tag{A26}$$

where equality holds, if all eigenvalues are identical. Denoting the eigenvalue decompositions of the matrices $\boldsymbol{C}$ and $(\boldsymbol{I}_r+\boldsymbol{D})^{-1}$ as

$$\boldsymbol{C} = \sum_{i=1}^{\mathrm{rank}(\boldsymbol{C})} \rho_i(\boldsymbol{C})\boldsymbol{u}_i\boldsymbol{u}_i^{\mathrm{H}}, \qquad (\boldsymbol{I}_r+\boldsymbol{D})^{-1} = \sum_{j=1}^{r}\frac{1}{1+\rho_j(\boldsymbol{D})}\boldsymbol{v}_j\boldsymbol{v}_j^{\mathrm{H}},$$

one obtains

$$\mathrm{tr}\left((\boldsymbol{I}_r+\boldsymbol{D})^{-1}\boldsymbol{C}\right) = \sum_{i=1}^{\mathrm{rank}(\boldsymbol{C})}\sum_{j=1}^{r}\frac{\rho_i(\boldsymbol{C})}{1+\rho_j(\boldsymbol{D})}\mathrm{tr}(\boldsymbol{u}_i\boldsymbol{u}_i^{\mathrm{H}}\boldsymbol{v}_j\boldsymbol{v}_j^{\mathrm{H}}) = \sum_{i=1}^{\mathrm{rank}(\boldsymbol{C})}\sum_{j=1}^{r}\frac{\rho_i(\boldsymbol{C})}{1+\rho_j(\boldsymbol{D})}|\boldsymbol{v}_j^{\mathrm{H}}\boldsymbol{u}_i|^2 \tag{A27}$$

As $\boldsymbol{D}$ is positive semi-definite, all its eigenvalues $\rho_j(\boldsymbol{D})$ can be upper bounded by its trace, i.e. $\rho_j(\boldsymbol{D}) \leq \mathrm{tr}(\boldsymbol{D})$. Hence, we the expression in (A27) can be lower bounded as

$$\sum_{i=1}^{\mathrm{rank}(\boldsymbol{C})}\sum_{j=1}^{r}\frac{\rho_i(\boldsymbol{C})}{1+\rho_j(\boldsymbol{D})}|\boldsymbol{v}_j^{\mathrm{H}}\boldsymbol{u}_i|^2 \geq \sum_{i=1}^{\mathrm{rank}(\boldsymbol{C})}\sum_{j=1}^{r}\frac{\rho_i(\boldsymbol{C})}{1+\mathrm{tr}(\boldsymbol{D})}|\boldsymbol{v}_j^{\mathrm{H}}\boldsymbol{u}_i|^2 = \sum_{i=1}^{\mathrm{rank}(\boldsymbol{C})}\frac{\rho_i(\boldsymbol{C})}{1+\mathrm{tr}(\boldsymbol{D})}\sum_{j=1}^{r}|\boldsymbol{v}_j^{\mathrm{H}}\boldsymbol{u}_i|^2.$$

As the vectors $\boldsymbol{v}_j$ form an orthonormal basis, we obtain

$$\sum_{i=1}^{\mathrm{rank}(\boldsymbol{C})}\frac{\rho_i(\boldsymbol{C})}{1+\mathrm{tr}(\boldsymbol{D})}\sum_{j=1}^{r}|\boldsymbol{v}_j^{\mathrm{H}}\boldsymbol{u}_i|^2 = \sum_{i=1}^{\mathrm{rank}(\boldsymbol{C})}\frac{\rho_i(\boldsymbol{C})}{1+\mathrm{tr}(\boldsymbol{D})} = \frac{\mathrm{tr}(\boldsymbol{C})}{1+\mathrm{tr}(\boldsymbol{D})},$$

which leads together with (A26) to the the desired result.

A7. Derivation of the Large System Limit for the Optimum Weighted Sum Rate in MISO Systems with Infinite Number of Transmit Antennas and Users

To derive the large system limit for the optimum weighted sum rate in MISO Systems with infinite number of transmit antennas and users, the Karush-Kuhn-Tucker conditions of Problem (5.7) are considered, which read as

$$\sum_{j=n(k)}^{N} \Delta\mu_j \frac{C}{\ln 2\sigma_{\mathrm{n}}^2} \boldsymbol{H}_k \left(\boldsymbol{I}_{CN_{\mathrm{T}}} + \frac{C}{\sigma_{\mathrm{n}}^2} \left[\sum_{m=1}^{\hat{K}_j} \boldsymbol{H}_{\hat{\pi}(m)}^{\mathrm{H}} \boldsymbol{W}_{\hat{\pi}(m)} \boldsymbol{H}_{\hat{\pi}(m)} \right] \right)^{-1} \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{W}_k = \hat{\lambda} \boldsymbol{W}_k, \quad \forall k,$$
$$\hat{\lambda} \left(\sum_{k=1}^{K} \mathrm{tr}(\boldsymbol{W}_k) - P_{\mathrm{Tx}} \right) = 0, \quad \hat{\lambda} \geq 0, \quad \boldsymbol{W}_k \succeq 0, \forall k, \tag{A28}$$

where $n(k)$ denotes the index of the subgroup user k belongs to and

$$\hat{K}_j = \sum_{n=1}^{j} \beta_n K.$$

As the problem of weighted sum rate maximization in the dual MAC is convex, the KKT are necessary and sufficient and solving (A28) leads to a global optimum. As it has been shown in Section 3.1, the covariance matrices optimum for (5.7) in OFDM systems, automatically fulfill the block-diagonality constraints, which implies for single antenna user terminals that the optimum uplink covariance matrices $\boldsymbol{W}_k \in \mathbb{C}^{C \times C}$ become diagonal matrices, i.e.,

$$\boldsymbol{W}_k = \mathrm{diag}\left(w_{k,1}, \ldots, w_{k,C}\right).$$

Exploiting this fact and the block-diagonal structure of the channel matrices $\boldsymbol{H}_k$, the KKT conditions in (A28) can be written as

$$\sum_{j=n(k)}^{N} \Delta\mu_j \frac{C}{\ln 2\sigma_{\mathrm{n}}^2} \hat{\boldsymbol{h}}_{k,c}^{\mathrm{H}} \left(\boldsymbol{I}_{N_{\mathrm{T}}} + \frac{C}{\sigma_{\mathrm{n}}^2} \left[\sum_{m=1}^{\hat{K}_j} \hat{\boldsymbol{h}}_{\hat{\pi}(m),c} \hat{\boldsymbol{h}}_{\hat{\pi}(m),c}^{\mathrm{H}} w_{\hat{\pi}(m),c} \right] \right)^{-1} \hat{\boldsymbol{h}}_{k,c} w_{k,c} = \hat{\lambda} w_{k,c}, \quad \forall k, c$$
$$\hat{\lambda} \left(\sum_{k=1}^{K} \sum_{c=1}^{C} w_{k,c} - P_{\mathrm{Tx}} \right) = 0, \quad \hat{\lambda} \geq 0, \quad w_{k,c} \geq 0, \forall k, c, \tag{A29}$$

where

$$\boldsymbol{H}_{k,c}^{\mathrm{H}} = \boldsymbol{S}_{c,\mathrm{T}} \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{S}_{c,k}^{\mathsf{T}} =: \hat{\boldsymbol{h}}_{k,c}$$

becomes a column vector in the MISO case. Note that for QoS constrained optimization problems, instead of (A29) the KKT of the corresponding optimization problem need to be considered. They can be analyzed in the large system limit with the same tools as applied in the following to the weighted sum rate maximization problem.

Similarly to [127], the covariance matrices $\boldsymbol{W}_k$ are considered to be deterministic in the following. Under this assumption the KKT conditions (A29) are evaluated in the large system limit. First, the terms

$$\hat{\boldsymbol{h}}_{k,c}^{\mathrm{H}} \left(\boldsymbol{I}_{N_{\mathrm{T}}} + \frac{C}{\sigma_{\mathrm{n}}^2} \left[\sum_{m=1}^{\hat{K}_j} \hat{\boldsymbol{h}}_{\hat{\pi}(m),c} \hat{\boldsymbol{h}}_{\hat{\pi}(m),c}^{\mathrm{H}} w_{\hat{\pi}(m),c} \right] \right)^{-1} \hat{\boldsymbol{h}}_{k,c}$$

are evaluated in the large system limit for $K \to \infty$ and $N_{\mathrm{T}} \to \infty$ using standard methods from the large system analysis (e.g. [112], [110]). First the matrix inversion lemma is applied so that the inverse matrices are independent of the vector $\hat{\boldsymbol{h}}_{k,c}$ and

$$\hat{\boldsymbol{h}}_{k,c}^{\mathrm{H}}\left(\boldsymbol{I}_{N_\mathrm{T}}+\frac{C}{\sigma_\mathrm{n}^2}\left[\sum_{m=1}^{\hat{K}_j}\hat{\boldsymbol{h}}_{\hat{\pi}(m),c}\hat{\boldsymbol{h}}_{\hat{\pi}(m),c}^{\mathrm{H}}w_{\hat{\pi}(m),c}\right]\right)^{-1}\hat{\boldsymbol{h}}_{k,c}=\frac{m_{k,c,j}}{1+w_{k,c}\frac{C}{\sigma_\mathrm{n}^2}m_{k,c,j}}, \tag{A30}$$

where

$$m_{k,c,j}=\hat{\boldsymbol{h}}_{k,c}^{\mathrm{H}}\left(\boldsymbol{I}_{N_\mathrm{T}}+\frac{C}{\sigma_\mathrm{n}^2}\boldsymbol{B}_{k,c,j}\right)^{-1}\hat{\boldsymbol{h}}_{k,c},$$

and

$$\boldsymbol{B}_{k,c,j}=\sum_{m=1}^{\hat{K}_j}\hat{\boldsymbol{h}}_{\hat{\pi}(m),c}\hat{\boldsymbol{h}}_{\hat{\pi}(m),c}^{\mathrm{H}}w_{\hat{\pi}(m),c}-\hat{\boldsymbol{h}}_{k,c}\hat{\boldsymbol{h}}_{k,c}^{\mathrm{H}}w_{k,c}=\hat{\boldsymbol{H}}_{k,c,j}\boldsymbol{W}_{k,c,j}\hat{\boldsymbol{H}}_{k,c,j}^{\mathrm{H}}\sigma_\mathrm{H}^2N_\mathrm{T}. \tag{A31}$$

The matrix $\hat{\boldsymbol{H}}_{k,c,j}\in\mathbb{C}^{N_\mathrm{T}\times\hat{K}_j-1}$ contains the vectors $\frac{1}{\sqrt{N_\mathrm{T}}\sigma_\mathrm{H}}\hat{\boldsymbol{h}}_{\hat{\pi}(m),c}$, $m=1,\ldots,\hat{K}_j$, $\hat{\pi}(m)\neq k$, as columns, where the normalization by $\frac{1}{\sqrt{N_\mathrm{T}}\sigma_\mathrm{H}}$ is done so that the entries in $\hat{\boldsymbol{H}}_{k,c,j}$ are i.i.d. with zero mean and variance $\frac{1}{N_\mathrm{T}}$, a property that will be required in the following. The matrix $\boldsymbol{W}_{k,c,j}\in\mathbb{C}^{\hat{K}_j-1\times\hat{K}_j-1}$ is diagonal and contains the powers $w_{\hat{\pi}(m),c}$, $m=1,\ldots,\hat{K}_j$, $\hat{\pi}(m)\neq k$. The matrix $\left(\boldsymbol{I}_{N_\mathrm{T}}+\frac{C}{\sigma_\mathrm{n}^2}\boldsymbol{B}_{k,c,j}\right)^{-1}$ is independent of $\hat{\boldsymbol{h}}_{k,c}$, which is due to the fact that the powers $w_{k,c}$ are assumed to be deterministic and the vectors $\hat{\boldsymbol{h}}_{\hat{\pi}(m),c}$ belong to other users than user k. As furthermore the vector $\hat{\boldsymbol{h}}_{k,c}$ contains i.i.d. entries, Lemma 2.7 from [137] can be applied, so that $m_{k,c,j}$ converges for $N_\mathrm{T}\to\infty$ according to

$$m_{k,c,j}\underset{N_\mathrm{T}\to\infty}{\longrightarrow}\sigma_\mathrm{H}^2\,\mathrm{tr}\left(\boldsymbol{I}_{N_\mathrm{T}}+\frac{C}{\sigma_\mathrm{n}^2}\boldsymbol{B}_{k,c,j}\right)^{-1},$$

where the factor σ_H^2 takes into account that the elements in $\hat{\boldsymbol{h}}_{k,c}$ do not have variance 1 as required in [137, Lemma 2.7]. Given the asymptotic eigenvalue distribution $f_{\boldsymbol{B}_{k,c,j}}^{(\infty)}(x)$ of the matrix $\boldsymbol{B}_{k,c,j}$, the trace of $\left(\boldsymbol{I}_{N_\mathrm{T}}+\frac{C}{\sigma_\mathrm{n}^2}\boldsymbol{B}_{k,c,j}\right)^{-1}$ can be computed according to

$$\begin{aligned}\mathrm{tr}\left(\boldsymbol{I}_{N_\mathrm{T}}+\frac{C}{\sigma_\mathrm{n}^2}\boldsymbol{B}_{k,c,j}\right)^{-1}&=\sum_{i=1}^{N_\mathrm{T}}\frac{1}{1+\frac{C}{\sigma_\mathrm{n}^2}\rho_i\left(\boldsymbol{B}_{k,c,j}\right)}\\&\underset{\substack{N_\mathrm{T},k\to\infty,\\ \frac{N_\mathrm{T}}{K}=\alpha}}{\longrightarrow}N_\mathrm{T}\int_0^\infty\frac{1}{1+\frac{C}{\sigma_\mathrm{n}^2}x}f_{\boldsymbol{B}_{k,c,j}}^{(\infty)}(x)\,\mathrm{d}\,x=N_\mathrm{T}\eta_{\boldsymbol{B}_{k,c,j}}\left(\frac{C}{\sigma_\mathrm{n}^2}\right)=\frac{m_{k,c,j}}{\sigma_\mathrm{H}^2}\end{aligned} \tag{A32}$$

