Antti Tölli

RESOURCE MANAGEMENT IN COOPERATIVE MIMO-OFDM CELLULAR SYSTEMS

FACULTY OF TECHNOLOGY, DEPARTMENT OF ELECTRICAL AND INFORMATION ENGINEERING, UNIVERSITY OF OULU
ANTTI TÖLLI

RESOURCE MANAGEMENT IN
COOPERATIVE MIMO-OFDM
CELLULAR SYSTEMS

Academic dissertation to be presented, with the assent of
the Faculty of Technology of the University of Oulu, for
public defence in Raahensali (Auditorium L10), Linnanmaa,
on April 11th, 2008, at 12 noon

OULUN YLIOPISTO, OULU 2008
Tölli, Antti, Resource management in cooperative MIMO-OFDM cellular systems
Faculty of Technology, Department of Electrical and Information Engineering, University of Oulu,
P.O.Box 4500, FI-90014 University of Oulu, Finland
Oulu, Finland

Abstract

Radio resource management techniques for broadband wireless systems beyond the existing cellular systems are developed while considering their special characteristics such as multi-carrier techniques, adaptive radio links and multiple-input multiple-output (MIMO) antenna techniques. Special focus is put on the design of linear transmission strategies in a cooperative cellular system where signal processing can be performed in a centralised manner across distributed base station (BS) antenna heads.

A time-division duplex cellular system based on orthogonal frequency division multiplexing (OFDM) with adaptive MIMO transmission is considered in the case where the received signals are corrupted by non-reciprocal inter-cell interference. A bandwidth efficient closed-loop compensation algorithm combined with interference suppression at the receiver is proposed to compensate for the interference and to guarantee the desired Quality of Service (QoS) when the interference structure is known solely at the receiver.

A greedy beam ordering and selection algorithm is proposed to maximise the sum rate of a multiuser MIMO downlink (DL) with a block zero forcing (ZF) transmission. The performance of the block-ZF transmission combined with the greedy scheduling is shown to approach the sum capacity as the number of users increases. The maximum sum rate is often found to be achieved by transmitting to a smaller number of users or beams than the spatial dimensions allow. In addition, a low complexity algorithm for joint user, bit and power allocation with a low signalling overhead is proposed.

Different linear transmission schemes, including the ZF as a special case, are developed for the scenario where the cooperative processing of the transmitted signal is applied to users located within a soft handover (SHO) region. The considered optimisation criteria include minimum power beamformer design; balancing the weighted signal-to-interference-plus-noise ratio (SINR) values per data stream; weighted sum rate maximisation; and balancing the weighted rate per user with additional QoS constraints such as guaranteed bit rate per user. The method can accommodate supplementary constraints, e.g., per antenna or per BS power constraints, and upper/lower bounds for the SINR values of the data streams. The proposed iterative algorithms are shown to provide powerful solutions for difficult non-convex transceiver optimisation problems.

System level evaluation is performed in order to assess the impact of a realistic multi-cell environment on the performance of a cellular MIMO-OFDM system. The users located in the SHO region are shown to benefit from greatly increased transmission rates. Consequently, significant overall system level gains result from cooperative SHO processing. The proposed SHO scheme can be used for providing a more evenly distributed service over the entire cellular network.

Keywords: capacity, cellular systems, convex optimisation, cooperative communication, linear transceiver design, multiuser MIMO-OFDM, non-reciprocal interference, quality of service, radio resource management, scheduling
To my grandmother Aune
Preface

This thesis represents a culmination of learning and research work that has taken place over a period of four years (2003-2007) at the Centre for Wireless Communications (CWC), Department of Electrical and Information Engineering, University of Oulu, Oulu, Finland. I want to thank Dr. Ian Oppermann, Prof. Matti Latva-aho and Lic. Tech. Ari Pouttu, the directors of CWC during the course of the thesis work, for giving me the opportunity to work in such a magnificent and inspiring working environment.

I have derived enormous personal and scientific benefit from the time spent at the CWC, both from the people who work there and the working environment created by them. Especially, I am grateful to my supervisor Prof. Markku Juntti, first, for convincing me to pursue towards a doctoral degree at the CWC, and second, for his invaluable scientific support and the opportunity to write this thesis under his guidance. I wish to express my gratitude to the reviewers Professor Wolfgang Utschick from Technical University of Munich, Germany, and Associate Professor Mats Bengtsson from Royal Institute of Technology, Stockholm, Sweden, for thorough reading and constructive criticism on the thesis manuscript. Samuli Kangaslampi is acknowledged for proofreading the manuscript.

I was fortunate to be part of two exceptional research projects, namely, the Future Radio Access (FUTURA) and Packet Access Networks with Flexible Spectrum Use (PANU). I thank the project managers of these projects, Prof. Jari Iinatti, Dr. Kari Hooli, Dr. Kimmo Kansanen and Dr. Pekka Pirinen, as well as all fellow researchers and representatives of the industrial partners for their support, and for providing sufficient degrees of freedom to pursue towards my academic goals. I have greatly enjoyed the time spent with my long-term office mate, Dr. Marian Codreanu, whom I would like to thank especially for his friendship, help and the excellent co-operation during these years.

I thank all past and present colleagues at the CWC and Telecommunication Laboratory for the joyful working atmosphere and the scientific, and most often maybe not so scientific, discussions. In particular, I would like to express my thanks to Prof. Carlos Pomalaza-Ráez, Dr. Attaphongse (Pong)
Taparugssanagorn, Davide Macagnano, Esa Kunnari, Dr. Giuseppe Abreu, Jari Ylioinas, Juha Karjalainen, Kaveh Ghaboosi, Dr. Kenta Umebayashi, Markus Myllylä, Mehdi Bennis, Petri Komulainen, Dr. Robert Hsieh and Dr. Toshiyuki Nakanishi, with whom I also shared many memorable moments outside the regular office hours. I have also had the privilege to guide the master’s theses of Mikko Alatossava and Maria Kangas. The administrative support of Elina Komminaho, Sari Luukkonen, Laila Kuhalampi, Mari Lehmikangas, Tero Suutari, Timo Äikäs, Antero Kangas and Jari Sillanpää is gratefully recognised.

The thesis work has been financially supported by Infotech Oulu Graduate School during the period of 2003-2007. Partial funding through the projects was provided by the Finnish Funding Agency for Technology and Innovation, the Finnish Defence Forces, Nokia, Elektrobit and Instrumentointi, which I highly acknowledge. I was also fortunate to receive personal grants for doctoral studies from the following Finnish foundations: Nokia Oyj:n säätiö, Oulun yliopiston tukisäätiö, Tauno Tönningin säätiö, Riitta ja Jorma J. Takasen säätiö, Seppo Säynäjänkunaan tiedesäätiö and Kaupallisten ja teknillisten tieteiden säätiö. These contributions further stimulated me to maintain the focus on the research work, which I greatly appreciate.

I would like to express my deepest gratitude to my closest network, including my parents Marja-Leena and Juhani, grandfather Olavi, all brothers, sisters, siblings and other relatives, who are too numerous to be acknowledged individually. Thank you for creating such a diverse, and sometimes even chaotic, atmosphere of mutual caring. My warmest appreciation belongs to my beloved life-companion Hanna, whose cheerful and brisk presence has helped me to disconnect from the work when I needed it the most.

Oulu, March 14, 2008

Antti Tölli
Symbols and abbreviations

\(\cdot^*\) complex conjugate of the argument
\(\cdot^*\) a solution of an optimisation problem
\(|\cdot|\) absolute of the argument
\(\cdot^T\) transpose of the argument
\(\cdot^H\) complex conjugate transpose (Hermitian) of the argument
\(\hat{\cdot}\) estimate of the argument
\(\otimes\) Kronecker product
\(|\mathcal{A}|\) cardinality of the set \(\mathcal{A}\)
\(|X|\) determinant of the matrix \(X\)
\(X^{-1}\) inverse of the matrix \(X\)
\(X^{1/2}\) Hermitian square root of the Hermitian matrix \(X\), i.e., \(X^{1/2}X^{1/2} = X\)
\([X]_{i,j}\) the \((i,j)\)th entry of the matrix \(X\)
\(CN(m, C)\) complex circularly symmetric Gaussian vector distribution with mean \(m\) and covariance matrix \(C\)
\(0\) zero matrix; a subscript can be used to indicate the dimension
\(I\) identity matrix; a subscript can be used to indicate the dimension
\(\text{blockdiag}\{\cdot\}\) block diagonal matrix of the argument matrices
\(\text{diag}(\mathbf{x})\) diagonal matrix with the elements of vector \(\mathbf{x}\) on the main diagonal
\(E(\cdot)\) expectation
\(\text{Im}(\cdot)\) imaginary part
\(\log(\cdot)\) logarithm in base 2
\(\log_{10}(\cdot)\) logarithm in base 10
\(\max(\cdot)\) maximum
\(\min(\cdot)\) minimum
\(\text{Re}(\cdot)\) real part
\(\text{vec}(\cdot)\) vec-operator: if \(\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_n]\), then \(\text{vec}(\mathbf{X}) = [\mathbf{x}_1^T, \ldots, \mathbf{x}_n^T]^T\)
\((x)^+\) positive part of \(x\), i.e., \(\max(0, x)\)
\(|x|\) absolute value (magnitude) of the scalar \(x\)
\(\sim\) distributed according to
\(\|X\|_2\) Frobenius norm of matrix \(X\)
\text{tr}(X) \quad \text{trace of matrix } X
\\triangleq \quad \text{defined as }
\succeq \quad \text{generalised inequality in a proper cone \cite[Sect. 2.4]{15}; between real vectors, it represents component-wise inequality; between Hermitian matrices, } A \succeq B \text{ means that } A - B \text{ is positive semidefinite}
\succeq_{SOC} \quad \text{inequality with respect to the second-order cone \cite{15}, i.e., for any } x \in \mathbb{R} \text{ and } y \in \mathbb{R}^n, [x, y^T]^T \preceq_K 0 \text{ is equivalent to } x \succeq \|y\|_2
\mathbb{R}^n_+ \quad \text{the set of } m \times n \text{ complex matrices}
\mathbb{R}^n \quad \text{the cone of nonnegative } n \text{-dimensional real vectors (the } n\text{-dimensional nonnegative orthant)}
\mathbf{a}_c \quad \mathbb{C}^{d-1} \text{ vector to form all possible linear combinations of the equivalent channel vectors of the } l - 1 \text{ beams for sub-carrier } c
\mathbb{A}_n \quad \text{arbitrary subset of transmit antennas}
\mathbf{a}_{b,k} \quad \text{large scale fading coefficient between BS } b \text{ and user } k
\mathbf{b}_{l,c} \quad l\text{th beam index for sub-carrier } c
\mathbf{c}_k \quad \text{multiplicative constant for monomial local approximation for user } k
\mathbf{C}_{k,c} \quad \mathbb{C}^{N_t M_k \times N_r M_k} \text{ covariance matrix of the signal transmitted to user } k \text{ on the } c\text{th sub-carrier, } \mathbf{C}_{k,c} = E\left[\tilde{x}_{k,c}\tilde{x}_{k,c}^H\right] = \mathbf{M}_{k,c}\mathbf{M}_{k,c}^H
\mathbf{C}_{k,c}^{eq} \quad \mathbb{C}^{m_{k,c} \times m_{k,c}} \text{ equivalent transmit covariance matrix, } \mathbf{Q}_{k,c}^{eq} = \mathbf{V}_{k,c}^H\mathbf{C}_{k,c}\mathbf{V}_{k,c}^T
\mathbf{d}_{k,c} \quad \text{vector of normalised complex data symbols transmitted to user } k \text{ at subcarrier } c, \mathbf{d}_{k,c} = [d_{1,k,c}, \ldots, d_{m_{k,c},k,c}]^T
\mathbf{d}_{k,l,c} \quad \text{vector of normalised complex data symbols excluding stream } l \text{ transmitted to user } k \text{ at subcarrier } c, \mathbf{d}_{k,l,c} = [d_{k,1,c}, \ldots, d_{k,l-1,c}, d_{k,l+1,c}, \ldots, d_{k,m_{k,c},k,c}]^T
D \quad \text{number of user drops in system simulations}
\mathbf{e}_s \quad \text{exponent of the best monomial local approximation for stream } s
\mathbf{f}_0(\cdot) \quad \text{signomial function to be approximated by a monomial function}
\mathbf{F}_{k,c} \quad \mathbb{C}^{N_n k \times m_{k,c}} \text{ matrix of receive beamformers for iterative BD algorithm}
g_{s,i} \quad \text{coupling coefficient between data streams } s \text{ and } i, g_{s,i} = |\mathbf{w}_s^H\mathbf{H}_k\mathbf{v}_i|^2
\( H_{b,k,c} \) normalised channel matrix from BS \( b \) to user \( k \)

\( H_{\text{noise}}^{b,k,c} \) estimation noise matrix

\( \tilde{H}_{k,c}^{b,k,c} \) global channel matrix from \( M_k \) BSs to user \( k \),

\( \tilde{H}_{k,c}^w \) whitened global channel matrix, \( \tilde{H}_{k,c}^w = R_{k,c}^{-\frac{1}{2}} \tilde{H}_{k,c} \)

\( \tilde{H}_{k,c}^{\text{eq}} \) equivalent channel matrix, \( \tilde{H}_{k,c}^{\text{eq}} = R_{k,c}^{-\frac{1}{2}} U_{k,c} \Lambda_{k,c}^\frac{1}{2} \)

\( \tilde{H}_{k,c}^\dagger \) equivalent channel matrix used in BD algorithm, \( \tilde{H}_{k,c}^\dagger = F_{k,c}^\dagger \tilde{H}_{k,c}^w \)

\( \bar{H}_{c} \) stacked channel matrix of all users for sub-carrier \( c \),

\( \bar{H}_{k,c} \) equivalent channel matrix after step \( l-1 \) including the equivalent channel vectors \( \bar{h}_{k,c} \) of the previously selected \( l-1 \) users or beams for each sub-carrier \( c \)

\( \tilde{H}_{k,c} \) stacked channel matrix of all users excluding user \( k \), \( \tilde{H}_{k,c} = [\tilde{H}_{k,c}^1, \ldots , \tilde{H}_{k,c}^{l-1,c}, \tilde{H}_{k,c}^{l+1,c}, \ldots , \tilde{H}_{k,c}^{K,c}]^T \)

\( \tilde{H}_{k,c}^{\text{inter}} \) equivalent channel matrix for user \( k \), \( \tilde{H}_{k,c}^{\text{inter}} = \tilde{H}_{k,c}^{\dagger} \tilde{V}_{k,c}^{(0)} \)

\( \tilde{H}_{k,c}^{\text{intra}} \) intra-cell interference received by user \( k \) at subcarrier \( c \)

\( I_0^{\text{(inst)}} \) average interference power density

\( I_k^{\text{(inst)}} \) instantaneous mutual information for user \( k \)

\( I_k^{\text{(BD)}} \) instantaneous sum rate of the BD method

\( I_k \) lower bound of instantaneous mutual information for user \( k \)

\( I \) ordered set of indices

\( I_{l-1,c} \) set of previously selected \( l-1 \) beam indices for sub-carrier \( c \)

\( K \) total number of users

\( l_{\min} \) minimum code word length for the bit and power loading algorithm

\( m(\cdot) \) monomial function approximating a signomial function

\( m_{k,c} \) number of active data streams for user \( k \) at subcarrier \( c \), \( m_{k,c} \leq \min(N_T M_k, N_{R_k}) \)

\( m_{\text{tot}} \) total number of data streams at subcarrier \( c \), \( m_{\text{tot}} = \sum_k m_{k,c} \)

\( m_{k,l,c} \) TX beamforming vector for user \( k \), beam \( l \) at subcarrier \( c \)

\( m_s \) TX beamforming vector for stream \( s \)
\[ m_s^{[n]} \] TX beamforming vector for stream \( s \) associated with antenna group \( n \)

\[ M_{k,c}^{[n]} \] \( q^{NT} \) pre-coding matrix for user \( k \) at subcarrier \( c \)

\[ M_{k,c}^{[n]} \] \( q^{NT \times m_k,c} \) pre-coding matrix for user \( k \) at subcarrier \( c \) that corresponds to \( n \)th base station belonging to \( S_k \)

\( M_k \) soft handover active set size \( |S_k| \) for user \( k \)

\( MCS_{k,l,c} \) modulation and coding scheme assigned to user \( k \), beam \( l \) at subcarrier \( c \)

\( \mathbf{n}_{k,c} \) \( \mathcal{CN}(0, N_0 \mathbf{I}_{N_t}) \) additive noise sample vector for user \( k \) at subcarrier \( c \)

\( N_0 \) noise power spectral density

\( N_B \) number of base stations

\( N_C \) number of sub-carriers

\( N_{R_k} \) number of receive antennas for user \( k \)

\( N_{R_{tot}} \) total number of receive antennas, \( N_{R_{tot}} = \sum_{k=1}^{K} N_{R_k} \)

\( N_T \) number of transmit antennas

\( p_{k,l,c} \) power allocated to user \( k \), beam \( l \) at subcarrier \( c \)

\( p'_{k,l,c} \) power allocated to user \( k \), beam \( l \) at subcarrier \( c \)

\( p_s \) power allocated to stream \( s \)

\( P_{offset} \) power offset value for user \( k \)

\( P_{step} \) step size for the power offset

\( \mathbf{P}_{k,c} \) \( \mathbb{R}^{m_k,c \times m_k,c} \) power allocation matrix for user \( k \) at subcarrier \( c \),

\( \mathbf{P}_{k,c} = \text{diag}(p_{k,1,c}, \ldots, p_{k,m_k,c,c}) \)

\( \mathbf{P}'_{k,c} \) \( \mathbb{R}^{m_k,c \times m_k,c} \) power allocation matrix for user \( k \) at subcarrier \( c \)

\( P_n \) the maximum transmit power for \( n \)th antenna set

\( P_{sum} \) sum/pooled power constraint, \( P_{sum} = MP_T \)

\( P_R \) signal power received by user \( k \) from BS \( b \), \( P_R = a_{b,k} P_T \)

\( P_T \) base station transmit power

\( \mathcal{P}_k \) subset of data streams that correspond to user \( k \)

\( Q \) diagonal covariance matrix of uncertain noise, \( Q = \text{diag}(q_1, \ldots, q_1, q_2, \ldots, q_2, \ldots, q_M, \ldots, q_M) \), where each \( q_n \) is repeated \( N_T \) times

\( r_{max} \) maximum rate limit per stream

\( r_s \) rate of data stream \( s \)

\( r_k \) rate of user \( k \)
\( r_{\text{min}}^k \) minimum rate requirement for user \( k \)

\[ r_o = \min_{k \in \mathcal{U}} \beta_k^{-1} \sum_{s \in \mathcal{P}} \log_2 (1 + \gamma_s) \]

\( \hat{r}_o \) reference rate

**\( R(\beta) \)** weighted sum rate of the data streams, \( \sum_{s=1}^{S} \beta_s r_s \)

**\( R_{k,c} \)** \( \mathbb{C}^{N_k \times N_k} \) covariance matrix of the inter-cell interference-plus-noise for user \( k \) at subcarrier \( c \)

**\( R_{\text{MIMO}} \)** \( \mathbb{C}^{N_T \times N_T} \) bi-spatial correlation matrix, \( R_{\text{MIMO}} = R_{\text{TX}} \otimes R_{\text{RX}} \)

**\( R_{\text{RX}} \)** \( \mathbb{C}^{N_k \times N_k} \) receiver correlation matrix

**\( R_{\text{TX}} \)** \( \mathbb{C}^{N_T \times N_T} \) transmit correlation matrix

**\( S \)** set of all base stations, \( S = \{1, \ldots, N_B\} \)

**\( \mathcal{S}_k \)** soft handover active set of BSs for user \( k \)

**\( t_{k,l,c} \)** \( \mathbb{C}^{N_k} \) \( l \)th column vector of \( T_{k,c} \)

**\( T_{k,c} \)** \( \mathbb{C}^{N_k \times m_{k,c}} \) equivalent channel matrix of the desired signal for user \( k \) at subcarrier \( c \), \( T_{k,c} = \tilde{H}_{k,c} M_{k,c} \)

**\( \tilde{T}_{k,l,c} \)** equivalent channel matrix of interfering beams for beam \( l \) at subcarrier \( c \), \( \tilde{T}_{k,l,c} = [t_{k,1,c}, \ldots, t_{k,l-1,c}, t_{k,l+1,c}, \ldots, t_{k,m_{k,c},c}] \)

**\( u_{k,l,c} \)** \( \mathbb{C}^{N_k} \) \( l \)th column vector of \( U_{k,c} \)

**\( u'_{k,l,c} \)** \( \mathbb{C}^{N_k} \) \( l \)th column vector of \( U'_{k,c} \)

**\( U_{k,c} \)** \( \mathbb{C}^{N_k \times m_{k,c}} \) left singular vectors of \( \tilde{\mathbf{H}}_{k,c}^{w} \)

**\( U'_{k,c} \)** \( \mathbb{C}^{N_k \times m_{k,c}} \) left singular vectors of \( \tilde{\mathbf{H}}_{k,c}^{l} \)

**\( \hat{\mathbf{U}}_{k,c} \)** left singular vectors of \( \mathbf{H}_{k,c} \)

**\( \hat{\mathbf{U}}_{k,c}^{(1)} \)** first \( m_{k,c} \) left singular vectors of \( \mathbf{H}_{k,c} \)

**\( \mathcal{U} \)** set of users

**\( \mathcal{U}_b \)** set of users allocated to BS \( b \)

**\( \mathcal{U}_{\text{GBR}} \)** subset of the full user set \( \mathcal{U} \) including users with minimum bit rate requirements \( r_k \)

**\( v_{k,l,c} \)** \( \mathbb{C}^{N_T M_k} \) normalised TX beamforming vector for user \( k \), beam \( l \) at subcarrier \( c \)

**\( v_{k,l,c}^{[n]} \)** \( \mathbb{C}^{N_T} \) normalised TX beamforming vector for the \( l \)'th stream of user \( k \) from BS \( S_k(n) \)

**\( V_{k,c} \)** \( \mathbb{C}^{N_T M_k \times m_{k,c}} \) normalised pre-coding matrix for user \( k \) at subcarrier \( c \), \( V_{k,c} = [v_{k,1,c}, \ldots, v_{k,m_{k,c},c}] \)
\( V'_{k,c} \) normalised pre-coding matrix for user \( k \) at subcarrier \( c \), includes \( m_k \) first columns of \( \tilde{V}'_{k,c} \)

\( \hat{V}_{k,c} \) right singular vectors obtained by SVD of the pre-whitened channel matrix \( \tilde{H}_{k,c} \)

\( \hat{V}'_{k,c} \) right singular vectors obtained by SVD of the channel matrix \( \tilde{H}_{k,c} \)

\( \tilde{x}(0)_{k,c} \) rightmost singular vectors of \( \tilde{H}_{k,c} \)

\( \tilde{x}(1)_{k,c} \) leftmost singular vectors of \( \tilde{H}_{k,c} \)

\( w_s \) normalised LMMSE weight vector for stream \( s \)

\( W_{k,c} \) LMMSE weight matrix for user \( k \) at subcarrier \( c \)

\( X_{b,k} \) matrix of transmitted symbols from BS \( b \) to user \( k \)

\( x_{b,k,c} \) signal vector transmitted from BS \( b \) to user \( k \) at subcarrier \( c \)

\( x'_{b,c} \) total transmitted signal vector from BS \( b \) at subcarrier \( c \)

\( \tilde{x}_{k,c} \) global signal vector transmitted to user \( k \) at subcarrier \( c \) from \( M_k \) BSs being in SHO active set \( S_k \), \( \tilde{x}_{k,c} = [X_{S(1),k,c}^T, \cdots, X_{S(M_k),k,c}^T]^T \)

\( Y_k \) matrix of symbols received by user \( k \)

\( y_{k,c} \) signal vector received by user \( k \) at subcarrier \( c \)

\( Z_{k,c} \) covariance matrix of the intra-cell interference for user \( k \) at subcarrier \( c \)

\( \alpha \) power reduction factor

\( \beta_k, \beta_s \) priority weight for user \( k \), or data stream \( s \)

\( \gamma_{k,l,c} \) SINR for spatial sub-channel \( l \) of user \( k \) at subcarrier \( c \)

\( \gamma_{\text{max}} \) maximum value for the SINR range

\( \gamma_{\text{min}} \) minimum value for the SINR range

\( \gamma_s \) SINR value for stream \( s \)

\( \gamma_o \) common weighted SINR value, \( \gamma_o = \gamma_s / \beta_s \)

\( \Gamma \) SNR gap to the channel capacity

\( \delta \) hysteresis value

\( \epsilon \) accuracy of the stopping criteria

\( \zeta \) spectral efficiency value

\( \theta_k \) weights of GBR users for the feasibility check, \( \theta_k = r_{k}^{\text{min}} / r_{\text{ref}} \)

\( \kappa \) target BER value (after the second turbo decoding iteration)
measured BER value (after the second turbo decoding iteration)

\[ \tilde{\kappa} \]

\[ \lambda_{k,l,c} \]  
\[ l \text{th eigenvalue of user } k \text{ at sub-carrier } c \]

\[ \dot{\lambda}_{k,l,c} \]  
\[ l \text{th eigenvalue of user } k \text{ at sub-carrier } c \]

\[ \hat{\lambda}_{k,l,c} \]  
\[ l \text{th diagonal element of } \hat{\Lambda}_{k,c} \]

\[ \hat{\lambda}_{\text{max}} \]  
\[ \text{maximum eigenvalue of } \mathbf{R}_{k,c} \text{ } \forall c, \hat{\lambda}_{\text{max}} = \max_{l,c} \hat{\lambda}_{k,l,c} \]

\[ \tilde{\hat{\lambda}}_{k,l,c} \]  
\[ l \text{th eigenvalue of } \mathbf{R}_{k,c} \]

\[ \tilde{\Lambda}_{k,c} \]  
\[ C_{N_R k} \times M_k \text{ eigenvalue matrix of user } k \text{ at sub-carrier } c \]

\[ \tilde{\Lambda}'_{k,c} \]  
\[ C_{N_R k} \times M_k \text{ eigenvalue matrix of user } k \text{ at sub-carrier } c \]

\[ \Lambda_{k,c} \]  
\[ \mathbb{C}_{m_k \times m_k} \text{ eigenvalue matrix of user } k \text{ at sub-carrier } c \text{ including the first } m_{k,c} \text{ eigenvalues of } \tilde{\Lambda}_{k,c}, \Lambda_{k,c} = \text{diag} \left( \lambda_{k,1,c}, \ldots, \lambda_{k,m_{k,c},c} \right) \]

\[ \Lambda'_{k,c} \]  
\[ \mathbb{C}_{m_k \times m_k} \text{ eigenvalue matrix of user } k \text{ at sub-carrier } c \text{ including the first } m_{k,c} \text{ eigenvalues of } \tilde{\Lambda}'_{k,c} \]

\[ \tilde{\bar{\Lambda}}_{k,c} \]  
\[ \text{eigenvalues of } \tilde{\hat{\mathbf{H}}}_{k,c} \]

\[ \tilde{\hat{\Lambda}}_{k,c} \]  
\[ m_{k,c} \times m_{k,c} \text{ diagonal eigenvalue matrix of } \tilde{\hat{\mathbf{H}}}_{k,c} \]

\[ \mu \]  
\[ \text{"water level" for water-filling} \]

\[ \mu' \]  
\[ \text{"water level" for water-filling without full interference knowledge} \]

\[ \bar{\mu} \]  
\[ \text{"water level" for heuristic water-filling with per BS power constraints} \]

\[ \rho \]  
\[ \text{power reduction factor, } \rho = \sqrt{\alpha} \]

\[ \sigma_{\text{est}} \]  
\[ \text{standard deviation of the estimation noise} \]

\[ \phi \]  
\[ \text{trust region parameter for signomial approximation} \]

\[ \psi \]  
\[ \text{arbitrary rate reward vector where the weights } \psi = [\psi_1, \ldots, \psi_K]^T \text{ are ordered in a descending order} \]

\[ v \]  
\[ \text{multiplier for } P_{\text{step}} \]

\[ \varphi \]  
\[ \text{phase mismatch value} \]

\[ \omega_{b,k} \]  
\[ \text{power ratio between the averaged received interference from the interference source } b \text{ and } N_0 \text{ for user } k, \omega_{b,k} = P_T a_{b,k}^2 / N_0 \]

\[ b \]  
\[ \text{base station index} \]

\[ c \]  
\[ \text{sub-carrier index} \]

\[ i \]  
\[ \text{interferer index} \]

\[ j \]  
\[ \text{iteration index} \]

\[ k \]  
\[ \text{user index} \]

\[ k_s \]  
\[ \text{user index associated with stream } s \]

\[ l \]  
\[ \text{spatial stream index, per-user processing} \]
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>base station or antenna group index in the SHO active set</td>
</tr>
<tr>
<td>$r$</td>
<td>receive antenna index</td>
</tr>
<tr>
<td>$s$</td>
<td>stream index, per-stream processing</td>
</tr>
<tr>
<td>$t$</td>
<td>transmit antenna index</td>
</tr>
<tr>
<td>$z$</td>
<td>time index</td>
</tr>
</tbody>
</table>

- **2G**: second generation cellular systems
- **3G**: third generation cellular systems
- **3GPP**: Third Generation Partnership Project
- **4G**: fourth generation cellular systems
- **ACK**: acknowledge character
- **AWGN**: additive white Gaussian noise
- **BER**: bit error rate
- **BC**: broadcast channel
- **BD**: block diagonalisation
- **BLAST**: Bell Labs space-time architecture
- **BS**: base station
- **CDF**: cumulative distribution function
- **CDMA**: code division multiple-access
- **CP**: cyclic prefix
- **CSI**: channel state information
- **DFE**: decision feedback equalizer
- **DL**: downlink
- **DPC**: dirty-paper coding
- **DSL**: digital subscriber line
- **DVB**: digital video broadcasting
- **EGPRS**: enhanced general packet radio service
- **ETSI**: European telecommunications standards institute
- **FB**: feedback
- **FDD**: frequency division duplex
- **FDMA**: frequency division multiple access
- **FFT**: fast Fourier transform
- **FL**: fully loaded
- **FUTURA**: Future Radio Access
- **GP**: geometric program
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>GBR</td>
<td>guaranteed bit rate</td>
</tr>
<tr>
<td>GS</td>
<td>greedy scheduling</td>
</tr>
<tr>
<td>HCC</td>
<td>high correlated channel</td>
</tr>
<tr>
<td>HH</td>
<td>Hughes-Hartogs</td>
</tr>
<tr>
<td>HIPERLAN</td>
<td>High Performance Radio LAN</td>
</tr>
<tr>
<td>HSPA</td>
<td>high-speed packet access</td>
</tr>
<tr>
<td>IEEE</td>
<td>Institute of Electrical and Electronics Engineers</td>
</tr>
<tr>
<td>IFFT</td>
<td>inverse fast Fourier transform</td>
</tr>
<tr>
<td>IID</td>
<td>independent and identically distributed</td>
</tr>
<tr>
<td>IP</td>
<td>internet protocol</td>
</tr>
<tr>
<td>ISI</td>
<td>inter-symbol interference</td>
</tr>
<tr>
<td>ITU</td>
<td>International Telecommunication Union</td>
</tr>
<tr>
<td>IWF</td>
<td>iterative waterfilling</td>
</tr>
<tr>
<td>LA</td>
<td>link adaptation</td>
</tr>
<tr>
<td>LAN</td>
<td>local area network</td>
</tr>
<tr>
<td>LOS</td>
<td>line-of-sight</td>
</tr>
<tr>
<td>LSO</td>
<td>low signalling overhead</td>
</tr>
<tr>
<td>LTE</td>
<td>long-term evolution</td>
</tr>
<tr>
<td>MAC</td>
<td>multiple access channel</td>
</tr>
<tr>
<td>MAI</td>
<td>multiple access interference</td>
</tr>
<tr>
<td>MC</td>
<td>multi-carrier</td>
</tr>
<tr>
<td>MCS</td>
<td>modulation and coding scheme</td>
</tr>
<tr>
<td>MF</td>
<td>matched filter</td>
</tr>
<tr>
<td>ME</td>
<td>maximum eigenvalue (scheduling)</td>
</tr>
<tr>
<td>MIMO</td>
<td>multiple-input multiple-output</td>
</tr>
<tr>
<td>MISO</td>
<td>multiple-input single-output</td>
</tr>
<tr>
<td>MMSE</td>
<td>minimum mean square error</td>
</tr>
<tr>
<td>M-QAM</td>
<td>multilevel quadrature-amplitude modulation</td>
</tr>
<tr>
<td>MSE</td>
<td>mean square error</td>
</tr>
<tr>
<td>MU</td>
<td>multi-user</td>
</tr>
<tr>
<td>NAK</td>
<td>negative-acknowledge character</td>
</tr>
<tr>
<td>NGBR</td>
<td>non-guaranteed bit rate</td>
</tr>
<tr>
<td>NLOS</td>
<td>non-line-of-sight</td>
</tr>
<tr>
<td>OFDM</td>
<td>orthogonal frequency division multiplexing</td>
</tr>
<tr>
<td>OFDMA</td>
<td>orthogonal frequency division multiple access</td>
</tr>
</tbody>
</table>

17
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDF</td>
<td>probability density function</td>
</tr>
<tr>
<td>PL</td>
<td>partially loaded</td>
</tr>
<tr>
<td>QAM</td>
<td>quadrature-amplitude modulation</td>
</tr>
<tr>
<td>QoS</td>
<td>quality of service</td>
</tr>
<tr>
<td>QPSK</td>
<td>quadrature phase-shift keying</td>
</tr>
<tr>
<td>RF</td>
<td>radio frequency</td>
</tr>
<tr>
<td>RX</td>
<td>receiver</td>
</tr>
<tr>
<td>SCM</td>
<td>spatial channel model</td>
</tr>
<tr>
<td>SDMA</td>
<td>space division multiple access</td>
</tr>
<tr>
<td>SDP</td>
<td>semidefinite program</td>
</tr>
<tr>
<td>SE</td>
<td>spectral efficiency</td>
</tr>
<tr>
<td>SHO</td>
<td>soft handover</td>
</tr>
<tr>
<td>SIC</td>
<td>successive interference cancellation</td>
</tr>
<tr>
<td>SINR</td>
<td>signal-to-interference-plus-noise ratio</td>
</tr>
<tr>
<td>SIMO</td>
<td>single-input multiple output</td>
</tr>
<tr>
<td>SISO</td>
<td>single-input single-output</td>
</tr>
<tr>
<td>SNR</td>
<td>signal-to-noise ratio</td>
</tr>
<tr>
<td>SOC</td>
<td>second-order cone</td>
</tr>
<tr>
<td>SOCP</td>
<td>second-order cone program</td>
</tr>
<tr>
<td>SVD</td>
<td>singular value decomposition</td>
</tr>
<tr>
<td>TCM</td>
<td>trellis-coded modulation</td>
</tr>
<tr>
<td>TDD</td>
<td>time division duplex</td>
</tr>
<tr>
<td>TDMA</td>
<td>time division multiple access</td>
</tr>
<tr>
<td>TX</td>
<td>transmitter</td>
</tr>
<tr>
<td>UL</td>
<td>uplink</td>
</tr>
<tr>
<td>WF</td>
<td>water-filling</td>
</tr>
<tr>
<td>WCDMA</td>
<td>wideband code division multiple access</td>
</tr>
<tr>
<td>WLAN</td>
<td>wireless local-area network</td>
</tr>
<tr>
<td>WMAN</td>
<td>wireless metropolitan-area network</td>
</tr>
<tr>
<td>ZF</td>
<td>zero forcing</td>
</tr>
</tbody>
</table>
Contents

Preface 7
Symbols and abbreviations 9

1 Introduction 23
   1.1 Requirements for future communication systems ................. 23
   1.2 Systems design .................................................. 24
   1.3 Cooperative MIMO-OFDM cellular system .......................... 27
   1.4 Scope and objectives of the thesis ................................ 29
   1.5 Outline of the thesis ............................................ 30
   1.6 Author’s contributions to the publications ....................... 31

2 Literature review 33
   2.1 Adaptive modulation for SISO fading channels .................... 33
   2.2 MIMO channel capacity .......................................... 36
       2.2.1 Point-to-point MIMO communications ....................... 36
       2.2.2 Multiuser MIMO communications ............................ 40
   2.3 Resource allocation for multiuser MIMO systems .................. 42
   2.4 MIMO communications with co-channel interference .............. 47
   2.5 Distributed MIMO systems with cooperation between BS antennas ............................................. 49
   2.6 MIMO transceiver design with linear processing ................ 52

3 System model 59
   3.1 Cooperative MIMO-OFDM cellular system with distributed base station antennas ........................... 59
   3.2 Simplified system models ....................................... 63
   3.3 Model for channel estimation errors ............................ 64

4 Compensation of non-reciprocal inter-cell interference 65
   4.1 Single user transmitter design .................................. 65
       4.1.1 Optimal single user eigenmode transmission ............... 67
       4.1.2 Interference structure known only at receiver ............ 68
       4.1.3 Bit and power loading algorithm ........................ 71
   4.2 Closed-loop interference non-reciprocity compensation algorithm ... 72
4.3 Numerical examples ................................................ 75
  4.3.1 Mutual information results .................................. 77
  4.3.2 Link level results ........................................... 78
  4.3.3 System level evaluation ..................................... 83
4.4 Summary and discussion ........................................... 89

5 Resource allocation with block-ZF transmission ...............91
  5.1 Multi-user zero forcing with iterative block diagonalisation .......... 92
  5.2 Scheduling algorithms ........................................... 95
    5.2.1 Greedy beam selection and ordering algorithm .............. 96
    5.2.2 Selection of optimal number of beams ...................... 98
    5.2.3 Greedy beam allocation with Hughes-Hartogs loading .... 99
    5.2.4 Low complexity bit, power and beam allocation with low
         signalling overhead ...................................... 102
    5.2.5 Compensation for estimation deficiencies at the
         transmitter ............................................... 103
  5.3 Numerical examples ............................................. 104
    5.3.1 Performance of block-ZF transmission with greedy
         scheduling ............................................. 104
    5.3.2 User, bit and power allocation with low signalling overhead .. 108
  5.4 Summary and discussion ......................................... 114

6 BS cooperation with linear transceiver processing .............117
  6.1 Joint design of linear TX and RX beamformers with per BS or
       antenna power constraints .................................. 118
    6.1.1 SINR balancing ......................................... 119
    6.1.2 Weighted sum rate maximisation .......................... 122
    6.1.3 User rate balancing ..................................... 127
    6.1.4 Sum power minimisation with individual user rate
           constraints ............................................ 129
    6.1.5 Multiuser MIMO-OFDM system with different quality of
           service classes ......................................... 130
  6.2 Multiuser ZF solution with per BS or antenna power constraints . 136
    6.2.1 Weighted sum rate maximisation with cooperative ZF
           transmission ............................................ 136
    6.2.2 Cooperative ZF transmission with QoS constraints .......... 137
    6.2.3 Practical considerations .................................. 137
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.3 Symmetric capacity subject to per BS or antenna power constraints</td>
<td>139</td>
</tr>
<tr>
<td>6.4 Numerical examples</td>
<td>140</td>
</tr>
<tr>
<td>6.4.1 Mutual information results</td>
<td>140</td>
</tr>
<tr>
<td>6.4.2 Link level results</td>
<td>156</td>
</tr>
<tr>
<td>6.4.3 Impact of non-synchronisation between BS antenna heads</td>
<td>157</td>
</tr>
<tr>
<td>6.4.4 System level evaluation</td>
<td>158</td>
</tr>
<tr>
<td>6.5 Summary and discussion</td>
<td>165</td>
</tr>
</tbody>
</table>

7 Conclusion and future work                                           167


1 Introduction

The fast development of wireless communication services is reflected in the transition from the second generation (2G) systems to the third generation (3G) cellular systems. Advanced features of the 2G systems, such as enhanced general packet radio service (EGPRS) [16], and the 3G systems [17, 18] are in active commercial use. Yet, expectations are constantly growing, and the existing networks will not be able to fulfil the high data rate and low latency requirements of future communication services. Therefore, the spectral efficiency of future wireless networks needs to be further improved allowing for increased flexibility to serve a large number of simultaneous users and different services.