[c.f. (5.5)]. Thus, $m_{k,c,j}$ is a function of the η-transform $\eta_{\boldsymbol{B}_{k,c,j}}(\gamma)$ of the matrix $\boldsymbol{B}_{k,c,j}$. Exploiting the structure of the matrix $\boldsymbol{B}_{k,c,j}$ as defined in (A31) and the fact that the matrices $\hat{\boldsymbol{H}}_{k,c,j}$ contain i.i.d. entries with zero mean and variance $\frac{1}{N_\mathrm{T}}$, Theorem 2.39 from [109] can be applied to obtain the η-transform $\eta_{\boldsymbol{B}_{k,c,j}}\left(\frac{C}{\sigma_\mathrm{n}^2}\right)$ and thus $m_{k,c,j}$ from the implicit equation

$$\sum_{n=1}^{j}\beta_n\frac{1}{\alpha}=\frac{1-\frac{m_{k,c,j}}{N_\mathrm{T}\sigma_\mathrm{H}^2}}{1-\frac{1}{\hat{K}_j-1}\sum\limits_{\substack{m=1\\ \hat{\pi}(m)\neq k}}^{\hat{K}_j}\frac{1}{1+m_{k,c,j}w_{\hat{\pi}(m),c}\frac{C}{\sigma_\mathrm{n}^2}}}. \tag{A33}$$

Furthermore in (A33) the facts have been used that for for large N_T and K

$$\frac{\hat{K}_j - 1}{N_\mathrm{T}} = \frac{\sum_{n=1}^{j} \beta_n K - 1}{N_\mathrm{T}} \underset{\substack{N_\mathrm{T}, K \to \infty \\ \frac{N_\mathrm{T}}{K} = \alpha}}{\longrightarrow} \sum_{n=1}^{j} \beta_n \frac{1}{\alpha}$$

and the η-transform of the diagonal matrix $\boldsymbol{W}_{k,c,j} \sigma_\mathrm{H}^2 N_\mathrm{T}$ is given by

$$\eta_{\boldsymbol{W}_{k,c,j}\sigma_\mathrm{H}^2 N_\mathrm{T}}(\gamma) = \frac{1}{\hat{K}_j - 1} \sum_{\substack{m=1 \\ \hat{\pi(m)} \neq k}}^{\hat{K}_j} \frac{1}{1 + \gamma w_{\hat{\pi}(m),c} \sigma_\mathrm{H}^2 N_\mathrm{T}}.$$

Using (A30) and adding the implicit conditions (A33) to the KKT conditions from (A29) leads to the following system of equations to determine the optimum powers $w_{k,c}$

$$\begin{aligned}
&\sum_{j=n(k)}^{N} \Delta\mu_j \frac{C}{\ln 2\sigma_\mathrm{n}^2} \frac{m_{k,c,j}}{1 + w_{k,c} m_{k,c,j} \frac{C}{\sigma_\mathrm{n}^2}} w_{k,c} = \hat{\lambda} w_{k,c}, \quad \forall k, c \\
&\hat{\lambda} \left(\sum_{k=1}^{K} \sum_{c=1}^{C} w_{k,c} - P_\mathrm{Tx} \right) = 0, \quad \hat{\lambda} \geq 0, \quad w_{k,c} \geq 0, \forall k, c, \\
\sum_{n=1}^{j} \beta_n \frac{1}{\alpha} &= \frac{1 - \frac{m_{k,c,j}}{N_\mathrm{T}\sigma_\mathrm{H}^2}}{1 - \frac{1}{\hat{K}_j - 1} \left(\sum_{m=1}^{\hat{K}_j} \frac{1}{1 + m_{k,c,j} w_{\hat{\pi}(m),c} \frac{C}{\sigma_\mathrm{n}^2}} - \frac{1}{1 + m_{k,c,j} w_{k,c} \frac{C}{\sigma_\mathrm{n}^2}} \right)}, \quad \forall j = n(k), \ldots, N, \forall k, c.
\end{aligned} \tag{A34}$$

The first and the last line in (A34) are identical for all carriers c and for all users, which belong to the same subgroup of users. Thus, the variables $m_{k,c,j}$ are therefore identical for all carriers and all users in the same subgroup and those users receive the same power, which is in turn equally distributed over the carriers of each user. Denoting w_n the total power allocated to the nth subgroup of users, the $w_{k,c}$ are therefore given by

$$w_{k,c} = \frac{w_n}{C\beta_n K}.$$

For the special case of sum rate maximization, i.e., when all users' weights are equal, this implies that an equal power allocation $w_{k,c} = \frac{P_\mathrm{Tx}}{CK}$ is optimum in the dual uplink. For unequal weights, by using the new variables

$$w_n = w_{k,c} C \beta_n K, \qquad m_n = \frac{m_{k,c,j}}{N_\mathrm{T}}, \forall k, c \tag{A35}$$

and $\lambda = \frac{\hat{\lambda}\ln 2\sigma_{\rm n}^2}{N_{\rm T}C}$, the system of equations (A34) can be reduced to

$$\sum_{j=n}^{N}\Delta\mu_j \frac{m_j}{1+w_n m_j \frac{\alpha}{\sigma_{\rm n}^2\beta_n}} w_n = \lambda w_n, \quad n=1,\ldots,N$$
$$\lambda\left(\sum_{n=1}^{N} w_n - P_{\rm Tx}\right)=0, \quad \lambda\geq 0, \quad w_n\geq 0, n=1,\ldots,N$$
$$\sum_{n'=1}^{n}\beta_{n'}\frac{1}{\alpha} = \frac{1-\frac{m_n}{\sigma_{\rm H}^2}}{1-\left(\sum_{m=1}^{n}\frac{\beta_m}{\sum_{n'=1}^{n}\beta_{n'}}\frac{1}{1+m_n w_m\frac{\alpha}{\sigma_{\rm n}^2\beta_m}}\right)} \quad n=1,\ldots,N,$$

which is identical to the implicit systems of equations (5.9) and where in the last line the asymptotic equivalence

$$\frac{1}{\hat{K}_j-1}\left(\sum_{m=1}^{\hat{K}_j}\frac{1}{1+m_j w_m\frac{\alpha}{\sigma_{\rm n}^2}\beta_m} - \frac{1}{1+m_j w_k\frac{\alpha}{\sigma_{\rm n}^2\beta_k}}\right)$$
$$\underset{K\to\infty}{\longrightarrow}\frac{1}{\hat{K}_j}\sum_{m=1}^{\hat{K}_j}\frac{1}{1+m_j w_m\frac{\alpha}{\sigma_{\rm n}^2\beta_m}} = \sum_{m=1}^{j}\frac{\beta_m}{\sum_{n'=1}^{j}\beta_{n'}}\frac{1}{1+m_j w_m\frac{\alpha}{\sigma_{\rm n}^2\beta_m}}$$

has been used. As (5.9) stems from the KKT of a concave optimization problem, the system of equations has a unique solution. Proceeding similarly to the KKT expression in (A29), the weighted sum rate expression in (5.7) can be written as

$$R_{\text{WSR,opt}} = \sum_{n=1}^{N}\Delta\mu_n\sum_{c=1}^{C}\log_2\left|\boldsymbol{I}_{N_{\rm T}}+\frac{C}{\sigma_{\rm n}^2}\left[\sum_{m=1}^{\hat{K}_n}\hat{\boldsymbol{h}}_{\hat{\pi}(m),c}\hat{\boldsymbol{h}}_{\hat{\pi}(m),c}^{\rm H}w_{\hat{\pi}(m),c}\right]\right|$$
$$= \sum_{n=1}^{N}\Delta\mu_n\sum_{c=1}^{C}\log_2\left|\boldsymbol{I}_{N_{\rm T}}+\frac{C}{\sigma_{\rm n}^2}\hat{\boldsymbol{B}}_{c,n}\right|,$$

where

$$\hat{\boldsymbol{B}}_{c,n} = \sum_{m=1}^{\hat{K}_n}\hat{\boldsymbol{h}}_{\hat{\pi}(m),c}\hat{\boldsymbol{h}}_{\hat{\pi}(m),c}^{\rm H}w_{\hat{\pi}(m),c} = \boldsymbol{B}_{\hat{\pi}(1),c,n}+\hat{\boldsymbol{h}}_{\hat{\pi}(1),c}\hat{\boldsymbol{h}}_{\hat{\pi}(1),c}^{\rm H}w_{\hat{\pi}(1),c} \tag{A36}$$

and $\boldsymbol{B}_{\hat{\pi}(1),c,n}$ is given by (A31). The large system sum rate can then be computed via the Shannon transform [c.f. (5.6)] of the matrix $\hat{\boldsymbol{B}}_{c,n}$ according to

$$R_{\text{WSR,opt}} \underset{\substack{N_{\rm T},K\to\infty\\ \frac{N_{\rm T}}{K}=\alpha}}{\longrightarrow}\sum_{n=1}^{N}\Delta\mu_n\sum_{c=1}^{C}N_{\rm T}\mathcal{V}_{\hat{\boldsymbol{B}}_{c,n}}\left(\frac{C}{\sigma_{\rm n}^2}\right),$$

From [109, Theorem 2.39] $\mathcal{V}_{\hat{\boldsymbol{B}}_{c,n}}(\gamma)$ is given by

$$\mathcal{V}_{\hat{\boldsymbol{B}}_{c,n}}(\gamma) = \frac{\hat{K}_n}{N_{\rm T}}\mathcal{V}_{\hat{\boldsymbol{W}}_{c,n}\sigma_{\rm H}^2N_{\rm T}}\left(\eta_{\hat{\boldsymbol{B}}_{c,n}}(\gamma)\gamma\right)-$$
$$-\log_2\left(\eta_{\hat{\boldsymbol{B}}_{c,n}}(\gamma)\right)+\left(\eta_{\hat{\boldsymbol{B}}_{c,n}}(\gamma)-1\right)\log_2(\mathrm{e}), \tag{A37}$$

where e is Euler's number and

$$\hat{\boldsymbol{W}}_{c,n} = \mathrm{diag}\left(w_{\hat{\pi}(1),c}, \ldots, w_{\hat{\pi}(n),c}\right).$$

For $K \to \infty$, the term $\hat{\boldsymbol{h}}_{\hat{\pi}(1),c}\hat{\boldsymbol{h}}^{\mathrm{H}}_{\hat{\pi}(1),c} w_{\hat{\pi}(1),c}$ can be neglected compared to the matrix $\boldsymbol{B}_{\hat{\pi}(1),c,n}$ in (A36), as the latter matrix consists of an infinite sum of rank one matrices and $\boldsymbol{w}_{\hat{\pi}(1),c} = \frac{w_{\hat{\pi}(1)}}{C\beta_1 K}$ converges to zero for infinite K. Thus,

$$\eta_{\hat{\boldsymbol{B}}_{c,n}}(\gamma) = \eta_{\boldsymbol{B}_{\hat{\pi}(1),c,n}}(\gamma)$$

in the large system limit and $\eta_{\boldsymbol{B}_{\hat{\pi}(1),c,n}}(\gamma)$ can be computed from the implicit equations (A32) and (A33). Together with (A35) one obtains

$$\eta_{\hat{\boldsymbol{B}}_{c,n}}\left(\frac{C}{\sigma_{\mathrm{n}}^2}\right) = \eta_{\boldsymbol{B}_{\hat{\pi}(1),c,n}}\left(\frac{C}{\sigma_{\mathrm{n}}^2}\right) = \frac{m_{\pi(1),c,n}}{\sigma_{\mathrm{H}}^2 N_{\mathrm{T}}} = \frac{m_n}{\sigma_{\mathrm{H}}^2}. \tag{A38}$$

The Shannon transform $\mathcal{V}_{\hat{\boldsymbol{W}}_{c,n}\sigma_{\mathrm{H}}^2 N_{\mathrm{T}}}(\gamma)$ of the diagonal matrix $\hat{\boldsymbol{W}}_{c,n}\sigma_{\mathrm{H}}^2 N_{\mathrm{T}}$ reads as

$$\mathcal{V}_{\hat{\boldsymbol{W}}_{c,n}\sigma_{\mathrm{H}}^2 N_{\mathrm{T}}}(\gamma) = \frac{1}{\hat{K}_n}\sum_{j=1}^{\hat{K}_n} \log_2\left(1 + \gamma N_{\mathrm{T}}\sigma_{\mathrm{H}}^2 w_{\hat{\pi}(j),c}\right) \underset{K\to\infty}{\longrightarrow} \frac{1}{\hat{K}_n}\sum_{j=1}^{n} \beta_j K \log_2\left(1 + \gamma w_j \frac{\sigma_{\mathrm{H}}^2 \alpha}{\beta_j C}\right) \tag{A39}$$

Inserting (A38) and (A39) into (A37), leads to the asymptotic weighted sum rate

$$\frac{R_{\text{WSR,opt}}}{C N_{\mathrm{T}}} \underset{\substack{N_{\mathrm{T}},K\to\infty,\\ \frac{N_{\mathrm{T}}}{K}=\alpha}}{\longrightarrow} \sum_{n=1}^{N} \Delta\mu_n \left[\frac{1}{\alpha}\sum_{j=1}^{n}\beta_j \log_2\left(1 + \frac{\alpha}{\sigma_{\mathrm{n}}^2\beta_j} w_j m_n\right) + \right.$$
$$\left. + \log_2\left(\frac{\sigma_{\mathrm{H}}^2}{m_n}\right) + \left(\frac{m_n}{\sigma_{\mathrm{H}}^2} - 1\right)\log_2(\mathrm{e})\right],$$

which is identical to (5.8).

A8. Derivation of the Large System Lower Bound of the Weighted Sum Rate in MISO Systems Achievable with Successive Resource Allocation and Zero-Forcing with DPC

The channel gain of the data stream allocated in the ith step is given by

$$\rho_1\left(\hat{\boldsymbol{P}}^{(i)}_{\text{DPC}}\boldsymbol{H}^{\mathrm{H}}_{k(i)}\boldsymbol{H}_{k(i)}\hat{\boldsymbol{P}}^{(i)}_{\text{DPC}}\right) = \rho_1\left(\boldsymbol{H}_{k(i)}\hat{\boldsymbol{P}}^{(i)}_{\text{DPC}}\boldsymbol{H}^{\mathrm{H}}_{k(i)}\right) = \max_c \hat{\boldsymbol{h}}^{\mathrm{H}}_{k(i),c}\boldsymbol{S}_{c,\mathrm{T}}\hat{\boldsymbol{P}}^{(i)}_{\text{DPC}}\boldsymbol{S}^{\mathsf{T}}_{c,\mathrm{T}}\hat{\boldsymbol{h}}_{k(i),c}$$

[c.f. (3.35)], where the last equality follows from the block-diagonal structure of the matrices $\boldsymbol{H}_k$ and $\hat{\boldsymbol{P}}^{(i)}_{\text{DPC}}$ and the fact that in MISO systems the matrix $\boldsymbol{H}_{k(i)}\hat{\boldsymbol{P}}^{(i)}_{\text{DPC}}\boldsymbol{H}^{\mathrm{H}}_{k(i)}$ is diagonal.

$$\hat{\boldsymbol{h}}_{k(i),c} = \boldsymbol{H}^{\mathrm{H}}_{k(i),c} = \boldsymbol{S}_{c,\mathrm{T}}\boldsymbol{H}^{\mathrm{H}}_{k(i)}\boldsymbol{S}^{\mathsf{T}}_{c,k(i)}$$

denotes the Hermitian channel vector of user $k(i)$ on carrier c and contains circularly symmetric i.i.d. entries with zero mean and variance σ_{H}^2. Recursively applying (3.34), the projectors $\hat{\boldsymbol{P}}^{(i)}_{\text{DPC}}$ are given by

$$\hat{\boldsymbol{P}}^{(i)}_{\text{DPC}} = \boldsymbol{I}_{CN_{\mathrm{T}}} - \sum_{j=1}^{i-1}\boldsymbol{t}^{(j)}_{k(j)}\boldsymbol{t}^{(j),\mathrm{H}}_{k(j)}.$$

By furthermore exploiting the block-diagonal structure of the precoding vectors $\boldsymbol{t}_{k(j)}^{(j)}$, the matrices $\boldsymbol{S}_{c,\mathrm{T}}\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}\boldsymbol{S}_{c,\mathrm{T}}^{\mathsf{T}}$ can be written as

$$\boldsymbol{S}_{c,\mathrm{T}}\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}\boldsymbol{S}_{c,\mathrm{T}}^{\mathsf{T}} = \boldsymbol{I}_{N_\mathrm{T}} - \sum_{j \left| c_{k(j)}^{(j)}=c \right.} \boldsymbol{S}_{c,\mathrm{T}}\boldsymbol{t}_{k(j)}^{(j)}\boldsymbol{t}_{k(j)}^{(j),\mathrm{H}}\boldsymbol{S}_{c,\mathrm{T}}^{\mathsf{T}},$$

where the sum is taken over all precoding vectors on the same carrier, i.e., over all j with $c_{k(j)}^{(j)} = c$. To obtain an analytical expression for the large system weighted sum rate, in the following it will be assumed that the vectors $\boldsymbol{t}_{k(j)}^{(j)}$ are independent of the vector $\hat{\boldsymbol{h}}_{k(i),c}$. For the originally proposed algorithm this is not the case, as the user $k(i)$ is chosen such that the weighted sum rate becomes maximum, which implies that this user's channel gain in the nullspace of the vectors $\boldsymbol{t}_{k(j)}^{(j)}$ should be as large as possible. The assumption is however justified, when in each step a subchannel is allocated randomly to a user within the same subgroup of users having equal weights and only the order in which the subgroups can allocate subchannels is optimized. The large system weighted sum rate to be presented in the following therefore is a lower bound for the weighted sum rate achievable with the algorithm. With the assumption of $\boldsymbol{t}_{k(j)}^{(j)}$ being independent of $\hat{\boldsymbol{h}}_{k(i),c}$, Lemma 2.7 from [137] can be applied and the channel gains in the large system limit can be computed according to

$$\begin{aligned}\max_c \hat{\boldsymbol{h}}_{k(i),c}^{\mathrm{H}}\boldsymbol{S}_{c,\mathrm{T}}\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}\boldsymbol{S}_{c,\mathrm{T}}^{\mathsf{T}}\hat{\boldsymbol{h}}_{k(i),c} &\underset{N_\mathrm{T}\to\infty}{\longrightarrow} \max_c \sigma_\mathrm{H}^2 \operatorname{tr}\left(\boldsymbol{S}_{c,\mathrm{T}}\hat{\boldsymbol{P}}_{\mathrm{DPC}}^{(i)}\boldsymbol{S}_{c,\mathrm{T}}^{\mathsf{T}}\right) \\ &= \sigma_\mathrm{H}^2\left(N_\mathrm{T} - \min_c \sum_{j \left| c_{k(j)}^{(j)}=c \right.} \boldsymbol{t}_{k(j)}^{(j),\mathrm{H}}\boldsymbol{S}_{c,\mathrm{T}}^{\mathsf{T}}\boldsymbol{S}_{c,\mathrm{T}}\boldsymbol{t}_{k(j)}^{(j)}\right) = \sigma_\mathrm{H}^2\left(N_\mathrm{T} - \min_c n_c^{(i-1)}\right), \quad \text{(A40)}\end{aligned}$$

where $n_c^{(i-1)}$ denotes the number of subchannels allocated on carrier c during steps $1,\ldots,i-1$ and the last equality follows from the fact that the vectors $\boldsymbol{t}_{k(j)}^{(j)}$ are orthonormal and obey to the carrier separation constraint (3.24). The large system channel gains (A40) are independent of the user index $k(i)$ and monotonically decreasing in $\min_c n_c^{(i)}$. It is therefore asymptotically optimum to schedule first all users in this subgroup having the largest weight, then the users in the subgroup having the second largest weight and so on. The allocation stops, if N_T subchannels have been allocated on each carrier or all users have been served. Note that through the zero-forcing constraints in case $\alpha < 1$, i.e., $K > N_\mathrm{T}$ and for equal weights for all users, not all users are served, whereas with the optimum algorithm all users receive non-zero power (see Appendix A7). Furthermore the same scheduling is applied on all carriers and and the same weighted sum rates can be achieved on all carriers, as $n_c^{(i)}$ only depends on the number of allocated subchannels on carrier c. That implies that on each carrier a total power of $\frac{P_\mathrm{Tx}}{C}$ is used. The large system limit of the weighted sum rate can therefore be computed according to

$$R_{\mathrm{WSR},DPC} \underset{K,N_\mathrm{T}\to\infty}{\longrightarrow} C \sum_{j=1}^{\min(N_\mathrm{T},K)} \mu_{\tilde{\pi}(j)} \log_2\left(1 + p_{\pi(j)}\frac{C}{\sigma_\mathrm{n}^2}N_\mathrm{T}\sigma_\mathrm{H}^2\left(1 - \frac{j}{N_\mathrm{T}}\right)\right). \quad \text{(A41)}$$

$\tilde{\pi}(j)$ denotes the user to be encoded at jth place, where the encoding order is done in decreasing order of weights and

$$p_{\tilde{\pi}(j)} = \left[\hat{\eta}\mu_{\tilde{\pi}(j)} - \frac{\sigma_{\mathrm{n}}^2}{C N_{\mathrm{T}} \sigma_{\mathrm{H}}^2 \left(1 - \frac{j}{N_{\mathrm{T}}}\right)} \right]^+$$

is the power allocated to user $\tilde{\pi}(j)$ determined from weighted water-filling (see Appendix A2), which implies that $\hat{\eta}$ is chosen to fulfill the transmit power constraint with equality. As the channel gains are monotonically decreasing with the index j and the users are served in decreasing order of weights, there exists an encoding position $j_{\max}$ below which all users receive non-zero power by the water-filling algorithm and above which all users receive zero power. Thus, (A41) can be written as

$$\begin{aligned} R_{\mathrm{WSR},DPC} \underset{\substack{K,N_{\mathrm{T}}\to\infty \\ \frac{N_{\mathrm{T}}}{K}=\alpha}}{\longrightarrow} & C \sum_{j=1}^{j_{\max}} \mu_{\tilde{\pi}(j)} \log_2 \left(\eta \mu_{\tilde{\pi}(j)} \left(1 - \frac{j}{N_{\mathrm{T}}}\right)\right) = \\ &= C \left[\sum_{j=1}^{j_{\max}} \mu_{\tilde{\pi}(j)} \log_2(\eta\mu_{\tilde{\pi}(j)}) + \sum_{j=1}^{j_{\max}} \mu_{\tilde{\pi}(j)} \log_2 \left(1 - \frac{j}{N_{\mathrm{T}}}\right) \right] = \\ &= C \left[K \sum_{n=1}^{n_{\max}-1} \beta_n \mu_n \log_2(\eta\mu_n) + K \left(\frac{j_{\max}}{K} - \sum_{n=1}^{n_{\max}-1} \beta_n \right) \mu_{n_{\max}} \log_2(\eta \mu_{n_{\max}}) + \right. \\ & \left. + \sum_{n=1}^{n_{\max}-1} \mu_n \sum_{j=\sum_{n'=1}^{n-1}\beta_{n'}K+1}^{\sum_{n'=1}^{n}\beta_{n'}K} \log_2\left(1 - \frac{j}{N_{\mathrm{T}}}\right) + \sum_{j=\sum_{n'=1}^{n_{\max}-1}\beta_{n'}K+1}^{k_{\max}} \mu_{n_{\max}} \log_2 \left(1 - \frac{j}{N_{\mathrm{T}}}\right) \right] \end{aligned} \quad \text{(A42)}$$

with $n_{\max}$ denoting the index of the last subgroup with users receiving non-zero power and

$$\eta = \hat{\eta}\sigma_{\mathrm{H}}^2 N_{\mathrm{T}} \frac{C}{\sigma_{\mathrm{n}}^2} = \sigma_{\mathrm{H}}^2 \frac{N_{\mathrm{T}}}{K} \frac{\frac{P_{\mathrm{Tx}}}{\sigma_{\mathrm{n}}^2} + \sum_{j=1}^{j_{\max}} \frac{1}{N_{\mathrm{T}}\sigma_{\mathrm{H}}^2\left(1-\frac{j}{N_{\mathrm{T}}}\right)}}{\sum_{j=1}^{j_{\max}} \mu_{\tilde{\pi}(j)} \frac{1}{K}} = \sigma_{\mathrm{H}}^2 \alpha \frac{\frac{P_{\mathrm{Tx}}}{\sigma_{\mathrm{n}}^2} + \sum_{j=1}^{j_{\max}} \frac{1}{N_{\mathrm{T}}\sigma_{\mathrm{H}}^2\left(1-\frac{j}{N_{\mathrm{T}}}\right)}}{\sum_{n=1}^{n_{\max}-1} \beta_n \mu_n + \left(\frac{j_{\max}}{K} - \sum_{n=1}^{n_{\max}-1} \beta_n\right) \mu_{n_{\max}}} \quad \text{(A43)}$$

$j_{\max}$ is chosen so that

$$\eta \mu_{n_{\max}} \geq \frac{1}{1 - \frac{j_{\max}}{N_{\mathrm{T}}}}, \qquad \text{and} \qquad \frac{j_{\max}}{N_{\mathrm{T}}} \leq \min\left(1, \frac{1}{\alpha}\right).$$

For QoS constrained optimization problems the same asymptotic channel gains $N_{\mathrm{T}}\sigma_{\mathrm{H}}^2 \left(1 - \frac{j_{\max}}{N_{\mathrm{T}}}\right)$ can be achieved but a different user and power allocation has to be considered.