This thesis concentrates on developing efficient radio resource management techniques for future cellular networks beyond the existing 2G and 3G systems. Special focus is put on the design of transmission strategies in a cooperative cellular system where signal processing can be performed in a centralised manner across distributed base station (BS) antenna heads. In Section 1.1, the general requirements for future communication systems are presented. Section 1.2 introduces the principal radio techniques that have to be considered when designing future radio systems. Section 1.3 describes the concept of a cooperative cellular system with distributed antenna heads. Section 1.4 presents the scope and aims of the thesis, whereas the outline of the thesis is given in Section 1.5. Finally, Section 1.6 describes the author’s contributions to the original publications.

1.1 Requirements for future communication systems

The third generation partnership project (3GPP) has taken a step towards higher data rates through the introduction of high-speed packet access (HSPA) for 3G systems [18]. 3GPP is also carrying out a study that focuses on the long-term evolution (LTE) of 3G [19], which aims at an evolved radio access technology that can provide service performance comparable to that of current fixed line accesses [20]. Other evolution paths towards higher data rates may involve wireless local- or metropolitan-area network (WLAN, WMAN) [21, 22] type of solutions with increased capacity and coverage as well as improved support
for mobility and Quality of Service (QoS), such as the IEEE 802.16e WiMAX standard [23].

In parallel to the activities related to 3G evolution, there is also an increased research effort on future radio access beyond 3G, generally referred to as fourth generation (4G) radio access. The latter is aimed at providing a ubiquitous system, fully based on the internet protocol (IP), where a wide range of services can be offered at a reasonable cost with quality of service (QoS) comparable to wireline technologies. A 4G system is anticipated to provide data rates of up to 100 Mb/s with wide-area coverage and up to 1 Gb/s with local-area coverage, thus fulfilling the requirements set by the International Telecommunication Union (ITU) for beyond 3G systems [20, 24]. A large part of European research activities on 4G have been gathered into the WINNER project [25], while other parallel activities have been carried out, e.g. in the Wireless World Research Forum [26] or the Chinese FuTuRE project [27, 28].

1.2 Systems design

The primary source of performance degradation in wireless cellular communication systems is the multiuser interference received by the antennas [29]. Apart from external interference and noise, signals are corrupted by time-varying fading, which can be further divided into large- and small-scale fading [29–32]. Large-scale fading represents the change of average path loss due to motion over large geographical areas and it is usually modelled with an experimental model [30, 33]. Small-scale fading, on the other hand, refers to the fast changes in signal level caused by an order of half a wavelength changes in the distance between the transmitter and the receiver. The transmitted signal may propagate over multiple reflective paths which causes fluctuations in the amplitude, phase, and angle of arrival of the received signal. This phenomenon is generally referred to as "multipath fading" causing frequency selective fading [29, 31, 34].

Inter-symbol interference (ISI) imposed by multipath fading can be efficiently mitigated, for example, via the use of an equaliser, as is done in the existing 2G and 3G cellular systems [35, 36]. Weinstein and Ebert [37] proposed already in 1971 an ISI resilient transmission method, later named orthogonal frequency division multiplexing (OFDM), where the transmitted data is modulated via an inverse fast Fourier transformer (IFFT) element and demodulated at the receiver
with an FFT element. Together with cyclic prefix (CP) insertion, the OFDM modem transforms the frequency-selective fading channel into a set of flat fading orthogonal sub-channels and eliminates the ISI between subsequent data blocks as long as the CP is chosen to be longer than the maximum excess delay of the channel [38, 39]. This significantly simplifies the channel equalisation at the receiver in comparison with conventional single-carrier modulation. Thus, OFDM has developed into a popular scheme for both existing and future wideband communication systems [19–22, 25].

The main limitation for wireless communication systems, apart from the transmit power and the complexity of terminal devices, is the fact that the common bandwidth must be shared between multiple users [30, 32]. Traditionally, the scarce radio resources have been shared between users in an orthogonal manner in time, frequency and/or code domains, using time-, frequency- or code division multiple access (TDMA, FDMA or CDMA), respectively [30, 32]. The multicarrier transmission can be also combined with any type of multiple access to provide separation of multiple users. Orthogonal frequency division multiple access (OFDMA) combines FDMA with OFDM by assigning different numbers of orthogonal sub-channels to different users, thus providing flexibility to allocate differentiated QoS to different users.

Apart from OFDM(A) transmission, the principal radio techniques which have to be considered when developing future radio systems include multiple-input multiple-output (MIMO) communication based on multiple antennas both at the transmitters (TX) and the receivers (RX) [40–44], as well as, adaptive modulation and coding [45]. The spectral efficiency (SE) of MIMO transmission can be dramatically increased if some level of channel state information (CSI) is available at the transmitter [32, 42, 44], allowing the system to effectively adapt to and take advantage of the available spectrum and the radio channel. The main challenge is to make the CSI available at the transmitter. This can be achieved by conveying CSI as feedback information over the reverse link. However, providing full CSI via feedback may cause an excessive overhead on the reverse link, and hence, it is difficult to realise in practice [32].

A time division duplex (TDD) system uses the same carrier frequency alternately for transmission and reception, and thus, the CSI can be tracked at the transmitter provided that fading is sufficiently slow. In a TDD system with adaptive MIMO transmission, the modulation parameters in the communication
link from base station (BS) to terminal and vice versa, i.e. the downlink (DL) and the uplink (UL), can be adapted according to the channel conditions. Due to channel reciprocity, the DL channel can be estimated accurately during the previous UL frame assuming that the frame length is shorter than the channel coherence time; a frame refers here to a TDD frame which is divided into UL and DL transmission parts. This assumption is mostly valid in environments with low mobility, e.g. in pedestrian metropolitan environments [46]. In order to provide timely channel estimates for DL transmission, the served users should have at least some minimum signalling (training sequences) in the UL direction. Yet, at the same time, the signalling eats up valuable uplink resources (capacity, battery power), especially if the terminals have no data to transmit1. On the other hand, with transmitter CSI, a large part of the signal processing load can be shifted to the BS both in the UL and the DL. This allows for a simple terminal design.

Even though the CSI can be made available at the transmitter via time duplexing, the actual interference structure (interference level, frequency, time and space selectivity) observed by the multi-antenna terminal receiver in the DL can be very different from that measured by the BS receiver in the UL. In other words, the interference in the UL and the DL is usually non-reciprocal, unlike the communication channel itself. In such a case, the obtained QoS at the receiver may differ significantly from the one desired if the transmission parameters are assigned on the basis of the reverse link measurements only. In theoretical studies, the other-cell-interference seen at the receiver is often assumed to be perfectly known at the transmitter in addition to the perfect channel knowledge both at the transmitter and the receiver [47–51]. However, the required signalling feedback would make the ideal feedback approach unpractical in most applications. For example, in practical adaptive MIMO–OFDM cellular systems, the ideal feedback would require the interference covariance matrix to be reported to the transmitter for each subcarrier and for each transmitted data frame.

The allocation of resources among multiple users can be performed jointly across time, frequency and space domains using advanced multi-user MIMO resource allocation and scheduling techniques in both the UL and the DL. The

---

1The terminals, however, will send at least some acknowledgement packets in response to downlink traffic.
particular challenge of the DL direction is that while the BS has the ability to coordinate transmission from all of its antennas, the receiver antennas are grouped among different users that are typically unable to coordinate the transmission with each other. While dirty-paper coding (DPC) is known to be a capacity achieving albeit very complex non-linear precoding technique in the DL [52–56], linear precoding/beamforming [57, 58] is much simpler to implement to perform multi-user transmission. Hence, the linear beamforming is an important solution in practical system design. In general, a solution for any allocation problem with linear transmission can be divided into user selection or grouping for each orthogonal dimension (frequency/sub-carrier, time), and the linear transceiver optimisation for the selected set of users per orthogonal dimension subject to a power constraint. The optimal user selection per each orthogonal dimension is generally a difficult combinatorial problem with integer constraints [59–61]. Therefore, sub-optimal allocation algorithms are commonly used in practice.

1.3 Cooperative MIMO-OFDM cellular system

A network infrastructure which is based on cooperative processing across distributed base station antenna heads has received significant interest in recent literature [27, 28, 62–74]. Although BS cooperation naturally increases system complexity, it has potentially significant capacity and coverage benefits, making it worth a more detailed consideration. The assumption is that cooperative signal processing can be performed in a centralised manner so that the MIMO antenna heads are distributed over a larger geographical area (e.g. hundreds of meters), as illustrated in Fig. 1. The distributed antenna heads are connected to the central processing unit (base station) via radio over fibre technology or wireless microwave links [27, 28], for example. The actual radio frequency and base band processing as well as radio resource and network management functionalities are included in the base station units. The antenna elements can be grouped strategically so that both coverage and capacity needs can be fulfilled adaptively according to the specific needs.

In order to attain the full CSI between all users and BS antennas in the cellular network, the channels should be jointly estimated at each BS. In practical TDD
MIMO–OFDM cellular systems, however, UL transmissions from adjacent cells can be significantly more attenuated compared to the own cell users. Therefore, the joint channel estimation may be difficult if not impossible to implement in practice. In this thesis, a somewhat more practical case is considered where the joint cooperative processing of the transmitted signal from several MIMO BS antenna heads is restricted to an area where the users have comparable signal strengths from adjacent BS antenna heads. Similarly to the soft handover (SHO) feature in wideband code division multiple access (WCDMA) systems [17], the SHO region is defined for users with similar received power levels from adjacent distributed BS antenna heads (a ±3 dB SHO window, for example). Since the signal processing of the BS antenna heads is concentrated at one central controller, joint resource allocation and beamforming from all the antennas belonging to the "active set" can be performed to the user(s) within the SHO region, as seen in Fig. 2 on page 61. The transmissions for other users outside the SHO region are seen as interference.

It is possible to serve several users having identical SHO active sets in the
same time-frequency transmission slot using space division multiple access (SDMA) methods. SDMA can be used to improve the utilisation of the physical resources (space, time, frequency) by exploiting the available spatial degrees of freedom in a downlink multi-user MIMO channel, at the expense of somewhat increased complexity.

Since each distributed antenna head is provided with a separate power amplifier, power constraints per antenna head have to be considered when designing cooperative transceivers for distributed antenna systems. In practice, each antenna element has a separate power amplifier in general, and hence, the power constraints per antenna need to be considered as well.

1.4 Scope and objectives of the thesis

The scope of this thesis in a broad sense is to develop methods to efficiently convey wireline packets over the wireless last segment between the BS’s and the mobile devices. In particular, issues pertaining to resource management for future wireless systems are examined while considering their special characteristics such as multi-carrier techniques, adaptive radio links and MIMO antenna techniques. When developing new wireless systems, it is essential to study the requirements and connections of both the physical layer (radio interface) and the data link layer (resource allocation) together. The objective here is to study the system level performance of future wireless systems and services in a realistic multi-cell environment by using both theoretical analysis and simulations.

Apart from the traditional cellular arrangements, special focus is put on how to apply multi-user (MU) resource allocation techniques with linear transceiver processing to the distributed MIMO antenna concept. Moreover, the impact of channel estimation uncertainties and a possible non-reciprocity between the UL and the DL interference to the scheduling decisions is studied, as well as the compensation of the increased inter-beam and/or inter-user interference. The air interface option considered is TDD with the orthogonal frequency division multiple access (OFDMA) concept which is suitable especially for low mobility urban environments with low-to-medium coverage (from tens to hundreds of meters).

Different cooperative transmission schemes with linear transceiver processing, including zero forcing (ZF) as a special case, are developed for a variety of
optimisation criteria subject to different QoS constraints per user and with per antenna or per BS power constraints. Furthermore, sub-optimal scheduling algorithms based on, e.g. best user selection, largest eigenvalue selection, greedy user or beam selection are considered.

1.5 Outline of the thesis

The thesis is written as a monograph for the sake of clarity, but parts of the contributions in Chapters 4–6 have been published or accepted for publication in thirteen original publications [1–13]. The rest of the thesis follows the organisation given below.

Chapter 2 contains the literature review of previous and parallel work related to the contributions of the thesis. The review includes several topics related to frequency selective MIMO broadcast channels with multiple users and with multiple transmitters, including the channel capacity, optimal and sub-optimal transmitter and receiver design subject to different quality of service criteria, and the allocation of available resources over different dimensions.

Chapter 3 presents a generic system model for the MIMO-OFDM cellular system with cooperative base station antenna heads. The special cases of the generic model including an adaptive MIMO-OFDM cellular system without cooperation between base stations and a single cell MIMO-OFDM system are also defined.

Chapter 4 considers an adaptive MIMO-OFDM system where the received signal is corrupted by non-reciprocal inter-cell interference and the downlink interference structure is known only at the receiver and not at the transmitter. The results have been included in part in [1–4]. A closed-loop compensation algorithm with simple scalar power offset feedback combined with interference suppression at the receiver is proposed. The performance of the proposed framework is analysed and compared to the ideal case where the interference structure per sub-carrier is perfectly known at the transmitter. Both theoretical analysis and more practical link and system level studies are carried out.

Chapter 5 focuses on the downlink of a multi-user MIMO-OFDM system, where the transmitter and receivers are equipped with multiple antennas and accurate TX CSI is available at the TX and the RX of all users. Generalised zero forcing processing at the transmitter is assumed in this chapter. The results have
been presented in part in [5, 6]. A greedy beam ordering and selection algorithm is proposed to maximise the downlink spectral efficiency of the considered system. In addition to the greedy algorithm, an efficient low complexity joint user, bit and power allocation algorithm with low signalling overhead (LSO) is proposed. Moreover, the impact of imperfect channel estimation at the transmitter and the imperfect orthogonalisation of allocated users on the system performance are studied.

In Chapter 6, part of which has been published in [7–14], the joint cooperative processing of the transmitted signal from several MIMO BS’s is considered for users located within a soft handover region. A downlink space-frequency resource allocation problem with per BS and per antenna power constraints is studied for the considered adaptive MIMO-OFDM system. The mathematical framework for the SHO based MIMO-OFDM system is derived and the joint design of linear TX and RX beamformers in a MIMO multi-user transmission is considered. A general method is proposed for the linear transceiver design subject to different optimisation criteria in addition to per BS and/or per antenna power constraints. The proposed joint transceiver optimisation algorithms are compared to corresponding generalised ZF transmission solutions, as well as to optimal non–linear transmission methods. Moreover, the link and system level gains and trade-offs from cooperative SHO processing are investigated.

Chapter 7 concludes the thesis. The main results and conclusions are summarised. Moreover, some remaining open questions and directions for future research are pointed out.

1.6 Author’s contributions to the publications

The thesis is in part based on thirteen original publications, including three journal papers [4, 9, 10] and ten conference papers [1–3, 5–8, 11–13]. The author has had the main responsibility in developing the original ideas, analysis as well as writing all the papers. The author has also implemented the software to perform all the numerical analysis and computer simulations.

The mutual information analysis in [4] was sketched in cooperation with the second author. The optimisation framework in [8, 9, 12, 13] was developed in cooperation with the second author. Section III in [9] was also written partly by the second author. Furthermore, the second author provided the bit and
power loading algorithm utilised in [1–4, 8, 9]. The third author has provided guidance, ideas and criticism during the writing process. In addition to the original publications [1–13], the author coauthored a journal paper [14] describing part of the optimisation framework utilised also in [8–13].

All the results where turbo coding has been applied, have utilised the turbo encoding and decoding software developed by Mikko Vehkaperä. All the simulations with a space-time correlated channel model have been produced with the channel generator created by Esa Kunnari [75]. The spatial channel model (SCM) [76] used in the system level simulations is based on the code developed in the WINNER project [77] and further extended by Mikko Alatossava. All the simulation results involving constrained optimisation utilised the disciplined convex optimisation software CVX developed by Grant, Boyd & Ye [78].
2 Literature review

This chapter reviews the existing literature and parallel work related to the scope of the thesis. The review begins with relevant background information on adaptive modulation for single-input single-output (SISO) fading channels, after which the capacity of MIMO systems is surveyed in Section 2.2. More precisely, Section 2.2.1 focuses on point-to-point MIMO and Section 2.2.2 concentrates on multiuser MIMO communications. Section 2.3 considers the problem of allocating resources among multiple users across the space, frequency and time dimensions. Strategies for designing a MIMO system with co-channel interference are introduced in Section 2.4, while Section 2.5 focuses on distributed MIMO systems with cooperation between BS antennas. Finally, Section 2.6 reviews the relevant literature related to transceiver optimisation for multiuser MIMO DL with linear processing both at the transmitter and the receiver.

2.1 Adaptive modulation for SISO fading channels

Information theory was introduced by Shannon in [79], where he derived the general theory on reliable communication and, in particular, the capacity of the additive white Gaussian (AWGN) channel. In [80], Shannon also showed that the capacity-achieving power allocation for a time-invariant spectrally shaped channel corresponds to a water-filling (WF) distribution [81]. The capacity of time-variant frequency-flat fading channels with full CSI was later analysed by Goldsmith and Varaiya in [82], where they showed that the Shannon capacity of a fading channel can be achieved by varying both the transmission rate and the power [82]. Caire and Shamai [83] demonstrated that the capacity can be achieved with a constant rate Gaussian code by varying the transmit power alone, when the power allocation is of the water-filling type. In the low SNR range, adaptive systems exhibit great performance enhancements compared to systems that do not exploit channel knowledge at the transmitter [82, 83]. With a high SNR, however, the gain from water-filling is reduced to zero. These capacity results can also be extended to the underspread (channel coherence time much larger than the delay spread) time-variant frequency-selective case.
with multi-carrier modulation, where the optimal solution is water-filling over time and frequency [32].

For practical communication systems, however, an arbitrarily small bit error rate (BER) and random infinitely long coding schemes with Gaussian codebooks assumed for the capacity optimal adaptive scheme are non-realisable [84]. Adaptive modulation for a flat-fading single-input single-output (SISO) case under more practical constraints has been extensively studied, e.g. in [85–93] and the references therein. Practical adaptive transmission with CSI feedback was pioneered by Hayes [85] and Cavers [86], who proposed the first adaptive schemes for variable power transmission with a fixed QoS target, and symbol rate variation with a fixed transmit power, respectively. After 20 years of inactivity, several proposals were made to adapt, e.g. the coding rate or scheme [87, 88], or the constellation size [89, 90] according to the instantaneous channel conditions. Goldsmith and Chua [91] developed an optimal uncoded variable-rate variable-power $M$-ary quadrature amplitude modulation (MQAM) scheme for maximising spectral efficiency subject to an average power and BER constraints. They showed that the power adaptation for both the capacity optimal transmission scheme [82] and the adaptive MQAM is the same. Furthermore, the spectral efficiency of the adaptive MQAM technique has the same slope as the capacity, with a constant BER-target dependent power gap to the capacity. It is noteworthy that essentially the same result for the same optimisation criterion had earlier been reported by Kalet [94] in the context of a wired multitone channel, with the exception that the water-filling is performed over a time-invariant multitone channel instead of a time-variant flat-fading channel. The variable MQAM technique was extended also to the case with coded modulation in [92, 93], where the gap to the capacity was further reduced.

Optimal adaptive loading algorithms developed in [91, 94] assumed continuous rate allocation (infinite granularity in constellation size) over sub-carriers or fading states. In practical systems, on the other hand, the available constellations are discrete, having finite granularity imposed by a finite set of modulation and coding schemes (MCS). An optimal discrete bit and power loading algorithm was developed by Hughes-Hartogs [95], while a simplified close-to-optimal loading algorithm was proposed in [96]. The loading algorithms were originally applied to wired (twisted pair lines) frequency selective channels with multicarrier modulation. However, the very same adaptive loading principles can be applied
to wireless communication with OFDM modulation. In general, adaptive loading algorithms employ different MCS’s for the individual OFDM subcarriers based on their corresponding signal-to-noise ratios (SNR), while excluding some of the most faded subcarriers. A comprehensive survey on adaptive techniques applied to wireless multicarrier communication was carried out by Keller and Hanzo [97].

Until recently, the fading in a wireless communication channel had been considered as an adverse phenomenon. This is due to the fact that the error probability in a fading channel is significantly higher than in a non-fading AWGN channel [29, 31, 98]. Different techniques utilising time, frequency or space diversity were traditionally used to mitigate the detrimental effect of fading. Time diversity is obtained via the interleaving of coded symbols over transmission blocks, whereas frequency diversity is achieved from multipath combining by using, e.g. the rake receiver or the equaliser. Multiple antennas at the receiver and/or the transmitter have traditionally been used to provide space or spatial diversity, i.e. redundancy across independently fading antennas [30–32].

When the transmitter(s) and the receiver(s) have the ability to track the wireless channel, the system throughput can be greatly increased by opportunistic communications. In such a case, channel fading is actually beneficial for communications, providing multiuser diversity [99, 100]. This is achieved by the adaptive allocation of resources among multiple users, transmitting at a high rate when the channel is good and vice versa [99, 101, 102]. For single-antenna systems, the optimum strategy for maximising the capacity is to allow only the user with the best instantaneous channel gain to transmit at any time [99, 100, 103]. This extends straightforwardly to the frequency-selective fading channel as well. The optimal sum-capacity achieving strategy is OFDMA where the user with the best instantaneous channel gain transmits on each sub-carrier and the optimal power allocation is water-filling over time and frequency [32, 103]. Similarly, discrete loading algorithms, such as the Hughes-Hartogs (HH) algorithm [95], can be applied to this setup to maximise the spectral efficiency. Another useful optimisation criterion is to find an optimal sub-carrier, bit and power allocation to minimise the total transmission power while satisfying a minimum rate constraint per user [102]. This is, however, a far more complicated combinatorial problem with integer constraints. Wong et al. [102] reformulated the original problem as a convex problem with relaxed non-integer constraints, and provided a close-to-optimal allocation algorithm based on the achieved lower bound
2.2 MIMO channel capacity

The pioneering work by Winters et al. [104–106] proposed the use of space division multiple access (SDMA) to boost up the capacity of wireless communication systems. The multiple antennas in the UL enable spatial separation of the signals from the different users, and hence, allow several users to simultaneously communicate with the BS. By the mid-nineties the very same idea was transferred to point-to-point communication with multiple antennas at both the transmitter and the receiver [40, 44, 107]. It was noticed that SDMA with single transmit antennas is in fact similar to point-to-point MIMO communications without CSI at the transmitter, i.e. users/antennas cannot co-operate. This resulted in large potential for the use of the angular or space domain to convey multiple independent data streams from a single user to the BS. In MIMO communications, the signals transmitted from co-located antennas can still be separated at the receiver provided that the scattering environment is rich enough. A minimum mean square error (MMSE) receiver with successive interference cancellation (SIC) was shown to be an information theoretically optimal solution for both SDMA with single transmit antennas and MIMO without CSI [32].

2.2.1 Point-to-point MIMO communications

The research on point-to-point MIMO communications was pioneered by Telatar [44, 107] and Foschini [40, 108]. Foschini considered the case where the CSI of the MIMO channel is only available at the receiver and not at the transmitter. For such a case, he also proposed a capacity achieving transmission architecture called the Bell Labs space-time architecture (BLAST) in [108]. Both Foschini and Telatar also provided an outage capacity analysis for the slow fading MIMO channel and demonstrated that in ideal conditions the capacity can grow linearly with the minimum of transmit and receive antennas. In addition to the analysis without the CSI at the transmitter, Telatar [44, 107] showed that with full CSI at the transmitter, the MIMO channel can be converted to parallel, non-interfering SISO sub-channels through a singular value decomposition (SVD) of the channel matrix. The number of parallel sub-channels or data streams is
dictated by the rank of the MIMO matrix, and the maximum rank is equal to
the minimum of transmit and receive antennas. The optimal capacity achieving
transmission is then carried out by pre and post combining each stream with
the right and left singular vectors of the channel matrix, respectively. Optimal
transmit power allocation is achieved via water-filling [81] over the parallel
SISO sub-channels with gains corresponding to the eigenvalues of the channel
matrix [44, 107]. Furthermore, the capacity-optimal transmission strategy
requires a Gaussian codebook with continuous rate allocation among the parallel
sub-channels [44, 81, 107].

Extensions of the MIMO capacity to multipath channels were provided
in [42]. The SVD based transmission has a clear similarity to the OFDM system,
where the frequency selective (multipath) channel is transformed into a set of
parallel independent sub-channels. Capacity optimal power allocation for a
MIMO-OFDM system is achieved via simultaneous water-filling over space and
frequency [42, 81]. This achieves the capacity of the frequency selective MIMO
channel as the number of sub-carriers approaches infinity [42]. An overview of
the Shannonian capacity limits of MIMO channels is provided in [109]. Other
general overviews on MIMO communications are included in [41, 43, 110–112]

The MIMO capacity depends heavily on the available channel knowledge at
both the receiver and the transmitter, the signal-to-noise-plus-interference ratio
(SINR) and the underlying channel properties, e.g, the correlation between the
antenna elements. The large capacity gains associated with MIMO channels are
based on the assumption that each transmit-receive antenna pair experiences
independent identically distributed (IID) fading. This can be only achieved
in a rich scattering environment [109]. MIMO transmission and the capacity
available in non-ideal conditions, e.g. correlated fading channels, were studied
in [113–115]. For MIMO-OFDM systems, Bölcskei et al. [116] analysed the
influence of propagation parameters and system parameters on capacity and
demonstrated the beneficial impact of delay spread and angle spread on capacity.

While optimal transceiver design with ideal CSI at the transmitter is rather
simple [44, 107], the case without CSI at the transmitter is less straightforward.
In general, multiple antennas can be used for increasing the amount of diversity
or the number of spatial multiplexing dimensions in wireless communication
systems [32]. There is a multitude of different techniques designed for extracting
the maximal diversity gain or the maximal spatial multiplexing gain of a channel.
Based on the original BLAST scheme by Foschini [108], several other spatial multiplexing schemes were introduced, e.g. in [116–118]. Different multi-antenna schemes aiming at maximising the available diversity have been proposed, e.g. in [119–124]. Alamouti [119] proposed a simple but elegant space-time coding technique, which turned out to be optimal from both the diversity and multiplexing perspectives for the case with two transmit and one receive antennas [125]. The Alamouti scheme was generalised to orthogonal designs with any number of transmit antennas in [122]. Unlike the Alamouti scheme, the coding structure from the orthogonal design [122], while indeed achieving the full diversity order, reduces the achievable spatial multiplexing gain [125]. Zheng and Tse [125] proved that a part of the diversity and multiplexing gains can be obtained simultaneously. Furthermore, they characterised the optimal trade-off between the type of gain achievable by any coding scheme. Since then, several trade-off optimal space-time codes have been proposed in the literature. See, for example, [126–128], and the references therein.

Perfect CSI at the transmitter is often difficult to achieve in real networks. Ideally, full CSI at the transmitter can be achieved in TDD systems, where the reciprocal UL and DL channels are time-multiplexed on the same physical wireless channel [110, 111]. The transceiver can obtain the CSI while receiving information in the current time slot and can use the same CSI to transmit in the next time slot, as long as the TDD frame length is shorter than the channel coherence time. In general, this can be guaranteed in low mobility environments for practical system parameters [46]. In high mobility environments, on the other hand, the CSI becomes quickly outdated as the terminal velocity increases due to a time delay between the estimation of the channel and the transmission of the data. CSI is typically obtained via channel estimation. A channel estimate is always a noisy version of the real channel. Training based channel estimation schemes and the losses due to estimation errors and the pilot overhead were studied in [129]. Furthermore, the impact of the channel estimation errors at both the transmitter and the receiver on the capacity was studied in [130]. Transmitter optimisation with noisy channel estimates is still largely an unresolved research problem. In some solutions it has been proposed to use worst case design criteria to guarantee robust performance for any realisation of the actual and estimated channels, e.g. worst-case mean square error (MSE) precoder design [131, 132], or to combine robust beamforming with space-time coding [133].
In frequency division duplex (FDD) systems, the full CSI knowledge at the transmitter requires an ideal feedback channel from the terminal(s). This, however, results in an enormous overhead due to the number of channel coefficients that need to be quantised and sent back to the transmitter over a limited bandwidth feedback channel. Hence, an ideal feedback is impractical for FDD systems with any mobility. When only some statistical information is available on the MIMO channel (distribution, mean, covariance) at the transmitter, the transmission strategy must be designed based on the statistical information instead of the instantaneous information [101, 109, 134–141]. It was discovered in [101, 109, 136–138, 140, 141] that the capacity-achieving eigenvectors should correspond to the eigenvectors of the statistical covariance matrix of the channel. Not surprisingly, this finding is very similar to the solution with instantaneous channel knowledge at the transmitter [44, 107]. However, the optimal power allocation among the transmit directions is not achieved by water-filling, but requires numerical solutions that resemble water-filling where the inter-stream interference due to non-orthogonal transmission is considered [134, 136, 139, 140]. In order to improve the reliability of systems with statistical or noisy channel information, some authors have proposed to couple the statistical beamforming with space–time block codes [141–145].

A multitude of solutions have also been proposed in the literature for a FDD system utilising a very low rate feedback from the receiver. A simple solution is to select a subset of transmit antennas to maximise the available rate or to minimise the rate with fixed rate requirements [146, 147], or to switch between spatial multiplexing and transmit diversity depending on the instantaneous channel characteristics (rank, correlation between antennas, etc.) [148, 149]. Another solution is to report an index of a transmission strategy (precoder) that has the best match with the instantaneous channel state. The transmit precoder is chosen from a finite set of predefined precoding matrices known to both the receiver and the transmitter [150–153]. Recently, Kim and Skoglund [154] characterised the diversity-multiplexing tradeoff in MIMO channels with quantised CSI at the transmitter.
2.2.2 Multiuser MIMO communications

The capacity region for the Gaussian multiple access channel (MAC), i.e. the UL channel with multiple users/transmitters and a single receiver, has been known for quite a while [155]. Following the pioneering work on the use of multiple receive antennas in the UL by Winters [104–106], the scalar Gaussian MAC capacity region was extended to ISI [103] and MIMO channels in [44, 156].

The capacity region for the Gaussian degraded broadcast channel (BC) has been also known for more than thirty years [81, 157]. The degraded BC channel implies that the Gaussian channel has a scalar input and scalar outputs, i.e. a single-antenna transmitter and several receivers. For such a case, the capacity region is achieved by using superposition coding at the transmitter and interference subtraction at the receivers [81, 157]. However, for the non-degraded BC channel, where the transmitter has a vector input, i.e. multiple transmit antennas, the superposition coding no longer achieves the capacity.

The foundation for the information theoretic studies on the capacity of the non-degraded Gaussian MIMO BC channel was laid in Costa’s landmark paper [158], where he studied the capacity of a BC channel with both additive Gaussian noise and additive Gaussian interference, and where the interference is non-causally known at the transmitter but not at the receiver(s). Costa concluded that the effect of the interference can be completely cancelled out at the transmitter by using specific precoding called dirty paper coding. Consequently, the capacity is identical to the case when the interference was known also at the receivers.