The expressions in (A42) and (A43) still depend on infinite sums of functions $g\left(\frac{j}{N_{\mathrm{T}}}\right)$ over j. In the large system limit these sums can be replaced by integrals over finite intervals according to

$$\sum_{j=a}^{b} g\left(\frac{j}{N_{\mathrm{T}}}\right) \longrightarrow \int_a^b g\left(\frac{j}{N_{\mathrm{T}}}\right) \mathrm{d}\, j \longrightarrow N_{\mathrm{T}} \int_{a/N_{\mathrm{T}}}^{b/N_{\mathrm{T}}} g\left(\rho\right) \mathrm{d}\, \rho.$$

Consequently, the asymptotic weighted sum rate at least achievable with Algorithm 3.2 is given by

$$\frac{R_{\mathrm{WSR},DPC}}{N_{\mathrm{T}}C} \underset{\substack{K,N_{\mathrm{T}}\to\infty \\ \frac{N_{\mathrm{T}}}{K}=\alpha}}{\longrightarrow} \frac{1}{\alpha}\sum_{n=1}^{n_{\max}-1}\beta_n\mu_n\log_2(\eta\mu_n)+\left(\frac{j_{\max}}{N_{\mathrm{T}}}-\frac{1}{\alpha}\sum_{n=1}^{n_{\max}-1}\beta_n\right)\mu_{n_{\max}}\log_2(\eta\mu_{n_{\max}})+ \\ +\sum_{n=1}^{n_{\max}-1}\mu_n\int\limits_{1-\hat{\beta}_{n-1}}^{1-\hat{\beta}_n}\log_2(1-\rho)\,\mathrm{d}\,\rho+\mu_{n_{\max}}\int\limits_{1-\hat{\beta}_{n_{\max}-1}}^{\frac{j_{\max}}{N_{\mathrm{T}}}}\log_2(1-\rho)\,\mathrm{d}\,\rho,$$

where $\hat{\beta}_n = 1-\sum\limits_{n'=1}^{n}\frac{\beta_{n'}}{\alpha}$ and with the asymptotic water-level

$$\eta=\sigma_{\mathrm{H}}^2\alpha\frac{\frac{P_{\mathrm{Tx}}}{\sigma_{\mathrm{n}}^2}+\int\limits_0^{\frac{j_{\max}}{N_{\mathrm{T}}}}\frac{1}{\sigma_{\mathrm{H}}^2(1-\rho)}\,\mathrm{d}\,\rho}{\sum\limits_{n=1}^{n_{\max}-1}\beta_n\mu_n+\left(\frac{j_{\max}}{K}-\sum\limits_{n=1}^{n_{\max}-1}\beta_n\right)\mu_{n_{\max}}}.$$

Using the results of the integrals

$$\int\limits_{1-\hat{\beta}_{n-1}}^{1-\hat{\beta}_n}\log_2(1-\rho)\,\mathrm{d}\,\rho=\hat{\beta}_{n-1}\log_2\left(\hat{\beta}_{n-1}\right)-\hat{\beta}_n\log_2\left(\hat{\beta}_n\right)+\frac{1}{\ln 2}\left(\hat{\beta}_n-\hat{\beta}_{n-1}\right),$$

$$\int\limits_{1-\hat{\beta}_{n_{\max}-1}}^{\frac{j_{\max}}{N_{\mathrm{T}}}}\log_2(1-\rho)\,\mathrm{d}\,\rho=-\left(1-\frac{j_{\max}}{N_{\mathrm{T}}}\right)\log_2\left(1-\frac{j_{\max}}{N_{\mathrm{T}}}\right)-\frac{j_{\max}}{N_{\mathrm{T}}\ln 2}+\hat{\beta}_{n_{\max}-1}\log_2\left(\hat{\beta}_{n_{\max}-1}\right)+ \\ +\frac{1-\hat{\beta}_{n_{\max}-1}}{\ln 2},$$

$$\int\limits_0^{\frac{j_{\max}}{N_{\mathrm{T}}}}\frac{1}{\sigma_{\mathrm{H}}^2(1-\rho)}\,\mathrm{d}\,\rho=-\frac{1}{\sigma_{\mathrm{H}}^2}\ln\left(1-\frac{j_{\max}}{N_{\mathrm{T}}}\right)$$

leads to the system of equations (5.10).

A9. Weighted Sum Rate Computation in the Large System Limit for Spatial Zero-Forcing in MISO Systems

In this appendix a method to solve Problem (5.13) will be presented. For this purpose it is first shown that the function

$$\hat{R}^{(\infty)}_{\text{WSR,lin}}(\alpha,\rho) = \frac{1}{\alpha}\left[\sum_{n=1}^{n(\rho)-1}\beta_n\mu_n\log_2(\mu_n) + \left(\rho - \sum_{n=1}^{n(\rho)-1}\beta_n\right)\mu_{n(\rho)}\log_2(\mu_{n(\rho)})\right.$$
$$\left. + \left(\sum_{n=1}^{n(\rho)-1}\mu_n\beta_n + \left(\rho - \sum_{n=1}^{n(\rho)-1}\beta_n\right)\mu_{n(\rho)}\right)\log_2\left(\frac{\frac{P_{\text{Tx}}}{\sigma_{\text{n}}^2}\sigma_{\text{H}}^2(\alpha-\rho)+\rho}{\sum_{n=1}^{n(\rho)-1}\beta_n\mu_n + \left(\rho - \sum_{n=1}^{n(\rho)-1}\beta_n\right)\mu_{n(\rho)}}\right)\right]$$

is piecewise concave in the intervals $0 \le \rho \le \beta_1, \ldots, \beta_n \le \rho \le \min(1,\alpha)$. In each of these intervals $n(\rho)$ is constant and the derivative of $\hat{R}^{(\infty)}_{\text{WSR,lin}}(\alpha,\rho)$ with respect to ρ can be computed according to

$$\frac{\partial \hat{R}^{(\infty)}_{\text{WSR,lin}}(\alpha,\rho)}{\partial\rho} = \frac{1}{\alpha}\left[\mu_{n(\rho)}\log_2(\mu_{n(\rho)}) + \mu_{n(\rho)}\log_2\left(\frac{\frac{P_{\text{Tx}}}{\sigma_{\text{n}}^2}\sigma_{\text{H}}^2(\alpha-\rho)+\rho}{\sum_{n=1}^{n(\rho)-1}\beta_n\mu_n + \left(\rho - \sum_{n=1}^{n(\rho)-1}\beta_n\right)\mu_{n(\rho)}}\right)\right.$$
$$\left. + \frac{\left(\sum_{n=1}^{n(\rho)-1}\mu_n\beta_n + \left(\rho - \sum_{n=1}^{n(\rho)-1}\beta_n\right)\mu_{n(\rho)}\right)\left(1 - \frac{P_{\text{Tx}}}{\sigma_{\text{n}}^2}\sigma_{\text{H}}^2\right)}{\ln 2\left(\frac{P_{\text{Tx}}}{\sigma_{\text{n}}^2}\sigma_{\text{H}}^2(\alpha-\rho)+\rho\right)} - \frac{\mu_{n(\rho)}}{\ln 2}\right]$$

and the second derivative is given by

$$\frac{\partial^2 \hat{R}^{(\infty)}_{\text{WSR,lin}}(\alpha,\rho)}{\partial^2\rho} = \frac{1}{\alpha\ln 2}\left[\mu_{n(\rho)}\frac{2\left(1 - \frac{P_{\text{Tx}}}{\sigma_{\text{n}}^2}\sigma_{\text{H}}^2\right)}{\frac{P_{\text{Tx}}}{\sigma_{\text{n}}^2}\sigma_{\text{H}}^2(\alpha-\rho)+\rho} - \frac{\mu_{n(\rho)}^2}{\sum_{n=1}^{n(\rho)-1}\beta_n\mu_n + \left(\rho - \sum_{n=1}^{n(\rho)-1}\beta_n\right)\mu_{n(\rho)}}\right.$$
$$\left. - \frac{\left(\sum_{n=1}^{n(\rho)-1}\mu_n\beta_n + \left(\rho - \sum_{n=1}^{n(\rho)-1}\beta_n\right)\mu_{n(\rho)}\right)\left(1 - \frac{P_{\text{Tx}}}{\sigma_{\text{n}}^2}\sigma_{\text{H}}^2\right)^2}{\left(\frac{P_{\text{Tx}}}{\sigma_{\text{n}}^2}\sigma_{\text{H}}^2(\alpha-\rho)+\rho\right)^2}\right] =$$
$$= -\frac{\left[\left(\sum_{n=1}^{n(\rho)-1}\mu_n\beta_n + \left(\rho - \sum_{n=1}^{n(\rho)-1}\beta_n\right)\mu_{n(\rho)}\right)\left(1 - \frac{P_{\text{Tx}}}{\sigma_{\text{n}}^2}\sigma_{\text{H}}^2\right)\frac{1}{\left(\frac{P_{\text{Tx}}}{\sigma_{\text{n}}^2}\sigma_{\text{H}}^2(\alpha-\rho)+\rho\right)} - \mu_{n(\rho)}\right]^2}{\alpha\ln 2\left(\sum_{n=1}^{n(\rho)-1}\beta_n\mu_n + \left(\rho - \sum_{n=1}^{n(\rho)-1}\beta_n\right)\mu_{n(\rho)}\right)} \le 0.$$

Thus, the optimum solution to Problem (5.13) can be found by determining the maximum large system weighted sum rate in each interval $\beta_{n-1} \leq \rho \leq \min(\alpha, \beta_n)$, $n = 1, \ldots, N$, which can be computed for example by the bisection method [33, Chapter 8.2], and then choosing the maximum of these interval maxima.

A10. Derivation of An Implicit Equation for the Large System Analysis of Successive Resource Allocation and Spatial Zero-Forcing with DPC

In this section Equation (5.20) stating that

$$\int_0^{\hat{\lambda}_{j-K}} \frac{f_{\boldsymbol{C}_{j-K}}(x)}{1 - \tilde{\beta}_j + (x - z) m_{\boldsymbol{C}_j}(z)} \,\mathrm{d}\, x = \frac{1 - (j-K)\delta\xi}{1 - (j-K-1)\delta\xi}$$

will be derived together with the optimum scheduling scheme. The matrices $\boldsymbol{C}_j$ are given by

$$\boldsymbol{C}_j = \boldsymbol{V}_{\text{DPC}}^{((j-1)\delta m+1),\text{H}} \boldsymbol{H}_{\hat{k}(j)}^{\text{H}} \boldsymbol{H}_{\hat{k}(j)} \boldsymbol{V}_{\text{DPC}}^{((j-1)\delta m+1)}$$

[c.f. (5.18)]. As the nullspace of $\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)} = \boldsymbol{V}_{\text{DPC}}^{(i)} \boldsymbol{V}_{\text{DPC}}^{(i),\text{H}}$ is enlarged by one dimension with each subchannel allocation as a consequence of (3.34),

$$\text{span}\left\{\hat{\boldsymbol{P}}_{\text{DPC}}^{(i)}\right\} = \text{span}\left\{\boldsymbol{V}_{\text{DPC}}^{(i)}\right\} \subset \text{span}\left\{\hat{\boldsymbol{P}}_{\text{DPC}}^{(n)}\right\}, \quad \forall n < i. \tag{A44}$$

Of special interest is the case $i = (j-1)\delta m + 1$ and $n = (\ell_j - 1)\delta m + 1$, where ℓ_j is the highest index of a subgroup that has been allocated to user $\hat{k}(j)$ before the allocation round j. For notational convenience in the remainder the following short notations will be used

$$\tilde{\boldsymbol{V}}_{\text{DPC}}^{(j)} := \boldsymbol{V}_{\text{DPC}}^{((j-1)\delta m+1)}, \qquad \tilde{\boldsymbol{V}}_{\text{DPC}}^{(\ell_j)} = \boldsymbol{V}_{\text{DPC}}^{((\ell_j-1)\delta m+1)}, \qquad \tilde{\boldsymbol{P}}_{\text{DPC}}^{(\ell_j)} = \hat{\boldsymbol{P}}_{\text{DPC}}^{((\ell_j-1)\delta m+1)}.$$

From (A44) it follows that

$$\tilde{\boldsymbol{V}}_{\text{DPC}}^{(j)} = \tilde{\boldsymbol{P}}_{\text{DPC}}^{(\ell_j)} \tilde{\boldsymbol{V}}_{\text{DPC}}^{(j)}$$

and consequently

$$\boldsymbol{C}_j = \tilde{\boldsymbol{V}}_{\text{DPC}}^{(j),\text{H}} \boldsymbol{H}_{\hat{k}(j)}^{\text{H}} \boldsymbol{H}_{\hat{k}(j)} \tilde{\boldsymbol{V}}_{\text{DPC}}^{(j)} = \tilde{\boldsymbol{V}}_{\text{DPC}}^{(j),\text{H}} \tilde{\boldsymbol{P}}_{\text{DPC}}^{(\ell_j)} \boldsymbol{H}_{\hat{k}(j)}^{\text{H}} \boldsymbol{H}_{\hat{k}(j)} \tilde{\boldsymbol{P}}_{\text{DPC}}^{(\ell_j)} \tilde{\boldsymbol{V}}_{\text{DPC}}^{(j)}.$$

A reduced eigenvalue decomposition of the matrix $\tilde{\boldsymbol{P}}_{\text{DPC}}^{(\ell_j)} \boldsymbol{H}_{\hat{k}(j)}^{\text{H}} \boldsymbol{H}_{\hat{k}(j)} \tilde{\boldsymbol{P}}_{\text{DPC}}^{(\ell_j)}$ can be stated as

$$\tilde{\boldsymbol{P}}_{\text{DPC}}^{(\ell_j)} \boldsymbol{H}_{\hat{k}(j)}^{\text{H}} \boldsymbol{H}_{\hat{k}(j)} \tilde{\boldsymbol{P}}_{\text{DPC}}^{(\ell_j)} = \boldsymbol{V}_1^{(\ell_j)} \boldsymbol{\Sigma}_1^{(\ell_j)} \boldsymbol{V}_1^{\text{H},(\ell_j)} + \boldsymbol{V}_{\text{r}}^{(\ell_j)} \boldsymbol{\Sigma}_{\text{r}}^{(\ell_j)} \boldsymbol{V}_{\text{r}}^{\text{H},(\ell_j)}, \tag{A45}$$

where $\boldsymbol{\Sigma}_1^{(\ell_j)} \in \mathbb{C}^{\delta m \times \delta m}$ is a diagonal matrix containing the δm strongest eigenvalues of the matrix $\tilde{\boldsymbol{P}}_{\text{DPC}}^{(\ell_j)} \boldsymbol{H}_{\hat{k}(j)}^{\text{H}} \boldsymbol{H}_{\hat{k}(j)} \tilde{\boldsymbol{P}}_{\text{DPC}}^{(\ell_j)}$ and $\boldsymbol{V}_1^{(\ell_j)} \in \mathbb{C}^{N_{\text{T}} \times \delta m}$ contains the corresponding eigenvectors. Due to the multiplication with projection matrices, $\tilde{\boldsymbol{P}}_{\text{DPC}}^{(\ell_j)} \boldsymbol{H}_{\hat{k}(j)}^{\text{H}} \boldsymbol{H}_{\hat{k}(j)} \tilde{\boldsymbol{P}}_{\text{DPC}}^{(\ell_j)}$ contains at least $(\ell_j - 1)\delta m$ zero eigenvalues. Omitting these zero eigenvalues, the remaining $N_{\text{T}} - \ell_j \delta m$ eigenvalues besides the δm strongest ones are subsumed in the matrix $\boldsymbol{\Sigma}_{\text{r}}^{(\ell_j)} \in \mathbb{C}^{N_{\text{T}} - \ell_j \delta m \times N_{\text{T}} - \ell_j \delta m}$ and $\boldsymbol{V}_{\text{r}}^{(\ell_j)} \in \mathbb{C}^{N_{\text{T}} \times N_{\text{T}} - \ell_j \delta m}$

contains the corresponding eigenvectors. As $\boldsymbol{V}_1^{(\ell_j)}$ contains the transmit vectors for the $(\ell_j - 1)\delta m + 1$th to the $\ell_j \delta m$th data stream, the span of $\boldsymbol{V}_\mathrm{r}^{(\ell_j)}$ is composed as

$$\mathrm{span}\left\{\boldsymbol{V}_\mathrm{r}^{(\ell_j)}\right\} = \mathrm{span}\left\{\tilde{\boldsymbol{P}}_\mathrm{DPC}^{(\ell_j)} - \boldsymbol{V}_1^{(\ell_j)}\boldsymbol{V}_1^{(\ell_j),\mathrm{H}}\right\} = \mathrm{span}\left\{\tilde{\boldsymbol{P}}_\mathrm{DPC}^{(\ell_j+1)}\right\}. \tag{A46}$$

Thus $\boldsymbol{V}_\mathrm{r}^{(\ell_j)}$ is a basis of $\tilde{\boldsymbol{P}}_\mathrm{DPC}^{(\ell_j+1)}$. Furthermore, as all transmit vectors of data streams chosen before step i, $\boldsymbol{V}_1^{(\ell_j)}$ lies in $\mathrm{null}\left\{\tilde{\boldsymbol{V}}_\mathrm{DPC}^{(j)}\right\}$ and therefore

$$\boldsymbol{C}_j = \tilde{\boldsymbol{V}}_\mathrm{DPC}^{(j),\mathrm{H}}\tilde{\boldsymbol{P}}_\mathrm{DPC}^{(\ell_j)}\boldsymbol{H}_{\hat{k}(j)}^\mathrm{H}\boldsymbol{H}_{\hat{k}(j)}\tilde{\boldsymbol{P}}_\mathrm{DPC}^{(\ell_j)}\tilde{\boldsymbol{V}}_\mathrm{DPC}^{(j)} = \tilde{\boldsymbol{V}}_\mathrm{DPC}^{(j),\mathrm{H}}\boldsymbol{V}_\mathrm{r}^{(\ell_j)}\boldsymbol{\Sigma}_\mathrm{r}^{(\ell_j)}\boldsymbol{V}_\mathrm{r}^{\mathrm{H},(\ell_j)}\tilde{\boldsymbol{V}}_\mathrm{DPC}^{(j),\mathrm{H}}.$$

The columns of $\tilde{\boldsymbol{V}}_\mathrm{DPC}^{(j)}$ are completely contained in $\mathrm{span}\left\{\boldsymbol{V}_\mathrm{r}^{(\ell_j)}\right\}$ due to (A44) and (A46). By introducing the matrix $\hat{\boldsymbol{V}}^{(j)} \in \mathbb{C}^{N_\mathrm{T}-\ell_j\delta m \times N_\mathrm{T}-(j-1)\delta m}$ representing $\tilde{\boldsymbol{V}}_\mathrm{DPC}^{(j)}$ in the basis $\boldsymbol{V}_\mathrm{r}^{(\ell_j)}$, the matrix $\tilde{\boldsymbol{V}}_\mathrm{DPC}^{(j)}$ is given by

$$\tilde{\boldsymbol{V}}_\mathrm{DPC}^{(j)} = \boldsymbol{V}_\mathrm{r}^{(\ell_j)}\hat{\boldsymbol{V}}^{(j)}$$

and thus

$$\boldsymbol{C}_j = \hat{\boldsymbol{V}}^{\mathrm{H},(j)}\boldsymbol{\Sigma}_\mathrm{r}^{(\ell_j)}\hat{\boldsymbol{V}}^{(j)}.$$

The matrix $\hat{\boldsymbol{V}}^{(j)}$ is uniformly distributed over the manifold of $N_\mathrm{T}-\ell_j\delta m \times N_\mathrm{T}-(j-1)\delta m$ complex matrices with $\hat{\boldsymbol{V}}^{\mathrm{H},(j)}\hat{\boldsymbol{V}}^{(j)} = \boldsymbol{I}_{N_\mathrm{T}-(j-1)\delta m}$. This is the case, if the subspace spanned by $\tilde{\boldsymbol{V}}_\mathrm{DPC}^{(j)}$ is uniformly distributed within the subspace spanned by $\boldsymbol{V}_\mathrm{r}^{(\ell_j)}$, which is equal to $\mathrm{span}\left\{\tilde{\boldsymbol{P}}_\mathrm{DPC}^{(\ell_j+1)}\right\}$ [c.f. (A46)]. To proof the latter statement, it has to be shown that in each step of the algorithm, the transmit vector for the newly allocated data stream is uniformly taken from the set of unit norm vectors orthogonal to the transmit vector of the previously allocated data streams. The transmit vector in the nth step is chosen to be the eigenvector corresponding to the principal eigenvalue of the matrix

$$\hat{\boldsymbol{P}}_\mathrm{DPC}^{(n)}\boldsymbol{H}_{k(n)}^\mathrm{H}\boldsymbol{H}_{k(n)}\hat{\boldsymbol{P}}_\mathrm{DPC}^{(n)} = \boldsymbol{V}_\mathrm{DPC}^{(n)}\boldsymbol{V}_\mathrm{DPC}^{\mathrm{H},(n)}\boldsymbol{H}_{k(n)}^\mathrm{H}\boldsymbol{H}_{k(n)}\boldsymbol{V}_\mathrm{DPC}^{(n)}\boldsymbol{V}_\mathrm{DPC}^{\mathrm{H},(n)}$$

and this transmit vector is uniformly distributed within $\mathrm{span}\left\{\boldsymbol{V}_\mathrm{DPC}^{(n)}\right\}$, if the matrix of eigenvectors of the matrix $\boldsymbol{V}_\mathrm{DPC}^{\mathrm{H},(n)}\boldsymbol{H}_{k(n)}^\mathrm{H}\boldsymbol{H}_{k(n)}\boldsymbol{V}_\mathrm{DPC}^{(n)}$ is Haar-distributed, which applies, if the latter matrix is unitarily invariant [109, Lemma 2.6]. This can be proved by induction. Suppose that in step $n-1$ the matrices $\boldsymbol{V}_\mathrm{DPC}^{\mathrm{H},(n-1)}\boldsymbol{H}_k^\mathrm{H}\boldsymbol{H}_k\boldsymbol{V}_\mathrm{DPC}^{(n-1)}$ are unitarily invariant for all users k. Then the transmit vector for the data stream allocated in the $n-1$th step is uniformly distributed within the subspace of unit-norm vectors orthogonal to the transmit vectors determined in previous steps. This implies that the representation $\tilde{\boldsymbol{V}}^{(n)} \in \mathbb{C}^{N_\mathrm{T}-(n-1)\times N_\mathrm{T}-(n-2)}$ of $\boldsymbol{V}_\mathrm{DPC}^{(n)}$ in the basis $\boldsymbol{V}_\mathrm{DPC}^{(n-1)}$ so that

$$\boldsymbol{V}_\mathrm{DPC}^{(n)} = \boldsymbol{V}_\mathrm{DPC}^{(n-1)}\tilde{\boldsymbol{V}}^{(n)},$$

is uniformly distributed in the manifold of $N_\mathrm{T}-(n-1)\times N_\mathrm{T}-(n-2)$ orthonormal matrices. Thus, $\tilde{\boldsymbol{V}}^{(n)}$ and consequently