Caire and Shamai [52] proposed to use Costa’s precoding for transmitting over the multi-antenna BC. By comparing DPC to Sato’s cooperative upper bound [159] in a simple case of two users with single-antenna receivers, they showed that DPC achieves the sum capacity. For the case with more than two users, [52] proposed also a sub-optimal precoder design for DPC based on zero forcing (ZF-DPC). For any given user, the ZF part completely suppresses interference caused by subsequently encoded users while the DPC coding is applied to the previously encoded users. This strategy was shown to be asymptotically optimal for both high and low SNR regions. The results of [52] sparked off a number of new studies [53, 54, 56], which generalised the results of [52] to the case of any number of users and an arbitrary number of receive
The duality between the DPC region for the MIMO BC and the capacity region of the MIMO MAC was established in [53, 54]. Both papers utilised the reciprocity of the UL and DL channels [44] combining it with the previously established duality between the scalar Gaussian BC and MAC [160]. They showed that the DPC region is exactly equal to the MAC capacity region, where the users have the same sum power constraint as the MIMO BC. Unlike in the case of the DPC region, the MIMO MAC rate maximisation can be formulated as a concave function of the covariance matrices. Therefore, the DPC-MAC duality allows the DPC region to be found using a standard convex optimisation problem solver [78, 161]. Furthermore, the explicit transformations of transmit covariance matrices from MAC to BC and vice versa were provided in [53].

A somewhat different approach to prove the same result as in [53, 54] was taken in [56], where the sum-rate optimal precoding structure was shown to correspond to a decision-feedback equaliser. In addition, the results in [56] laid the foundation for the studies on the more general case with arbitrary convex input constraints [55, 162, 163]. Yu [162] established a connection between the duality approach [53, 54] and the decision-feedback approach [56], and generalised UL-DL duality to BC channels under arbitrary linear covariance constraints, e.g. per-antenna power constraints. The original DL optimisation problem with linear covariance constraints was transformed into a dual UL minimax optimisation with uncertain noise. The sum capacity result of [162] was extended to the entire BC capacity region with per-antenna power constraints in [163].

The DPC region was finally shown to be indeed the capacity region of the entire non-degraded Gaussian MIMO BC by Weingarten et al. [55]. This result proved the conjecture of the DPC optimality which was already reported in [164, 165]. The MIMO BC capacity region was also characterised in [55] under various constraints on input covariance matrices, including as special cases, the sum and per-antenna power constraints. Thus, [55] parallels the work on the sum capacity and the capacity region with linear covariance constraints reported in [162, 163].

Even though the sum capacity and any point on the boundary of the MIMO BC capacity region can be computed by any standard interior point convex optimisation solver [78, 161], there are a few methods proposed in the literature
that allow for an iterative computation of the sum capacity [166–168] or any point at the boundary [169, 170] without an explicit need to use convex optimisation tools. Based on the iterative water-filling algorithm for MIMO MAC [156], Jindal et al. [167] proposed a sum power iterative water-filling algorithm for Gaussian MIMO BC. This simple algorithm exploits the structure of the dual sum power MAC problem and provides the optimum transmission policies for the MAC, which can easily be mapped to the optimal BC policies by the MAC-BC transformation [165]. An alternative approach based on dual decomposition was proposed by Yu [168]. Despite being less complex than the standard interior point solutions, the convergence of these algorithms can be rather slow, especially for a large number of users. Codreanu et al. [166] proposed a random user pairing technique which was shown to greatly improve the convergence of the iterative water-filling algorithm. Viswanathan et al. proposed a steepest descent method which makes it possible to calculate the weighted rate sum for a given set of weights, i.e. to find any point on the boundary of the BC capacity region [170]. The iterative water-filling approach was also later extended to a more general case with weighted rates by Kobayashi and Caire [169].

Jindal and Goldsmith [171] analysed the asymptotic gains from the optimal DPC compared to the single user capacity with TDMA in a multiple-antenna broadcast channel. They showed that, for Rayleigh fading channels, the gain from DPC largely depends on the ratio between the transmit and the receive antennas, and the total number of users. As a conclusion, the highest DPC gain is achieved with a large number of users, and when the ratio between the transmit and receive antennas is high. On the other hand, the DPC gain converges to unity for both low and high SNRs when the number of receive antennas is higher than or equal to the number of transmit antennas. Furthermore, they demonstrated that transmit beamforming always performs better than or equal to the TDMA, and its achievable rate is upper bounded by the DPC rate.

2.3 Resource allocation for multiuser MIMO systems

A major challenge for wireless communication systems is how to allocate resources among users across the space (including different cells), frequency and time dimensions with different system optimisation objectives. In MIMO-OFDM systems, this leads to an OFDMA solution with a three-dimensional sub-channel
and power allocation problem, i.e. how many sub-channels should be allocated to each user in different dimensions. The problem remains unresolved for a large variety of optimisation criteria, especially when combined with practical modulation and coding schemes [172]. For a single-cell multiuser MIMO system, the optimal sum capacity achieving allocation of resources across different dimensions (users, space, frequency) is given by the actual sum rate capacity of the frequency-selective MIMO BC as discussed in [53, 55, 56]. However, the computation of the sum capacity achieving covariance matrices requires the solving of a convex optimisation problem in dual MAC and transforming the solution back to BC [53]. Especially for frequency selective case with OFDM, the computational complexity becomes very high for an increasing number of sub-carriers, users and antennas. Therefore, sub-optimal but less complex allocation techniques have been proposed [52, 173, 174].

Multiuser diversity for the dirty paper approach with ZF-DPC was studied by Tu and Blum in [173], where they proposed a greedy scheduling algorithm for the selection of users and their encoding order for maximising the sum rate. The greedy user selection and ordering algorithm combined with ZF-DPC was shown to have a sum rate very close to the capacity. Tejera et al. [174] investigated different sub-channel allocation methods that aim at maximising the sum rate of the multiuser MIMO BC. They extended the sub-optimal ZF-DPC with single-antenna users from [52] to a more general case with multiple receive antennas per user. In addition, the greedy approach was extended to the case with multiple receive antennas and it was utilised for allocating user beams on different sub-channels in the space and frequency domains.

The sum rate maximising solutions can occasionally result in a very non-uniform rate allocation between users. Therefore, other transmitter design criteria should be considered in order to guarantee, for example, the instantaneous QoS for all users. The symmetric (balanced) capacity providing absolute fairness between users becomes an important performance metric for delay constrained applications [175, 176]. The weighted symmetric capacity refers to the situation where the weighted user rates are equal, while their rates belong to the boundary of the capacity region [175]. This enables the system to control the rates assigned to users that belong to distinct service priority classes. An iterative algorithm aiming at finding the weighted symmetric capacity for MIMO BC with a sum power constraint was proposed in [175].
The capacity achieving schemes generally require very complex nonlinear precoding based on the DPC [54, 55]. Therefore, research on sub-optimal, but less complex transmission techniques is justified. Linear beamforming [57, 58], also traditionally called as SDMA, is a sub-optimal transmission strategy which enables the spatial separation of several concurrent users. Each user stream is encoded independently and spread over multiple antennas by a beamforming weight vector. Mutual interference between multiple streams is controlled or even completely eliminated by the proper selection of weight vectors. Unlike the sum-rate capacity of MIMO BC using the DPC, the sum rate achieved by optimal beamforming cannot be written as a convex optimisation problem [60]. Therefore, the throughput comparison between the DPC and beamforming is computationally intensive, especially for a large number of users. Yet, despite its sub-optimality, beamforming combined with a proper selection of users has been shown to have the same asymptotic sum-rate as the DPC, when the number of users approaches infinity [59–61]. This is due to a multiuser diversity effect [99, 100, 103], i.e. the probability of finding a set of close-to-orthogonal users with large channel gains increases for a large number of users.

In general, the solution for any sub-optimal allocation problem can be divided into two phases. Firstly, a set of users is selected for each orthogonal dimension (frequency/sub-carrier, time). Secondly, the transceiver is optimised for the selected set of users per orthogonal dimension. The optimal user selection per each orthogonal dimension is a difficult non-convex combinatorial problem with integer constraints [59–61]. Consequently, finding the optimal solution requires an exhaustive search over the entire user set which is computationally prohibitive for a large number of users. A more detailed treatment on the design of the linear transceiver with different optimisation criteria is given in Section 2.6.

Several scheduling algorithms based on, e.g. best user selection, largest eigenvalue selection and greedy user/beam selection, have been proposed for DL beamforming, e.g. in [61, 177–181] and by the author of this thesis in Chapter 5. Dimic and Sidiropoulos [181] utilised the sub-optimal greedy user selection algorithm from [173] for the ZF beamforming with single antenna receivers. Yoo and Goldsmith [61] also used the greedy algorithm with an additional semi-orthogonality test between users and showed that the performance of the ZF beamforming with sub-optimal user selection is still asymptotically optimal. Since the performance of the ZF beamforming is always inferior to optimal

44
linear beamforming, the result in [61] proves the asymptotic optimality of linear beamforming in general.

Often, the allocation problems have been addressed for systems with users having a single receive antenna. When the users are equipped with multiple receive antennas, receiver antenna coordination further enhances the data rates. The signal space of each user has multiple dimensions, allowing for multiple beams to be allocated per user. Therefore, the receiver signal space has to be considered when selecting the optimal sets of users, as well as the dimension and orientation of the signal subspace used by each selected user, for each orthogonal dimension. This further complicates the optimisation problem. Since the transmitter vectors, and, thus, the corresponding receiver vectors at each user are affected by the set of selected users, it is impossible to know the actual receiver structure at the transmitter before the final beam allocation. An obvious candidate for an intelligent initial guess of the receiver matrix is the optimum single user receiver, i.e. the left singular vectors of the channel matrix. This decomposes the system into a MIMO BC with virtual single-antenna users with corresponding channel gains. This type of approach has been taken by the author of this thesis in Chapter 5, as well as in parallel work such as [61, 179]. It was shown in [61] that the performance penalty from non-coordination decreases as the number of users approaches infinity. Thus, antenna coordination is not obligatory for achieving the asymptotically optimal sum rate.

There is a multitude of papers available that provide sub-optimal solutions for the general MIMO-OFDMA resource allocation problem with somewhat relaxed conditions, e.g. [102, 172, 182–190]. The relaxation may involve, for example, zero-forcing transmission which allows for separating the beamformer design from the power allocation making it much easier to solve [182]. Another idea aiming at simplifying the allocation problem is to group users according to the spatial separability or comparability of the user channels for maximising the system throughput [185, 188–190].

Several authors have recognised the importance of the availability of transmitter CSI in multiuser MIMO BC channels [59, 171, 191–193]. Hochwald and Marzetta [191] studied the multiuser diversity gains achievable from MIMO channels without the transmitter CSI and with a simple best user scheduling based on TDMA. They showed that, in such a scenario, the scheduling gains (multiuser diversity) decrease rapidly in relation to an increasing number of
antennas due to the fact that the mutual information fluctuation between the channels of different users is reduced. In contrast to the point-to-point MIMO capacity, where the transmitter CSI availability only affects the SNR offset to the capacity and not the multiplexing gain \[44, 107\], both the multiplexing gain and the rate achievable from the multiuser MIMO broadcast channel are greatly affected by the level of CSI available at the transmitter. Jindal \[192\] demonstrated that the throughput of the multiuser MIMO DL with linear transmit beamforming becomes saturated with imperfect or noisy transmitter CSI. This is due to increased multiuser interference. However, full multiplexing gain can be achieved if the quality of CSI is increased linearly as a function of SNR \[192\].

Random opportunistic beamforming and nulling is a simple but remarkable limited feedback strategy for MIMO DL channels \[59, 193\]. Multiple random orthonormal beams are formed at the BS, and multiple users are simultaneously scheduled on these beams. Each user reports the channel quality metric, i.e. SINR for the strongest beam(s), and the users with the highest instantaneous metric value are scheduled at a time. The opportunistic transmission strategy relies on the fact that with a large number of users, the probability of finding a set of nearly orthogonal users with high channel gains is high \[59, 193\]. This has been shown to achieve asymptotically the performance of linear beamforming \[193\] and to have the same capacity scaling obtained with perfect CSI using the DPC \[59, 60\] as the number of users approaches infinity. However, it may result in very poor performance from both the capacity and fairness point of views, when applied in a system with a low or medium number of users. Therefore, the transmitter CSI is important for systems with a low to moderate number of users, and especially with a large number of transmit antennas \[192\].

In a realistic network with multiple users, the assumption of having full CSI from all users may be overly optimistic. This is due to the excessive overhead required for providing the transmitter with instantaneous CSI. The combination of opportunistic beamforming for the initial user selection from a finite user set and the use of supplementary CSI feedback for the selected users has been proposed in \[14, 194\] allowing for improved optimisation of linear transmit and receive beamformers at the BS and mobile stations, respectively. It was shown in \[14\], co-authored by the author of this thesis, that the supplementary CSI for the selected users greatly improves the performance of the opportunistic
2.4 MIMO communications with co-channel interference

In cellular MIMO systems with interference, there are multiple users located in different cells that suffer from co-channel interference which originates from the transmissions to users in other cells. A similar scenario may arise also in ad hoc networks where each transmitter-receiver pair suffers from the interference originating from other users using the same channel. In general, the interfering signals are assumed to be unknown in systems where co-operation between transmitters is not allowed. In such a case, neither joint transmission from several transmitters nor multiuser detection at the receivers is possible. In the theoretical analysis, both the interfering signals as well as the desired signal are generally modelled as Gaussian distributed which is the usual form of the optimal MIMO signalling [81, 107].

As discussed earlier in Section 2.2.1, the optimum transmit strategy for a single link MIMO system with transmitter CSI but without interference is water-filling over non-interfering SISO sub-channels achieved through a SVD of the channel matrix. Assuming that the interference structure is known at the transmitter, the same result is generalised to a single-link MIMO system with a fixed interference structure, i.e. covariance matrix. Consequently, the parallel sub-channels are obtained by the SVD of the pre-whitened channel matrix [47, 49, 51]. However, when the interference is not fully known at the transmitter, the transceiver design is considerably more challenging.

Point-to-point transmission with an imperfectly known interference structure was studied in [195–197]. These papers considered the worst-case capacity analysis subject to different constraints on the interference structure, such as, the trace constraint on the interference covariance matrix. The results in [196] demonstrated that in some extreme cases with worst case interference, the channel knowledge at the transmitter does not increase the capacity of the MIMO link. On the other hand, Blum et al. [198, 199] studied the MIMO channel capacity with neither interference nor channel knowledge at the transmitter. In such a case, the optimum transmit strategies for the weak and the strong interference regimes are equal power allocation across antennas and single-antenna allocation, respectively.
In cellular systems, the interference originating from intra- and inter-cell users can be controlled by frequency, time and space allocation by a centralised unit in charge of all resources, provided that the CSI and the interference structure are known at the transmitter(s). However, transmit strategies selected at one user/transmitter affect the interference experienced by the other users/transmitters, and hence, their receive strategies. This in turn affects the transmit strategies of the other transmitters. The system capacity can be formulated as neither as a convex nor concave function with respect to the transmit covariance matrices of the users [51]. Thus, due to the non-convexity of the problem, analytical or even numerical solutions are difficult to find.

Sub-optimal signalling methods have been developed in the literature to optimise the mutual information of a MIMO system where the users in different cells suffer from co-channel interference from the users in other cells [48, 50, 51]. Ye and Blum [51] extended the result of [199] and considered a MIMO system with multiple users and perfect CSI (including the interference structure) at both sides. In a sufficiently low interference regime, the users are decoupled and the optimisation problem becomes a concave function of the transmit covariance matrices. On the other hand, the optimum transmission strategy in the large interference regime is to assign a single beam for all users [51]. For the intermediate interference region, the authors in [51] proposed an efficient numerical algorithm based on the gradient projection method.

All studies with transmitter CSI, e.g. [47–51], assumed that the inter-cell interference seen at the receiver is perfectly known at the transmitter in addition to perfect channel knowledge at both the transmitter and the receiver. However, the required signalling feedback would make the ideal feedback approach unpractical in most applications. For example, in practical adaptive MIMO–OFDM cellular systems, the ideal feedback would require that the interference covariance matrix is reported to the transmitter for each subcarrier and for each transmitted data frame. Cepeda et al. [200] recognised the problem of non-reciprocal interference in the case of adaptive modulation in general. A simple feedback method was developed to compensate for it in a flat fading single-antenna transmission scenario. In Chapter 4 of this thesis, an adaptive MIMO–OFDM cellular system is considered, where the received signal is corrupted by non-reciprocal inter-cell interference and the DL interference structure is known only at the receiver and not at the transmitter.
2.5 Distributed MIMO systems with cooperation between BS antennas

Since the introduction of the relaying concept [81, 201, 202], user cooperation has received significant attention in improving the capacity or coverage of wireless communications. In the classic relaying problem, users may cooperate in routing or relaying each others’ data packets. A large number of papers have addressed the relaying problem with different assumptions on the transceiver capabilities. See, e.g. [81, 201–204] and the references therein.

The user cooperation diversity builds upon the classical relay channel model [81, 201, 202], where the space diversity gains can be achieved both at the transmitters and the receivers using a collection of distributed antennas belonging to multiple terminals. Each node is responsible for transmitting not only their own information, but also the information of the other nodes that they receive and detect. Thus, spatial diversity is achieved through the joint use of all the antennas. Sendonaris et al. [205, 206] considered the beamforming approach where different nodes with the transmitter CSI adjust their transmissions so that the transmitted signals add up coherently at the destination. On the other hand, Laneman et al. [207, 208] assumed no CSI at the transmitters and proposed a variety of low-complexity, cooperative diversity protocols that enable wireless terminals to fully exploit spatial diversity in the channel. However, the user cooperation is complicated due to the fact that the (wireless) channel between the nodes is noisy [205, 206]. Therefore, the user cooperation often increases the interference level, the protocol overhead and the complexity of the transceivers [64, 205].

In the ideal case with a noiseless link between the receiving or the transmitting nodes, the node cooperation is simplified to MIMO MAC or MIMO BC with per node power constraints, respectively [205]. In cellular networks, this can be achieved for example by a wired backbone connection between the distributed BS antenna heads, or by highly directional wireless microwave links [27, 28]. In conventional cellular systems, each BS processes the signals of in-cell users independently, treating other users as inter-cell interference. This leads to interference limited behaviour where the interference, experienced especially by the cell-edge users, must be mitigated by sharing and reusing the degrees of freedom available to the network. Conversely, by deploying joint processing
over all the distributed antenna heads, the impact of inter-cell interference can theoretically be fully eliminated.

Network infrastructure based cooperative processing across distributed BS antenna heads [27, 28, 62–74] or fixed relay stations [209–211] has received significant interest in the recent literature. These types of systems have also been considered specifically for voice oriented CDMA communications [212–214]. Wyner [73] considered the UL of a cellular network with joint processing at the global receiver which has access to all the received signals, and optimally decodes all the transmitted data in the entire network. In this simplistic model, each cell senses only the signal radiated from a limited number of neighbouring cells which yields closed form expressions for the achievable rates and allows, to a certain extent, the analytical treatment of the distributed antenna systems with joint-processing. It was shown in [73] that a cellular network with such a joint-processing receiver significantly outperforms a traditional network with individual processing per BS. Joint receiver processing for the UL was further considered in [62, 63, 215–219].

Shamai et al. [70, 71] were among the first to consider the DL sum rate and spectral efficiency optimisation for cooperative MIMO systems with perfect data cooperation between BS’s. They applied the ZF-DPC to the multi-cell joint processing in Wyner’s scenario [73] with an average system power constraint and showed significant capacity enhancements from BS cooperation. For a single-cell scenario with co-located multi-antenna arrays, the power constraints are generally imposed on the total power radiated by all the elements of the array. Conversely, per antenna head power constraints have to be enforced for distributed antenna systems, since each antenna head is provided with a separate power amplifier. In practice, each antenna element may also have a separate power amplifier.

Jafar et al. [215, 216, 220] considered a multi-cell DL channel with perfect CSI at both ends, where an individual power constraint per BS is imposed. They proposed a heuristic but efficient sub-optimal method based on iterative water-filling aimed at maximising the sum-rate throughput of the cooperative system while meeting the individual BS power constraints. The sum capacity and the entire capacity region of the MIMO DL with per antenna or per BS power constraints were discovered in [162] and [55, 163], respectively. These important findings can be utilised for finding the maximum achievable user rates
from the rate region of a cooperative cellular MIMO system with practical peak power constraints per antenna or per BS.

Inspired by the pioneering work of Shamai, Jafar et al., BS cooperation has been studied by several other authors [66, 68, 69, 72, 74, 219, 221]. Somekh et al. [72, 219, 221] provided an information theoretic analysis of distributed antenna systems under the circular Wyner model [73], and derived bounds for the sum-rate capacity supported by the multi-cell DL under per BS power constraints. Karakayali et al. [68, 69] and Dawod et al. [66] studied the cooperative cellular DL using different ZF transmission schemes. In [68, 69], the symmetric (or common) rate maximisation in the cooperative cellular DL with ZF transmission subject to per BS power constraints was formulated as a convex optimisation problem which can be solved efficiently. It must be noted that the theoretical studies above mostly assume perfect and complete CSI for all the transmitters, which is difficult to accomplish in practical cellular networks. A somewhat more practical case is considered by the author of this thesis in Chapter 6, where the joint cooperative processing of the transmitted signal from several MIMO BS antenna heads is restricted to an SHO region.

In order to perform joint transmission from all the distributed BS antenna heads, the baseband signals must have a common carrier phase reference for the base baseband processing. The radio frequency (RF) impairments and the impact of the propagation delay from each of the transmitters to the intended user must be compensated for at the transmitter, for example, by using some feedback from terminals. The BS antenna heads cannot fully synchronise the desired and the interfering signals received by different users due to different propagation times between the BS antenna heads and the user terminals. Zhang et al. [222] showed that significant performance degradation may follow if the asynchronous nature of the multiuser interference is not taken into account when designing the precoder for the cooperative DL. In MIMO-OFDM systems, however, this problem can be handled efficiently as long as the received signal paths are within the guard interval. Specifically, the increase of the delay spread is not necessarily so dramatic if the cooperative processing is limited to an SHO region. The impact of imperfect phase synchronisation between the BS antenna heads on the achievable gains is also addressed with some practical examples in Chapter 6.
2.6 MIMO transceiver design with linear processing

This section reviews the relevant literature related to transceiver optimisation for a multiuser MIMO DL with linear processing both at the transmitter and the receiver. Recall from the previous sections that the optimal sum capacity achieving schemes for the DL require non-linear precoding based on dirty paper coding [54, 55]. On the other hand, linear precoding/beamforming is usually remarkably simpler to implement, and, hence, it is an important solution in practical system design. The maximum number of data streams that can be allocated to the linear transmission system is limited by the number of transmit antennas at the BS, while the number of streams per user is limited by the number of receive antennas at the terminal. The number of streams assigned per user may vary depending on the channel realisation. If the total number of allocated streams exceeds the number of BS antennas, the system becomes interference limited [223]. The assumption in this section is that a fixed set of users has been selected from a possibly larger group of users for each orthogonal dimension prior to the transceiver optimisation.

Convex optimisation methods [15], such as second-order cone programming [224], semidefinite programming [225] and geometric programming [226, 227], are very important and powerful tools which allow for efficient numerical solutions to many signal processing and communications problems – especially optimal linear transceiver design problems. Some communications problems solved via convex optimisation can be found in the tutorial papers [226–229].

The joint design of a linear precoder at the transmitter and an equaliser at the receiver for point-to-point MIMO systems with perfect CSI, according to a variety of design criteria, has received significant attention in the literature. In terms of the spectral efficiency, the system should be designed in such a manner that it achieves the channel capacity, which leads to a diagonalised MIMO channel with water-filling power allocation on the spatial subchannels [44, 107]. Several other design criteria have been proposed in the recent literature in order to consider the quality of the communication link between the transmitter and the receiver given in terms of MSE, SINR, or BER.

In [230–232], the minimisation of the sum of the MSE of all the channel sub-streams (the trace of the MSE matrix) under a sum power constraint was considered in the context of cross-coupled wired communication, e.g. digital...
subscriber lines (DSL). This criterion was later extended to the wireless MIMO system in [233, 234], and generalised to weighted MSE in [235]. It was shown in [230–235] that the optimum structure for the linear precoder and decoder diagonalises the MIMO channel into eigen sub-channels, for any set of error weights. Thus, it differs from the capacity optimal solution only by the power allocated to each sub-channel. In fact, the capacity optimal and the MMSE solutions are connected through the MSE matrix; the first design minimises the determinant and the second minimises the trace of the MSE matrix [234, 236].

In addition to the aforementioned criteria, several other optimisation criteria have been considered for single user MIMO communication [237–243]. Palomar et al. provided a general framework based on convex optimisation for a large variety of MSE-, SINR- and BER-based design criteria subject to different power constraints [238–242]. Moreover, they considered the power minimisation problem subject to QoS constraints given in terms of MSE, SINR, or BER for fixed modulation and coding schemes [243].

In general, the joint design of a linear transceiver for a multiuser MIMO scenario is a much more demanding optimisation problem than the point-to-point scenario. Due to the non-cooperative user terminals, it is often difficult to simplify the optimisation problems to an easily solvable form unlike in the single user case [242]. The optimisation problems employed in the linear beamformer design are not convex in general. Therefore, the problem of finding the global optimum for most of the optimisation criteria is intrinsically nontractable.

The minimum power beamforming design, also known as sum power minimisation under the minimum SINR constraint per stream was the first problem to be solved for the multiuser MIMO DL. The uplink-downlink SINR duality was utilised in [244, 245] to solve this problem for single-antenna terminals. It was shown that the minimum sum power required for satisfying the minimum SINR constraint per stream is equal both in the DL and in the dual UL. The duality property was utilised in [244, 245] to develop iterative algorithms for calculating the optimal beamformers and power allocations. Schubert and Boche [246] provided another solution for the minimum power beamformer design that accounts for the feasibility of the problem setting, as well as an optimal solution for the worst SINR per user maximisation problem under a sum power constraint. They utilised the property of the Perron-Frobenius theory [247] that the UL power assignment for maximising the minimum SINR per user is closely related
to the Perron-Frobenius eigenvalue of the interference cross-coupling matrix between the users. These optimisation criteria were extended to the case of non-linear processing with single-antenna receivers in [248, 249].

Bengtsson and Ottersten [250, 251] presented a different solution for the minimum power beamforming problem of [244–246] based on the framework of the convex optimisation theory. They considered a general case with a statistical DL channel model, where the rank of the estimated channel covariance matrix per user can have a value higher than one due to multi-path propagation. In such a case, the beamforming problem with rank one beamformers is non-convex. However, the rank constraint can be relaxed so that semidefinite programming [225] can be applied. Despite the non-convexity of the problem, SDP relaxation achieves the global optimum of the problem [250, 251]. The relaxed problem can also be modified by introducing additional constraints, e.g. to limit the dynamic range of the transmitted signal or to increase the robustness against channel estimation errors. The beamforming problem was later generalised to cope with indefinite quadratic side constraints in [252]. It was also noticed in [250, 251] that the beamforming problem can be formulated as a second order cone problem [224] for the special case of rank one channels, e.g. for an instantaneous channel realisation where the users are equipped with a single receive antenna or with fixed receiver beamformers. Wiesel et al. [223] studied the precoder design for a multiuser MIMO system with fixed linear multi-antenna receivers or, equivalently, with single-antenna users. The two optimisation criteria considered in [223] were the worst SINR maximisation subject to a sum power constraint, and minimisation of the required sum power to satisfy the minimum SINR constraints per stream. Similarly to [250, 251], the power minimisation problem was cast as an SOCP in [223]. On the other hand, the SINR optimisation can be carried out via power optimisation using one-dimensional bisection to search for the maximum SINR value which can be satisfied for all users/streams. A similar approach based on bisection was also proposed in [253].

In [245, 254, 255], the power minimisation problem was extended to the situation where the terminals are also equipped with antenna arrays. In addition, the worst SINR maximisation per user was considered in [254]. However, the problems are no more convex and convergence of the algorithm to the global optimum cannot be guaranteed. A sub-optimal iterative approach based on a
coordinate ascent method, i.e. optimisation with respect to one set of variables at a time while keeping the other sets fixed, was proposed for providing a local minimum for the problem.

Transceiver design for linear multiuser MIMO systems was considered in [14, 256–260], where iterative algorithms were proposed for solving the sum-MSE minimisation criterion. Luo et al. [261] recognised that the optimal MMSE transceiver design problem for the UL can be reformulated as a semidefinite program which can be solved using highly efficient interior point methods [15]. By using the UL-DL duality, the solution from [261] was later extended to the DL case with a sum power constraint in [14, 262]. The sum-MSE minimisation problem is convex only when the total number of terminal antennas is lower than or equal to the number of transmit antennas. Otherwise, the non-convex rank constraints become active. The rank constraints can be relaxed to obtain a lower bound for the sum MSE [14, 258]. Linear transceiver optimisation for the sum-rate maximisation criterion is a non-convex problem that has so far eluded a simple solution. Sub-optimal solutions were considered, e.g. in [253, 263] for single-antenna terminals.

In [14], co-authored by the author of this thesis, a general method for the joint design of linear transmit and receive beamformers, according to different optimality criteria was provided. By exploiting the UL-DL SINR duality [53, 54], the original optimisation problems were decomposed into a series of remarkably simpler optimisation subproblems which can be efficiently solved by using standard geometric program solvers [226]. Even though each subproblem is optimally solved, the global optimum cannot be guaranteed due to the nonconvexity of the original problems. The algorithms were shown to converge fast to a solution, which can be a local optimum, but remains efficient. In contrast to the previously proposed solutions [223, 244–246, 250, 251, 253, 254, 263], the method proposed in [14] can handle multiple antennas at the BS and at the mobile users with an arbitrary number of data streams per scheduled user. The optimisation criteria considered include the sum power minimisation under the minimum SINR constraint per data stream, the balancing of SINR values under the sum power constraint, the weighted sum rate maximisation under the sum power constraint and the weighted sum MSE minimisation under the sum power constraint. At each step of the iterative algorithm, the non-convex objectives of the weighted sum rate maximisation and the weighted sum MSE
minimisation problems were solved via signomial programming [226], i.e. solving a sequence of geometric programs which locally approximate the original problem. Other optimisation problems utilising signomial programming can be found in [226, 227, 264].

The linear multiuser transceiver design can be greatly simplified by imposing the constraint that all interuser interference terms must be zero, i.e. the multiuser channel becomes diagonal. The simplest approach to achieve this is via channel inversion [52, 106]. However, when users are equipped with multiple antennas, channel inversion is sub-optimal since each user is able to coordinate the processing of his own receiver outputs [265]. One solution to this problem is to use block channel inversion or block diagonalisation (BD) as proposed by Spencer et al. [266], which essentially generalises the channel inversion for a group of receive antennas rather than a single-antenna. This is achieved by forcing the precoder for a given user to lie in the null space of the other users’ channels. Other solutions with some variations on the ZF constraint were proposed in [267–269]. The basic block-ZF method in [266] relies on the condition that the sum of the terminal antennas is smaller than or equal to the number of TX antennas. The method can be easily extended to operate with any number of users by coordinating the processing between the transmitter and the receivers. This can be achieved by estimating the receivers’ linear combiners at the transmitter and applying the ZF solution [266] to a set of equivalent channels representing the transfer functions between the transmitter and the output of the receivers’ linear combiners [270]. The design problem of jointly selecting the transmit and receive beamformers for providing interference free data streams is difficult to solve in closed form. This is due to the coupling between the optimal transmit and receive beamformers. Generalised iterative ZF solutions were proposed in [270–272], where the BS and MS antenna weights are iteratively recomputed until they satisfy a convergence criterion.

The ZF methods are generally power inefficient because beamforming weights are not matched to the user channels. Inverting an ill-conditioned channel matrix results in a large power penalty that will dramatically reduce the SNR at the receivers. Allowing a limited amount of interference at each receiver generally provides a better solution than forcing the interference to be zero [14, 253, 258, 263]. A simple non-iterative approach utilising this fact was proposed in [273] for the case of single-antenna terminals. In [273], the channel
inversion is regularised with a term that depends on the SNR, and hence, the transmitter structure resembles an MMSE receiver. Despite its sub-optimality, the ZF transmission can still perform well when some multiuser diversity is available, i.e. semi-orthogonal users can be grouped together for ZF processing from a larger pool of users as will be shown in Chapter 5. The main benefit of the ZF method is that the solutions to different transceiver optimisation problems can be simplified into simple power allocation problems. This results from the fact that the ZF processing decouples the beamformer design and the power allocation. However, some performance penalty is caused by the ZF constraint.

It was shown already in [242, 250, 251] that convex MIMO transceiver optimisation problems can be extended with additional convex constraints such as limiting the dynamic range of the power amplifier at the transmitter. In real systems, a maximum sum power for any subset of the TX antennas can be imposed. Such power constraints are useful in the case of distributed MIMO architectures, where a central BS is connected via a high speed backbone to multiple antenna heads which are distributed over a large geographical area [9, 69]. A per antenna power constraint may also be required in practice, e.g. in a case where each TX antenna is equipped with a power amplifier and the linearity of the power amplifier is the limiting performance factor [163]. Yu and Lan investigated the minimum-power beamformer design for MISO DL under per antenna power constraints in [163], where the original DL problem was transformed into a dual UL minimax optimisation problem with an uncertain noise covariance.

For single-antenna receivers, the received SINR uniquely defines the user rate, and hence, the minimum user rate maximisation is equal to minimum SINR per stream maximisation, i.e. rate balancing equals SINR balancing. With multiple receive antennas, the equal or weighted SINR per stream requirement is too restrictive, since it does not allow for the power to be freely allocated over the multiple streams that are assigned to a user, unlike in the case of the rate maximisation. User rate balancing with MMSE receivers was investigated by Shi et al. in [274] under a sum power constraint. They utilised the UL-DL duality, providing an iterative algorithm where the power allocations and linear MMSE receive filters were updated in an alternating manner in both virtual UL and DL channels. The UL power control step was formulated as a geometric program...
where the rate requirements per user are handled via minimising the product of MSEs per stream. Furthermore, the authors in [274] provided a similar solution for minimising the total transmitted power with individual rate requirements per user. It must be noted that the MSE duality UL and DL, as presented in [274], may only hold in the case of a sum power constraint. The rate balancing for multiuser MIMO with ZF transmission subject to per BS (or per antenna) power constraints was studied in [68, 69].

A general method for joint design of the linear transmit and receive beamformers for several optimisation criteria is proposed by the author of this thesis in Chapter 6. The method can accommodate a variety of supplementary constraints, e.g. the lower bounds for the SINR values of data streams or the minimum rate constraints per user as well as per antenna group power constraints.
3 System model

The generic system model used in the rest of the thesis, and especially in Chapter 6, for the MIMO-OFDM cellular system with cooperative base station antenna heads is presented in Section 3.1. The special cases of the generic model including an adaptive MIMO-OFDM cellular system without cooperation between base stations and a single cell MIMO-OFDM system are defined in Section 3.2. A simple model for channel estimation errors is introduced in Section 3.3. In the rest of this thesis, the distributed BS antenna heads are called base stations for the sake of simplicity.