$$\boldsymbol{V}_\mathrm{DPC}^{\mathrm{H},(n)}\boldsymbol{H}_k^\mathrm{H}\boldsymbol{H}_k\boldsymbol{V}_\mathrm{DPC}^{(n)} = \tilde{\boldsymbol{V}}^{\mathrm{H},(n)}\boldsymbol{V}_\mathrm{DPC}^{\mathrm{H},(n-1)}\boldsymbol{H}_k^\mathrm{H}\boldsymbol{H}_k\boldsymbol{V}_\mathrm{DPC}^{(n-1)}\tilde{\boldsymbol{V}}^{(n)}$$

is unitarily invariant for all users k, if $\boldsymbol{V}_\mathrm{DPC}^{\mathrm{H},(n-1)}\boldsymbol{H}_k^\mathrm{H}\boldsymbol{H}_k\boldsymbol{V}_\mathrm{DPC}^{(n-1)}$ is unitarily invariant for all k. To complete the proof, it has therefore to be shown that the matrices $\boldsymbol{V}_\mathrm{DPC}^{\mathrm{H},(1)}\boldsymbol{H}_k^\mathrm{H}\boldsymbol{H}_k\boldsymbol{V}_\mathrm{DPC}^{(1)}$ are unitarily

invariant, which follows from the facts that $\boldsymbol{V}_{\text{DPC}}^{(1)} = \boldsymbol{I}_{N_\text{T}}$ and the matrices $\boldsymbol{H}_k$ contain Gaussian i.i.d. entries.
Thus, it has been shown that $\hat{\boldsymbol{V}}^{(j)}$ is uniformly distributed over the manifold of $N_\text{T} - \ell_j \delta m \times N_\text{T} - (j-1)\delta m$ complex matrices with $\hat{\boldsymbol{V}}^{\text{H},(j)} \hat{\boldsymbol{V}}^{(j)} = \boldsymbol{I}_{N_\text{T}-(j-1)\delta m}$, which will be used in the following to derive the η-transform of the matrix $\boldsymbol{C}_j = \hat{\boldsymbol{V}}^{\text{H},(j)} \boldsymbol{\Sigma}_\text{r}^{(\ell_j)} \hat{\boldsymbol{V}}^{(j)}$. For that purpose first Lemma 2.28 from [109] is applied, which states that

$$\eta_{\boldsymbol{C}_j}(\gamma) = 1 - \frac{1}{\tilde{\beta}_j} + \frac{1}{\tilde{\beta}_j} \eta_{\boldsymbol{\Sigma}_\text{r}^{(\ell_j)} \hat{\boldsymbol{V}}^{(j)} \hat{\boldsymbol{V}}^{\text{H},(j)}}(\gamma), \tag{A47}$$

where

$$\tilde{\beta}_j = \frac{N_\text{T} - (j-1)\delta m}{N_\text{T} - \ell_j \delta m} = \frac{1 - (j-1)\delta\xi}{1 - \ell_j \delta\xi}. \tag{A48}$$

The matrices $\boldsymbol{\Sigma}_\text{r}^{(\ell_j)}$ and $\hat{\boldsymbol{V}}^{(j)} \hat{\boldsymbol{V}}^{\text{H},(j)}$ are asymptotically free [109, Definition 2.19], as they are independent and $\hat{\boldsymbol{V}}^{(j)} \hat{\boldsymbol{V}}^{\text{H},(j)}$ is unitarily invariant, from which the asymptotic freeness can be derived with the proof of the theorem in [138]. From Theorem 2.68 and Example 2.51 in [109], the η-transform $\eta_{\hat{\boldsymbol{V}}^{(j)} \hat{\boldsymbol{V}}^{\text{H},(j)} \boldsymbol{\Sigma}_\text{r}^{(\ell_j)}}(\gamma)$ is given by the implicit equation

$$\eta_{\hat{\boldsymbol{V}}^{(j)} \hat{\boldsymbol{V}}^{\text{H},(j)} \boldsymbol{\Sigma}_\text{r}^{(\ell_j)}}(\gamma) = \eta_{\boldsymbol{\Sigma}_\text{r}^{(\ell_j)}} \left(\gamma + \gamma \frac{\tilde{\beta}_j}{\eta_{\hat{\boldsymbol{V}}^{(j)} \hat{\boldsymbol{V}}^{\text{H},(j)} \boldsymbol{\Sigma}_\text{r}^{(\ell_j)}}(\gamma)} \right).$$

By using (A47), the η-transform $\eta_{\boldsymbol{C}_j}(\gamma)$ is given implicitly by

$$\tilde{\beta}_j \eta_{\boldsymbol{C}_j}(\gamma) - \tilde{\beta}_j + 1 = \eta_{\boldsymbol{\Sigma}_\text{r}^{(\ell_j)}} \left(\frac{\gamma \tilde{\beta}_j \eta_{\boldsymbol{C}_j}(\gamma)}{\tilde{\beta}_j \eta_{\boldsymbol{C}_j}(\gamma) - \tilde{\beta}_j + 1} \right). \tag{A49}$$

The diagonal elements of the matrix $\boldsymbol{\Sigma}_\text{r}^{(\ell_j)}$ are given by the eigenvalues of the matrix $\boldsymbol{C}_{\ell_j}$ except the largest δm ones, which correspond to the channel gains from step $(\ell_j - 1)\delta m + 1$ to step $\ell_j \delta m$ [c.f. (A45)]. The asymptotic eigenvalue distribution of the matrix $\boldsymbol{\Sigma}_\text{r}^{(\ell_j)}$ is therefore given by the a.e.d. of the matrix $\boldsymbol{C}_{\ell_j}$ truncated at $\hat{\lambda}_{\ell_j}$, where $\hat{\lambda}_{\ell_j}$ is defined in (5.19), and normalized by $\frac{N_\text{T} - (\ell_j - 1)\delta m}{N_\text{T} - \ell_j \delta m}$. Thus, the η-transform of the matrix $\boldsymbol{\Sigma}_\text{r}^{(\ell_j)}$ can be written as

$$\eta_{\boldsymbol{\Sigma}_\text{r}^{(\ell_j)}}(\gamma) = \frac{N_\text{T} - (\ell_j - 1)\delta m}{N_\text{T} - \ell_j \delta m} \int_0^{\hat{\lambda}_{\ell_j}} \frac{f_{\boldsymbol{C}_{\ell_j}}^{(\infty)}(x)}{1 + \gamma x} \,\mathrm{d}\,x, \tag{A50}$$

Inserting (A50) into (A49) and using the relationship (5.5) between the η- and the Stieltjes transform leads to

$$\int_0^{\hat{\lambda}_{\ell_j}} \frac{f_{\boldsymbol{C}_{\ell_j}}^{(\infty)}(x)}{1 - \tilde{\beta}_j + (x - z) m_{\boldsymbol{C}_j}(z)} \,\mathrm{d}\,x = \frac{N_\text{T} - \ell_j \delta m}{N_\text{T} - (\ell_j - 1)\delta m}.$$

By using this implicit relationship to test in each step, when a group of subchannels is allocated to the same user, which user leads to the strongest increase in sum rate, it turns out that, as in the first K allocation rounds, it is optimum to serve each user every K allocations. Consequently $\ell_j = j - K$ so that

$$\tilde{\beta}_j = \frac{1-(j-1)\delta\xi}{1-\ell_j\delta\xi} = \frac{1-(j-1)\delta\xi}{1-(j-K)\delta\xi}$$

and

$$\frac{N_{\mathrm{T}}-\ell_j\delta m}{N_{\mathrm{T}}-(\ell_j-1)\delta m} = \frac{1-(j-K)\delta\xi}{1-(j-K-1)\delta\xi},$$

which leads to the desired result in (5.20).

Bibliography

[1] G.J. Foschini and M.J. Gans. On Limits of Wireless Communications in a Fading Environment using Multiple Antennas. *Wireless Personal Communications*, 6:311–355, 1998.

[2] I.E. Telatar. Capacity of Multi-Antenna Gaussian Channels. *European Transactions on Communications*, 10:585–595, 1999.

[3] N. Jindal and A. Goldsmith. Dirty-Paper Coding versus TDMA for MIMO Broadcast Channels. *IEEE Transactions on Information Theory*, 51(5):1783–1794, May 2005.

[4] W. Yu, W. Rhee, S. Boyd, and J.M. Cioffi. Iterative Water-filling for Gaussian Vector Multiple Access Channels. *IEEE Transactions on Information Theory*, 50:145–151, 2004.

[5] S. Vishwanath, N. Jindal, and A. Goldsmith. Duality, Achievable Rates, and Sum-Rate Capacity of Gaussian MIMO Broadcast Channels. *IEEE Transactions on Information Theory*, 49:2658–2668, 2003.

[6] N. Jindal, W. Rhee, S. A. Vishwanath, S. Jafar, and A. Goldsmith. Sum Power Iterative Water-filling for Multi-antenna Gaussian Broadcast Channels. *IEEE Transactions on Information Theory*, 51(4):1570–1580, April 2005.

[7] H. Viswanathan, S. Venkatesan, and H. Huang. Downlink Capacity Evaluation of Cellular Networks With Known-Interference Cancellation. *IEEE Journal on Selected Areas in Communications*, 21(6):802–811, June 2003.

[8] G. Wunder and T. Michel. Minimum Rates Scheduling for MIMO-OFDM Broadcast Channels. In *Proc. of 9th IEEE International Symposium on Spread Spectrum Techniques and Applications (ISSSTA 2006)*, pages 510–514, August 2006.

[9] J. Lee and N. Jindal. Symmetric Capacity of MIMO Downlink Channels. In *IEEE International Symposium on Information Theory (ISIT)*, pages 1031–1035, July 2006.

[10] P. Tejera, W. Utschick, G. Bauch, and J. A. Nossek. Subchannel Allocation in Multiuser Multiple Input Multiple Output Systems. *IEEE Transactions on Information Theory*, 52(10):4721–4733, October 2006.

[11] P. Tejera, W. Utschick, J.A. Nossek, and G. Bauch. Rate Balancing in Multiuser MIMO OFDM Systems. *IEEE Transactions on Communications*, 57(5):1370–1380, May 2009.

[12] G. Dimić and N.D. Sidoropoulos. On Downlink Beamforming with Greedy User Selection. *IEEE Transactions on Signal Processing*, 53(10):3857–3868, October 2005.

[13] T.M. Cover and J.A. Thomas. *Elements of Information Theory*. John Wiley & Sons, 2nd edition, 2006.

[14] H. Weingarten, Y. Steinberg, and S. Shamai. The Capacity Region of the Gaussian Multiple-Input Multiple-Output Broadcast Channel. *IEEE Transactions on Information Theory*, 52(9):3936–3964, September 2006.

[15] M.H.M. Costa. Writing on Dirty Paper. *IEEE Transactions on Information Theory*, 29(3):439–441, May 1983.

[16] D. Schmidt, M. Joham, and W. Utschick. Minimum Mean Square Error Vector Precoding. *European Transactions on Telecommunications*, 19:219–231, 2007.

[17] S. ten Brink and U. Erez. A Close-to-Capacity Dirty Paper Coding Scheme. In *Proc. of International Symposium on Information Theory (ISIT)*, page 533, July 2004.

[18] D.J. Love, R.W. Heath, V.K.N. Lau, D. Gesbert, B.D. Rao, and M.Andrews. An Overview of Limited Feedback in Wireless Communication Systems. *IEEE Journal on Selected Areas in Communications*, 26(8):1341–1365, October 2008.

[19] H. Zhang, Y. Li, V. Stoplman, and N. van Waes. A Reduced Feedback Approach for Precoded MIMO-OFDM Systems. *IEEE Transactions on Wireless Communications*, 6(1):55–58, 2007.

[20] N. Khaled, B. Mondal, G. Leus, R.W. Heath, and F. Petré. Interpolation-Based Multi-Mode Precoding for MIMO-OFDM Systems with Limited Feedback. *IEEE Transactions on Wireless Communications*, 6(3):1003–1013, March 2007.

[21] C. Guthy, W. Utschick, and G. Dietl. Finite Rate Feedback Schemes for the MIMO OFDM Broadcast Channel. In *Proc. of ITG Workshop on Smart Antennas*, 2008.

[22] C. Guthy, A. Krebs, and W. Utschick. Reduced Entropy Parameterization of MIMO Channel Matrices for Finite-Rate Feedback. In *Proc. of 10th IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 374–378, 2009.

[23] Y. Hara, L. Brunel, and K. Oshima. Spatial Scheduling With Interference Cancellation in Multiuser MIMO Systems. *IEEE Transactions on Vehicular Technology*, 57(2):893–905, March 2008.

[24] P. Tejera, W. Utschick, G. Bauch, and J.A. Nossek. Efficient Implementation of Successive Encoding Schemes for the MIMO OFDM Broadcast Channel. In *IEEE International Conference on Communications (ICC)*, Istanbul, June 2006.