3.1 Cooperative MIMO-OFDM cellular system with distributed base station antennas

The cellular adaptive MIMO-OFDM system consisting of $N_B$ base stations has $N_C$ sub-carriers, each BS has $N_T$ transmit antennas and a user $k$ is equipped with $N_{R_k}$ receive antennas. A set $\mathcal{U}$ includes all users active at the given time instant, while a subset $\mathcal{U}_b \subseteq \mathcal{U}$ includes the users allocated to BS $b$, $k \in \mathcal{U}_b$. At each time instant\(^2\), the preprocessor of base station $b$, comprising channel coding, bit loading of the information data bit stream and linear pre-combining, generates an $N_T \times N_C$ dimensional matrix $X_{b,k}$ for each user $k$. The element of $t$'th row and $c$'th column of matrix $X_{b,k}$ is associated with the symbol transmitted at the transmit antenna $t$ and sub-carrier $c$. The OFDM modulator at the transmitter comprises the IFFT element, CP insertion and parallel-to-serial conversion. The CP is chosen to be longer than the maximum excess delay of the channel to avoid inter-symbol and inter-carrier interference. At the receiver side, the CP is first removed, and the resulted signals are then serial-to-parallel converted, and passed through the FFT transformers. For each time instant, the OFDM demodulator generates an $N_{R_k} \times N_C$ dimensional matrix $Y_k$, each row in the matrix corresponding to a receiving antenna. The length of the entire OFDM frame is chosen to be shorter than the coherence time of the channel, so that the fading can be modelled constant over one frame.

\(^2\)The time index is omitted in order to simplify the notation.
Each user $k$ can be served by $M_k$ BS’s which define the soft handover active set $S_k$ for the user $k$, and $S_k \subset S = \{1, \ldots, N_B\}$. Assuming perfect frequency and sample clock synchronisation between the transmitted and the received signals, the signal vector $y_{k,c} \in \mathbb{C}^{N_{R_k}}$ received by the user $k$ at the subcarrier $c$, $c = 1, \ldots, N_C$ can be expressed as

$$y_{k,c} = \sum_{b \in S_k} a_{b,k} H_{b,k,c} x_{b,k,c} + \sum_{b \in S_k} a_{b,k} H_{b,k,c} \sum_{i \neq k} x_{b,i,c}$$

$$+ \sum_{b \in S \setminus S_k} a_{b,k} H_{b,k,c} x'_{b,c} + n_{k,c}$$

$$= \tilde{H}_{k,c} \tilde{x}_{k,c} + i_{k,c}^{\text{intra}} + i_{k,c}^{\text{inter}} + n_{k,c}$$

(1)

where $x_{b,k,c}$ and $y_{k,c}$ denote the $c$’th columns in the matrices $X_{b,k}$ and $Y_k$, respectively. The vector $x_{b,k,c} \in \mathbb{C}^{N_{T}}$ is the transmitted signal from the $b$’th base station to user $k$, $x'_{b,c} \in \mathbb{C}^{N_{T}}$ denotes the total transmitted signal vector from BS transmitter $b$, $n_{k,c} \sim \mathcal{CN}(0, N_0 I_{N_{R_k}})$ represents the additive noise sample vector with noise power density $N_0$, and $a_{b,k} H_{b,k,c} \in \mathbb{C}^{N_{R_k} \times N_{T}}$ is the channel matrix from BS $b$ to user $k$ with large scale fading coefficient $a_{b,k}$. The entry $[H_{b,k,c}]_{r,t}$ represents the complex channel gain between TX antenna $t$ and RX antenna $r$ at sub-carrier $c$. The elements of $H_{b,k,c}$ are normalised to have unitary variance, i.e., $E\{[H_{b,k,c}]_{r,t}^2\} = 1$.

At each time instant, the active users $k \in U_b$ have identical SHO active set composition, $S_k = S_i$, $\forall$ $k, i \in U_b$. Similarly, all $b \in S_k \mid k \in U_b$ have identical sets of allocated users. The SHO system model is further illustrated in Fig. 2. The total transmitted vector $x'_{b,c}$ from BS $b$ consists of transmissions to all the users in the user set $U_b$, $x'_{b,c} = \sum_{k \in U_b} x_{b,k,c}$. The transmitted signal $\tilde{x}_{k,c} = [x_{S_k(1),k,c}^T, \ldots, x_{S_k(M_k),k,c}^T]^T \in \mathbb{C}^{M_k N_{T}}$ intended for user $k$ is distributed over $M_k$ base stations in the SHO active set $S_k$. The global channel matrix $\tilde{H}_{k,c} \in \mathbb{C}^{N_{R_k} \times M_k N_{T}}$ for user $k$ from all $M_k$ BS’s is defined as

$$\tilde{H}_{k,c} = [a_{S_k(1),k} H_{S_k(1),k,c}, \ldots, a_{S_k(M_k),k} H_{S_k(M_k),k,c}]$$

(2)

The interference vectors

$^3$Note that users having different SHO active set compositions cannot be served at the same time instant, and hence, they need to be served in different time and/or frequency slots. Single-cell processing is carried out if $M_k = |S_k| = 1 \ \forall \ k \in U_b$. 

60
Fig 2. SHO system model example

The transmitted vector for user $k$ is generated as

$$\tilde{x}_{k,c} = M_{k,c}d_{k,c} = V_{k,c}P_{k,c}^{1/2}d_{k,c}$$

where $M_{k,c} \in \mathbb{C}^{N_T \times m_{k,c}}$ is the pre-coding matrix, $d_{k,c} = [d_{1,k,c}, \ldots, d_{m_{k,c},k,c}]^T$ is the vector of normalised complex data symbols transmitted at sub-carrier $c$, and $m_{k,c} \leq \min(N_T M_k, N_{R_k})$ denotes the number of active data streams. $M_{k,c}$ is further split into $M_{k,c} = V_{k,c}P_{k,c}^{1/2}$, where $V_{k,c} = [v_{k,1,c}, \ldots, v_{k,m_{k,c},c}]$ contains the normalised TX beamforming vectors and $P_{k,c} = \text{diag}(p_{k,1,c}, \ldots, p_{k,m_{k,c},c})$ controls the powers allocated to each of the $m_{k,c}$ streams.

The receiver is equipped with a linear minimum mean square error filter and the decision variables are generated as $\hat{d}_{k,c} = W_{k,c}^{H}y_{k,c}$. The weight matrix
The weight matrix \( W_{k,c} \in \mathbb{C}^{N_h \times m_{k,c}} \) of the LMMSE filter is found by minimising

\[
W_{k,c} = \arg \min_{W_{k,c}} E \left[ \| d_{k,c} - W_{k,c}^H y_{k,c} \|^2_2 \right]
\]

and is given as \([238, 275]\)

\[
W_{k,c}^H = M_{k,c}^H \tilde{H}_{k,c}^H \left( \tilde{H}_{k,c} M_{k,c} \tilde{H}_{k,c}^H + Z_{k,c} + R_{k,c} \right)^{-1}
\]

where

\[
Z_{k,c} = \sum_{i \neq k} \tilde{H}_{k,c} M_{i,c} \tilde{H}_{k,c}^H
\]

is the covariance matrix of the intra-cell interference. It consists of transmissions to the users \( i \) that have an identical SHO active set composition with user \( k \), \( S_i = S_k \). The inter-cell interference-plus-noise covariance matrix \( R_{k,c} \) can be expressed as

\[
R_{k,c} = \sum_{b \in S \setminus S_k} \left( a_{b,k}^2 H_{b,k,c} E[ x_{b,c} x_{b,c}^H ] H_{b,k,c}^H \right) + N_0 I_{N_h}
\]

\[
= \sum_{b \in S \setminus S_k} \left( a_{b,k}^2 \sum_{i \in U_b} H_{b,k,c} E[ x_{b,i,c} x_{b,i,c}^H ] H_{b,k,c}^H \right) + N_0 I_{N_h}.
\]

\( R_{k,c} \) is assumed to be known at the receiver and to remain unchanged during the transmission of a frame. In practical TDD MIMO–OFDM cellular systems, the ideal knowledge of \( R_{k,c} \) at the transmitter would require it to be reported to the transmitter for each subcarrier and for each transmitted data frame. Therefore, in the system level studies a more practical case where \( R_{k,c} \) is known only at the receiver is considered. The channel matrix \( H_{b,k,c} \) is assumed to be known at the transmitter in all cases. Furthermore, the TX signals are assumed to have a common carrier phase reference and the differences in the propagation delays from all the transmitters to the intended users are assumed to stay within the guard interval. A more detailed analysis of the impact of imperfect phase synchronisation between the BS antenna heads will be carried out in Section 6.4.3.
The total power transmitted by the BS $S_k(n)$, $n = 1, \ldots, M_k \mid k \in \mathcal{U}_{S_k(n)}$ is

$$
\sum_{c=1}^{N_C} \text{Tr} \left( E \left[ x'_{S_k(n),c} x'_{S_k(n),c}^H \right] \right) = \sum_{c=1}^{N_C} \text{Tr} \left( \sum_{k \in \mathcal{U}_{S_k(n)}} M_{k,c}^{[n]} M_{k,c}^{[n]} H \right) = \sum_{k \in \mathcal{U}_{S_k(n)}} \sum_{c=1}^{N_C} \sum_{l=1}^{m_k,c} \|v_{k,l,c}^{[n]}\|_2^2 p_{k,l,c} \tag{10}
$$

where $M_{k,c}^{[n]} \in \mathbb{C}^{N_T \times m_k,c}$ is the pre-coder matrix of user $k$ that corresponds to the $n$'th base station belonging to $S_k$, i.e.,

$$
M_{k,c}^{[n]} = [M_{k,c}^{[n]}]_{(n-1)N_T+1:nN_T, :}, \quad n = 1, \ldots, M_k. \tag{11}
$$

Similarly, $v_{k,l,c}^{[n]} \in \mathbb{C}^{N_T}$ is the transmit vector for the $l$'th stream of user $k$ from BS $S_k(n)$, i.e.,

$$
v_{k,l,c}^{[n]} = [v_{k,l,c}^{[n]}]_{(n-1)N_T+1:nN_T,}, \quad n = 1, \ldots, M_k. \tag{12}
$$

Per antenna power constraints can be easily incorporated into (10) by assuming $n = 1, \ldots, M_k N_T$ and $M_{k,c}^{[n]} \in \mathbb{C}^{1 \times m_k,c}$.

### 3.2 Simplified system models

If the SHO feature is disabled, each user is served only by a single base station. In such a case, the system model is further simplified so that $M_k = |S_k| = 1 \forall k$ and $\tilde{H}_{k,c} = a_{S_k(1),k} H_{S_k(1),k,c}$. The single link SNR is defined as $\text{SNR} = P_T a_{S_k(1),k}^2 / N_0$, where $P_T$ is the BS transmit power.

In single-cell studies, the focus is restricted to a single set of users $\mathcal{U}$ with $|\mathcal{U}| = K$ users. Also, the inter-cell interference is assumed to be fixed and ideally incorporated into the whitened channel matrix $\tilde{H}_{k,c}^w = R_{k,c}^{-\frac{1}{2}} \tilde{H}_{k,c} \forall k, c$. Per antenna group power constraints can still be considered in the single-cell case. Let $\mathcal{A}_n$ be an arbitrary subset of transmit antennas. If $P_n$ is the maximum transmit power for the antenna set $\mathcal{A}_n$, then the sum power constraint for $\mathcal{A}_n$ can be expressed as $\sum_{k=1}^{K} \sum_{c=1}^{N_C} \sum_{l=1}^{m_k,c} \|v_{k,l,c}^{[n]}\|_2^2 p_{k,l,c} \leq P_n$, where $v_{k,l,c}^{[n]} \in \mathbb{C}^{|\mathcal{A}_n|}$ is the part of $v_{k,l,c}$ that contains the weights of the antennas belonging to $\mathcal{A}_n$.  

63
3.3 Model for channel estimation errors

A simple channel estimation error model is adopted to consider the impact of imperfect channel knowledge at the transmitter on system performance. It is assumed that the receiver performs linear MMSE estimation of the channel, and the channel estimate and the estimation error are uncorrelated due to the orthogonality principle [129, 130]. For simplicity, the channel estimation error is modelled as Gaussian IID noise, which is the worst case estimation error assumption [129]. The normalised estimated channel matrix $\hat{H}_{b,k,c}$ is given as

$$\hat{H}_{b,k,c} = \frac{1}{\sqrt{1 + \sigma_{\text{est}}^2}} (H_{b,k,c} + H_{b,k,c}^{\text{noise}})$$ (13)

where the entries of the estimation noise matrix $H_{b,k,c}^{\text{noise}}$ between each TX antenna $t$ and RX antenna $r$ are IID and generated as $[H_{b,k,c}^{\text{noise}}]_{r,t} \sim \mathcal{CN}(0, \sigma_{\text{est}}^2)$. 
4 Compensation of non-reciprocal inter-cell interference

This chapter focuses on the evaluation of adaptive MIMO-OFDM system performance in the presence of non-reciprocal inter-cell interference when only the receiver has knowledge of the DL interference structure. The concept of non-reciprocal interference was described in more detail in Section 1.2. The BS transmitters are assumed to be unaware of the transmission parameters like beamforming weights, transmission powers, etc., used in other cells for the interfering co-channel users. A linear MMSE filter is applied at the receiver to suppress the effect of structured inter-cell interference together with a simple and bandwidth efficient closed-loop compensation algorithm. The link and system level performance of the proposed system with the interference suppression and the closed-loop power offset adaptation algorithm is studied for different antenna configurations. The results are compared to those obtained in the ideal case in which the interference structure is perfectly known also at the transmitter. Furthermore, a simple analysis of the mutual information with different levels of channel and interference information at the transmitter is carried out.

The chapter is organised as follows. Section 4.1 describes how the transmitter can be designed to cope with non-reciprocal inter-cell interference with different levels of interference knowledge at the transmitter. In addition, the mutual information bounds for the studied cases are derived. The closed-loop adaptation algorithm is introduced in Section 4.2. Section 4.3 describes the simulation environment, assumptions, and both link and system level simulation results. Finally, conclusions are drawn in Section 4.4.

4.1 Single user transmitter design

This section presents the design of the pre-coder $M_{k,c}$ with different levels of interference information available at the transmitter. The focus in this chapter is limited to single user transmitter design for simplicity, and hence, a single user is allocated per cell and the intra-cell interference is assumed to be non-existent, i.e. $i_{k,c}^{\text{intra}} = 0$. By incorporating (5) and $i_{k,c}^{\text{intra}} = 0$ into (1), the received signal at
sub-carrier $c$ is simplified to:

$$y_{k,c} = T_{k,c}d_{k,c} + i_{\text{inter}}^{k,c} + n_{k,c}$$  \hfill (14)

where $T_{k,c} \in \mathbb{C}^{N_R \times m_{k,c}}$ represents the equivalent channel matrix of the desired signal at subcarrier $c$ and is defined as

$$T_{k,c} = \tilde{H}_{k,c}M_{k,c}. \hfill (15)$$

It includes the accumulated effect of signal processing at the transmitter side and channel propagation on the transmitted data signal. Matrix $T_{k,c}$ is assumed to be perfectly known at the receiver regardless of $M_{k,c}$ used at the transmitter. By utilizing the matrix inversion lemma

$$(A + UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1}, \quad \text{where } A, U, C \text{ and } V \text{ all denote matrices of the correct size},$$

the LMMSE receiver in (7) can be expressed as

$$W_{k,c} = T_{k,c}H_{k,c}^{-1} = T_{k,c}R_{k,c}^{-1} - T_{k,c}R_{k,c}^{-1}T_{k,c} + I^{-1}T_{k,c}R_{k,c}^{-1}.$$

The design of the pre-coder $M_{k,c}$ with different levels of interference information available at the transmitter are studied for the following two cases:

- optimal eigenmode transmission with full knowledge of $\tilde{H}_{k,c}$ and $R_{k,c}$ per sub-carrier at the transmitter
- sub-optimal eigenmode transmission with more practical assumptions where $R_{k,c}$ is not known at (not reported to) the transmitter and only $\tilde{H}_{k,c}$ is available at the transmitter together with a scalar feedback value

The instantaneous mutual information [bit/s/Hz] of the general single user ($Z_{k,c} = 0 \ \forall \ c$) MIMO-OFDM link with inter-cell interference is [49, 238]

$$I_{k}^{(\text{inst})} = \log_2 \left| I + R_{k}^{-1}\tilde{H}_{k}C_{k}\tilde{H}_{k}^{H} \right|$$  \hfill (17)

where $\tilde{H}_{k} = \text{blockdiag} \left( \tilde{H}_{k,1}, \ldots, \tilde{H}_{k,N_c} \right) \in \mathbb{C}^{N_R \times N_T \times N_c}$, $C_{k} \in \mathbb{C}^{N_R \times N_T \times N_c}$ is the covariance matrix of the TX signal and the interference-plus-noise covariance matrix $R_{k} \in \mathbb{C}^{N_R \times N_T \times N_c}$ can be expressed as

$$R_{k} = \sum_{b \in S_k} \left( a_{b,k}^2 \mathbf{H}_{b,k}\mathbf{E}[\mathbf{x}_b^H\mathbf{H}_{b,k}^H] \right) + N_0 I_{N_R \times N_c} \hfill (18)$$
where $H_{b,k} = \text{blockdiag}(H_{b,k,1}, \ldots, H_{b,k,N}) \in \mathbb{C}^{N_{b,k} \times N_T N_C}$ and the total transmitted signal from BS $b$ is $x'_b = [x'_{b,1}^T, \ldots, x'_{b,N}^T]^T \in \mathbb{C}^{N_T N_C \times 1}$.

Assuming that the interfering signals in the adjacent base stations are also generated as OFDM signals (each data symbol transmitted on a separate orthogonal sub-carrier), the term $E[x'_b x'^H_b]$ becomes block-diagonal. Thereby, $R_k$ gets a block-diagonal form as well, i.e. $R_k = \text{blockdiag}(R_{k,1}, \ldots, R_{k,N})$.

In such a case, the instantaneous mutual information (bit/s/Hz) of the single user MIMO-OFDM link ($Z_{k,c} = 0$) with OFDM type of interference becomes:

$$I_{(\text{inst})}^k = \frac{1}{N_T N_C} \sum_{c=1}^{N_C} \log_2 \left| I + R_{k,c}^{-\frac{1}{2}} \tilde{H}_{k,c} C_{k,c} \tilde{H}_{k,c}^H R_{k,c}^{-\frac{1}{2}} \right|$$

where $C_{k,c} = E[\tilde{x}_{k,c} \tilde{x}_{k,c}^H] = M_{k,c} M_{k,c}^H \in \mathbb{C}^{N_T M_k \times N_T M_k}$ is the covariance matrix of the signal transmitted to user $k$ on the $c$th sub-carrier.

In the analysis of the following sub-sections, the system model is further simplified so that the SHO feature is disabled and each user is served only by a single base station, i.e. $M_k = |S_k| = 1 \forall k$ and $\tilde{H}_{k,c} = a_{S_k(1),k} H_{S_k(1),k,c}$.

### 4.1.1 Optimal single user eigenmode transmission

With $\tilde{H}_{k,c}$ and $R_{k,c}$ known at the transmitter the optimum pre-combiner $M_{k,c}$ in (5) which maximises (19) is given by $M_{k,c} = V_{k,c} \sqrt{P_{k,c}}$, where the matrix $V_{k,c} = [v_{k,1,c}, \ldots, v_{k,m_{k,c},c}]$ contains the first $m_{k,c} = \text{rank}(\tilde{H}_{k,c})$ columns of the unitary matrix $\tilde{V}_{k,c} \in \mathbb{C}^{N_T \times N_T}$, which is obtained by singular value decomposition (SVD) of the pre-whitened channel matrix [42, 49, 51, 234, 235]

$$\tilde{H}_{k,c} = R_{k,c}^{-\frac{1}{2}} \tilde{H}_{k,c} = U_{k,c} \tilde{A}_{k,c} \tilde{V}_{k,c}^H.$$

In this way, a set of $m_{k,c}$ orthogonal spatial sub-channels are obtained at each sub-carrier and the diagonal matrix $P_{k,c} = \text{diag}(p_{k,1,c}, \ldots, p_{k,m_{k,c},c})$ controls the powers allocated for each of the $m_{k,c}$ eigenmodes. The diagonal matrix $A_{k,c} = \text{diag}(\lambda_{k,1,c}, \ldots, \lambda_{k,m_{k,c},c})$ includes the first $m_{k,c}$ eigenvalues of the Hermitian matrix $\tilde{H}_{k,c}^H R_{k,c}^{-\frac{1}{2}} \tilde{H}_{k,c}$.

If the term $T_{k,c}^H R_{k,c}^{-1} T_{k,c} = \Lambda_{k,c} P_{k,c}$ is now incorporated into (16), the MMSE filter gets the form

$$W_{k,c}^H = (\Lambda_{k,c} P_{k,c} + I)^{-1} T_{k,c}^H R_{k,c}^{-1} \Lambda_{k,c} P_{k,c}$$

67
and the decision variables are given by:

\[
\hat{d}_{k,c} = (A_{k,c}P_{k,c} + I)^{-1} \left( A_{k,c}P_{k,c}d_{k,c} + T_{k,c}^H R_{k,c}^{-1}(i_{k,c} + n_{k,c}) \right). \tag{22}
\]

The signal-to-interference-plus-noise ratio (SINR) for each spatial sub-channel \(l\), denoted as \(\gamma_{k,l,c}\), becomes

\[
\gamma_{k,l,c} = \frac{[A_{k,c}P_{k,c}]_{l,l}^2}{[T_{k,c}^H R_{k,c}^{-1}T_{k,c}]_{l,l}} = \frac{(\lambda_{k,l,c}p_{k,l,c})^2}{[A_{k,c}P_{k,c}]_{l,l}} = \lambda_{k,l,c}p_{k,l,c}. \tag{23}
\]

The SINR values \(\gamma_{k,l,c}\) can be independently controlled for each sub-channel \(l\), for example by some practical bit and power loading algorithm.

The transmit covariance matrix is now

\[
C_{k,c} = V_{k,c}P_{k,c}V_{k,c}^H \tag{24}
\]

and the instantaneous mutual information in (19) is maximised by water-filling power allocation between the sub-channels \([44, 81]\]

\[
p_{k,l,c} = \left( \mu - \frac{1}{\lambda_{k,l,c}} \right)^+ \tag{25}
\]

where \((x)^+\) denotes the positive part of \(x\), and where the "water level", \(\mu\), comes from the constraint of total transmitted power \([44, 81]\]

\[
\sum_{c=1}^{NC} \sum_{l=1}^{m_{k,c}} p_{k,l,c} = P_T. \tag{26}
\]

The instantaneous mutual information (19) for user \(k\) can be now simplified to

\[
I_{k}^{(\text{inst})} = \frac{1}{NC} \sum_{c=1}^{NC} \sum_{l=1}^{m_{k,c}} \log_2 (1 + \lambda_{k,l,c}p_{k,l,c}). \tag{27}
\]

### 4.1.2 Interference structure known only at receiver

Let us now consider the more practical case where the interference covariance matrix \(R_{k,c}\) is known only at the receiver. The channel matrix \(\hat{H}_{k,c}\) is assumed still to be known at the transmitter. Without \(R_{k,c}\) at the transmitter the optimum combiner \(M_{k,c}\) which maximises the mutual information (19) cannot be
computed. Therefore, a sub-optimal but still efficient design of the pre-coder is proposed. Specifically, a lower bound of $I_{k}^{\text{(inst)}}$ is found which does not depend on the statistical interference structure in frequency or space and then a pre-coder is designed which maximises this bound. It is easy to observe that

$$\hat{\lambda}_{\text{max}} \geq \mathbf{R}_{k,c} \Rightarrow \frac{1}{\hat{\lambda}_{\text{max}}} \preceq \mathbf{R}_{k,c}^{-1}$$

where $\hat{\lambda}_{\text{max}} = \max_{k,l,c} \hat{\lambda}_{k,l,c}$ are eigenvalues of $\mathbf{R}_{k,c}$ and $\mathbf{A} \succeq \mathbf{B}$ implies that $\mathbf{A} - \mathbf{B}$ is positive semidefinite.

By combining (28) with (19) the instantaneous mutual information of the considered system can be now bounded for any $\mathbf{C}_{k,c} \succeq 0$ as

$$I_{k}^{\text{(inst)}} \geq I_{k}^{\text{(lb)}} = \frac{1}{N_{c}} \sum_{c=1}^{N_{c}} \log_{2} \det \left( \mathbf{I} + \frac{1}{\hat{\lambda}_{\text{max}}} \tilde{\mathbf{H}}_{k,c} \mathbf{C}_{k,c} \tilde{\mathbf{H}}_{k,c}^{H} \right)$$

(29)

where the fact that log is a strictly increasing function has been utilised, and, the following implications are valid $\forall \mathbf{A}, \mathbf{B}, \mathbf{C} \succeq 0 : \mathbf{B} \succeq \mathbf{C} \Rightarrow \mathbf{A} \mathbf{B} \succeq \mathbf{A} \mathbf{C}$ and $\mathbf{A} \succeq \mathbf{B} \Rightarrow \det \mathbf{A} \geq \det \mathbf{B}$.

The lower bound (29) is used to find the sub-optimal transmit covariance matrix $\mathbf{C}_{k,c}$ for the case when only the channel knowledge and a single scalar $\hat{\lambda}_{k}^{\text{max}}$ are available at the transmitter. The lower bound is maximised by choosing

$$\mathbf{C}_{k,c} = \mathbf{V}_{k,c}^{H} \mathbf{P}_{k,c} \mathbf{V}_{k,c}^{H}$$

(30)

where $\mathbf{V}_{k,c}$ are the first $m_{k,c}$ columns of $\tilde{\mathbf{V}}_{k,c}$. SVD of the channel matrix $\tilde{\mathbf{H}}_{k,c} = \mathbf{U}_{k,c}^{H} \tilde{\Lambda}_{k,c} \tilde{\mathbf{V}}_{k,c}$ is used to obtain $\tilde{\mathbf{V}}_{k,c}$ and the transmit power is optimised by water-filling (25) over the scaled eigenvalues $\tilde{\lambda}_{k,l,c}^{\text{max}}/\hat{\lambda}_{k}$ as

$$p'_{k,l,c} = \left( \mu' - \frac{\tilde{\lambda}_{k}^{\text{max}}}{\hat{\lambda}_{k,l,c}} \right)^{+}$$

(31)

where $\mathbf{A'}_{k,c} = \text{diag} \left( \lambda_{k,1,c}', \ldots, \lambda_{k,m_{k,c},c}' \right)$ include the first $m_{k,c}$ eigenvalues of $\tilde{\mathbf{A}}_{k,c}$. Similarly to the optimal case but using the sub-optimal $\mathbf{C}_{k,c}$, the mutual information of such a system can be then calculated from (19).

Both optimum signalling and sub-optimal signalling with only the channel knowledge at the transmitter provide the same number of eigenmodes in the high SINR range, and, hence, both have the same slope of increase with increasing
The difference with the optimum signalling capacity case decreases as the SINR increases. This can be observed by rewriting (19) as

$$I^{(\text{inst})}_k = \frac{1}{N_C} \sum_{c=1}^{N_C} \log_2 \det \left( \mathbf{I} + \mathbf{H}^{eq}_{k,c} \mathbf{C}^{eq}_{k,c} \mathbf{H}^{eq H}_{k,c} \right),$$

where

$$\mathbf{H}^{eq}_{k,c} = \mathbf{R}^{\frac{1}{2}}_{k,c} \mathbf{U}^{\dagger}_{k,c} \Lambda^{\frac{1}{2}}_{k,c}$$

and

$$\mathbf{C}^{eq}_{k,c} = \mathbf{V}^{\dagger}_{k,c} \mathbf{C}_{k,c} \mathbf{V}_{k,c},$$

and maximising it with respect to $\mathbf{C}^{eq}_{k,c}$. This is equivalent to maximising the capacity of a MIMO system with channel matrix $\mathbf{H}^{eq}_{k,c} \in \mathbb{C}^{N_{R,k} \times m_{k,c}}$, $N_{R,k} \geq m_{k,c}$ and transmit covariance matrix $\mathbf{C}^{eq}_{k,c} \in \mathbb{C}^{m_{k,c} \times m_{k,c}}$ under transmit power constraint $P_T$. It is well known that the optimum solution in the high SNR region approaches

$$\mathbf{C}^{eq}_{k,c} = \frac{P_T}{m_{k,c} N_C} \mathbf{I}.$$  

The same solution is obtained with a high SINR by using sub-optimal transmit covariance matrix $\mathbf{C}_{k,c} = \mathbf{V}^{\dagger}_{k,c} \mathbf{P}^{\dagger}_{k,c} \mathbf{V}^{\dagger}_{k,c}$ from (30) since the water-filling power allocation in (31) converges to equal power allocation.

Based on the reasoning above, a sub-optimal eigenmode transmission method is chosen which relies on the channel knowledge only. The pre-combining matrix at subcarrier $c$ is then derived as $\mathbf{M}_{k,c} = \mathbf{V}^{\dagger}_{k,c} \mathbf{P}^{\dagger}_{k,c}$. Due to the non-reciprocal interference at the receiver, inter-stream interference is introduced by the linear MMSE matrix operation. The linear MMSE filter (7) is used at the receiver at each subcarrier to maximise the SINR for each $m_{k,c}$ eigenmodes. The SINR for each sub-channel $l$ can be derived as follows:

$$\gamma_{k,l,c} = \frac{\mathbb{E} \left[ \left| \mathbf{w}^{H}_{k,l,c} \mathbf{d}_{k,l,c} \right|^2 \right]}{\mathbb{E} \left[ \left| \mathbf{w}^{H}_{k,l,c} \left( \mathbf{T}_{k,l,c} \mathbf{d}_{k,l,c} + \mathbf{i}_{k,c} + \mathbf{n}_{k,c} \right) \right|^2 \right]} = \rho_{k,l,c} \frac{\mathbf{w}^{H}_{k,l,c} \mathbf{u}^{\dagger}_{k,l,c} \mathbf{u}^{H}_{k,l,c} \mathbf{w}_{k,l,c}}{\mathbf{w}^{H}_{k,l,c} \left( \mathbf{R}_{k,c} + \mathbf{\hat{T}}_{k,l,c} \mathbf{\hat{H}}_{k,l,c} \right) \mathbf{w}_{k,l,c}}$$

Note that this assumption is not necessarily true in the low SINR range, where the water-filling may choose to use a different numbers of streams depending on the structure of $\mathbf{R}_{k,c}$.
where $w_{k,l,c}, t_{k,l,c}$ and $u'_{k,l,c}$ are the $k$th column vectors of matrices $W_{k,c}, T_{k,c}$ and $U'_{k,c}$, respectively. Matrices $\tilde{T}_{k,l,c}$ and $\tilde{d}_{k,l,c}$ are defined as

$$\tilde{T}_{k,l,c} = [t_{k,1,c}, \ldots, t_{k,l-1,c}, t_{k,l+1,c}, \ldots, t_{k,m_{k,c}}]$$ (37)

and

$$\tilde{d}_{k,l,c} = [d_{k,1,c}, \ldots, d_{k,l-1,c}, d_{k,l+1,c}, \ldots, d_{k,m_{k,c}}]^T,$$ (38)

respectively. The SINR values derived in (23) and (36) are used as input parameters to the decoding process at the receiver. It can be seen from (36) that with the presence of non-reciprocal interference, the SINR per sub-channel can no longer be controlled at the transmitter as in (23) but is affected by the structure of $R_{k,c}$. If the SINR values are set to some fixed value at the transmitter by the bit and power loading algorithm to achieve a certain frame error rate (FER) at the receiver [95, 276, 277], these targets cannot necessarily be met. Therefore, some additional mechanisms are needed to maintain the quality of service at the receiver. In the special case, where the interference term $i^{(\text{inter})}_{k,c}$ is white, the SINR of sub-channel $l$ at sub-carrier $c$ is simplified to

$$\gamma_{k,l,c} = \frac{\lambda_{k,l,c} p'_{k,l,c}}{I_0 + N_0}$$ (39)

where $I_0$ is the average interference power density.

4.1.3 **Bit and power loading algorithm**

A low complexity bit and power loading algorithm requiring a low signalling overhead was introduced in [276, 277]. The throughput degradation to the optimal discrete HH loading algorithm for a fixed target FER was shown to be negligible while the signalling overhead was $N_C$ times reduced. The single user loading algorithm from [276, 277] is used in the simulations to achieve maximum throughput at certain target FER. Both the link and system level simulations are performed assuming $N_T \geq N_{R_h}$ which is a realistic assumption given the complexity and size limitations of mobile terminals. All $N_{R_h} N_C$ eigenmodes are grouped into $N_{R_h}$ clusters. The first cluster contains the strongest eigenmodes from each subcarrier, the second cluster contains the second strongest eigenmodes from each subcarrier and so on. Except for some eigenmodes corresponding to the most faded sub-carriers, a small difference between the eigenmodes’
gains belonging to the same cluster is experienced due to spatial diversity. Consequently, by skipping some of the most faded eigenmodes, the same MCS can be used for all selected eigenmodes belonging to the same cluster and the required signalling overhead is $N_C$ times reduced.

For each transmitted frame, first, the algorithm divides the total power between the clusters according to their capacities, and then, the MCS and the selected eigenmodes in each cluster are independently optimised to maximise the throughput subject to an equal SNR constraint for all the selected eigenmodes. Target FER is maintained by using a look-up-table which contains the SNR required by each of the available MCS in order to achieve the same FER in the AWGN channel. More details can be found in [277]. Turbo codes were used for channel coding [278]. The encoding is performed jointly in the time and frequency domains, a code word covering the selected eigenmodes from one cluster during the whole transmitted frame.

### 4.2 Closed-loop interference non-reciprocity compensation algorithm

In this section, a simple and bandwidth efficient closed-loop scalar power offset feedback method for compensating for the non-reciprocity between uplink and downlink interference structures is introduced. The proposed algorithm can also equally be used to compensate for the impact of channel estimation errors jointly with the non-reciprocal interference. The basic idea of the proposed method is to apply a power offset value at the transmitter which compensates for the difference in frequency, time and space selectivity between TDD DL and UL. The block chart of the proposed method is shown in Fig. 3. The method described here is for the downlink case, but it can also be used in the uplink.

In the beginning, the BS determines the initial link adaptation (LA) transmission parameters (MCS, power loading) on the basis of channel estimation measured in the previous frame(s) so that the target FER would not be exceeded in the receiving end. The MCS selection and the transmission power per sub-channel $p'_{k,l,c}$ are computed for the desired FER target based on the total transmission power $P_T$ available, the eigenvalues $\lambda'_{k,l,c}$ estimated from the channel information only and short-term average (over time, space and frequency
dimensions) scalar interference-plus-noise \((I_0 + N_0)\) level information feedback received from the terminal. This can be formulated as

\[
\{p'_{k,l,c}, MCS_{k,l,c}\} = f \left(I_0 + N_0, \lambda'_{k,l,c}, P_T \right)
\]

where \(f\) denotes the function for the bit and power loading algorithm \([95, 277]\). The BS sends a frame to the terminal using the given parameters. The terminal estimates a signal quality metric from the received frame. The metric depicts the signal quality degradation caused by non-reciprocal interference. The signal quality metric can be based on, for example, erroneous frames, i.e. acknowledgement messages (ACK/NAK), the FER or some other metric such as BER, raw BER, etc. A metric based on FER measurements, may be too slow for fast adaptation, especially if the FER target is low. Therefore, a metric based on BER after a second turbo decoding iteration corresponding to the selected 10% FER target was chosen for the simulation study in this chapter. It was shown to perform more reliably than the BER after the first decoding iteration, while it still provided enough erroneous bits per frame for sufficient statistical accuracy for the selected 10% FER target.
The terminal now compares the measured BER $\hat{\kappa}$ after the second decoding iteration to the target BER value $\kappa$. If the measured value is higher than the target value, the BS should either use more transmission power or more robust MCS so that the target is reached, and vice versa. On the basis of this comparison, an offset value illustrating the difference between the measured and target values is generated and transmitted to the BS. Similarly to the power control algorithm in CDMA systems [17], a simple implementation requiring only two bit feedback (up, down or do nothing) from the receiver is used in this thesis. The transmitter adjusts the power offset value $P_{\text{offset}}^k$ [dB] based on the feedback (FB) command as shown in Table 1, where $z$ indicates time index and $\delta$ is the hysteresis value used at the receiver to avoid a ping-pong effect of the power offset value at the transmitter. If the target metric and the measured metric differ from each other by less than a predetermined hysteresis value, the offset value is not changed at the BS. The step size $\nu P_{\text{step}}$ [dB] for the offset can be different depending on whether the offset value needs to be decreased or increased. The value $\nu = 2$ was used in the simulation studies in this thesis to guarantee fast recovery in the case the received signal quality suddenly deteriorates.

Table 1. Feedback strategy

<table>
<thead>
<tr>
<th>Receiver</th>
<th>Feedback</th>
<th>Transmitter</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\kappa} &gt; \kappa + \delta$</td>
<td>up</td>
<td>$P_{\text{offset}}^k(z) = P_{\text{offset}}^k(z-1) + \nu P_{\text{step}}$</td>
</tr>
<tr>
<td>$\hat{\kappa} &lt; \kappa - \delta$</td>
<td>down</td>
<td>$P_{\text{offset}}^k(z) = P_{\text{offset}}^k(z-1) - P_{\text{step}}$</td>
</tr>
<tr>
<td>$\kappa - \delta &lt; \hat{\kappa} &lt; \kappa + \delta$</td>
<td>-</td>
<td>$P_{\text{offset}}^k(z) = P_{\text{offset}}^k(z-1)$</td>
</tr>
</tbody>
</table>

The BS performs the required compensation of the interference non-reciprocity for the next transmitted frame based on the newly computed offset value. The offset value can be used either to adjust the transmission rate or the transmitted power per sub-channel. The way to use the feedback information depends on the set of the modulation and coding schemes used and the status of available resources, e.g. power and bandwidth, at the transmitter. In this thesis, the total

1-bit feedback would also suffice in most of the situations. It would create unnecessary fluctuation when the interference is slowly varying, though.

Note that $P_{\text{offset}}^k = 0$ dB if the interference term $I_{\text{inter}}^k$ is white Gaussian.
maximum transmission power at the transmitter is fixed and the power offset value feedback is used to modify the waterfilling reference level \( I_0 + N_0 + P_{\text{offset}}^k \) in the bit and power loading algorithm [95, 277]

\[
\{ p_k', I_{k,l,c}, MCS_{k,l,c} \} = f \left( I_0 + N_0 + P_{\text{offset}}^k, \lambda_k', P_T \right). \tag{41}
\]

The initial settings such as the target metric \( \kappa \) for the signal quality comparison are also communicated to the terminal prior to the transmission or any time the settings are changed. Note that adding the offset \( P_{\text{offset}}^k \) [dB] to \( I_0 + N_0 \) is equal to adjusting the values \( \lambda_k', l, c \) by the corresponding \( P_{\text{offset}}^k \) [lin] as in (31). For example, the transmission rate is reduced (\( P_{\text{offset}}^k \) increased) if the received quality is less than the target quality, and vice versa.

### 4.3 Numerical examples

The simulated system is based on HIPERLAN/2 [279] and IEEE 802.11a assumptions, where the number of subcarriers in the OFDM air interface is \( N_C = 64 \) and the cyclic prefix length is 16 samples. In all cases, the channel is assumed to be static during one coded OFDM frame. One coded OFDM frame consists of 16 OFDM symbols.\(^7\) Modulation schemes used in the simulations are QPSK, 16QAM and 64QAM. Half rate turbo code defined in [278] is used for channel coding. The minimum code word length for the bit and power loading algorithm with low signalling overhead [277] is \( l_{\min} = 500 \) bits and the SNRs required by each MCS to achieve the desired target FER=10% in the AWGN channel are: 1.8 dB, 7.1 dB and 11.6 dB, respectively. The number of TX antennas is changed between 2, 4 and 8, while the number of RX antennas is varied between 2 and 4. Power offset step \( P_{\text{step}} \) used in the simulation is 0.25 dB. The main parameters used in the simulation are summarised in Table 2.

Both link and system level performance evaluations are carried out by simulations. Interfering base stations have the same number of transmit antennas as the desired link. All \( N_B - 1 \) interfering signals \( x_{b,c}', b \notin S_k \) including both the precoding vectors and transmitted data are assumed to be unknown to the receiver.

\(^7\)The forthcoming 4G systems are anticipated to have up to 100MHz bandwidth. In such a case, one coded OFDM frame could be fit into four OFDM symbols, thus allowing for better mobility support.
Table 2. Main simulation parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>System bandwidth</td>
<td>20MHz</td>
</tr>
<tr>
<td>Modulation schemes</td>
<td>QPSK, 16QAM and 64QAM</td>
</tr>
<tr>
<td>FER target</td>
<td>10%</td>
</tr>
<tr>
<td>SNR per MCS</td>
<td>1.8dB (QPSK), 7.1dB (16QAM) and 11.6dB (64QAM)</td>
</tr>
<tr>
<td>Channel coding rate</td>
<td>R=1/2 for all MCS</td>
</tr>
<tr>
<td>Power delay profile</td>
<td>ETSI BRAN A1 [279]</td>
</tr>
<tr>
<td>NT</td>
<td>2, 4, 8</td>
</tr>
<tr>
<td>NRk</td>
<td>2, 4</td>
</tr>
<tr>
<td>P_step</td>
<td>2×0.25 dB (Up), 0.25 dB (Down)</td>
</tr>
<tr>
<td>Hysteresis value</td>
<td>0.5%</td>
</tr>
</tbody>
</table>

Moreover, they are modelled as white Gaussian noise \( x'_{b,c} \sim \mathcal{CN}(0, N_0/N_T I_{N_T}) \), for simplicity. This means that each interfering base station is assumed to transmit the maximum number of independent streams (= NT) with equal power per stream. The factor \( \omega_{b,k} = P_T a_{b,k}^2/N_0 \) represents the power ratio between the averaged received interference from the interference source b and \( N_0 \). The factor \( I_0/N_0 = \sum_{b \not\in S_k} \omega_{b,k} \) models the ratio between the total average interference power density \( I_0 \) and the noise power density \( N_0 \). Now, the quasi-static interference-plus-noise covariance matrix \( R_{k,c} \) with \( E \left[ x'_{b,c} x_{b,c}^H \right] = N_0/N_T I_{N_T} \) is simplified to

\[
R_{k,c} = \sum_{b \not\in S_k} \omega_{b} N_0 \left( H_{b,k,c} H_{b,k,c}^H \right) + N_0 I_{N_R_k} \tag{42}
\]

and it is assumed to be perfectly known at the receiver. The matrices \( H_{b,k,c} \) are assumed to remain unchanged during the transmission of a frame.

A fixed interference scenario with one dominant interference source was assumed in the link level evaluation as well as in the capacity analysis, in order to better quantify the gains from the interference suppression. This resembles the situation where the interference is coming from a single dominant adjacent cell, and the remaining inter-cell interference is modelled as AWGN. The same channel characteristics were used for both the desired signal and for the interference. In this fixed scenario, the power spectrum density of the remaining white Gaussian interference was assumed to be 9dB below the dominant interferer \( (I_0/N_0 = 9dB) \). The channel’s delay taps were considered independent of each
other with a power delay profile specified by ETSI BRAN Channel A [279]. The link level performance was examined with both independently fading and spatially correlated antennas. Correlated channel coefficients were generated as described in [280], where the complex correlation matrices for the transmitter \((N_T \times N_T)\) and the receiver \((N_{R_k} \times N_{R_k})\) are \(R_{TX}\) and \(R_{RX}\), respectively, and the bi-spatial correlation matrix for the \(N_T \times N_{R_k}\) MIMO channel is given by their Kronecker product \(R_{MIMO} = R_{TX} \otimes R_{RX}\). One correlated channel type presented in [281] was simulated: a high correlated channel (HCC) which corresponds to an outdoor link with the base station antenna located above surrounding scatterers.

A realistic multi-cell environment with a standardised geometric stochastic channel model [76] was used for the system level evaluation. The aim was to assess the impact of realistic varying interference scenarios and channel characteristics on the system level performance.

### 4.3.1 Mutual information results

Fig. 4 shows the ergodic mutual information of the optimal and sub-optimal methods versus average signal-to-interference-plus-noise ratio with the presence of structured interference for several MIMO scenarios. For the \(2 \times 2\), \(4 \times 2\) and \(4 \times 4\) cases, the HCC channel model is used for both the desired signal and for the interference, while an uncorrelated antenna scenario is used for the \(8 \times 2\), \(8 \times 4\) and \(16 \times 4\) cases. The mutual information expressions derived in Section 4.1 are used for the evaluation of the cases with different levels of channel state information available at the transmitter. For the sub-optimal method, the lower bound of \(I_{k}^{(\text{inst})}\) is computed by finding \(\hat{\lambda}_{k}^{\text{max}}\) for each channel realisation as in (28) and using it for the transmit power optimisation (31).

The results indicate that knowledge of the interference structure \(R_{k,c}\) at the transmitter does not necessarily provide a significant additional gain as compared to the case where only \(\hat{H}_{k,c}\) is known at the transmitter, as long as knowledge of \(R_{k,c}\) is utilised at the receiver side. In the high SINR region, the mutual information of the sub-optimal case approaches one of the optimal case as already discussed in Section 4.1.2.

The effect of unknown interference is more pronounced when the number of
antennas at the transmitter is equal to the number of receive antennas. When the number of interfering antennas increases (with either more transmit antennas in the dominant interferer as in Fig. 4 or more independent interfering adjacent cells with similar average received power) and the channel correlation is reduced, the interference summed up at the receiver approaches white Gaussian. Thus, the mutual information loss from the interference non-reciprocity is small.

4.3.2 Link level results

The impact of structured (time, frequency and space selective) non-reciprocal interference on system performance was studied by link level simulations. The performance of the sub-optimal receiver with interference suppression (labelled as ‘MMSE’) was compared to the one without interference suppression at the
receiver (labelled as 'MF'), and to the optimal case where $R_{k,c}$ is known both at
the transmitter and at the receiver (labelled as 'R known at TX/RX'). In the
case without interference suppression, both the transmitter and the receiver
assume the interference to be white ($R_{k,c} = I$), and thus, the receiver (16) is
reduced to the matched filter (MF) as

$$W_{k,c}^H k_c = \left( P'_{k,c} k_c V'_{k,c} k_c H_{k,c} H_{k,c} V'_{k,c} k_c + I \right)^{-1} M_{k,c} H_{k,c}$$

and the SINR per stream $\gamma_{k,l,c}$ is defined as in (36). Note that the diagonal part
of (43) does not affect the SINR and can be neglected. The two sub-optimal
cases were simulated both with fast and slow power offset adaptation, denoted
as fast and slow tuning in the figures. The results from all cases with structured
interference were compared to that with white Gaussian interference.

In the cases labelled as "slow tuning", it is assumed that the transmitter is
aware of only the long term average impact of the structured interference and
cannot follow the interference selectivity as time evolves. The power offset remains
fixed for a long time (tens of seconds, for example) as the terminal measures the
long term performance and tries to maintain the long-term QoS target (10%
FER). In the simulations, the fixed power offset at the transmitter was increased
between the simulation runs and used to modify the water-filling reference level
in the bit and power loading algorithm as in (41) until the impact of interference
was fully compensated for and the long-term average FER target was achieved.
A block fading channel model for both the desired and interference channels
was used to avoid excessively long simulations. "Fast tuning" means that the
power offset is varied at the transmitter as a function of time-continuous changes
in the interference structure so that the QoS target is preserved over shorter
time instances. Time-continuous fading was modelled in these simulations. The
case without the interference non-reciprocity compensation cannot be depicted
together with the aforementioned cases with a fixed FER target, since in that
case the FER at the receiver cannot be controlled by the transmitter.

The spectral efficiency versus the SINR at 10% FER target for the $2 \times 2$ and
for the $4 \times 2$ MIMO-OFDM system with a high correlated channel model are
shown in Figs. 5 and 6. The SINR is defined as $P_R / (I_0 + N_0)$, where $P_R = a_{b,k} P_T$
is the average received power of the desired signal. The SINR difference between the cases with AWGN type interference and coloured interference shows the average power offset difference required to compensate for the time, frequency and space selectivity of the interference. In all cases with structured interference and without interference suppression at the receiver, higher SINR is required on average than in the AWGN type interference case for the same spectral efficiency. For example, about 2 dB higher SINR is required in the low SINR region with MF and with slow tuning, as seen from Fig. 6.

The MMSE receiver yields clearly better performance than the matched filter without interference suppression, especially in the low SINR region. This is due to the fact that the transmitter does not work in the full eigenmode excitation mode but performs beamforming by concentrating all the transmitted power

---

Footnote: The spectral efficiency is measured as bits/s/Hz, and it is the maximum throughput of a point-to-point link with a given set of modulation and coding schemes.
into the strongest eigenmode. Thereby, the MMSE receiver structure can utilise the remaining degree of freedom to perform additional inter-cell interference suppression.

The MMSE receiver combined with the simple scalar power offset feedback to the transmitter (fast tuning) shows almost as good performance as the case where $R_{k,c}$ is perfectly known at the transmitter. The gain compared to the worst case can be up to 6.5 dB and 4 dB for the $2 \times 2$ and $4 \times 2$ cases, respectively. The performance of the ideal case can be better still, especially, in the high SINR region with full eigenmode excitation, as seen in Fig. 6. However, the required signalling feedback ($R_{k,c}$ reported for each subcarrier) would make the ideal approach rather unpractical.

Low spatial correlation makes the interference less selective, and thus, less relative gains are available from both the interference suppression and the tuning algorithm. In the case of uncorrelated transmitter and receiver antennas, the
curves (not shown) would look similar to those in Figs. 5 and 6, except that the difference between the worst and the best case would be reduced to 4 dB and 2 dB for the $2 \times 2$ and $4 \times 2$ cases, respectively, in the low SINR region.

An example plot of the transmission power offset behavior for $I_0/N_0 = 9 \text{ dB}$ with 4 dB SINR is shown in Fig. 7 for the $2 \times 2$ case with the HCC channel model and with the MMSE receiver. It can be seen from the plot that the fast adaptation is able to follow the time-continuous changes in the interference structure, while the long term 10% FER target is preserved. The gain from the fast power offset adaptation, i.e. the average power offset difference between fast and slow tuning, is mostly between 1 and 2 dB. This is demonstrated in Fig. 8, where the power offset behaviour in different simulation cases is compared.

![Plot of power offset behavior](image)

**Fig 7.** Example plot of closed-loop adaptation of the transmitted power at the BS as a function of time (data frame), MMSE receiver, 4 dB SINR, $I_0/N_0 = 9\text{ dB}$.
4.3.3 System level evaluation

System level evaluation assesses the impact of a realistic multi-cell environment (realistic interference power profiles, spatial properties, etc.) on the system performance with and without non-reciprocal interference compensation. The gain from both interference suppression and fast adaptation observed from the link level simulation results can be reduced in situations where the number of interference sources is high (either more transmit antennas or transmitted streams in the dominant interferer or more independent interfering adjacent cells with similar average received power) and/or the $I_0/N_0$ ratio decreases, i.e. the interference becomes more white.

A 57-cell scenario (19 3-sector sites) was used for the multi-cell system level evaluation. Fig. 9 illustrates the simulation scenario. The Okumura-Hata propagation model [33, 76] is used for modelling the path loss and the effect of
sectorised antenna patterns is included in the path loss calculations [76]. In addition to the path loss, the signal suffers from the shadowing caused by large obstacles usually around the mobile. The shadowing factor is a log-normal random variable with a mean of 0 dB and a standard deviation of 6 dB.

The independent time-continuous fast fading process is simulated for each MIMO antenna transmitter-receiver pair including both the desired link $H_{S_k(1),k,c}$ and the most dominant interference links $H_{b,k,c}$, $b \notin S_k$ for each terminal $k$ dropped in the system. The channel coefficients are generated by a standardised geometric stochastic channel model, the 3GPP/3GPP2 SCM model [76] with the urban micro scenario. This approach requires deterministic scatterers to be incorporated into the model, and the small scale fading is given implicitly by the superposition of the paths/rays as the terminal moves along a predefined route [76].

The simulation run consists of $D$ user drops where $K$ users are randomly uniformly distributed within the geographic area of the system in the beginning of each drop [76]. In order to speed up the simulation time and to avoid border effects, only the center site users’ data is recorded for the statistics. Large scale mobility is not modelled, but the fast fading is emulated for the stationary users during a drop. An example distribution of some of the traced user locations is also shown in Fig. 9 depicting the simulation scenario. A traced user refers here to a user whose data is recorded for the user performance statistics.

Frequency reuse one is assumed. The multiple access scheme is time division multiple access, where each user is assigned a fixed length transmission slot within the downlink TDD frame. In these simulations, the number of DL slots is set to 12, each slot corresponding to 16 OFDM symbols. The number of users, $K$, is adjusted so that 40% average downlink load (time slot occupation) is achieved. Such a loading scenario was selected in order to create a sufficient amount of interference while avoiding the system becoming too interference limited. A simple random transmission slot allocation scheme is selected for the system level evaluation. In each cell/sector, the users are randomly allocated to the transmission slots within the frame and the user allocation tables are maintained fixed during one simulation drop. The base station transmission power is fixed to 33 dBm. Only those center site users with an average SINR of more than 0 dB are traced for the statistics while the others are declared to be
in the outage. The outage probability in this scenario is 10%. The user locations and allocation tables (interference profiles) are identical between the simulation cases in order to make the results comparable.

The closed-loop interference non-reciprocity compensation algorithm "fast tuning" capable of following the time-continuous changes in the interference structure, is used in the system level simulation cases; it is labelled as "tuning". The slow tuning option would have required an iterative search for the optimum fixed power offset value to reach the long-term FER target for each terminal. Thus, it was not modelled to simplify the simulation setup. The power offset value feedback is used to adjust the reference level in the bit and power loading algorithm as in (41), that is, the lower the received power offset value the higher the transmission rate and vice versa. In all cases, it is assumed that the transmitter is at least aware of the average SINR at the receiver.

Fig. 10 illustrates the CDF of the average user FER for a simulation case with
4 × 2 antenna structure. It can be seen from the results that the performance without interference suppression or power offset tuning is poor, and the target FER cannot be achieved. For example, the median of the FER is more than 40% while the target was set at 10% in Fig. 10. The case with interference suppression but without power offset tuning results in significantly better performance than the case without interference suppression. However, the resulting FER cannot be controlled at the transmitter due to the varying SINR per sub-carrier as shown in (36), and thus, the FER variance is rather large. The SINR value per sub-carrier depends on the interference structure (the number of independent interference sources, $I_0/N_0$, correlation between antennas, etc.) and the average SINR of the given user. If the interference is very selective and the average SINR is low for the given user, i.e. single eigenmode transmission is likely, the MMSE filter is able to effectively suppress the interference, and, thus, the resulting FER can be significantly lower than the target, as seen in Fig. 10. This means that the capacity of the channel is not fully utilised and a higher transmission rate could be used. On the other hand, a user with a high SINR (full eigenmode transmission) and/or less selective interference may experience FER higher than the target value, and, hence, a lower transmission rate should be used to reach the target FER. In the cases with power offset tuning, the average user FER is maintained close to the target. The variation is almost as small as in the optimum case where the interference structure is perfectly known both at the transmitter and the receiver.

Figs. 11(a) and 11(b) show the cumulative density function of the average user spectral efficiency [bits/s/Hz] at the receiver for different antenna configurations. Due to high FER, the spectral efficiency of the case without interference suppression or tuning is considerably lower than in the other cases. In general, the cases labelled as 'MMSE, no tuning' and 'MF, tuning' have very similar performance with all antenna configurations. However, the users with high spectral efficiency (high SINR, full eigenmode transmission) gain more from fast power offset adaptation especially if the number of transmit antennas of the interference source is high, as seen from Fig. 11(b).

The system performance is further improved if both interference suppression and fast power offset tuning are enabled for all the users. The gain as compared
to the cases labelled as 'MMSE, no tuning' and 'MF, tuning' is about 15% at a 10% outage probability for the 4 × 2 case depicted in Fig. 11(a), for example. Moreover, the 4 × 2 scenario shows almost as good performance as the optimal case where \( R_{k,c} \) is perfectly known at the transmitter, labelled as 'R known at TX/RX'. In the 8 × 2 case, the spatial selectivity of the interference is reduced and at the same time the majority of the users operate in the full eigenmode excitation mode. Therefore, the benefits from the MMSE receiver are reduced and the gap to the ideal feedback case is somewhat increased.

The results indicate that even the users closer to the base station have in general some interference sources above the noise level in the simulated scenario. Thus, the interference still has some detrimental effect on the experienced user quality. However, the average error rates for the users closer to the BS with no compensation tend to be somewhat lower than for the users on the edge of the cell. This can be either due to the fact that the received interference is a sum of several interference sources or that the level of interference is closer to the noise level. This can be seen also by looking for example at Fig. 11(b).
Fig 11. CDF of average spectral efficiency per user.
When comparing the curves 'MF, no tuning' and 'R known at Tx/Rx', the relative difference is lower at high spectral efficiency, implying that the impact of interference is less than with a lower SEFF. The bit and power loading algorithm minimises the transmitted power when the spectral efficiency is saturated. Since there is unused tx power available, both cases with tuning can also achieve the saturated spectral efficiency with a high SINR.

4.4 Summary and discussion

The performance of an adaptive MIMO-OFDM system was studied in the presence of non-reciprocal inter-cell interference when the downlink interference structure is known only at the receiver. A linear MMSE filter is applied at the receiver to suppress the effect of structured interference together with a simple and bandwidth efficient closed-loop compensation algorithm proposed in this thesis. The results were compared to those in the ideal case where the interference structure per sub-carrier is also known at the transmitter. The results with the MMSE receiver show significantly better performance than the cases without interference suppression, especially in the low SINR region where the transmitter does not utilise all available eigenmodes.

The proposed closed-loop compensation algorithm with a simple scalar power offset feedback combined with interference suppression at the receiver results in nearly the same performance as the ideal case where the interference structure per sub-carrier is perfectly known at the transmitter in both link and system level studies. The performance of the ideal case was found to be better still, especially in the high SINR region with full eigenmode excitation. However, the required signalling feedback would make this approach unpractical in most applications. The results also demonstrate that in the presence of non-reciprocal inter-cell interference, the quality of service at the receiver cannot be controlled if the transmission parameters are defined based on the reverse link measurements only. Therefore, some feedback to the transmitter is always needed in order to make a cellular adaptive TDD MIMO-OFDM system to function properly.

In the numerical evaluation, a sort of worst case assumption about the interference characteristics was taken. The interfering BS’s were assumed to transmit always the maximum number of independent streams (= $N_T$) with equal power per stream. In reality, the number of streams per transmission
slot in the interfering BS’s can often be less than $N_T$ due to decisions by the loading algorithm. In such a case, the inter-cell interference would become more selective, thus resulting in larger differences between the simulated cases than with the worst case assumptions.

An interesting future extension is to evaluate the effect of bursty packet traffic and the impact of scheduling decisions on the system performance. The fact that a higher initial power adaptation value is needed in the beginning of packet transmissions reduces the gains available from the fast power offset adaptation for short packets, and, thus, it can have a negative impact on the system level performance. In practical cellular networks, the inter-cell interference experienced by a given user may sometimes vary even faster than the channel itself, depending on the fast frame-by-frame channel allocations in the neighbouring cells. Different users may be served in the subsequent frames. This affects the precoder design at the neighbouring BS transmitters, which in turn changes the structure of the interference received by the given user.

Another future research topic is the further refinement of the feedback scheme. Since the scalar feedback affects only the power (and the MCS) allocated to different eigenmodes/subcarriers, the beamforming vectors are not affected. Thus, the impact on the structure of interference experienced by other users is significantly smaller than in the case where the interference covariance matrices are ideally available at the transmitters. Obviously, the feedback may still cause some oscillation in the adaptive process due to the distributed nature of the algorithm. One possible solution for such a problem is to distribute the feedback information between all the affected nodes via a backbone network and to design the feedback strategies jointly across the nodes.
Chapter 5: Resource allocation with block-ZF transmission

This chapter considers the DL of a multi-user MIMO-OFDM system, where the transmitter and receivers are equipped with multiple antennas and accurate TX CSI is available. The fundamental idea of this approach is that with full transmitter CSI of all same cell users, the precoding can be designed so that inter-user interference is minimised. Iterative block-ZF processing is utilised in this chapter to guarantee interference free reception for all scheduled users.

An efficient greedy beam ordering and selection algorithm is proposed to maximise the DL sum rate or sum spectral efficiency (SE) of the block-ZF method for any number of users and RX antennas. It is found that due to the inherent noise amplification problem with the linear block-ZF method, the maximum sum rate is often achieved by transmitting to less users/beams than the spatial dimensions available, especially in the low SNR region and with a low number of users. The performance is compared to the sum rate available with several other scheduling algorithms and to the sum rate capacity. In addition, an efficient low complexity joint user, bit and power allocation algorithm with low signalling overhead (LSO) is proposed. Moreover, the impact of imperfect channel estimation at the transmitter and imperfect orthogonalisation of allocated users on system performance is studied. The focus is on the optimisation of the usage of single cell resources, restricted to a single set of users $U$ with $|U| = K$ users. Each user is served only by a single base station, i.e. $M_k = |S_k| = 1 \forall k$ and $\mathbf{H}_{k,c} = a_{S_k(1),k} \mathbf{H}_{S_k(1),k,c}$.

The inter-cell interference is assumed to be fixed and it is ideally incorporated into the whitened channel matrix $\mathbf{H}_{k,c}^w = \mathbf{R}_k^{-1/2} \mathbf{H}_{k,c}$. The chapter has the following organisation. The iterative block-ZF processing based on block diagonalisation is introduced in Section 5.1. Section 5.2 describes the proposed scheduling and resource allocation algorithms, and the numerical analysis is carried out in Section 5.3. Finally, a summary is given in Section 5.4.
5.1 Multi-user zero forcing with iterative block diagonalisation

Block diagonalisation of multiple user channels combined with coordinated TX-RX processing and scheduling between users is a simple but efficient block-ZF method [266]. The design problem of jointly selecting the transmit and receive beamformers for providing interference free data streams is difficult to solve in closed form. This is due to the coupling between the optimal transmit and receive beamformers. Thus, iterative solutions have been proposed, e.g. [270–272]. An iterative BD method was originally proposed in [270, 271], and it is here extended to include joint user, bit and power allocation and a method to compensate for the impact of a finite number of iterations.

The basic BD method in [266] relies on the condition that \( N_T \geq N_R^\text{tot} \), where \( N_R^\text{tot} = \sum_{k=1}^{K} N_{R_k} \). In general, the transmitter can send up to \( N_T \) interference free data streams, regardless of the number of users. The BD algorithm can be extended to operate with any number of users by coordinating the processing between TX and RX. This is achieved by using a new equivalent channel matrix \( \hat{H}_{k,c} = F_{k,c}^H \tilde{H}_{w_{k,c}}^c \in \mathbb{C}^{m_{k,c} \times N_T} \) in the BD algorithm instead of \( \tilde{H}_{w_{k,c}}^c \), where \( F_{k,c} \) is an \( N_{R_k} \times m_{k,c} \) matrix consisting of \( m_{k,c} \) receive beamformers (and spatial dimensions) [266]. The matrix \( \hat{H}_{k,c} \) can be used in the BD algorithm instead of \( \tilde{H}_{w_{k,c}}^c \), as long as \( \sum_{k=1}^{K} m_{k,c} \leq N_T \).

In order to eliminate the multi-access interference (MAI), the constraint \( \hat{H}_{i,c} M_{k,c} = 0 \) for \( i \neq k \) is imposed. If the stacked channel matrix for all other users except the user \( k \) is defined as

\[
\hat{H}_{k,c} = [\hat{H}_{1,c}^T \ldots \hat{H}_{k-1,c}^T \hat{H}_{k+1,c}^T \ldots \hat{H}_{K,c}^T]^T
\] (44)

the zero interference constraint forces \( M_{k,c} \) to lie in the null space of \( \hat{H}_{k,c} \). Following the procedure from [266, 271], the last \( m_{k,c} \) right singular vectors from the SVD of

\[
\hat{V}_{k,c} = \hat{U}_{k,c} \hat{\Lambda}_{k,c}^{1/2} (\hat{V}_{k,c}^{(1)} \hat{V}_{k,c}^{(0)})^H
\] (45)

are chosen. The selected matrix \( \hat{V}_{k,c}^{(0)} \) forms an orthogonal basis for the null space of \( \hat{H}_{k,c} \) so that \( \hat{H}_{k,c} \hat{V}_{i,c}^{(0)} = 0 \) for \( i \neq k \).
The SVD of \( \tilde{H}_{k,c} = \tilde{H}_{k,c}^{\text{w}} \tilde{V}_{k,c}^{(0)} \) is now determined individually for each user as:

\[
\tilde{H}_{k,c} = [\tilde{U}_{k,c}^{(1)} \tilde{U}_{k,c}^{(0)}] \begin{bmatrix}
\Lambda_{k,c}^{2} & 0 \\
0 & 0
\end{bmatrix} [\tilde{V}_{k,c}^{(1)} \tilde{V}_{k,c}^{(0)}]^H,
\] (46)

where \( \tilde{U}_{k,c}^{(1)} \) and \( \tilde{V}_{k,c}^{(1)} \) represent the first \( m_{k,c} \) left and right singular vectors of \( \tilde{H}_{k,c} \), and \( \Lambda_{k,c} \) is the \( m_{k,c} \times m_{k,c} \) diagonal matrix of eigenvalues. The pre-coding matrix \( M_{k,c} \) for user \( k \) is now defined as

\[
M_{k,c} = \tilde{V}_{k,c}^{(0)} \tilde{V}_{k,c}^H \Lambda_{k,c} \tilde{V}_{k,c}^H P_{k,c}^2 = V_{k,c} P_{k,c}^2
\] (47)

where the diagonal matrix \( P_{k,c} = \text{diag}(p_{1,c}, \ldots, p_{m_{k,c},c}) \) controls the powers allocated for each of the \( m_{k,c} \) eigenmodes. The transmitter and receiver matrices are successively recomputed by assigning \( F_{k,c}(j) = \tilde{U}_{k,c}^{(1)}(j-1) \), where \( j \) denotes the iteration index. The initial set of \( F_{k,c}(1) \) is generated by selecting the \( m_{k,c} \) dominant left singular vectors (columns) \( U_{k,c} \) of each \( \tilde{H}_{k,c}^{\text{w}} = U_{k,c} \Lambda_{k,c} V_{k,c}^H \) [266]. The iterative BD algorithm is now briefly summarised in Algorithm 1.

**Algorithm 1 Iterative BD decomposition**

1. Let a scheduling algorithm define the number of streams \( m_{k,c} \) allocated for each user \( k \in \mathcal{U} \). Initialise the \( F_{k,c}(j) \) matrix to include the dominant \( m_{k,c} \) left singular vectors of \( \tilde{H}_{k,c}^{\text{w}} \). Let \( j = 1 \).

2. Set \( \tilde{H}_{k,c} = F_{k,c}^H(j) \tilde{H}_{k,c}^{\text{w}} \), define \( \mathcal{U}_k = \mathcal{U}\setminus\{k\} \) and \( \tilde{H}_{k,c} = [\tilde{H}_{k,c}^{\text{w}}(\mathcal{U}_k), \ldots, \tilde{H}_{k,c}^{\text{w}}(\mathcal{U}_k \cup \{k\})]^T \).

3. Set \( \tilde{V}_{k,c}^{(0)} \) to be the orthogonal basis for the null space of \( \tilde{H}_{k,c} \) so that \( \tilde{V}_{i,c}^{(0)} = 0 \) for \( i \neq k \).

4. Perform SVD of \( \tilde{H}_{k,c} \) as in (46), and let \( \tilde{U}_{k,c}^{(1)} \) and \( \tilde{V}_{k,c}^{(1)} \) represent the first \( m_{k,c} \) left and right singular vectors of \( \tilde{H}_{k,c} \), respectively. Set \( F_{k,c}(j+1) = \tilde{U}_{k,c}^{(1)} \) and check the stopping criterion. If it is not satisfied, let \( j = j + 1 \) and go to Step 2, otherwise STOP.

As a result of the iterative process, the mismatch between \( F_{k,c} \) and \( \tilde{U}_{k,c}^{(1)} \) disappears, i.e. \( F_{k,c}(j) = \tilde{U}_{k,c}^{(1)}(j) \). Thus, the MAI is eliminated between users, i.e. \( F_{k,c}^H \tilde{H}_{k,c}^{\text{w}} M_{i,c} = 0, \ \forall \ i \neq k \). Similarly to (21), the receiver (7) can be reduced after some manipulation to

\[
W_{k,c}^H = (\Lambda_{k,c} P_{k,c} + I)^{-1} M_{k,c}^H \tilde{H}_{k,c}^{\text{w}} R_{k,c}^{-1}
\] (48)
since
\[ M_{k,c}^H \tilde{H}_{k,c} R_{k,c}^{-1} \tilde{H}_{k,c} M_{i,c} = 0, \quad \forall i \neq k \] (49)
and
\[ M_{k,c}^H \tilde{H}_{k,c} R_{k,c}^{-1} \tilde{H}_{k,c} M_{k,c} = \hat{\Lambda}_{k,c} P_{k,c} \] (50)
where the diagonal matrix \( \hat{\Lambda}_{k,c} = \text{diag}(\hat{\lambda}_{k,1,c}, \ldots, \hat{\lambda}_{k,m_{k,c},c}) \) includes the first \( m_{k,c} \) eigenvalues of \( \tilde{H}_{k,c} \).

The SINR for each sub-channel \( l \) is now reduced to:
\[
\gamma_{k,l,c} = \frac{E \left[ |w^H_{k,l,c} \tilde{H}_{k,c} m_{k,l,c} d_{k,l,c}|^2 \right]}{E \left[ w^H_{k,l,c} \left( \tilde{H}_{k,c} \left( \sum_{i \neq k} M_{i,c} d_{i,c} + \sum_{i \neq l} m_{k,i,c} d_{k,i,c} \right) + i_{k,c} + n_{k,c} \right) \right]^2} = \frac{\hat{\lambda}_{k,l,c} P_{k,l,c}}{m_{k,c}^H \tilde{H}_{k,c} R_{k,c}^{-1} \tilde{H}_{k,c} m_{k,l,c}} = \hat{\lambda}_{k,l,c} p_{k,l,c}
\] (51)
where \( w_{k,l,c} \) and \( m_{k,l,c} \) are the \( l \)th column vectors of the matrices \( W_{k,c} \) and \( M_{k,c} \), respectively.

Finally, the instantaneous sum rate of the BD method can be expressed as
\[
I_{\text{BD}}^{\text{(inst)}} = \max_{p_{k,l,c}} \frac{1}{N_C} \sum_{k \in U} \sum_{c=1}^{N_c} m_{k,c} \sum_{l=1}^{N_L} \log_2 \left( 1 + \hat{\lambda}_{k,l,c} p_{k,l,c} \right)
\] (52)
which can be maximised with respect to \( p_{k,l,c} \) subject to sum power constraint \( P_T \). Note that in the case of \( |U| = 1 \) the above procedure is replaced by the single user MIMO-OFDM precoder design from Section 4.1.1 [32, 40, 44].