[25] C.-B. Chae, D. Mazzarese, T. Inoue, and R.W. Heath. Coordinated Beamforming for the Multiuser MIMO Broadcast Channel With Limited Feedforward. *IEEE Transactions on Signal Processing*, 56(12):6044–6056, Dec. 2008.

[26] M. Kobayashi and G. Caire. An Iterative Water-Filling Algorithm for Maximum Weighted Sum-Rate of Gaussian MIMO-BC. *IEEE Journal on Selected Areas in Communications*, 24(8):1640–1646, August 2006.

[27] D.N.C. Tse and S. V. Hanly. Multiaccess Fading Channels–Part I: Polymatroid Structure, Optimal Resource Allocation and Throughput Capacities. *IEEE Transactions on Information Theory*, 44(7):2796–2815, November 1998.

[28] S. Vishwanath, S. A. Jafar, and A. Goldsmith. Optimum Power and Rate Allocation Strategies for Multiple Access Fading Channels. In *Proc. of 53rd Vehicular Technology Conference (VTC Spring)*, pages 546 – 550, May 2001.

[29] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2008.

[30] R. Hunger, D. Schmidt, M. Joham, and W. Utschick. A General Covariance-Based Optimization Framework Using Orthogonal Projections . In *Proc. of IEEE Conference on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 76–80, July 2008.

[31] R. Böhnke and K.-D. Kammeyer. Weighted Sum Rate Maximization for the MIMO-Downlink Using a Projected Conjugate Gradient Algorithm. In *Proc. of International Workshop on Cross Layer Design (IWCLD)*, pages 82–85, September 2007.

[32] J. Liu, Y. T. Hou, and H. D. Sherali. On the Maximum Weighted Sum-Rate of MIMO Gaussian Broadcast Channels. In *Proc. of International Conference on Communications (ICC)*, pages 3664 – 3668, May 2008.

[33] M.S. Bazaraa, H.D. Sherali, and C.M. Shetty. *Nonlinear Programming – Theory and Applications*. Wiley Interscience, 3rd edition, 2006.

[34] R. Hunger. Floating Point Operations in Matrix-Vector Calculus. Technical report, Technische Universität München, 2007.

[35] G. H. Golub and C. F. van Loan. *Matrix Computations*. The John Hopkins University Press, 1989.

[36] P. Tejera. *Principles and Algorithms for Transmission in Multiple-Input Multiple-Output Systems*. PhD thesis, Technische Universität München, 2008.

[37] M. Codreanu, M. Juntti, and M. Latva-aho. Low-Complexity Iterative Algorithm for Finding the MIMO-OFDM Broadcast Channel Sum Capacity. *IEEE Transactions on Communications*, 55(1):48–53, January 2007.

[38] R. Böhnke, V. Kühn, and K.-D. Kammeyer. Fast Sum Rate Maximization for the Downlink of MIMO-OFDM Systems. In *Proc. of Canadian Workshop on Information Theory (CWIT)*, June 2005.

[39] W. Yu. Sum-Capacity Computation for the Gaussian Vector Broadcast Channel. *IEEE Transactions on Information Theory*, 52(2):754–759, February 2006.

[40] M. Codreanu, M. Juntti, and M. Latva-aho. On the Dual-Decomposition Based Sum Capacity Maximization for Vector Broadcast Channels. *IEEE Transactions on Vehicular Technology*, 56(6):3577–3581, November 2007.

[41] R. Hunger, D. Schmidt, and W. Utschick. Sum-Capacity and MMSE for the MIMO Broadcast Channel without Eigenvalue Decompositions. In *International Symposium on Information Theory*, June 2007.

[42] G. Caire and S. Shamai. On the Achievable Throughput of Multiantenna Gaussian Broadcast Channel. *IEEE Transactions on Information Theory*, 49(7):1691–1706, July 2003.

[43] Z. Tu and R.S. Blum. Multiuser Diversity for a Dirty Paper Approach. *IEEE Communications Letters*, 7(8):370–372, August 2003.

[44] J. Dai, C. Chang, Z. Ye, and Y.S. Hung. An Efficient Greedy Scheduler for Zero-Forcing Dirty-Paper Coding. *IEEE Transactions on Communications*, 57(7):1939–1943, July 2009.

[45] A.D. Dabbagh and D.J. Love. Precoding for Multiple Antenna Gaussian Broadcast Channels with Successive Zero-Forcing. *IEEE Transactions on Signal Processing*, 57(7):3837–3850, July 2007.

[46] L. Tran and E. Hong. Multiuser Diversity for Successive Zero-Forcing Dirty-Paper Coding: Greedy Scheduling Algorithms and Asymptotic Performance. *IEEE Transactions on Signal Processing*, 58(6):3411–3416, June 2010.

[47] S. Sigdel and W.A. Krzymien. Simplified Transmit Covariance Optimization and User Ordering Algorithm for Successive Zero-Forcing Precoding. In *Proc. of Signal Processing Advances in Wireless Communications (SPAWC)*, pages 235–239, June 2009.

[48] S. Sigdel and W.A. Krzymien. Efficient User Selection and Ordering Algorithms for Successive Zero-Forcing Precoding for Multiuser MIMO Downlink. In *Proc. of 69th Vehicular Technology Conference (VTC Spring)*, April 2009.

[49] C. Chen. *Optimized Resource Allocation for MIMO Multi-Carrier Multi-User Communications Systems*. PhD thesis, Stanford University, 2008.

[50] J. Wang, D.L. Love, and M. Zoltowski. User Selection for MIMO Broadcast Channel with Sequential Water-Filling. In *Proc. of 44th Annual Allerton Conf. on Communications, Control, and Computing*, 2006.

[51] M. Tomlinson. New Automatic Equalizer Employing Modulo Arithmetic. *Electronic Letters*, 7:138–139, March 1971.

[52] H. Harashima and H. Miyakawa. Matched-Transmission Technique for Channels with Intersymbol Interference. *IEEE Transactions on Communications*, 20(4):774–780, August 1972.

[53] W. Yu, D. P. Varodayan, and J. M. Cioffi. Trellis and Convolutional Precoding for Transmitter-Based Interference Presubtraction. *IEEE Transactions on Communications*, 53(7):1120–1230, July 2005.

[54] E. Jorswieck and E.G. Larsson. Linear Precoding in Multiple Antenna Broadcast Channels: Efficient Computation of the Achievable Rate Region. In *Proc. of IEEE/ITG Workshop on Smart Antennas (WSA)*, pages 21–28, February 2008.

[55] S. Shi, M. Schubert, and H. Boche. Rate Optimization for Multiuser MIMO Systems with Linear Processing. *IEEE Transactions on Signal Processing*, 56(8):4020–4030, August 2008.

[56] S. Shi, M. Schubert, and H. Boche. Downlink MMSE Transceiver Optimization for Multiuser MIMO Systems: Duality and Sum-MSE Minimization. *IEEE Transactions on Signal Processing*, 55(11):5436–5446, 2007.

[57] M. Codreanu, A. Tölli, M. Juntti, and M. Latva-aho. Joint Design of Tx-Rx Beamformers in MIMO Downlink Channel. *IEEE Transactions on Signal Processing*, 55(9):4639–4655, September 2007.

[58] R. Böhnke and K.-D. Kammeyer. Weighted Sum Rate Maximization for MIMO-OFDM Systems with Linear and Dirty Paper Precoding. In *Proc. of 7th International ITG Conference on Source and Channel Coding (SCC 08)*, page 5 pages, Ulm, Germany, January 2008.

[59] R. Hunger, D. A. Schmidt, and M. Joham. A Combinatorial Approach to Maximizing the Sum Rate in the MIMO BC with Linear Filtering. In *Proc. of 42nd Asilomar Conference on Signals, Systems, and Computers 2008*, page 5 pages, October 2008.

[60] R.C. Elliot, S. Sigdel, W.A. Krzymien, M. Al-Shalash, and A.C.K. Soong. Genetic and Greedy User Scheduling for Multiuser MIMO Systems with Successive Zero-Forcing. In *Proc. of 70th Vehicular Technology Conference (VTC Fall)*, 2009.

[61] 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*, 52(2):461–471, February 2004.

[62] Z. Shen, R. Chen, J.G. Andrews, R.W. Heath, and B.L. Evans. Low Complexity User Selection Algorithms for Multiuser MIMO Systems with Block Diagonalization. *IEEE Transactions on Signal Processing*, 54(9):3658 – 3663, September 2006.

[63] M. Fuchs, G. Del Galdo, and M. Haardt. Low-Complexity Space -Time - Frequency Scheduling for MIMO Systems With SDMA. *IEEE Transactions on Vehicular Technology*, 56(5):2775 – 2784, September 2007.

[64] J. Wang, D.J. Love, and M.D. Zoltowski. User Selection With Zero-Forcing Beamforming Achieves the Asymptotically Optimal Sum Rate. *IEEE Transactions on Signal Processing*, 56(8):3713–3726, August 2008.

[65] T. Yoo and A. Goldsmith. On the Optimality of Multiantenna Broadcast Scheduling Using Zero-Forcing Beamforming. *IEEE Journal on Selected Areas in Communications*, 24(3):528–541, March 2006.

[66] T. Yoo and A. Goldsmith. Sum-Rate Optimal Multi-Antenna Downlink Beamforming Strategy Based on Clique Search. In *Proc. of Global Telecommunications Conference (GLOBECOM)*, 2005.

[67] J. Wang, D.J. Love, and M. Zoltowski. User Selection for the MIMO Broadcast Channel with a Fairness Constraint. In *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages III–9 – III–12, 2007.

[68] L. Sun and M.R. McKay. Eigen-Based Transceivers for the MIMO Broadcast Channel With Semi-Orthogonal User Selection. *IEEE Transactions on Signal Processing*, 58(10):5246–5261, October 2010.

[69] A. Bayesteh and A.K. Khandani. On the User Selection for MIMO Broadcast Channel. *IEEE Transactions on Information Theory*, 54(3):1086–1107, March 2008.

[70] Z. Shi, W. Xu, S. Jin, C. Zhao, and Z. Ding. On Wireless Downlink Scheduling of MIMO Systems with Homogeneous Users. *IEEE Transactions on Information Theory*, 56(7):3369–3377, July 2010.

[71] F. Boccardi and H. Huang. A Near-optimum Technique Using Linear Precoding for the MIMO Broadcast Channel. In *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 3, pages III–17 – III–20, April 2007.

[72] R. Chen, Z. Shen, J.G. Andrews, and R.W. Heath. Multimode Transmission for Multiuser MIMO Systems with Block Diagonalization. *IEEE Transactions on Signal Processing*, 56(7):3294–3302, July 2008.

[73] L. Wei and M. Latva-aho. An Efficient Joint User and Antenna Selection for MIMO Broadcast Communications Using Generalized Zero Forcing Precoder. In *Proc. of IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pages 1312–1316, 2009.

[74] B.C. Lim, W.A. Krzymien, and C. Schlegel. Efficient Sum Rate Maximization and Resource Allocation in Block-Diagonalized Space-Division Multiplexing. *IEEE Transactions on Vehicular Technology*, 58(1):478–484, January 2009.

[75] G. Primolevo, O. Simeone, and U. Spagnolini. Channel Aware Scheduling for Broadcast MIMO Systems with Orthogonal Linear Precoding and Fairness Constraints. In *Proc. of International Conference on Communications (ICC)*, pages 493–497, 2005.

[76] R. Hunger, P. de Kerret, and M. Joham. An Algorithm for Maximizing a Quotient of Two Hermitian Form Determinants with Different Exponents. In *Proc. of International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 3346–3349, March 2010.

[77] C. Swannack, E. Uysal-Biyikoglu, and Gregory Wornell. Low Complexity Multiuser Scheduling for Maximizing Throughput in the MIMO Broadcast Channel. In *Proc. of 42nd Annual Allerton Conf. on Communications, Control, and Computing*, pages 440–449, 2004.

[78] C. Guthy, W. Utschick, R. Hunger, and M. Joham. Weighted Sum Rate Maximization in the MIMO MAC with Linear Transceivers: Algorithmic Solutions . In *Proc. of 43rd Asilomar conference on Signals, Systems, and Computers*, pages 1474 – 1478, 2009.

[79] R. Hunger and M. Joham. A General Rate Duality of the MIMO Multiple Access Channel and the MIMO Broadcast Channel. In *Proc. of IEEE Global Telecommunications Conference (GLOBECOM)*, pages 1–5, December 2008.

[80] C. Guthy, W. Utschick, R. Hunger, and M. Joham. Efficient Weighted Sum Rate Maximization With Linear Precoding. *IEEE Transactions on Signal Processing*, 58(4):2284 – 2297, April 2010.