The iterative solution may require in some extreme cases even tens of iterations to achieve the zero interference constraint \( \tilde{H}_{k,c} M_{i,c} = 0 \) for \( i \neq k \). In practice, the iterations can be terminated when the inter-stream interference terms \( \tilde{H}_{k,c} M_{i,c} \) for \( i \neq k \) are sufficiently small, e.g. 20 dB below the noise level [270]. It will be demonstrated by simulations that even a single reiteration round suffices to achieve most of the gains from the iterative solution. A linear MMSE filter (7) can be applied at the receiver to compensate for the remaining inter-stream interference. In such a case, however, the SINR per sub-carrier in (51) cannot be fully controlled at the transmitter. If the SINR values per eigenmode are fixed at the transmitter by the bit and power loading algorithm
in order to achieve certain FER, they cannot necessarily be guaranteed at the receiver due to the remaining interference. Therefore, some additional mechanisms are needed to maintain the quality of service at the receiver. This will be discussed in more detail in Section 5.2.5.

A simple non-iterative method was also presented in [266] to choose the proper beamforming matrices $\mathbf{F}_{k,c}$ at the transmitter for any $m_{k,c}$. The procedure is otherwise identical to Algorithm 1, except that only a single iteration is carried out and $\hat{\mathbf{H}}_{k,c}$ in (46) is defined as $\hat{\mathbf{H}}_{k,c} \tilde{\mathbf{V}}_{k,c}^{(0)}$ instead of $\hat{\mathbf{H}}_{w}^{w} \tilde{\mathbf{V}}_{w}^{(0)}$. The interference generated at the transmitter by using a fixed $\mathbf{F}_{k,c}$ is eliminated at the receiver beamformer outputs since it is steered into the nulls of the $\mathbf{F}_{k,c}$ beampatterns. In such a case, the receiver producing zero MAI is [266]

$$
\mathbf{W}_{k,c}^H = \hat{\mathbf{U}}_{k,c}^{(1)H} \mathbf{P}_{k,c}^H R_{k,c}^{-\frac{1}{2}}
$$

and hence $\hat{\mathbf{U}}_{k,c}^{(1)H} \mathbf{P}_{k,c}^H R_{k,c}^{-\frac{1}{2}} \tilde{\mathbf{H}}_{k,c} \mathbf{M}_{k,c} = \mathbf{0}$, $\forall i \neq k$ and $\hat{\mathbf{U}}_{k,c}^{(1)H} \mathbf{P}_{k,c}^H R_{k,c}^{-\frac{1}{2}} \tilde{\mathbf{H}}_{k,c} \mathbf{M}_{k,c} = \hat{\mathbf{A}}_{k,c} \mathbf{P}_{k,c}^H$. The MMSE filter (7) can be also used as a receiver for the non-iterative case, and in fact it always provides performance at least equal to (53). However, the control of the SINR per values per stream is more difficult due to the remaining inter-stream interference terms. The non-iterative method results in somewhat smaller sum rate than the iterative solution.

### 5.2 Scheduling algorithms

The number of beams assigned per user may vary depending on the channel realisation. The maximum number of beams for the linear transmission system is limited by the number of transmit antennas $N_T$ at the BS while the number of beams $m_k$ per user is limited by the number of receive antennas at the terminal. When $N_T = \sum N_{R_k}$, the value $m_k$ can get any value between zero and $N_{R_k}$ for each user $k$ depending on how good channels are available for the given user. If, for example, the transmitter has eight antennas, a maximum of eight orthogonal beams can be transmitted simultaneously. If the total number of receive antennas is larger than eight, multiuser or scheduling diversity is also available. Scheduling strategies may vary. For example, only the users’ beams with the best SNR can be selected resulting in maximum system throughput. However, fairness issues should also be considered in practical systems in order to avoid the starvation of some distant users. The optimal user selection per
each orthogonal dimension is generally a difficult combinatorial problem and it requires an exhaustive search over all possible combinations of user allocations. This is clearly computationally prohibitive for a large number of users. Therefore, sub-optimal allocation algorithms are considered in this section.

5.2.1 Greedy beam selection and ordering algorithm

The number of beams assigned per user and per sub-carrier may vary depending on the channel realisation so that

\[
0 \leq \sum_k m_{k,c} \leq N_T, \quad 0 \leq m_{k,c} \leq N_{R_k} \quad \forall c. \tag{54}
\]

A greedy user scheduling algorithm based on successive projections has been utilised in several studies [173, 177, 181] for the case where \(N_{R_k} = 1\). In this section, a greedy beam selection algorithm is proposed to maximise the DL sum rate of the linear block-ZF transmission (BD method) with multiple antennas at the receivers, \(N_{R_k} > 1\).

Since the transmitter vectors, and thus, the corresponding receiver vectors at each user are affected by the set of selected users, it is impossible to know the actual receiver structure \(W_{k,c}\) at the transmitter before the final beam allocation. An obvious candidate for an intelligent initial guess of the receiver matrix \(F_{k,c}\), and the one used in the proposed algorithm, is the optimum single user receiver, i.e. the left singular vectors of \(\tilde{H}_{k,c}\). For each user \(k\), \(F_{k,c}\) includes all the \(N_{R_k}\) left singular vectors of \(U_{k,c}\).

Let us define a 2-dimensional \(N_T \times N_C\) ordered set \(\mathcal{I}\) which includes the indices of the selected \(N_T\) beams for all the sub-carriers. The set \(\mathcal{I}\) can also be considered as a resource allocation table. The aim of the greedy beam selection algorithm is to select such beams that they create as little interference as possible to each other while providing a large beamforming gain. The proposed algorithm selects \(N_T\) beams for each sub-carrier \(c\) by comparing each row \(\tilde{h}_{i,c}\) of the stacked \(N_{R_k}^{\text{tot}} \times N_T\) row matrix

\[
\tilde{H}_c = \begin{bmatrix}
F_{1,c}^H \tilde{H}_{1,c}^w \\
\vdots \\
F_{K,c}^H \tilde{H}_{K,c}^w
\end{bmatrix}
\tag{55}
\]

\(\tilde{H}_c\) to the set of previously selected beams, and chooses the one having the highest
orthogonal distance to the space formed by the set of previously selected beams.

This can be formulated as the following maximisation problem where the \( l \)th beam index \( b_{l,c} \) for sub-carrier \( c \) is selected as:

\[
b_{l,c} = \arg \max_{1 \leq i \leq N^{\mathrm{tot}}, \text{i.e.} I_{l-1,c}} \min_{\mathbf{a}_c} \| \tilde{h}_{i,c} - \mathbf{a}_c \mathbf{H}_{I_{l-1},c,c} \|_2^2
\]  

(56)

where the inner minimisation defines the closest distance of \( \tilde{h}_{i,c} \) to the set of previously selected beams [282, p. 435]. The vector \( \mathbf{a}_c = (a_{1,c}, \ldots, a_{l-1,c}) \in \mathbb{C}^{l-1} \) is used to form all possible linear combinations of the equivalent channel vectors of the \( l-1 \) beams selected earlier by the algorithm after step \( l-1 \) and \( I_{l-1,c} \) is the set of indices of the previously selected \( l-1 \) beams for sub-carrier \( c \).

The equivalent channel matrix \( \tilde{H}_{I_{l-1},c,c} \) after step \( l-1 \) includes the equivalent channel vectors \( h_{i,c} \) of the previously selected \( l-1 \) users or beams for each sub-carrier \( c \).

The inner minimisation leads to

\[
\mathbf{a}^{\text{min}}_c = \arg \min_{\mathbf{a}_c} \| \tilde{h}_{i,c} - \mathbf{a}_c \tilde{H}_{I_{l-1},c,c} \|_2^2.
\]  

(57)

After differentiating and minimising the argument with respect to \( \mathbf{a} \), the solution can be expressed in closed form as

\[
\mathbf{a}^{\text{min}}_c = \tilde{h}_{i,c} \tilde{H}_{I_{l-1},c,c}^H \left( \tilde{H}_{I_{l-1},c,c} \tilde{H}_{I_{l-1},c,c}^H \right)^{-1} \tilde{H}_{I_{l-1},c,c}.
\]  

(58)

By incorporating (58) into (56), it is reduced to:

\[
b_{l,c} = \arg \max_{1 \leq i \leq N^{\text{tot}}, \text{i.e.} I_{l-1,c}} \| \tilde{h}_{i,c} - \left( \mathbf{I} - \mathbf{B}_{I_{l-1},c,c} \right) \|_2^2,
\]  

(59)

where the projection matrix \( \mathbf{B}_{I_{l-1},c,c} \) is defined as

\[
\mathbf{B}_{I_{l-1},c,c} = \tilde{H}_{I_{l-1},c,c} \tilde{H}_{I_{l-1},c,c}^H \left( \tilde{H}_{I_{l-1},c,c} \tilde{H}_{I_{l-1},c,c}^H \right)^{-1} \tilde{H}_{I_{l-1},c,c}.
\]  

(60)

The first beam is selected based on the largest channel eigenvalue, which may not be the optimal choice (full search) in all occasions. When \( N^{\text{tot}} = \sum N_{R_k} > N_T \), \( m_{k,c} \) per user \( k \) can get any value between zero and \( N_{R_k} \) depending on how strong channels are available for the given user and how compatible they are with other users and/or beams.
5.2.2 Selection of optimal number of beams

After the beams are ordered according to the previously described ordering criteria, the optimal number of beams per sub-carrier is selected for scheduling. Due to the inherent noise amplification property of the BD method, the power penalty can be significant if all $N_T$ spatial dimensions are utilised, especially in the low SNR region and with a low number of users. Therefore, it is often beneficial to allocate less than $N_T$ beams per sub-carrier, $m_{\text{tot}} = \sum_k m_{k,c} < N_T$, at a time to maximise the mutual information. Let us begin with a simple single carrier case with flat fading ($N_C = 1$, subscript $c$ omitted). A heuristic but still efficient solution for such a case is simply to find a subset of the first indices from the full ordered set of indices so that the mutual information is maximised:

$$l^* = \arg \max_{1 \leq l \leq N_T} \left[ \max_{P_{Z_l}} \log_2 \left| \frac{\mathbf{I} + \hat{\Lambda}_l P_{Z_l}}{N_0} \right| \right]$$

(61)

where $\mathbf{I}_l$ is the subset of the $l$ first indices of $\mathbf{I}$, $\hat{\mathbf{A}}_l \in \mathbb{R}^{l \times l}$ includes diagonal elements of $\text{blockdiag} \left( \hat{\mathbf{A}}_1, \ldots, \hat{\mathbf{A}}_K \right)$ corresponding to indices $\mathbf{I}_l$ and $P_{Z_l} \in \mathbb{R}^{l \times l}$ controls the powers allocated for each of the $l$ selected eigenmodes. Note that $m_k$ and $\hat{\mathbf{A}}_k$ computed as in (46) are updated for each $l$.

An OFDM system adds another dimension to the optimisation problem above. The optimal solution for the 2D space-frequency allocation problem would require an exhaustive search over all possible combinations of beam allocations for each sub-carrier. The number of computations required is clearly prohibitive. Therefore, a sub-optimal method is proposed below. A subset of indices from the full $N_T \times N_C$ set of indices $\mathbf{I}$ resulting in the maximum sum rate is found. The size of $\mathbf{I}$ is reduced one-by-one until the peak sum rate has been reached. The heuristic iterative algorithm for the selection of the optimal number of beams per sub-carrier is described in Algorithm 2.

The procedure described in Algorithm 2 does not necessarily increase the objective monotonically as a function of the iteration index $j$. Even though the eigenmode gains $\hat{\mathbf{A}}_c$ in the remaining beams are generally increased as the least significant beam in $\mathbf{I}_c$ is discarded, the increase does not necessarily always compensate for the loss of the discarded beam. Therefore, the parameter $\delta \geq 1$ is used to prevent the algorithm from stagnating into a local maximum. The simulations show that $\delta \geq 5$ is enough to avoid such local maxima.
Algorithm 2 Greedy beam allocation

1. Compute the (iterative) BD decomposition (Algorithm 1) for the selected beams in allocation table $I$, and compose an $N_T \times N_C$ eigenvalue matrix $\hat{\Lambda}$, where the columns $\hat{\Lambda}_c$ are diagonal elements of $\text{blockdiag}(\hat{\Lambda}_1,c, \ldots, \hat{\Lambda}_K,c)$ corresponding to indices $I_c$. Initialise the iteration index $j = 1$.

2. Compute the mutual information $I_{BD}^{(\text{inst})}$ as in (52).

3. Look for the sub-carrier $c$ with the lowest eigenmode $\{c,l\}^* = \arg \min_{c=1,\ldots,N_C, l=1,\ldots,|I_c|} \hat{\Lambda}_c$.

4. Discard the last index from the set $I_c^*$ and update $m_{k,c^*}$ of the affected user $k$ accordingly.

5. Compute Algorithm 1 for the affected sub-carrier with the reduced set of beams, and update $\hat{\Lambda}_{c^*}$.

6. Goto Step 2. and iterate until $I_{BD}^{(\text{inst})}(j) < I_{BD}^{(\text{inst})}(j-\delta)$, where parameter $\delta \geq 1$ is used to avoid local maxima.

7. Select $\tau^* = \arg \max_{\tau=1,\ldots,\delta} I_{BD}^{(\text{inst})}(j-\tau)$ as the maximum sum rate with final allocation table $I(j-\tau^*)$, where $\tau^* = \arg \max_{\tau=1,\ldots,\delta} I_{BD}^{(\text{inst})}(j-\tau)$.

Figs. 12(a) and 12(b) illustrate the evolution of $I_{BD}^{(\text{inst})}$ for one channel realisation with different SNR ($= P_k / N_0$) values versus the number of iterations for $\{N_T, N_R_k, K, N_C\} = \{4, 2, 2, 64\}$ and $\{N_T, N_R_k, K, N_C\} = \{4, 4, 2, 64\}$ systems, respectively. Furthermore, the behaviour of the algorithm is plotted for the cases with and without iterative BD processing. The maximum sum rate $I_{BD}^{(\text{inst})}$ for each SNR is indicated with asterisk and triangle markers. It can be seen from the figures that achieving the maximum $I_{BD}^{(\text{inst})}$ with the BD method requires partial loading in the spatial domain, especially in with a low SNR. Moreover, significant gains from iterative BD processing are available when $m_{k,c} < N_{R_k}$.

5.2.3 Greedy beam allocation with Hughes-Hartogs loading

In this section, the allocation problem introduced in the previous sections is extended to a case with practical bit and power loading algorithms. Greedy loading algorithms are considered that try to maximise the achievable sum spectral efficiency for certain quality of service criteria, such as the target FER. Again, the idea is to look for the best allocation of beams resulting in
Fig 12. Evolution of sum rate vs. discarded sub-carriers / spatial modes.

(a) $N_T = 4, N_{R_k} = 2, K = 2, N_C = 64, \text{SNR} = 0 - 20 \text{ dB}$

(b) $N_T = 4, N_{R_k} = 4, K = 2, N_C = 64, \text{SNR} = 0 - 20 \text{ dB}$
the highest spectral efficiency by iteratively discarding one of the beams from the allocation table and recomputing the BD decomposition for the affected sub-carrier. An optimal discrete bit and power loading algorithm, the Hughes-Hartogs algorithm [95], is used in Algorithm 3 to compute the maximum sum spectral efficiency $\zeta$ for each 2D beam allocation.

**Algorithm 3** Greedy beam allocation with Hughes-Hartog loading

1. Equivalent to Step 1. in Algorithm 2.
2. Run HH algorithm on the values of $\hat{A}$ with power constraint $P_T$ to compute $\zeta$.
3. Equivalent to Step 3. in Algorithm 2.
4. Equivalent to Step 4. in Algorithm 2.
5. Equivalent to Step 5. in Algorithm 2.
6. Goto Step 2. and iterate until $\zeta(j) < \zeta(z - \delta)$.
7. Select $\max_{\tau=1,\ldots,\delta} \zeta(j - \tau)$ as the maximum sum SE with final allocation table $I(j - \tau^*)$, where $\tau^* = \arg \max_{\tau=1,\ldots,\delta} \zeta(j - \tau)$.

An example in Fig. 13 illustrates the SE evolution for one channel realisation with different SNR ($= P_R/N_0$) values versus the number of iterations for an $\{N_T, N_{R_k}, K, N_C\} = \{4, 2, 2, 64\}$ system (solid curves). The parameters for the modulation and coding scheme set and the corresponding SNRs required by each MCS to achieve the desired target FER are given in Table 2. The maximum available sum SE with the given parameter set is 12 bits/s/Hz. The maximum SE for each SNR is pointed out with an asterisk marker. One way to simplify the algorithm above is to one-by-one discard the last row entirely from the index matrix $I$ until the peak SE is reached. The spectral efficiency achieved with this method (dashed curves, max value marked with triangle) is also shown in Fig. 13. It can be seen from the figure that achieving the maximum SE with the BD method requires partial loading in the spatial domain, especially with a low SNR. In the high SNR region, the system becomes saturated due to the finite MCS set, and full spatial loading is required. Moreover, a simplified version of the algorithm can achieve nearly the same performance, especially with a low SNR.

The greedy allocation of resources among sub-carriers and spatial modes results in an OFDMA solution where the allocation table can be severely fragmented ($0 < m^\text{tot}_k \leq N_T$, $0 \leq m_{k,c} \leq N_{R_k}$). In addition, the MCS selected...
per sub-carrier per beam can be any from the given MCS set. Moreover, the iterative greedy algorithm proposed above can be still computationally expensive for practical applications. Therefore, less complex scheduling and loading algorithms with low signalling overhead are required.

5.2.4 Low complexity bit, power and beam allocation with low signalling overhead

A low complexity bit and power loading algorithm requiring a low signalling overhead was introduced in [277] for a single user adaptive MIMO-OFDM system. More details of the algorithm can be found in Section 4.1.3. Herein, the single user algorithm is extended to the considered multiuser MIMO-OFDM case. The aim is to avoid the fragmentation of user/beam allocation in the sub-carrier domain by imposing a condition $m_{k,c} = m_k \forall c$, i.e. the number of beams per
sub-carrier is the same for each selected user $k$. Consequently, the single user loading algorithm is computed for each user $k$ with $m_k$ clusters (allocated beams) as in Section 4.1.3. For each transmitted frame, the algorithm first divides the total power between $\sum_k m_k \leq \min(N_t^{\text{tot}}, N_T)$ clusters according to their eigenmode gains, and then, the MCS and the selected eigenmodes (sub-carriers) in each cluster are independently optimised to maximise the throughput subject to an equal SNR constraint for all the selected sub-carriers. The rest of the processing is identical to the single user loading algorithm in Section 4.1.3. The amount of signalling needed is significantly less than in the case of the greedy scheduling and loading algorithm introduced in Section 5.2.3. The signalling information required at the terminal is reduced to the number of beams/clusters allocated for each selected user and the MCS used on each cluster.

The allocation problem is now reduced to finding the optimum $m_k$ for each user $k$, and the optimum total number of beams $m_c^{\text{tot}} = m_t^{\text{tot}} \forall c$. The straightforward solution is to perform an exhaustive search over all combinations of possible user/beam allocations, where the BD decomposition and loading is computed for each allocation, and the best allocation with the highest spectral efficiency is selected. Some statistical fairness can also be easily included by limiting the maximum number of beams per user, for example. Obviously, the complexity of this method grows with increasing $K$, as the number of beam combinations (iterations) increases. In addition, the scheduling gain from an exhaustive search quickly saturates in highly frequency selective channels due to frequency diversity on vertically coded frames. For a relatively small $K$ ($\approx N_T$), however, the performance of the proposed method is very close to the greedy scheduling and loading algorithm (5.2.3), as will be demonstrated numerically in Section 5.3.2.

5.2.5 Compensation for estimation deficiencies at the transmitter

A simple closed-loop algorithm for compensating for the effect of interference non-reciprocity at the transmitter was introduced in Section 4.2. The basic idea of the method was to apply a single power offset value at the transmitter over all eigenmodes which compensates for the frequency, time and space selective
non-reciprocal interference between the TDD DL and the UL. The receiver estimates a signal quality metric (BER after 2nd turbo decoding iteration corresponding to the target FER) from the transmitted signal, compares the estimated metric to the target metric, and transmits the updated offset value to the transmitter. The offset value is then used to adjust the transmission rate per sub-channel so that the target signal quality metric is reached.

The same method is used herein to compensate for the residual MAI caused by either the limited number of iterations in the iterative BD decomposition and/or the channel estimation errors at the transmitter. The power offset value feedback is used to modify the waterfilling reference level \( (I_0 + N_0) \) [dB] in the bit and power loading algorithm

\[
\{p_{k,l,c}, MCS_{k,l,c}\} = f \left( I_0 + N_0 + P_{\text{offset}}^{k}, \lambda_{k,l,c}, P_T \right), \forall k, l, c
\]

where \( f \) denotes the function for the bit and power loading algorithm as in Section 4.2. Note that adding the offset \( P_{\text{offset}}^{k} \) [dB] to \( I_0 + N_0 \) is equal to adjusting the values \( \lambda_{k,l,c} \) \( \forall l, c \) by the corresponding \( P_{\text{offset}}^{k} \) [lin] as in (31).

5.3 Numerical examples

5.3.1 Performance of block-ZF transmission with greedy scheduling

The performance of the greedy scheduling algorithm is compared with other scheduling criteria. Three user scheduling criteria are used in the simulations:

- **Greedy scheduling (GS)**: At most \( N_T \) beams are selected such that they create a small amount of interference to each other while having large channel eigenvalues. The first beam is selected based on the largest channel eigenvalue.
- **Maximum eigenvalue (ME) scheduling**: The eigenvalues of the equivalent channel matrices \( \bar{H}_{k,c} \) of each user are simply sorted and at most \( N_T \) beams providing the maximum sum rate are selected for the transmission at any time instant. Spatial compatibility with other beams is not considered.
- **Best user scheduling**: A single user with the maximum channel norm is selected at any time instant.

104
The greedy scheduling algorithm follows the procedure described in Sections 5.2.1–5.2.2. All the algorithms above try to achieve the maximum system throughput without considering the fairness or the QoS of a single user. Both fairness and QoS issues are considered in more detail in Chapter 6.

The performance of the proposed algorithms is studied by simulations in a single-cell environment. In the simulations, the users’ channels are assumed to be constant during a transmission block, and changing randomly from block to block. The sum rate of the BD method with coordinated TX-RX processing and with the proposed allocation algorithms is compared to the sum rate capacity. The single user capacity with and without CSI at the transmitter is plotted for comparison (dotted curves in the figures). Two different spatial loading scenarios are considered in the simulations:

- **Fully loaded**: $\min(N_{\text{tot}}, N_T)$ beams are always allocated for each channel realization (plotted with dashed curves in the figures).
- **Partially loaded**: the optimal number of beams from the ordered set of beam indices $S_{NT}$ is selected for the scheduling based on the optimisation criteria in (61). These results are plotted with solid curves in the figures.

MIMO antenna configurations considered in the simulations consist of 4–16 transmit antennas at the base station and each mobile node equipped with 2–4 receive antennas. The simulation cases are denoted as $N_T \times N_{R_k}$, 4 × 2 and 16 × 2, for example. Each user is equipped with the same number of receive antennas. The number of users present in the cell is varied from 1 to 64. In this section, a flat fading scenario is considered, i.e. the number of sub-carriers is limited to $N_C = 1$, for simplicity. Moreover, only the non-iterative version of the coordinated TX-RX processing (53) is used in these simulations.

Fig. 14 shows an example of the sum rate for the 4 × 2 antenna configuration versus the SNR with 8 users and different scheduling algorithms. It can be seen in this particular scenario that the performance of the BD method with intelligent scheduling is rather close to the sum rate capacity even with 8 users. The loss from the noise amplification with full spatial loading is rather small with GS scheduling, and it is non-existent in the high SNR region. ME scheduling with full spatial loading performs even worse than the simple time-switched best user scheduling in the low SNR region.
Figs. 15(a) and 15(b) show the sum rate for the $4 \times 2$ MIMO configuration versus the number of users at 0 dB and 20 dB SNR values, respectively. The results indicate that the BD method with greedy scheduling approaches the sum rate capacity in the high SNR region as the number of users present in the system becomes large. However, in the low SNR region the capacity loss from the noise amplification can be significant as can be seen by comparing the curves labelled as 'fully loaded' to 'partially loaded' cases in Fig. 15(a). The BD method suffers from the noise amplification in the fully loaded case especially when $N_{R}^{\text{tot}} = N_T$ ($K = 2$ in Fig. 15(a)) and no scheduling gains are available. In such a case the sum rate of the BD method can fall even below the single user capacity. Therefore, it is often beneficial to allocate less beams than the spatial dimensions available allow in order to reduce the noise amplification, especially in the low SNR region and with a low number of users. In the high SNR region...
Fig 15. $\{N_T, N_R, N_C, K\} = \{4, 2, 1, 1 - 64\}$ MIMO system with scheduling.
(Fig. 15(b)), the difference between the fully loaded and partially loaded cases is small or non-existent in the studied $4 \times 2$ scenario.

Due to implementation limitations it is often difficult if not impossible to increase the number of antenna elements at the mobile receiver beyond certain limits. Therefore, it can be easier to increase the sum rate of the system by adding more antennas at the base station transmitter. In the following figures, $N_T$ is increased to 16 while $N_{R_k} = 2$. Again, the sum rate versus the number of users for the $16 \times 2$ MIMO configuration at fixed 0 dB and 20 dB SNR values is shown in Figs. 16(a) and 16(b), respectively.

The main difference to the $4 \times 2$ scenario is that the impact of noise amplification is far more significant in the $16 \times 2$ scenario. The sum rate of both ME and GS scheduling algorithms with full spatial loading can drop below the single user capacity in the low SNR region. Far better performance is achieved by allocating only an intelligently selected subset of beams at a time. In addition, in the high SNR region (Fig. 16(b)) the noise amplification is significant with a moderate number of users, and partial loading helps to smooth out the notch in the curves. In the case of the greedy algorithm, the difference between fully loaded and partially loaded cases disappears as the number of users becomes large.

Figs. 17(a) and 17(b) illustrate the number of beams allocated versus the total number of users present in the system for $4 \times 2$ and $16 \times 2$ MIMO configurations, respectively. The greedy scheduling with partial spatial loading is considered. It can be seen from the figure that the optimal number of allocated beams can be as low as half the number of spatial dimensions available in the low SNR region. With a high SNR, however, the number of allocated beams converges to the number of spatial dimensions $N_T$ as the number of users becomes large.

5.3.2 User, bit and power allocation with low signalling overhead

The performance of the scheduling algorithms proposed in Sections 5.2.3–5.2.5 is now studied in a more realistic environment by simulations. The simulated OFDM system is based on the parameters given in Table 2 on page 75. In all
Fig 16. \( \{N_T, N_{R_k}, N_C, K\} = \{16, 2, 1 \ldots 64\} \) MIMO system with scheduling.
cases, the users’ channels are assumed to be constant during one coded OFDM frame, and changing randomly from block to block. The channel’s delay taps are considered independent of each other with a power delay profile specified by ETSI BRAN Channel A [279]. The number of TX antennas is set to four, while the number of RX antennas were varied between 2 and 4. The fading in antenna elements is assumed to be independent of each other. The number of active users is kept relatively low varying from 1 to 8.

The sum spectral efficiency with the proposed LSO allocation algorithm and with linear non-iterative or iterative BD decomposition is depicted in Fig. 18 for \( \{N_T, N_{R_k}, N_C, K\} = \{4, 2-4, 2\} \). The maximum number of beams per user is limited to \( m_k \leq 2 \) in all cases for fair comparison. The performance of TDMA with best user scheduling is also depicted for reference. The best user scheduling means simply that the user with higher spectral efficiency is selected for scheduling from two users at each time instant. The performance of BD is shown to be superior to TDMA in all cases. The sum spectral efficiency of the BD method with \( N_T = 4 \) and with 10% FER target is saturated to \((1 - 0.1) \times 4 \times 3 = 10.8\) bits/s/Hz with a high SNR.

The rightmost curves with the square tick mark in Fig. 18 from the groups of curves with \( N_{R_k} = 2 \) and 4 correspond to the original fully loaded case \( \{m_1, m_2\} = \{2, 2\} \) with non-iterative BD processing [266] and with the LSO algorithm.
Fig 18. Non-iterative versus iterative BD decomposition with LSO loading algorithm, \(\{N_T, N_{Rx}, N_C, K\} = \{4, 2-4, 64, 2\}\).

from Section 5.2.4. The iterative search for optimum allocation (partial spatial loading) results in significant gains in the low SNR region, especially for \(N_{Rx} = 2\). The possible combinations for \(m_k\) used in the simulations were \(\{m_1, m_2\} = \{\{2, 2\}, \{2, 1\}, \{1, 2\}, \{1, 1\}\}\).

The iterative BD decomposition gives significant gains when \(m_k < N_{Rx}\), as seen clearly from the curves with \(N_{Rx} = 4\). For \(N_{Rx} = 2\), all the curves overlap with a high SNR as the maximum sum SE is achieved from full spatial loading \((m_k = N_{Rx})\), and there is no gain from the iterative loading algorithm nor iterative BD decomposition since \(N_{R}^{\text{tot}} = N_T\). Even a single reiteration round (labelled as 2 iters.) is sufficient to achieve most of the gains from iterative BD decomposition. This can be seen by comparing the curves with 2 iterations to the ones with 20 iterations, which was sufficient in nearly all of the channel realisations to converge.
Fig. 19. Comparison of LSO algorithm to greedy allocation with HH loading, $\{N_T, N_{R_k}, N_C, K\} = \{4, 2 − 4, 64, 2 − 8\}$.

Fig. 19 compares the sum spectral efficiency provided by the proposed LSO algorithm (Section 5.2.4) to the greedy algorithm with HH loading (Section 5.2.3). In addition, the single user performance (TDMA, best user scheduling) is plotted for comparison. The results show that the performance loss to the greedy algorithm, less than 10% smaller sum spectral efficiency, is rather insignificant for small $K$ considering the far less complicated processing and signalling. In the four user case with LSO, the number of beams per user was limited to one ($m_k = 1$) in order to reduce the number of allocation combinations. The optimisation was reduced to finding the optimum $r$, which was mostly 4 in this case, except in the very low SNR region. Allowing $m_k > 1$ for LSO, slightly better spectral efficiency could have been achieved with a high SNR but with increased complexity.

The performance loss to the greedy algorithm increases as $K$ becomes large,
as the scheduling gains in the sub-carrier domain cannot be fully utilised. If $K$ is doubled to 8, the sum SE of the greedy algorithm improves by 10% while the SE of the LSO algorithm is only increased by 5% (not shown in Fig. 19). However, the practical implementation of the greedy algorithm as such is difficult for large $K$. In addition to the large amount of signalling required and the high complexity, the allocation table becomes fragmented and it is difficult to construct data frames such that the target (frame) error rate can be controlled and the frames are long enough to achieve some coding gain.

As the basic principle of the BD method is to orthogonalise the scheduled beams/users, it is interesting to see what is the impact of the partially lost orthogonality due to imperfect transmitter CSI. Fig. 20 depicts the spectral efficiency with three different levels of channel estimation errors $\sigma^2_{\text{est}} = [-\infty, -20, -10]$ dB. It is seen from Fig. 20 that high channel estimation error destroys the orthogonality between beams/users and full spatial loading cannot be supported in the case where $N_{R}^{\text{tot}} = N_{T}$, i.e. saturated sum spectral efficiency cannot be reached even with a very high SNR.

A straightforward solution to overcome the problem is to increase the number of receive antennas, or to impose a limit $m_k < N_{R_k}$, to give more degrees of freedom for interference suppression at the receiver. Three receiver antennas are sufficient for reaching the saturated sum SE for $\sigma^2_{\text{est}} = -10$ dB, with a clear penalty to the full CSI case, though. The plot is not fully fair, however, since the channel estimation error does not scale along the SNR axis but is fixed for the whole SNR range. In reality, the channel estimation error could be less with a high SNR and vice versa.

When $N_{R_k} = N_{T}$, one clear advantage of MU-MIMO processing is lost since the TDMA solution also has an equal number of spatial modes to be utilised and it can reach the saturated SE with a high SNR. In Fig. 21, a comparison between 2-user BD and TDMA with the same number of TX and RX antennas is carried out. It is seen that the BD method clearly outperforms the TDMA with best user scheduling even with large channel estimation errors. With more users the advantage for BD would be even more clear.
Fig 20. Performance of BD MU MIMO-OFDM system with channel estimation errors, \(\{N_T, N_R, N_C, K\} = \{4, 2 − 4, 64, 2\}\).

5.4 Summary and discussion

Efficient scheduling algorithms were proposed to maximise the downlink sum rate of the multiuser MIMO system with generalised block-ZF transmission for any number of users and receive antennas. The performance was compared to the sum rate available with several other scheduling algorithms and to the sum-rate capacity. It was shown that the performance of the block-ZF method with coordinated transmitter-receiver processing and with greedy scheduling approaches the sum rate capacity with a high SNR as the number of users increases and the equivalent channel vectors become linearly more independent.

Due to the inherent noise amplification problem with linear block-ZF transmission, the maximum sum rate is often achieved by transmitting to less users/beams than the spatial dimensions available, especially in the low SNR region, with a
An efficient user, beam, bit and power allocation algorithm was also proposed to maximise the downlink sum spectral efficiency of the adaptive multiuser MIMO-OFDM system. The performance of the proposed allocation algorithm with low signalling overhead was shown to be close to the greedy allocation algorithm with Hughes-Hartogs loading. The performance loss was less than 10% for a small number of active users. The iterative block-ZF method with a single reiteration round was shown to provide significant gains to the non-iterative solution, especially if the number of beams assigned per user is less than the number of receive antennas. An LMMSE filter is applied at the receiver to suppress the remaining multi access interference from incomplete orthogonalisation together with a simple compensation algorithm with low rate scalar feedback to the transmitter. The proposed MU MIMO-OFDM system was also shown to be robust against imperfect channel estimation at the transmitter,
always providing superior performance to the single user TDMA solution.

The method proposed in this thesis for compensating for the imperfect TX
CSI is rather simple because it utilises neither statistics nor the behaviour of the
CSI imperfection. A refined approach considering an elaborated robust design
against imperfect CSI at the transmitter is a promising research line for future
work.
6 BS cooperation with linear transceiver processing

The purpose of this chapter is to analyse the BS cooperation with linear processing in more detail and propose practical radio resource allocation solutions. Several users having identical SHO active sets $S_k$ can be served in the same time-frequency transmission slot by separating their transmissions in the space domain, as illustrated in Fig. 2. Consequently, each group of users with identical SHO active set composition forms a distinct user set and can be optimised separately.

A general method for joint design of the linear TX and RX beamformers for several optimisation criteria subject to per BS power constraints is proposed. The optimisation criteria studied in this chapter are:

- Power minimisation subject to per stream SINR constraints
- Power minimisation subject to individual user rate constraints
- Minimum weighted SINR maximisation, i.e. SINR balancing
- Weighted sum rate maximisation
- Minimum weighted rate maximisation, i.e. weighted common rate maximisation, i.e. weighted rate balancing

An extension to the frequency selective DL channel with additional QoS constraints, such as a guaranteed bit rate (GBR) per user is also provided. The methods proposed can handle multiple antennas at BS’s and mobile users, and any number of data streams is allowed per scheduled user. Furthermore, it can be easily modified to accommodate supplementary constraints, e.g. per antenna power constraints or upper/lower bounds for the SINR values of data streams, and the feasibility of the resulting optimisation problems can easily be checked.