[81] C. Guthy, W. Utschick, and G. Dietl. Low Complexity Linear Zero-Forcing for the MIMO Broadcast Channel. *IEEE Journal of Selected Topics in Signal Processing, special issue on Managing Complexity in Multiuser MIMO Systems*, 3(6):1106–1117, December 2009.

[82] G. Dietl, C. Guthy, W. Utschick, G. Bauch, and J.A. Nossek. Subspace Beamforming and Scheduling for the Multiuser MIMO Broadcast Channel. In *Proc. of VDE/ITG conference on Source and Channel Coding (SCC), Ulm , Germany*, 2008.

[83] C. Guthy, W. Utschick, and G. Dietl. A User Grouping Method for Maximum Weighted Sum Capacity Gain. In *Proc. of International Conference on Communications (ICC)*, June 2009.

[84] C. Guthy, W. Utschick, J. A. Nossek, G. Dietl, and G. Bauch. Rate-Invariant User Preselection for Complexity Reduction in Multiuser MIMO Systems. In *Proc. of IEEE Vehicular Technology Conference (VTC)*, September 2008.

[85] C. Guthy, W. Utschick, G. Dietl, and P. Tejera. Efficient Linear Successive Allocation for the MIMO Broadcast Channel. In *Proc. of* 42 *nd Asilomar Conference on Signals, Systems, and Computers*, October 2008.

[86] K. Seong, M. Mohseni, and J. M. Cioffi. Optimal Resource Allocation for OFDMA Downlink Systems. In *Proc. of International Symposium on Information Theory (ISIT)*, July 2006.

[87] M. Mohensi, R. Zhang, and J.M. Cioffi. Optimized Transmission for Fading Multiple-Access and Broadcast Channels With Multiple Antennas. *IEEE Journal on Selected Areas in Communications*, 24(8):1627–1639, August 2006.

[88] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H.V. Poor. *MIMO Wireless Communications*. Cambridge University Press, 2007.

[89] S. Boyd. Lecture Slides on Convex Optimization II (EE364b), Lecture 7: Ellipsoid Method.

[90] T. Michel and G. Wunder. Achieving QoS and Efficiency in the MIMO Downlink with Limited Power. In *Proc. of ITG/IEEE Workshop on Smart Antennas (WSA)*, 2007.

[91] J. Brehmer. *Utility Maximization in Nonconvex Wireless Systems*. PhD thesis, Technische Universität München, 2010.

[92] C. Fung, W. Yu, and T. Lim. Multiantenna Downlink Precoding with Individual Rate Constraints: Power Minimization and User Ordering. In *International Conference on Communication Systems*, pages 45–49, September 2004.

[93] J. Oh, S.-J. Kim, R. Narasimhan, and J.M. Cioffi. Transmit Power Optimization for Gaussian Vector Broadcast Channels. In *Proc. of International Conference on Communications (ICC)*, pages 2712 – 2716, 2005.

[94] R.A. Johnson and C.R. Horn. *Matrix Analysis*. Cambridge University Press, 1985.

[95] R.G. Bland, D. Goldfarb, and M.J. Todd. The Ellipsoid Method: A Survey. *Operations Research*, 29(6):1039–1091, November-December 1981.

[96] J. Brehmer, Q. Bai, and W. Utschick. Time-Sharing Solutions in MIMO Broadcast Channel Utility Maximization. In *Proc. of ITG/IEEE Workshop on Smart Antennas (WSA)*, pages 153–156, 2008.

[97] M. Johannson and L. Xiao. Cross-layer optimization of wireless networks using nonlinear column generation. *IEEE Transactions on Wireless Communications*, 5(2):435–445, February 2006.

[98] J. Chen and A.L. Swindlehurst. Downlink Resource Allocation for Multi-user MIMO-OFDMA Systems: The Kalai-Smorodinsky Bargaining Approach. In *Proc. of 3rd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pages 380 – 383, 2009.

[99] I.C. Wong, Z. Shen, B.L. Evans, and J.G. Andrews. A Low Complexity Algorithm for Proportional Resource Allocation in OFDMA Systems. In *Proc. of IEEE Workshop on Signal Processing Systems (SIPS)*, 2004.

[100] Q.H. Spencer, A.L. Swindlehurst, and M. Haardt. Fast Power Minimization with QoS Constraints in Multi-user MIMO Downlinks. In *Proc. of International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 461–471, 2003.

[101] W.W.L. Ho and Y.-C. Liang. Efficient Power Minimization for MIMO Broadcast Channels with BD-GMD. In *Proc. of International Conference on Communications (ICC)*, pages 2791–2796, 2007.

[102] N. Ul Hassan and M. Assaad. Low Complexity Margin Adaptive Resource Allocation in Downlink MIMO-OFDMA System. *IEEE Transactions on Wireless Communications*, 8(7):3365–3371, July 2009.

[103] C. Hellings, M. Joham, and W. Utschick. Power Minimization in Parallel Vector Broadcast Channels with Zero-Forcing Beamforming. In *Proc. of Global Telecommunications Conference (GLOBECOM)*, 2010.

[104] G. Zheng, K.-K. Wong, and T.-S. Ng. Throughput Maximization in Linear Multiuser MIMO-OFDM Downlink Systems. *IEEE Transactions on Vehicular Technology*, 57(3):1993–1998, May 2008.

[105] A. Tölli, M. Codreanu, and M. Juntti. Linear Multiuser MIMO Transceiver Design With Quality of Service and Per-Antenna Power Constraints. *IEEE Transactions on Signal Processing*, 56(7):3049–3055, July 2008.

[106] A. Mezghani, M. Joham, R. Hunger, and W. Utschick. Transceiver Design for Multi-User MIMO Systems. In *Proc. of ITG/IEEE Workshop on Smart Antennas (WSA)*, 2006.

[107] C. Hellings, M. Joham, and W. Utschick. Gradient-Based Rate Balancing for MIMO Broadcast Channels With Linear Precoding. In *Proc. of ITG/IEEE Workshop on Smart Antennas (WSA)*, 2011.

[108] C. Guthy, W. Utschick, and G. Dietl. Spatial Resource Allocation for the Multiuser Multicarrier MIMO Broadcast Channel - A QoS Optimization Perspective. In *Proc. of International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2010.

[109] A.M. Tulino and S. Verdu. *Random Matrix Theory and Wireless Communications*. Number 1 in Foundations and Trends in Communications and Information Theory. Now Publishers, 2004.

[110] D.N.C. Tse and S.V. Hanly. Linear Multiuser Receivers: Effective Interference, Effective Bandwidth and User Capacity. *IEEE Transactions on Information Theory*, 45(2):641–657, 1999.

[111] J. Evans and D.N.C. Tse. Large System Performance of Linear Multiuser Receivers in Multipath Fading Channels. *IEEE Transactions on Information Theory*, 46(6):2059–2078, September 2000.

[112] M.J.M. Peacock, I.B. Collings, and M.L. Honig. *Advances in Multiuser Detection*, volume 1 of *Wiley Series in Telecommunications and Signal Processing*, chapter 4 Performance with Random Signatures. John Wiley & Sons, 2009.

[113] M.J.M. Peacock, I.B. Collings, and M.L. Honig. Unified Large-System Analysis of MMSE and Adaptive Least Squares Receivers for a Class of Random Matrices. *IEEE Transactions on Information Theory*, 52(8):3567–3600, August 2006.

[114] M.J.M. Peacock, I.B. Collings, and M.L. Honig. Eigenvalue Distributions of Sums and Products of Large Random Matrices Via Incremental Matrix Expansions. *IEEE Transactions on Information Theory*, 54(5):2123–2138, May 2008.

[115] A. Lozano and A.M. Tulino. Capacity of Multiple-Transmit Multiple-Receive Antenna Architectures. *IEEE Transactions on Information Theory*, 48(12):3117–3128, December 2002.

[116] P. Kazakopoulos, P. Mertikopolous, A.L. Moustakas, and G. Caire. Distribution of MIMO Mutual Information: A Large Deviations Approach. In *Proc. of IEEE Information Theory Workshop on Networking and Information Theory (ITW)*, pages 306–310, 2009.

[117] V.K. Nguyen and J.S. Evans. Multiuser Transmit Beamforming via Regularized Channel Inversion: A Large System Analysis. In *Proc. of Global Telecommunications Conference (GLOBECOM)*, 2008.

[118] C.S. Vaze and M.K. Varanasi. A large-system Analysis of the Imperfect-CSIT Gaussian Broadcast Channel with a DPC-based Transmission Strategy. In *Proc. of 44th Annual Conference on Information Sciences and Systems (CISS)*, 2010.

[119] W. Santipach and M.L. Honig. Capacity of Multiple-Antenna Fading Channel With a Quantized Precoding Matrix. *IEEE Transactions on Information Theory*, 55(3):1218–1234, March 2009.

[120] H. Huh, A.M. Tulino, and G. Caire. Network MIMO Large-System Analysis and the Impact of CSIT Estimation. In *Proc. of 44th Annual Conference on Information Sciences and Systems (CISS)*, 2010.

[121] S.-H. Moon, H. Huh, Y.-T. Kim, G. Caire, and I. Lee. Weighted Sum Rate of Multi-Cell MIMO Downlink Channels in the Large System Limit. In *Proc. of International Conference on Communications (ICC)*, 2010.

[122] H. Huh, G. Caire, S.-H. Moon, and I. Lee. Multi-cell MIMO Downlink with Fairness Criteria: the Large System Limit. In *Proc. of International Symposium on Information Theory (ISIT)*, pages 2058 – 2062, 2010.

[123] R. Zakhour and S.V. Hanly. Large System Analysis of Base Station Cooperation on the Downlink. In *Proc. of 48th Annual Allerton Conference on Communication, Control and Computing*, pages 270–277, 2010.

[124] H. Dai and H.V. Poor. Asymptotic Spectral Efficiency of Multicell MIMO Systems With Frequency-Flat Fading. *IEEE Transactions on Signal Processing*, 51(11):2976–2988, November 2003.

[125] A.M. Tulino, A. Lozano, and S. Verdu. Impact of Antenna Correlation on the Capacity of Multiantenna Channels. *IEEE Transactions on Information Theory*, 51(7):2491 – 2509, July 2005.

[126] W. Hachem, P. Loubaton, and J. Najim. Deterministic Equivalents for Certain Functionals of Large Random Matrices. *The Annals of Applied Probability*, 17(3):875–930, 2007.

[127] J. Dumont, W. Hachem, S. Lasaulce, P. Loubaton, and J. Najim. On the Capacity Achieving Covariance Matrix for Rician MIMO Channels: An Asymptotic Approach. *IEEE Transactions on Information Theory*, 56(3):1048–1069, 2010.

[128] V. Marčenko and L. Pastur. Distribution of Eigenvalues for Some Sets of Random Matrices. *Math USSR Sbornik*, 1:457–483, 1967.

[129] C. Guthy, W. Utschick, and M.L. Honig. Large System Analysis of Projection Based Algorithms for the MIMO Broadcast Channels. In *Proc. of International Symposium on Information Theory (ISIT)*, pages 2128 –2132, 2010.

[130] P. Tejera, W. Utschick, G. Bauch, and J.A. Nossek. Sum-Rate Maximizing Decomposition Approaches for Multiuser MIMO OFDM. In *Proc. of IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2005.

[131] F.B. Hildebrand. *Introduction to Numerical Analysis*. International Series in Pure and Applied Mathematics. McGrawHill Book Company, 1956.

[132] A. Dotzler, W. Utschick, and G.Dietl. Linear Successive User Allocation in the Multi-Cell MIMO Environment. In *Proc. of IEEE Wireless Communications and Networking Conference (WCNC)*, March 2011.

[133] A. Dotzler, W. Utschick, and G. Dietl. Zero-Forcing based Interference Coordination for MISO Networks. In *Proc. of 73rd Vehicular Technology Conference (VTC Spring)*, May 2011.

[134] A. Wiesel, Y.C. Eldar, and S. Shamai. Zero-Forcing Precoding and Generalized Inverses. *IEEE Transactions on Information Theory*, 56(9):4409–4418, September 2008.

[135] J.R. Silvester. Determinants of Block Matrices. *Maths Gazette*, 84:268–271, 2000.

[136] A.W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and Its Applications*, volume 143 of *Mathematics in Science and Engineering*. Academic Press, 1979.

[137] Z.D. Bai and J.W. Silverstein. No Eigenvalues Outside the Support of the Limiting Spectral Distribution of Large Dimensional Sample Covariance Matrices. *Annals of Probability*, 26(1):316–345, 1998.

[138] R. Speicher. Free Convolution and the Random Sum of Matrices. *Publ. Res. Inst. Math. Sc.*, 29:731–744, 1993.