Unlike the optimal non-linear transmission schemes, the optimisation problems employed in the linear multiuser MIMO transceiver design are not convex in general. However, an iterative solution is proposed where each sub-step can be efficiently solved by using convex optimisation tools [78]. Even though each subproblem is optimally solved, global optimality cannot be guaranteed due to the non-convexity of the original problem. However, the simulation results demonstrate that the achieved locally optimal solutions are very efficient in
several practically relevant scenarios. Furthermore, particular emphasis is put on
generalised ZF transmission due to its simplicity. It enables decoupling the data
streams, and, as a result, allows for efficient implementation of the bit and power
loading algorithms in practical systems.

System level evaluation is carried out in order to assess the impact of a realistic
multi-cell environment (including non-reciprocal inter-cell interference) on the
cellular MIMO-OFDM system performance. The impact of the size of the SHO
region, overhead from increased hardware and physical (time, frequency) resource
utilisation, as well as different non-reciprocal inter-cell interference distributions
due to the SHO are evaluated by system level simulations. Even though large link
and system level gains are available from cooperative processing, the importance
of providing a common phase reference for the baseband processing will be
emphasised. Feedback from terminals should be used to compensate for possible
RF impairments between the distributed antenna heads.

The rest of the chapter is organised as follows. Section 6.1 considers the joint
design of the linear transmit and receive beamformers for various optimisation
criteria subject to per BS or per antenna power constraints. The cooperative
processing in the SHO with ZF multiuser transmission is introduced in Section 6.2
and the relevant power optimisation problems are derived. Section 6.3 introduces
a method to compute the symmetric capacity subject to per BS or antenna power
constraints. Section 6.4 describes the simulation environment and assumptions,
and the results for theoretical mutual information studies and for more practical
link and system level simulation studies are presented. Finally, conclusions
are drawn in Section 6.5. In the derivation in Sections 6.1–6.3, the focus is
restricted to a single set of users $\mathcal{U}$, where all users $k \in \mathcal{U}$ have identical active
set composition, $\mathcal{S}_k = \mathcal{S}_i, \forall k, i \in \mathcal{U}$. The SHO active set size is denoted by
$M = |\mathcal{S}_k|$, which is common to all users $k \in \mathcal{U}$.

6.1 Joint design of linear TX and RX beamformers with
per BS or antenna power constraints

In this section, the focus is on linear TX schemes, where the transmitters send
$S$ independent streams, $S \leq \min(MN_T, \sum_{k \in \mathcal{U}} N_{R_k})$ per transmit dimension
(frequency, time). The subcarrier index $c$ is omitted to simplify the presentation
in this section, but the results can be straightforwardly extended to multiple subcarriers similarly to Section 6.1.5. Per data stream processing is considered, where for each data stream \( s, s = 1, \ldots, S \) the scheduler unit associates an intended user \( k_s \), with the channel matrix \( \tilde{H}_{k_s} \in \mathbb{C}^{M_{N_T} \times N_{R,s}} \). Note that more than one stream can be assigned to one user, i.e. the cardinality of the set of scheduled users, \( \mathcal{U} = \{ k_s | s = 1, \ldots, S \} \), is less than or equal to \( S \). In order to simplify the presentation in this section but without loss of generality, the interference is assumed to be non-existent and normalised to one, i.e. \( R_{k_s} = I \). In case of coloured inter-cell interference, \( \tilde{H}_{k_s} \) is replaced by the whitened channel matrix \( \tilde{H}_{k_s}^{w} = R_{k_s}^{-1/2}\tilde{H}_{k_s} \).

Let \( m_s \in \mathbb{C}^{M_{N_T}} \) and \( w_s \in \mathbb{C}^{N_{R,s}} \), \( \| w_s \|_2 = 1 \) be arbitrary transmit and receive beamformers for the stream \( s \). \( m_s \) can be further split into \( m_s = \sqrt{p_s}v_s \), where \( v_s \in \mathbb{C}^{M_{N_T}} \) and \( p_s \) are the normalised transmit beamformer and the allocated power for the stream \( s \), respectively. The SINR of the data stream \( s \) can be expressed as

\[
\gamma_s = \frac{|w_s^H\tilde{H}_{k_s}m_s|^2}{1 + \sum_{i=1,i\neq s}^S |w_s^H\tilde{H}_{k_i}m_i|^2} = \frac{p_s|w_s^H\tilde{H}_{k_s}v_s|^2}{1 + \sum_{i=1,i\neq s}^S p_i|w_s^H\tilde{H}_{k_i}v_i|^2}.
\]

Similarly to (10), the total power transmitted by the \( n \)th BS is given by \( \sum_{s=1}^S \| m_s^n \|_2^2 = \sum_{s=1}^S p_s \| v_s^n \|_2^2 \), where \( m_s^n \in \mathbb{C}^{N_{R,s}} \) and \( v_s^n \in \mathbb{C}^{N_T} \) are the unnormalised and normalised transmit vectors, respectively, for the data stream \( s \) associated with BS \( n \), i.e. \( m_s^n = [m_s]_{(n-1)N_T+1:nN_T}, n = 1, \ldots, M \) and \( v_s^n = [v_s^1]^T, \ldots, [v_M]^T]^T \).

### 6.1.1 SINR balancing

Suppose now that the system has to keep the SINR per data stream \( \gamma_s \) in fixed ratios in order to guarantee fairness between streams/users, i.e. \( \gamma_s/\beta_s = \gamma_o \), and \( \gamma_o \) has to be maximised subject to per BS power constraints. This can be
formulated as maximisation of the minimum weighted SINR per stream:

\[
\begin{align*}
\text{maximise} \quad & \min_{s=1,\ldots,S} \frac{\beta_s^{-1}|w_s^H \tilde{H}_{ks} m_s|^2}{1 + \sum_{i=1,i\neq s}^S |w_i^H \tilde{H}_{ks} m_i|^2} \\
\text{subject to} \quad & \begin{array}{l}
\sum_{s=1}^S \|m_s[n]\|^2_2 \leq P_n, \quad n = 1, \ldots, M \\
\|w_s\|^2_2 = 1, \quad s = 1, \ldots, S
\end{array}
\end{align*}
\]

(65)

where the variables are \( m_s \in \mathbb{C}^{MN_T} \) and \( w_s \in \mathbb{C}^{N_{Rks}} \), \( s = 1, \ldots, S \). The weights, \( \beta_s > 0 \), \( s = 1, \ldots, S \), are used to prioritise the data streams of different users differently and they can be chosen based on different criteria, e.g. by the preference class or by the latency requirements of the particular service. Obviously, when \( \beta_s = 1, \forall s \), (65) reduces to the classical worst SINR maximisation. Problem (65) is not jointly convex in variables \( m_s \) and \( w_s \). However, for a fixed \( m_s \), (65) has a unique solution given by the normalised LMMSE receiver [238]

\[
w_s = \frac{\tilde{w}_s}{\|\tilde{w}_s\|_2}, \quad \tilde{w}_s^H = m_s^H \tilde{H}_{ks}^H \left( \sum_{i=1}^S \tilde{H}_{ks} m_i m_i^H \tilde{H}_{ks}^H + I \right)^{-1}
\]

(66)

which provides the maximum SINR for stream \( s \). Furthermore, for a fixed \( w_s \), (65) is quasiconvex in \( m_s \) [223]. Thus, it can be solved by using the bisection method [15]. Note that the constraints are also separable, i.e. they act on distinct sets of the variables \( m_s \) and \( w_s \). The above observations suggest using a coordinate ascent method [283] for solving (65). At each iteration the objective is maximised with respect to one set of variables \( w_s \) (or \( m_s \)) by considering the other set fixed. This leads to Algorithm 4. Note that a similar method was also used for minimum power beamforming design in [245, 254, 255].

**Algorithm 4** SINR balancing under per BS power constraints

1. Initialise \( m_s^{(0)} \) such that BS power constraints are satisfied. Let \( j = 1 \).
2. Compute \( w_s^{(j)} \), \( s = 1, \ldots, S \) given by (66), where \( m_s = m_s^{(j-1)} \), \( s = 1, \ldots, S \).
3. Solve (65) for the variables \( m_s \), \( s = 1, \ldots, S \) by fixing \( w_s = w_s^{(j)} \), \( s = 1, \ldots, S \).

Denote the solution by \( m_s^* \) and update the transmit beamformers \( m_s^{(j)} = m_s^* \), \( s = 1, \ldots, S \). Test a stopping criterion. If it is not satisfied, let \( j = j + 1 \) and go to Step 2, otherwise STOP.
Subsequently, Step 3 of Algorithm 4 can be solved with any accuracy $\epsilon > 0$ by the bisection method [15] presented in Algorithm 5. In order to speed up Algorithm 5, SINR balancing for fixed receive beamformers:

1. Initialise $\gamma_{\text{min}} = \text{SINR}_{\text{min}}$ and $\gamma_{\text{max}} = \text{SINR}_{\text{max}}$, where $\text{SINR}_{\text{min}}$ and $\text{SINR}_{\text{max}}$ define the range of relevant SINRs. Let $\epsilon > 0$ be the desired accuracy.
2. Set $\gamma_0 = (\gamma_{\text{max}} + \gamma_{\text{min}})/2$
3. Solve the following feasibility problem

$$\begin{align*}
\text{find} & \quad m_s, \quad s = 1, \ldots, S \\
\text{subject to} & \quad \frac{|w_s^H\tilde{H}_{k_s}m_s|^2}{1 + \sum_{i=1, i \neq s}^S |w_s^H\tilde{H}_{k_s}m_i|^2} \geq \beta_s \gamma_o, \quad s = 1, \ldots, S \\
& \quad \sum_{s=1}^S \|m_s\|^2 \leq P_n, \quad n = 1, \ldots, M
\end{align*}$$

(67)

If the problem is feasible, then set $\gamma_{\text{min}} = \gamma_0$. Otherwise, set $\gamma_{\text{max}} = z$.
4) If $(\gamma_{\text{max}} - \gamma_{\text{min}}) > \epsilon$ then go to Step 2. Otherwise, return $m_s^* = m_s, \quad s = 1, \ldots, S$, where $m_s$ is the last feasible solution of (67) and STOP.

Algorithm 4, SINR$_{\text{min}}$ can be also replaced by $\gamma_o$ from the previous iteration, since $\gamma_o$ increases monotonically in each iteration.

Note that the constraints of problem (67) can be expressed as a generalised inequality with respect to the second-order cone [15, 223], and hence, it can be solved by using a second-order cone program (SOCP) solver [78]. By reformulating the approach presented in [223, Section IV.B] as a feasibility problem and accommodating additional per BS power constraints, (67) can be presented as the following SOCP

$$\begin{align*}
\text{find} & \quad m_s, \quad s = 1, \ldots, S \\
\text{s. t.} & \quad \begin{bmatrix}
\sqrt{1 + \frac{1}{\beta_s \gamma_o} w_s^H \tilde{H}_{k_s} m_s} \\
\frac{1}{M^H \tilde{H}_{k_s} w_s} \\
\sqrt{T_n} \\
\text{vec}(M[n])
\end{bmatrix} \succeq \text{SOC}_0, \quad s = 1, \ldots, S \\
& \quad \begin{bmatrix}
1 \\
\text{vec}(M[n])
\end{bmatrix} \succeq \text{SOC}_0, \quad n = 1, \ldots, M
\end{align*}$$

(68)

where the variables are $m_s \in \mathbb{C}^{MN_T}$, and where $M = [m_1, \ldots, m_S], M[n] = \ldots$
The inequality with respect to the second-order cone is denoted by \( \succeq_{\text{SOC}} \). i.e., for any \( x \in \mathbb{R} \) and \( y \in \mathbb{R}^n \), \( [x, y^T]^T \succeq_{\text{SOC}} 0 \) is equivalent to \( x \geq \|y\|_2 \) [15].

The following observation is made about the convergence of Algorithm 4. The block coordinate ascent method converges to the global optimum if the problems solved at each step have unique solutions [283]. The maximisation with respect to \( w_s, s = 1, \ldots, S \) (i.e., Step 2 of Algorithm 4) has a unique solution. The optimal objective value \( \gamma_o \) for fixed \( w_s \) (i.e., Step 3 of Algorithm 4) is indeed unique, but the resulting \( m_s, s = 1, \ldots, S \) is not guaranteed to have a unique solution in general, due to the quasi-convexity of the original problem [223, Section V.B]. Therefore, global optimality cannot be guaranteed.

### 6.1.2 Weighted sum rate maximisation

Now, the problem of joint design of the linear transmit and receive beamformers for maximising the weighted sum of the rates of the individual data streams subject to per BS power constraints is considered. Assuming a Gaussian codebook [81] for each data stream, the weighted sum rate can be expressed as

\[
R_\beta = \sum_{s=1}^{S} \beta_s r_s = \sum_{s=1}^{S} \beta_s \log(1 + \gamma_s) = \log \prod_{s=1}^{S} (1 + \gamma_s)^{\beta_s}
\]

where \( r_s \) and \( \gamma_s \) are the rate and the SINR of the data stream \( s \), respectively. The weight vector, \( \beta \triangleq [\beta_1, \ldots, \beta_S]^T \), \( \beta_s \geq 0 \), is used to give different relative importance to the data streams. The entire rate region achievable with linear processing can be found by varying \( \beta \) [15]. \( \beta = 1 \) corresponds to the usual sum rate maximisation or best effort service. Since \( R_\beta \) increases with respect to each \( \gamma_s \) and \( \log(\cdot) \) is an increasing function, the weighted sum rate maximisation problem with per BS power constraints can be formulated as follows

\[
\begin{align*}
\text{maximise} & \quad \prod_{s=1}^{S} (1 + \gamma_s)^{\beta_s} \\
\text{subject to} & \quad \gamma_s \leq \frac{p_s |w_s^H \tilde{H}_k v_s|^2}{1 + \sum_{i=1, i \neq s}^{S} p_i |w_s^H \tilde{H}_k v_i|^2}, \quad s = 1, \ldots, S \\
& \quad \sum_{s=1}^{S} p_s \|v_s^{[n]}\|^2 \leq P_n, \quad n = 1, \ldots, M \\
& \quad \|w_s\|_2 = 1, \quad \|v_s\|_2 = 1, \quad p_s \geq 0, \quad s = 1, \ldots, S,
\end{align*}
\]
where the variables are $v_s \in \mathbb{C}^{MN_T}$, $w_s \in \mathbb{C}^{N_R}$, $p_s \in \mathbb{R}$, $\gamma_s \in \mathbb{R}$. It is easy to observe that at the optimal point of (70), the first constraint holds with equality, thus the optimal value of $\gamma_s$ represents the SINR of the data stream $s$.

The optimisation problem (70) is not convex, and, hence, the problem of finding the global optimum is intrinsically non-tractable. However, (70) can be maximised with respect to different subsets of variables by considering the others fixed. For instance, the maximum SINR receiver given by

$$w_s = \frac{\tilde{w}_s}{\|\tilde{w}_s\|_2}, \quad \tilde{w}_s^H = \sqrt{p_s}v_s^H \tilde{H}_k^H \left( \sum_{i=1}^S p_i \tilde{H}_k^H v_i^H \tilde{H}_k^H + 1 \right)^{-1}$$ (71)

is optimal for any fixed $v_s$ and $p_s$. Furthermore, by fixing $v_s$ and $w_s$, (70) becomes a signomial problem in variable $p_s$ [14, 226, 227]. The problem is not convex as such, but there are efficient methods for approximating the solution by using geometric programming [226, 227]. The procedure consists of searching for a close local maxima by solving a sequence of geometric programs which locally approximate the original problem. This procedure is known to converge fast (in a few iterations) [226]. Finally, for fixed $w_s$ and $\gamma_s$ a maximum reduction factor is found which is common for all the per BS power constraints and preserves the SINR values $\gamma_s$ and, implicitly, the rate. This is given by the optimum $\alpha^*$ that solves the problem

$$\begin{align*}
\text{minimise} \quad & \alpha \\
\text{subject to} \quad & \gamma_s \leq \frac{p_s \|w_s^H \tilde{H}_k^H v_s\|^2}{1 + \sum_{i=1, i \neq s}^S p_i \|w_s^H \tilde{H}_k^H v_i\|^2}, \quad s = 1, \ldots, S \\
& \sum_{s=1}^S p_s \|v_s[n]\|^2 \leq \alpha P_n, \quad n = 1, \ldots, M \\
& \|v_s\|_2^2 = 1, \quad s = 1, \ldots, S
\end{align*}$$ (72)

where the variables are $\alpha \in \mathbb{R}_+$, $p_s \in \mathbb{R}_+$, $v_s \in \mathbb{C}^{MN_T}$, $s = 1, \ldots, S$. The solutions $v_s^*$ and $p_s^*$ do not directly increase the objective of (70). However, they increase the power margin for a fixed value of the objective, and hence, the saved power can be used to increase the objective. This is realised by updating $v_s$ and $p_s$ in (70) to the new values $v_s^*$ and $p_s^*/\alpha^*$, respectively, and increasing all $\gamma_s$ until all SINR constraints become tight. Note that this is an ascent step since $\alpha^* \leq 1$ for any $w_s$ and $\gamma_s$ that are feasible for (70).
The above observations suggest an iterative optimisation algorithm shown in Algorithm 6.

**Algorithm 6** Weighted sum rate maximisation under per BS power constraints
1. Initialise $v_s^{(0)}$ and $p_s^{(0)}$ such that the per BS power constraints are satisfied. Compute $w_s^{(0)}$ and $\gamma_s^{(0)}$ given by (71) and (64), where $v_s = v_s^{(0)}$ and $p_s = p_s^{(0)}$, $s = 1, \ldots, S$. Let $j = 1$.
2. Solve (70) for the variables $p_s$ and $\gamma_s$, by fixing $w_s = w_s^{(j-1)}$ and $v_s = v_s^{(j-1)}$, $s = 1, \ldots, S$. Denote the solutions by $p_s^\star$ and $\gamma_s^\star$.
3. Solve (72), where $\gamma_s = \gamma_s^\star$ and $w_s = w_s^{(j-1)}$, $s = 1, \ldots, S$. Denote the solutions by $\alpha^\star$, $p_s^\star$ and $v_s^\star$. Update $p_s^{(j)} = p_s^\star/\alpha^\star$ and $v_s^{(j)} = v_s^\star$, $s = 1, \ldots, S$.
4. Update $w_s^{(j)}$ and $\gamma_s^{(j)}$ according to (71) and (64), respectively. Test a stopping criterion. If it is not satisfied, let $j = j + 1$ and go to Step 2, otherwise STOP.

Even though Algorithm 6 increases the objective of (70) monotonically, there is no guarantee that the global optimum is found due to the non-convexity of (70). However, the simulation results in Sect. 6.4.1 show that Algorithm 6 converges to a solution that can be a local optimum but is still efficient.

Next, the algorithm used at Step 2 of Algorithm 6 is presented, which solves (70) for fixed $w_s$ and $v_s$. The objective of (70), $f_0(\gamma_1, \ldots, \gamma_S) = \prod_{s=1}^S (1 + \gamma_s)^{\beta_s}$, is approximated by a monomial function [226] $m(\gamma_1, \ldots, \gamma_S) = c \prod_{s=1}^S \gamma_s^{e_s}$, near the point $\hat{\gamma} = (\hat{\gamma}_1, \ldots, \hat{\gamma}_S)$. Now, the monomial function $m$ is the best local approximation of $f_0$ near the point $\hat{\gamma}$ if [226]

$$
\begin{align*}
\{ & m(\gamma_1, \ldots, \gamma_S) = f_0(\gamma_1, \ldots, \gamma_S) \\
& \nabla m(\gamma_1, \ldots, \gamma_S) = \nabla f_0(\gamma_1, \ldots, \gamma_S).
\end{align*}
$$

(73)

After solving the resulting system of equations, the parameters $c$ and $e_s$ of the best monomial local approximation are given by

$$
e_s = \beta_s \frac{\hat{\gamma}_s}{1 + \hat{\gamma}_s}, \quad c = \frac{f_0(\hat{\gamma}_1, \ldots, \hat{\gamma}_S)}{\prod_{s=1}^S \hat{\gamma}_s^{e_s}}.
$$

(74)

By using the local approximation in the objective of problem (70), and ignoring the multiplicative constant $c$ which does not affect the problem solution, the iterative algorithm shown in Algorithm 7 is obtained.
**Algorithm 7** Geometric optimisation step under per BS power constraints

1. Let the initial SINR guess be, \( \hat{\gamma} = (\hat{\gamma}_1^{(j-1)}, \ldots, \hat{\gamma}_S^{(j-1)}) \)
2. Solve the following geometric program,

\[
\begin{align*}
\text{maximise} & \quad \prod_{s=1}^{S} \gamma_s^{\beta_s} \\
\text{subject to} & \quad (1 - \phi) \hat{\gamma}_s \leq \gamma_s \leq (1 + \phi) \hat{\gamma}_s, \ s = 1, \ldots, S \\
& \quad g_{s,k}^{-1} p_s^{-1} \gamma_s + \sum_{k=1,k\neq s}^{S} g_{s,k} p_k g_{s,k}^{-1} \gamma_s \leq 1 , \ s = 1, \ldots, S \\
& \quad \sum_{s=1}^{S} p_s \| v_s^{[n]} \|_2^2 \leq P_n, \ n = 1, \ldots, M
\end{align*}
\] (75)

where the variables are \( \gamma_s, p_s \) and \( g_{s,k} = |w_s^H \tilde{H}_{k} v_k| \) are fixed values. Denote the solution by \( p^*_s \) and \( \gamma^*_s \). If \( \max_s |\gamma^*_s - \hat{\gamma}_s| > \epsilon \) set \( \hat{\gamma} = (\gamma^*_1, \ldots, \gamma^*_S) \) and go to Step 2, otherwise STOP.

The geometric program (75) approximates the original signomial problem (70) around the point \( \hat{\gamma} = (\hat{\gamma}_1, \ldots, \hat{\gamma}_S) \). The first set of inequality constraints of (75) are called trust region constraints [226] and they limit the domain of variables \( \gamma_s \) in a region where the monomial approximation is accurate enough. The constant \( \phi < 1 \) controls the desired approximation accuracy and a typical value is \( \phi = 0.1 \) [226].

The convergence of Algorithm 7 to a point, which is not necessarily the global solution, can be guaranteed as follows. Notice that the SINR guess \( \hat{\gamma}_s = \gamma_s^{(j-1)}, \forall s \) for each iteration \( j \) is always taken from the solution of the previous iteration \( j - 1 \), and the updated objective value can be easily compared to the objective value from the previous iteration. If the objective is decreased, the algorithm goes back to the solution given by the previous iteration. Then, either Algorithm 7 is terminated with a solution given by the previous iteration, or \( \phi \) is decreased until the objective is increased or it remains at the same value. Recall that the approximation becomes exact as \( \phi \) approaches zero. In the simulations, however, the objective value was never decreased during the execution of Algorithm 7.

Now, Step 3 of Algorithm 6 is explored. First, observe that the change of variable \( m_s = \sqrt{p_s} v_s \) defines a bijective mapping between the sets \( \{ (p_s, v_s) \mid p_s \in \mathbb{R}_+, \| v_s \|_2 = 1, v_s \in \mathbb{C}^{MN_T} \} \) and \( m_s \in \mathbb{C}^{MN_T} \). Thus, (72) can be solved for
\( \mathbf{m}_s \in \mathbb{C}^{M \times N_T} \), and then the optimal \( p_s \) and \( \mathbf{v}_s \) can be recovered. Furthermore, by replacing the positive variable \( \alpha \) by \( \rho^2 \) the following equivalent reformulation of (72) is obtained

\[
\begin{align*}
\text{minimise} & \quad \rho^2 \\
\text{subject to} & \quad \gamma_s \leq \frac{\left| \mathbf{w}_s^H \mathbf{\hat{H}}_k \mathbf{m}_s \right|^2}{1 + \sum_{s=1, s \neq s}^S \left| \mathbf{w}_s^H \mathbf{\hat{H}}_k \mathbf{m}_s \right|^2}, \quad s = 1, \ldots, S \\
& \quad \sum_{s=1}^S \left\| \mathbf{m}^{[n]}_s \right\|_2^2 \leq \rho^2 P_n, \quad \rho \geq 0, \quad n = 1, \ldots, M.
\end{align*}
\]

(76)

Note that the objective \( \rho^2 \) can be replaced by \( \rho \) since for \( \rho \geq 0 \), minimising \( \rho^2 \) is equivalent to minimising \( \rho \). Moreover, the constraints of (76) can be expressed as a generalised inequality with respect to the second-order cone [15, 223, 224]. Thus, (76) can be further reformulated as a SOCP. By modifying the approach presented in [223, Section IV.B] to accommodate per BS power constraints, the following equivalent SOCP formulation is obtained

\[
\begin{align*}
\text{minimise} & \quad \rho \\
\text{subject to} & \quad \begin{bmatrix} \sqrt{1 + \frac{1}{\gamma_s}} \mathbf{w}_s^H \mathbf{\hat{H}}_k \mathbf{m}_s \\\n\mathbf{M}^H \mathbf{\hat{H}}_k^H \mathbf{w}_s \\\n1 \end{bmatrix} \succeq \mathbf{0}, \quad s = 1, \ldots, S \\
& \quad \begin{bmatrix} \rho \sqrt{P_n} \\
\text{vec}(\mathbf{M}^{[n]}) \end{bmatrix} \succeq \mathbf{0}, \quad n = 1, \ldots, M
\end{align*}
\]

(77)

where the variables are \( \rho \in \mathbb{R}_+ \) and \( \mathbf{m}_s \in \mathbb{C}^{N_T} \), \( s = 1, \ldots, S \), and \( \mathbf{M} = [\mathbf{m}_1, \ldots, \mathbf{m}_S] \), \( \mathbf{M}^{[n]} = [\mathbf{m}^{[n]}_1, \ldots, \mathbf{m}^{[n]}_S] \). Let us denote the solution of (77) by \( \rho^*, \mathbf{m}^*_s, \quad s = 1, \ldots, S \). The solution of (72) is given by \( \alpha^* = \rho^2 \), \( \mathbf{v}_s^* = \mathbf{m}_s^*/\left\| \mathbf{m}_s^* \right\|_2, \quad p_s^* = \left| \mathbf{m}_s^* \right|^2_2, \quad s = 1, \ldots, S \). Notice that (77) provides a minimum-power beamformer design under per BS power constraints for a MIMO DL with fixed receivers. This is equivalent to the method proposed in [163], where the original DL problem was transformed into a dual UL minimax optimisation problem with an uncertain noise covariance.

The optimisation problem in (70) can be easily modified to accommodate supplementary constraints, e.g.

- per antenna or sum power constraints, and/or
minimum or maximum SINR values for some of the data streams,
\[\gamma_s \geq \gamma_s^{\text{min}} \quad \text{or} \quad \gamma_s \leq \gamma_s^{\text{max}}.\]

The modified problem with minimum SINR constraints can be directly solved using the proposed algorithm, if it is feasible under the initial beamformer configuration obtained in Step 1 of Algorithm 6, i.e. \(\gamma_s^{(0)} \geq \gamma_s^{\text{min}}\). The feasibility of the modified optimisation problem with any minimum SINR constraints can easily be checked by iterating between problem (72) with fixed \(\gamma_s = \gamma_s^{\text{min}}\) and (71). If the resulting \(\alpha \leq 1\), then the problem is feasible and the resulting beamformer configuration can be used as a feasible starting point for Algorithm 6.

### 6.1.3 User rate balancing

In this section, the problem of finding a maximum weighted common rate achievable for each scheduled user subject to per BS power constraints \(P_n\) is examined. This is equivalent to maximising the minimum of the weighted user rates

\[
r_o = \min_{k \in \mathcal{U}} \beta_k^{-1} \sum_{s \in \mathcal{P}_k} \log_2 (1 + \gamma_s) \tag{78}
\]

where \(\mathcal{P}_k\) is a subset of data streams that correspond to the user \(k\) and the weights \(\beta_k \geq 0, \forall k\) are used to prioritise the user data rates differently. Now, the problem of maximising the weighted common user rate can be formulated as

\[
\begin{align*}
\text{maximise} \quad & r_o \\
\text{subject to} \quad & \sum_{s \in \mathcal{P}_k} \log_2 (1 + \gamma_s) \geq \beta_k r_o \ \forall \ k \in \mathcal{U} \\
& \gamma_s \leq \frac{p_s |w_s^H \hat{H}_k v_s|^2}{1 + \sum_{i=1,i\neq s}^S p_i |w_s^H \hat{H}_k v_i|^2}, \quad s = 1, \ldots, S \\
& \sum_{s=1}^S p_s \|v_s^{[n]}\|^2 \leq P_n, \quad n = 1, \ldots, M \\
& \|w_s\|_2 = 1, \|v_s\|_2 = 1, \quad s = 1, \ldots, S
\end{align*}
\tag{79}
\]

where the variables are \(r_o \in \mathbb{R}_+, v_s \in \mathbb{C}^{MN_T}, w_s \in \mathbb{C}^{N_R^k}, p_s \in \mathbb{R}_+, \gamma_s \in \mathbb{R}_+\). Notice that the first and third set of constraints are relaxed with inequality in (79). However, they hold with equality at the optimal point of (79), and thus the optimal value of \(\gamma_s\) and \(r_o\) represent the SINR of the \(s\)th data stream and the achievable common rate, respectively. Similarly to (70), the optimisation
problem (79) is not convex, and hence, the problem of finding the global optimum is intrinsically non-tractable. However, a very efficient solution for (79) can be found by using the same iterative approach as in Algorithm 6. The resulting iterative algorithm is described in Algorithm 8.

**Algorithm 8** User rate balancing under per BS power constraints

1. Equivalent to Step 1. in Algorithm 6.
2. Solve (79) for the variables \( p_s \) and \( \gamma_s \), by fixing \( w_s = w_s^{(i-1)} \) and \( v_s = v_s^{(i-1)} \), \( s = 1, \ldots, S \). Denote the solutions by \( p^*_s \) and \( \gamma^*_s \).
3. Equivalent to Step 3. in Algorithm 6.
4. Equivalent to Step 4. in Algorithm 6.

It should be emphasised that Algorithm 8 increases the objective of (79) monotonically, but there is no guarantee that the global optimum is found due to the nonconvexity of the problem. Yet, the computer simulations in Section 6.4.1 will show that the algorithm converges to very efficient solutions.

Now, the algorithm used at Step 2 of Algorithm 8 is presented, which solves problem (79) for fixed \( w_s \) and \( v_s \). First, (79) is reformulated as the following signomial problem:

\[
\text{maximise } r_o \quad \text{subject to } \prod_{s \in P_k} \left( 1 + \gamma_s \right)^{\frac{\beta_k}{2}} \geq \tilde{r}_o \quad \forall k \in U \\
\frac{p_s g_{s,s}}{1 + \sum_{i=1, i \neq s}^{S} p_i g_{s,i}} \geq \gamma_s, \quad s = 1, \ldots, S \\
\sum_{s=1}^{S} p_s \left\| v_s^{[n]} \right\|_2^2 \leq P_n, \quad n = 1, \ldots, M
\]  

(80)

where the objective \( r_o \) has been replaced with \( \tilde{r}_o = 2^{r_o} \) and \( g_{s,i} = \left| w_s^H \tilde{H}_k v_i \right|^2 \), \( s, i = 1, \ldots, S \). Problem (80) is solved similarly to the weighted sum rate maximisation problem in (75). The first constraint of (80) \( f_0(\{\gamma_s|s \in P_k\}) = \prod_{s \in P_k} \left( 1 + \gamma_s \right)^{\frac{\beta_k}{2}}, \forall k \) is approximated by a monomial function [226] \( m(\{\gamma_s|s \in P_k\}) = c_k \prod_{s \in P_k} \gamma_s^{e_s}, \forall k \), near the point \( \hat{\gamma} = (\hat{\gamma}_1, \ldots, \hat{\gamma}_S) \), where the parameters \( c_k \) and \( e_s \) of the best monomial local approximation are given by (see similar
derivation for weighted sum rate maximisation in (73) - (74))

e_s = \frac{1}{\beta_k} \frac{\hat{\gamma}_s}{1 + \hat{\gamma}_s}, \forall s \in \mathcal{P}_k, k \in \mathcal{U},
\quad c_k = \frac{1}{\prod_{s \in \mathcal{P}_k} (1 + \hat{\gamma}_s)^{1/\beta_k}} \prod_{s \in \mathcal{P}_k} \hat{\gamma}_s^{e_s}, \forall k \in \mathcal{U}.
(81)

By using the local approximation in the first constraint of problem (80) and presenting it in the standard form geometric program [15, 226], an iterative algorithm shown in Algorithm 9 is obtained.

**Algorithm 9** Signomial optimisation step for user rate balancing

1. Let the initial SINR guess be \( \hat{\gamma} = (\hat{\gamma}_1^{(j-1)}, \ldots, \hat{\gamma}_S^{(j-1)}) \)
2. Update \( e_s, s = 1, \ldots, S \) and \( c_k, \forall k \in \mathcal{U} \) as in (81)
3. Solve the following geometric program,

\[
\begin{align*}
\text{minimise} & \quad \tilde{r}_o^{-1} \\
\text{subject to} & \quad (1 - \phi)\gamma_s \leq \gamma_s \leq (1 + \phi)\gamma_s, \ s = 1, \ldots, S \\
& \quad \tilde{r}_o c_k^{-1} \prod_{s \in \mathcal{P}_k} \gamma_s^{e_s} \leq 1 \forall k \in \mathcal{U} \\
& \quad g^{-1}_{s,s} p_s^{-1} \gamma_s + \sum_{i=1, i \neq s}^S g_{s,i} p_i g^{-1}_{s,i} p_s^{-1} \gamma_s \leq 1, \ s = 1, \ldots, S \\
& \quad P_n^{-1} \sum_{s=1}^S p_s ||v_n^s||_2^2 \leq 1, \ n = 1, \ldots, M
\end{align*}
\]

where the variables are \( \gamma_s, p_s, s = 1, \ldots, S \) and \( \tilde{r}_o \). Denote the solution by \( p_s^* \) and \( \gamma_s^* \). If \( \max_s |\gamma_s^* - \hat{\gamma}_s| > \epsilon \) set \( \hat{\gamma} = (\hat{\gamma}_1^*, \ldots, \hat{\gamma}_S^*) \) and go to Step 2, otherwise STOP.

The optimal points \( p_s^*, s = 1, \ldots, S \) represent the power allocation required in Step 2 of Algorithm 8. The problem (82) is the geometric program which approximates the original signomial problem (79) around the point \( \hat{\gamma} = (\hat{\gamma}_1, \ldots, \hat{\gamma}_S) \).

### 6.1.4 Sum power minimisation with individual user rate constraints

The objective is to minimise the total transmitted power while guaranteeing the minimum bit rate requirements \( r_{k,\min} \) for the users. The optimisation problem
can be formulated as follows:

\[
\begin{align*}
\text{minimise} & \quad \mathbf{p}^T \mathbf{1} \\
\text{subject to} & \quad \sum_{s \in P_k} \log_2 (1 + \gamma_s) \geq r_k^\text{min}, \quad k \in \mathcal{U} \\
& \quad \gamma_s \leq \frac{p_s |w^H \hat{H}_k v_s|^2}{1 + \sum_{i=1, i \neq s}^S p_i |w^H \hat{H}_k v_i|^2}, \quad s = 1, \ldots, S \\
& \quad \sum_{s=1}^S p_s ||v_s^{[n]}||_2^2 \leq P_n, \quad n = 1, \ldots, M \\
& \quad ||w_s||_2 = 1, ||v_s||_2 = 1, \quad s = 1, \ldots, S
\end{align*}
\]

(83)

where the variables are \(v_s \in \mathbb{C}^{MN}, w_s \in \mathbb{C}^{N \times k}, p_s \in \mathbb{R}_+,\) and \(\gamma_s \in \mathbb{R}_+\).

Problem (83) can be solved in an iterative fashion, similarly to the previous sections. The minimum rate constraints per user may be conflicting with some of the maximum power constraints per antenna group. Therefore, the feasibility of (83) must be checked first. This can be achieved by replacing the objective \(\mathbf{p}^T \mathbf{1}\) of (83) with a power reduction factor \(\alpha\), common to all per BS power constraints, as in (72). For fixed \(w_s\) and \(v_s\), (83) can be reformulated as a signomial program, and solved similarly to Algorithm 9. The feasibility problem can be solved via iterating between the signomial step, (72) for fixed \(w_s\) and \(\gamma_s\), and (71) for fixed \(v_s\) and \(p_s\). The problem is feasible if the resulting \(\alpha \leq 1\). For a feasible problem, the original objective \(\mathbf{p}^T \mathbf{1}\) and the per BS power constraints can be used in the signomial step.

### 6.1.5 Multiuser MIMO-OFDM system with different quality of service classes

The formulation in Section 6.1.3 is now extended to a frequency selective downlink channel with additional QoS constraints, such as the guaranteed bit rate per user. Assume a MIMO-OFDM system which transforms the frequency selective channel into \(N_C\) parallel independent MIMO sub-channels. A separate beamformer is assigned to each sub-channel \(c\) and each scheduled data stream \(s\) at the transmitter and at the receiver. Suppose now that the system has to provide different services to different users. Each user \(k \in \mathcal{U}\) may have different rate allocation priorities \(\beta_k\), and the subset \(\mathcal{U}_{\text{GBR}}\) from the full user set \(\mathcal{U}\) includes users with minimum bit rate requirements \(r_k^\text{min}\). A more practical model for
the rate provided by each data stream is also adapted. The achievable rate per stream is modelled as \( \min(\log_2(1 + \Gamma \gamma_{s,c}), r_{\text{max}}) \), where \( \Gamma \) describes the SNR gap to the channel capacity [163] and \( r_{\text{max}} \) is the maximum rate limit, both imposed by a set of some practical modulation and coding schemes.

Subsequently, the objective is to maximise the weighted common rate of all users while guaranteeing the minimum bit rate requirements for the GBR users. The optimisation problem can be formulated as follows:

\[
\begin{align*}
\text{maximise} & \quad r_o \\
\text{subject to} & \quad \sum_{c=1}^{N_C} \sum_{s \in P_{k,c}} \log_2(1 + \Gamma \gamma_{s,c}) \geq \beta_k r_o, \quad k \in U \\
& \quad \sum_{c=1}^{N_C} \sum_{s \in P_{k,c}} \log_2(1 + \Gamma \gamma_{s,c}) \geq \gamma_{k,\text{min}}, \quad k \in U_{\text{GBR}} \\
& \quad \gamma_{s,c} \leq \frac{p_{s,c} |\mathbf{w}_{s,c}^H \tilde{\mathbf{H}}_{k,c} v_{k,c}|^2}{1 + \sum_{i=1, i \neq s}^S p_{i,c} |\mathbf{w}_{i,c}^H \tilde{\mathbf{H}}_{k,c} v_{i,c}|^2}, \quad \forall \ s, c \\
& \quad \gamma_{s,c} \leq \gamma_{\text{max}}, \quad \forall \ s, c \\
& \quad \sum_{c=1}^{N_C} p_{s,c} \|\mathbf{w}_{s,c}\|^2 \leq P_n, \quad n = 1, \ldots, M \\
& \quad \|\mathbf{w}_{s,c}\|^2 = 1, \quad \|\mathbf{v}_{s,c}\|^2 = 1, \quad \forall \ s, c \\
\end{align*}
\]

where the variables are \( r_o \in \mathbb{R}_+ \), \( \mathbf{v}_{s,c} \in \mathbb{C}^{M N_T} \), \( \mathbf{w}_{s,c} \in \mathbb{C}^{N_{R_{k,c}}} \), \( p_{s,c} \in \mathbb{R}_+ \) and \( \gamma_{s,c} \in \mathbb{R}_+ \). \( P_{k,c} \) is a subset of data streams assigned to the user \( k \) at sub-carrier \( c \) and \( \gamma_{\text{max}} = (2^{r_{\text{max}}} - 1) / \Gamma \). Observe that a GBR user \( k \) can enjoy higher rates than \( r_{\text{min}}^k \), as long as \( \beta_k r_o > r_{\text{min}}^k \). In some cases, however, it may be more desirable to allocate all the available extra rate to non-guaranteed bit rate (NGBR) users \( k \in U \setminus U_{\text{GBR}} \). This can be easily achieved by assigning \( \beta_k = 0, \forall \ k \in U_{\text{GBR}} \).

Without GBR constraints, the problem is a straightforward extension to the frequency selective case from the flat fading case (Section 6.1.3), where (79)-(82) are extended to include the maximum SINR constraints and to cover also the sub-carrier domain. With the GBR constraints, (84) may become infeasible, if the rate requirements cannot be satisfied for the available power and the given channel (sub-carrier/stream-to-user) allocations. In addition, the initial beamformer configuration may be highly sub-optimal leading to infeasible solutions in the first steps of the iterative algorithm. Therefore, the problem is split into two separate phases. First, a feasibility check for the minimum rate
requirements is made. Second, the common rate maximisation is carried out. The feasibility of (84) is verified by using the weighted common rate maximisation criterion to check whether the minimum bit rates are achievable. One of the guaranteed bit rates \( r_{\text{min}}^k \) is selected as reference rate \( r_{\text{ref}} \) in order to define the relative weights for the weighted common rate maximisation. The relative weights \( \theta_k \) for the GBR users are then defined as 

\[
\theta_k = \frac{r_{\text{min}}^k}{r_{\text{ref}}}, \quad k \in \mathcal{U}_{\text{GBR}}
\]

If the achieved common rate \( r'_o \) is higher than or equal to the reference rate, \( r'_o \geq r_{\text{ref}} \), the problem is feasible. The optimum \( r'_o \) is found by solving the following weighted common rate maximisation problem:

\[
\begin{align*}
\text{maximise} & \quad r'_o \\
\text{subject to} & \quad \sum_{c=1}^{N_C} \sum_{s \in \mathcal{F}_k,c} \log_2 \left( 1 + \Gamma_{s,c} \gamma_{s,c} \right) \geq \theta_k r'_o, \quad k \in \mathcal{U}_{\text{GBR}} \\
& \quad \gamma_{s,c} \leq \frac{p_{s,c} \| \mathbf{w}_{s,c}^H \tilde{\mathbf{H}}_{k,c} \mathbf{v}_{s,c} \|_2^2}{1 + \sum_{i=1,i \neq s}^S p_{i,c} \| \mathbf{w}_{i,c}^H \tilde{\mathbf{H}}_{k,c} \mathbf{v}_{i,c} \|_2^2}, \quad \forall s, c \\
& \quad \gamma_{s,c} \leq \gamma_{\text{max}}, \quad \forall s, c \\
& \quad \sum_{c=1}^{N_C} \sum_{s=1}^S p_{s,c} \| \mathbf{v}_{s,c}[n] \|_2^2 \leq P_n, \quad n = 1, \ldots, M \\
& \quad \| \mathbf{w}_{s,c} \|_2 = 1, \| \mathbf{v}_{s,c} \|_2 = 1, \quad \forall s, c
\end{align*}
\]  

(85)

where the variables are \( r'_o \in \mathbb{R}, \mathbf{v}_{s,c} \in \mathbb{C}^{M \times 1}, \mathbf{w}_{s,c} \in \mathbb{C}^{N_{\text{ant}} \times 1}, p_{s,c} \in \mathbb{R}, \) and \( \gamma_{s,c} \in \mathbb{R}. \) The problem (85) can be solved similarly to (79)–(82) by extending them to the frequency selective fading case and including the maximum SINR constraints in the signomial optimisation (Algorithm 9).

After a feasible solution has been found for the GBR users, the original optimisation problem (84) is initialised with the feasible beamformer configuration. However, (85) tends to assign zero powers to the streams that correspond to the NGBR users. Therefore, the signomial optimisation (as in Algorithm 8 Step 3) cannot adjust the rates of NGBR users, since the initial SINR values, and hence, the exponents of their monomial approximations (81) are zero valued. Now, the powers of the non-guaranteed users’ streams and the corresponding SINR values must be initialised with some non-zero values that still maintain the GBR users satisfied. This can be achieved, for example, by maximizing the minimum SINR for the streams that correspond to the NGBR users while keeping the SINRs of
GBR users fixed at the values given by (85). The problem can be formulated as

\[
\begin{align*}
\text{max} & \quad \min_{s,c} \frac{|w_{s,c}^H \tilde{H}_{k,s,c} m_{s,c}|^2}{1 + \sum_{i=1,i\neq s}^S |w_{s,c}^H \tilde{H}_{k,s,c} m_{i,c}|^2}, \quad s \in \mathcal{P}_{k,c}, k \in \mathcal{U}, \forall c \\
\text{s.t} & \quad \gamma_{s,c} \leq \frac{|w_{s,c}^H \tilde{H}_{k,s,c} m_{s,c}|^2}{1 + \sum_{i=1,i\neq s}^S |w_{s,c}^H \tilde{H}_{k,s,c} m_{i,c}|^2}, \quad s \in \mathcal{P}_{k,c}, k \in \mathcal{U}_{\text{GBR}}, \forall c \\
& \quad \sum_{c=1}^{N_C} \sum_{s=1}^S |m_{s,c}[n]|^2 \leq P_n, \quad n = 1, \ldots, M
\end{align*}
\]

where the variables are \( m_{s,c} \in \mathbb{C}^{MN_T} \) \( \forall s, c \). The receiver vectors \( w_{s,c} \), \( \forall s, c \) and \( \gamma_{s,c} \), \( s \in \mathcal{P}_{k,c}, k \in \mathcal{U}_{\text{GBR}}, \forall c \) are kept at the fixed values obtained from (85). The problem can be solved similarly to Algorithm 5 by replacing (67) in Step 3 with the following SOCP feasibility problem

\[
\begin{align*}
\text{find} & \quad m_{s,c}, \forall s, c \\
\text{s.t} & \quad \sqrt{1 + \frac{1}{\gamma_{s,c}}} w_{s,c}^H \tilde{H}_{k,s,c} m_{s,c} \begin{bmatrix} M_{s,c}^H & w_{s,c} \end{bmatrix} \begin{bmatrix} \tilde{H}_{k,s,c} & 1 \\ 1 & \sqrt{1 + \frac{1}{\gamma_{s,c}}} w_{s,c}^H \tilde{H}_{k,s,c} m_{s,c} \end{bmatrix} \begin{bmatrix} M_{s,c}^H & w_{s,c} \end{bmatrix} \begin{bmatrix} \tilde{H}_{k,s,c} & 1 \\ 1 & \sqrt{P_n} \end{bmatrix} \begin{bmatrix} \mathbf{vec}(m_{s,c}) \end{bmatrix} \succeq \mathbf{0}, \quad s \in \mathcal{P}_{k,c}, k \in \mathcal{U}, \forall c \\
& \quad \sum_{c=1}^{N_C} \sum_{s=1}^S |m_{s,c}[n]|^2 \leq P_n, \quad n = 1, \ldots, M
\end{align*}
\]

where \( M_c = [m_{1,c}, \ldots, m_{S,c}] \) and \( M^{[n]} = [M^{[n]}_1, \ldots, M^{[n]}_{N_C}] \). Finally, all the steps explained above for solving the original optimisation problem (84) are summarised in Algorithm 10.

Algorithm 11 presents the signomial optimisation algorithm used in Step 2.2 of Algorithm 10, which solves (84) for fixed \( w_{s,c} \) and \( v_{s,c} \). Note that the step 1.2 in Algorithm 10 is also solved similarly to Algorithm 11, except that the second constraint is removed and the parameter \( \beta_k \) is replaced with \( \theta_k \), \( \forall k \in \mathcal{U}_{\text{GBR}} \).

It must be noted that the performance and also the feasibility of the proposed algorithm depends on the initial channel (sub-carrier/stream-to-user) allocations.
Algorithm 10 User rate balancing under minimum rate and per BS power constraints

1. Feasibility check phase

1.1. Let the scheduler decide the stream-to-user association for each sub-carrier. Initialise $v_{s,c}^{(0)}$ and $p_{s,c}^{(0)}$ so that the power constraints are satisfied, and compute $w_{s,c}^{(0)}$ and $\gamma_{s,c}^{(0)}$ given by (71) and (64) where $v_{s,c} = v_{s,c}^{(0)}$ and $p_{s,c} = p_{s,c}^{(0)}$, $s = 1, \ldots, S$. Let $j = 1$.

1.2. Solve (85) for the variables $p_{s,c}$ and $\gamma_{s,c}$, by fixing $w_{s,c} = w_{s,c}^{(j-1)}$ and $v_{s,c} = v_{s,c}^{(j-1)}$, $s = 1, \ldots, S$. Denote the solutions by $p_{s,c}^\ast$ and $\gamma_{s,c}^\ast$.

1.3. Solve the frequency selective version of problem (77), where $\gamma_{s,c} = \gamma_{s,c}^\ast$ and $w_{s,c} = w_{s,c}^{(j-1)}$, $s = 1, \ldots, S$. Denote the solutions by $\rho^\ast$, $p_{s,c}^\ast$ and $v_{s,c}^\ast$. Update $w_{s,c}^{(j)}$ and $\gamma_{s,c}^{(j)}$ according to (71) and (64), respectively. Calculate the rate $r_k = \sum_{c=1}^{NC} \sum_{s \in P_{k,c}} \log_2(1 + \gamma_{s,c}^{(j)})$, $\forall k \in \mathcal{U}_{\text{GBR}}$. Test a stopping criterion for the feasibility check. If it is satisfied, declare the problem infeasible and STOP, otherwise let $j = j + 1$. If $r_k > r_k^{\text{min}} \forall k \in \mathcal{U}_{\text{GBR}}$, go to Step 2.1, otherwise go to Step 1.2.

2. Common rate maximisation phase

2.1. Maximise the minimum SINR per streams for all users by solving problem (86), where the SINR values of GBR users are fixed, $\gamma_{s,c} = \gamma_{s,c}^{(j-1)}$ $\forall s \in \mathcal{F}_{k,c}, k \in \mathcal{U}_{\text{GBR}}$. Denote the solutions by $m_{s,c}^\ast$ and $\gamma_{s,c}^\ast$. Update $v_{s,c}^{(j)} = m_{s,c}^\ast / \|m_{s,c}^\ast\|_2$ and $\gamma_{s,c}^{(j)} = \gamma_{s,c}^\ast$ $\forall s \in \mathcal{F}_{k,c}, k \in \mathcal{U} \setminus \mathcal{U}_{\text{GBR}}$.

2.2. Solve problem (84) for the variables $p_{s,c}$ and $\gamma_{s,c}$, by fixing $w_{s,c} = w_{s,c}^{(j-1)}$ and $v_{s,c} = v_{s,c}^{(j-1)}$, $s = 1, \ldots, S$. Denote the solutions by $p_{s,c}^\ast$ and $\gamma_{s,c}^\ast$.

2.3. Equivalent to Step 1.3

2.4. Update $w_{s,c}^{(j)}$ and $\gamma_{s,c}^{(j)}$ according to (71) and (64), respectively. Test the stopping criterion. If it is not satisfied, let $j = j + 1$ and go to Step 2.2, otherwise STOP.

Finding the optimal stream-to-user allocation, i.e. how many streams per sub-carrier $P_{k,c}$ should be allocated to each user $k$, providing the highest weighted common rate while guaranteeing minimum bit rate constraints is a difficult non-convex combinatorial optimisation problem with integer constraints [172], and is an interesting topic for further studies. In the numerical examples, a simple random allocation scheme is used, where users are assigned with equal
Algorithm 11: Signomial optimisation for user rate balancing with minimum rate constraints

1. Let the initial SINR guess, \( \hat{\gamma}_{s,c} = \gamma_{s,c}^{(j-1)}, \forall s,c \)
2. Update \( e_{s,c}, \forall s,c \) and \( c_k, \forall k \in U \)
3. Solve the following geometric program,

\[
\begin{align*}
\text{minimise} & \quad \tilde{r}_o^{-1} \\
\text{subject to} & \quad (1 - \phi)\hat{\gamma}_{s,c} \leq \gamma_{s,c} \leq (1 + \phi)\hat{\gamma}_{s,c}, \quad \forall s,c \\
& \quad \tilde{e}_{s,c} \prod_{c=1}^{N_C} \prod_{s \in P_{k,c}} \gamma_{s,c}^{-e_{s,c}} \leq 1 \quad \forall k \in U \\
& \quad 2^\gamma_{s,c}^{\min} c_k^{-1} \prod_{c=1}^{N_C} \prod_{s \in P_{k,c}} \gamma_{s,c}^{-e_{s,c}} \leq 1 \quad \forall k \in U_{GBR}
\end{align*}
\]

\[
\begin{align*}
& \quad g_{s,i,c}^{-1} p_{s,c}^{-1} \gamma_{s,c} + \sum_{i=1, i \neq s}^{S} g_{s,i,c} p_{s,c} g_{s,s,c}^{-1} p_{s,c}^{-1} \gamma_{s,c} \leq 1, \quad \forall s,c \\
& \quad \gamma_{s,c}^{-1} \leq 1, \quad \forall s,c \\
& \quad \gamma_{\max}^{s,c} N_C S \sum_{c=1}^{N_C} \sum_{s=1}^{S} p_{s,c} \|v_{s,c}^n\|_2^2 \leq 1, \quad n = 1, \ldots, M
\end{align*}
\]

where \( g_{s,i,c} = \|w_{s,c}^{H} \hat{H}_{k_{s,c}} v_{i,c}^s\|^2, \quad s, i = 1, \ldots, S \), and the variables are \( \gamma_{s,c}, p_{s,c} \) and \( \tilde{r}_o \). The parameters \( e_{s,c} \) and \( c_k \) are derived as

\[
e_{s,c} = \Gamma \hat{\gamma}_{s,c} / (1 + \Gamma \hat{\gamma}_{s,c}), \quad \forall s,c
\]

and

\[
c_k = \prod_{c=1}^{N_C} \prod_{s \in P_{k,c}} (1 + \Gamma \hat{\gamma}_{s,c}) / \left( \prod_{c=1}^{N_C} \prod_{s \in P_{k,c}} \hat{\gamma}_{s,c} \right), \quad \forall k \in U
\]

respectively. Denote the solution by \( p_{s,c}^*, \gamma_{s,c}^* \) and \( \tilde{r}_o^* \). If \( \max_{s,c} |\gamma_{s,c}^* - \hat{\gamma}_{s,c}| > \epsilon \)

set \( \hat{\gamma}_{s,c} = \gamma_{s,c}^* \), \( \forall s,c \) and go to Step 2, otherwise STOP.

If the optimisation problem (84) is infeasible for a given allocation then a straightforward solution is to increase the number of spatial and spectral transmit dimensions allocated to GBR users until the problem becomes feasible or all the available dimensions are allocated to the GBR users.
6.2 Multiuser ZF solution with per BS or antenna power constraints

6.2.1 Weighted sum rate maximisation with cooperative ZF transmission

First, the power allocation for maximising the weighted sum rate with block-ZF processing under an individual power constraint for each BS is provided. In [242], a similar treatment was also provided for a single user MIMO transmission. With block-ZF processing (Algorithm 1), the problem of maximising the weighted sum of rates of the data streams under the per BS power constraint in (70) is reduced to

\[
\begin{aligned}
\text{maximise} & \quad N_C^{-1} \sum_{k \in U} \sum_{c=1}^{N_C} m_{k,c} \beta_{k,i} \log_2 (1 + \lambda_{k,i,c} p_{k,i,c}) \\
\text{s. t.} & \quad \sum_{k \in U} \sum_{c=1}^{N_C} \sum_{i=1}^{m_{k,c}} \|v_{k,i,c}^{[n]}\|^2 p_{k,i,c} \leq P_n, \quad n = 1, \ldots, M \\
& \quad p_{k,i,c} \geq 0, \quad k \in U, \ i = 1, \ldots, m_{k,c}, \ c = 1, \ldots, N_C \quad (91)
\end{aligned}
\]

where the variables are \( p_{k,i,c}, k \in U, i = 1, \ldots, m_{k,c}, c = 1, \ldots, N_C \) and \( P_n \) is the power constraint on BS \( n \). The weights, \( \beta_{k,i} \geq 0, \forall k, i \), are used to prioritise the data streams of different users differently. It is easy to observe that the objective function of (91) is concave and all the inequality constraints are affine. Thus, (91) is a convex, and it can be efficiently solved numerically by using standard optimisation software packages, e.g. CVX [78]. Under the sum power constraint, \( P_{\text{sum}} \) and with equal user priorities, the sum rate is maximised by the well known water-filling power allocation [81], \( p_{k,i,c} = (\hat{\mu} - 1/\lambda_{k,i,c})^+ \), where the "water level", \( \hat{\mu} \), is chosen such that the sum power constraint holds with equality, i.e. \( \sum_{k \in U} \sum_{c=1}^{N_C} \sum_{i=1}^{m_{k,c}} p_{k,i,c} = P_{\text{sum}} \).

Inspired by the earlier work in [220], a simple heuristic algorithm which finds a suboptimal, but still efficient, power allocation for the problem (91) with equal user priorities is also proposed. Similar to the sum power constraint case, a water-filling power allocation is imposed, \( p_{k,i,c} = (\hat{\mu} - 1/\lambda_{k,i,c})^+ \), but the water level \( \hat{\mu} \) is increased until one of the BS’s reaches its power constraint. The water level can be efficiently found by using, for example, a bisection method [15], where in each iteration all the per BS power constraints are checked. In the case of equal power constraints for all the BS, i.e. \( P_n = P_T, n = 1, \ldots, M \), the
final TX power is allocated such that the BS with the strongest reception at the receiver is using the full power $P_T$ while the other BS’s are using power less than $P_T$. It will be shown in Section 6.4.1 that the heuristic method results in almost the same mutual information as (91).

6.2.2 Cooperative ZF transmission with QoS constraints

Suppose now that the system has to keep the SINR per data stream $\gamma_{k,i} = \lambda_{k,i}p_{k,i}$ in some fixed ratios, i.e. $\gamma_{k,i}/\beta_{k,i} = \gamma_o$. With ZF processing, the SINR balancing problem in (65) is reduced to

$$\begin{align*}
\text{maximise} & \quad i=1,...,m_k, k \in \mathcal{U} \quad \beta_{k,i}^{-1}\lambda_{k,i}p_{k,i} \\
\text{subject to} & \quad \sum_{k \in \mathcal{U}} \sum_{i=1}^{m_k} ||\mathbf{v}_{k,i}||_2^2 p_{k,i} \leq P_n, \quad n = 1, \ldots, M
\end{align*}$$

where the variables are $p_{k,i} \in \mathbb{R}_+$, $k \in \mathcal{U}$, $i = 1, \ldots, m_k$. It is easy to observe that the objective function of (92) is concave and all the inequality constraints are affine. Thus, (92) is a convex optimisation problem.

Similarly, the balanced rate maximisation problem with GBR users in (84) with ZF processing can be reformulated as the following convex optimisation problem:

$$\begin{align*}
\text{maximise} & \quad r_o \\
\text{subject to} & \quad \sum_{c=1}^{N_C} \sum_{k \in \mathcal{U}} \sum_{i=1}^{m_k} \log_2 (1 + \Gamma \lambda_{k,i,c}p_{k,i,c}) \geq \beta_k r_o, \quad k \in \mathcal{U} \\
& \quad \sum_{c=1}^{N_C} \sum_{k \in \mathcal{U}} \sum_{i=1}^{m_k} \log_2 (1 + \Gamma \lambda_{k,i,c}p_{k,i,c}) \geq r_{k,\text{min}}, \quad k \in \mathcal{U}_{\text{GBR}} \\
& \quad \lambda_{k,i,c}p_{k,i,c} \leq \gamma_{\text{max}}, \quad \forall \ k, i, c \\
& \quad \sum_{c=1}^{N_C} \sum_{k \in \mathcal{U}} \sum_{i=1}^{m_k} ||\mathbf{v}_{k,i,c}||_2^2 p_{k,i,c} \leq P_n, \quad n = 1, \ldots, M
\end{align*}$$

where the variables are $r_o \in \mathbb{R}_+$ and $p_{k,i} \in \mathbb{R}_+$, $k \in \mathcal{U}$, $i = 1, \ldots, m_k$. Notice that (93) is reduced to SINR balancing problem (92) with additional minimum and maximum SINR constraints, when $\Gamma = 1$, $N_C = 1$ and $m_k = 1 \forall k$.

6.2.3 Practical considerations

Previous subsections dealt with maximising several optimisation criteria for ZF transmission with per BS power constraints. In practical systems, the
finite granularity imposed by the finite set of MCS’s makes the bit and power optimisation problems non-convex, and solutions similar to (91) or (93) are difficult if not impossible to obtain. Therefore, greedy bit and power loading algorithms, such as the LSO algorithm from Section 5.2.4 are considered in the numerical evaluations that try to maximise the achievable spectral efficiency for certain quality of service criteria, such as the target FER. In order to guarantee the per BS power constraints, the same heuristic solution as in Section 6.2.1 is proposed. Basically, the only difference with single link algorithms shown in Sections 5.2.3 and 5.2.4, is that the per BS power constraints, which are a function of $v_{k,i,c}^{[n]}$ as in (10), are included in the stopping criterion. Similarly to the previous section, the iterative process (bit and power loading) is continued until one of the BS’s in $S_k$ reaches its power constraint. Note that a solution based on joint TX-RX design similar to (70) would require a new design of the bit and power loading algorithm, and hence, is not considered in this thesis.

In previous sections, it was assumed that the inter-cell interference $R_{k,c}$ was perfectly known at the transmitter and it could be incorporated into the whitened channel matrix $\tilde{H}_{k,c} = R_{k,c}^{-\frac{1}{2}} \tilde{H}_{k,c}$. However, without $R_{k,c}$ known at the transmitter, the optimum pre-combiner obtained from the whitened channel matrix $\tilde{H}_{k,c}$ cannot be computed as shown in Section 4.1.2. Therefore, a sub-optimal but still efficient design of the pre-coder is used, which relies on the channel knowledge only. The procedure to compute the sub-optimal pre-combining matrix per user is exactly the same as that in Algorithm 1, except that $\tilde{H}_{k,c}$ is used in the algorithms instead of $\tilde{H}_{k,c}^w$. Due to the non-reciprocal interference at the receiver, inter-stream interference is not completely removed by the LMMSE receiver operation. The SINR per sub-channel can no longer be controlled at the transmitter but it is affected by the structure of $R_{k,c}$, similarly to the single user case shown in (36). Therefore, the simple closed-loop algorithm introduced in Section 4.2 is used for compensating for the effect of interference non-reciprocity at the transmitter.
6.3 Symmetric capacity subject to per BS or antenna power constraints

In this section, an absolute upper bound is provided for the scenario where $K = |\mathcal{U}|$ active users are served with equal weighted rates by $M$ BS's with individual power constraints $(P_1, \ldots, P_M)$. Let $\psi = [\psi_1, \ldots, \psi_K]^T$ denote an arbitrary rate reward vector where the weights $\psi_k$ are ordered in a descending order. The boundary point of the capacity region for the broadcast channel with per BS power constraints that corresponds to the weight reward vector $\psi$ is denoted as $C(\tilde{H}_1, \ldots, \tilde{H}_K, P_1, \ldots, P_M, \psi)$. This can be computed by solving the following dual uplink optimisation problem with a sum power constraint $\sum_{n=1}^M P_n$ and with uncertain noise with diagonal covariance matrix $Q = \text{diag}(q_1, \ldots, q_1, q_2, \ldots, q_2, \ldots, q_M, \ldots, q_M)$, where each $q_n$ is repeated $N_T$ times [163, Corollary 4]

$$
\begin{align*}
\min_{Q} \quad & \max_{C_i} \quad \sum_{k=1}^{K} \psi_k \log \left[ \frac{\sum_{i=1}^{k} \tilde{H}_i^H C_i \tilde{H}_i + N_0 Q}{\sum_{i=1}^{k-1} \tilde{H}_i^H C_i \tilde{H}_i + N_0 Q} \right] \\
\text{s. t.} \quad & \sum_{k=1}^{M} \text{tr}(C_i) \leq \sum_{n=1}^{M} P_n \\
& \sum_{n=1}^{M} q_n P_n \leq \sum_{n=1}^{M} P_n \\
& C_i, Q \succeq 0.
\end{align*}
$$

(94)

The weighted symmetric capacity is defined such that the weighted user rates are equal, i.e. $\beta_1^{-1} r_1 = \beta_2^{-1} r_2 = \cdots = \beta_K^{-1} r_K$ while $(r_1, \ldots, r_K)$ belong to the boundary of the capacity region, i.e. (95)

$$
\begin{align*}
\text{find} \quad & \psi \\
\text{s. t.} \quad & \beta_k^{-1} r_k = \beta_{k+1}^{-1} r_{k+1}, \ k = 1, \ldots, K - 1 \\
& (r_1, \ldots, r_K) = C(\tilde{H}_1, \ldots, \tilde{H}_K, P_1, \ldots, P_M, \psi).
\end{align*}
$$

The optimal $\psi$ providing equal weighted user rates can be found by using a similar iterative ellipsoid algorithm as proposed originally in [175] for finding the symmetric capacity with a sum power constraint.
6.4 Numerical examples

The simulated OFDM system is based on the parameters given in Table 2 on page 75. Except in the system level studies, the channel delay taps are considered independent of each other with a power delay profile specified by ETSI BRAN Channel A [279], and the elements of the channel matrices are modelled as IID Gaussian random variables. The number of both TX and RX antennas is fixed at 2, \( \{N_T, N_R\} = \{2, 2\} \). For simplicity, all the base stations are assumed to have an equal maximum power limit \( P_T \), i.e. \( P_n = P_T \ \forall \ n \). The impacts of the following three power constraints are studied.

- **Sum power constraint**: All \( M \) BS’s in \( S_k \) have perfect power cooperation in addition to the data cooperation. This provides an unrealistic upper bound, where the pooled maximum available power is always \( P_{\text{sum}} = MP_T \), while the antenna array gain from having \( MN_T \) transmit antennas depends on the RX power differences between BS’s.

- **Per BS power constraint**: The TX power (10) in all \( M \) BS’s belonging to \( S_k \) has a maximum limit \( P_T \). Total available power per user can be increased up to \( M \) times depending on the RX power difference between BS’s. In addition, the antenna array gain from having \( MN_T \) transmit antennas depends on the RX power difference.

- **Shared single BS power constraint**: The total TX power for all \( M \) BS’s in \( S_k \) is constrained by \( P_T \), i.e. \( P_{\text{sum}} = P_T \). Only antenna array gain is available from the SHO processing.

6.4.1 Mutual information results

The mutual information for 2-branch SHO with different power constraints and with different optimisation criteria is studied in this section. The impact of inter-cell interference is omitted for simplicity, i.e. \( R_{k,c} = N_0 I \). The simulation scenario is further illustrated in Fig. 22. Observe that with a 0 dB RX power imbalance this scenario is equivalent to a single link MIMO system with four transmitter antennas, but with a \( P_T \) power constraint per group of two antennas.
Fig. 22. Simulation scenario for SHO between 2-adjacent BS’s.

Weighted sum rate maximisation

Fig. 23 illustrates the single user ergodic mutual information for different received power imbalance values $\alpha = \frac{a^2_{S_k(2),k}}{a^2_{S_k(1),k}}$ and for 0 dB and 20 dB single link SNRs.\(^9\) $S_k(1)$ is the BS with the strongest reception at the terminal and the single link SNR is defined as $\text{SNR} = \frac{P_T a^2_{S_k(1)}}{N_0}$. As already shown in [242], the single user ($|U| = 1$) rate optimisation reduces to (91), where $\lambda_{k,i,c}$ and $\mathbf{v}_{k,i,c}$ are the squared singular values and the right singular vectors of $\mathbf{R}_{k,c}^{-1} \tilde{\mathbf{H}}_{k,c}$, respectively. Fig. 23 shows that the performance of the proposed heuristic method is close to the optimal solution (91) with per BS power constraints. Moreover, the gain from joint processing quickly diminishes as the imbalance between the received BS powers increases, especially with low SNR values. On the other hand, the highest SHO gains are achieved in the low SNR range, where the achievable rate can even be doubled. The achievable rate with a sum power constraint provides an unrealistic upper bound assuming that the BS’s can share

Note that the RX power imbalance, $\alpha = \frac{P_T a^2_{S_k(2),k}}{(P_T a^2_{S_k(1),k})}$ is equivalent to the channel gain imbalance, $\alpha = \frac{a^2_{S_k(2),k}}{a^2_{S_k(1),k}}$ since all the BS’s in $S_k$ are assumed to use all available power and $P_n = P_T \forall n$. \(^{141}\)
their TX powers. The case with a sum power constraint and with infinite power imbalance (-Inf) is equivalent to single link transmission from a single BS but with 3 dB higher TX power (SNR = 2 × PTa2S1(1)/N0).

Next, a multiuser case is considered where two SHO users (labelled as K = 2) are served simultaneously by two BS’s in a flat fading scenario, as illustrated in Fig. 22. Furthermore, the assumption is that they have identical large scale fading coefficients for simplicity, i.e. aS1(1,1) = aS1(1,2) and aS1(2,1) = aS1(2,2). First, the impact of beamformer initialisation on the behaviour of the linear weighted sum rate maximisation algorithm is studied. Due to the non-convexity of the original optimisation problem (70), different initial beamformer configurations {v(0)1,...,v(0)S} used in Algorithm 6 may end up in different local optima. The behaviour of Algorithm 6 is illustrated with 0 and 10 dB single link SNRs and -3 and -10 dB RX power imbalance values in Fig. 24 for a 2-user channel, where rate pairs corresponding to different weight vectors are plotted with both sum and per BS power constraints. Equal weights are assigned to all streams associated with one user. For each weight vector, ten rate pair points are

![Fig 23. Single user ergodic mutual information of \{N_T, N_{R_k}, N_C, K, M\} = \{2, 2, 64, 1, 2\} system at 0 dB and 20 dB single link SNR.](image)
Fig 24. Broadcast capacity region and rate region with linear processing for \( \{N_T, N_{Rx}, N_C, K, M\} = \{2, 2, 1, 2, 2\} \) system.
generated each with random linear TX beamformer initialisation. The rate region with linear processing is then plotted as a convex hull of the achievable rate pairs. Furthermore, the rate pairs corresponding to the equal user weights are indicated. The rate regions are plotted for a single random channel realisation per user. All the points deviating from the boundary of the convex hull are local optima. Note that the flat parts in the rate regions are only achievable via time sharing. Moreover, the capacity region with per BS and sum power constraints, computed as in [163], are plotted as the absolute upper bounds of the scenario.

Fig. 25 depicts the ergodic 2-user sum rate as a function of the number of random beamformer initialisations. The best one out of \( n_{\text{init}} \) random TX beamformer initialisations was selected for each channel realisation. The imbalance between BS’s is fixed at 0 dB and the single link SNRs are 0 dB and 10 dB. The ergodic sum rate is depicted for the weighted sum rate maximisation algorithm (Algorithm 6 labelled as ‘lin. max rate’) with sum power and per BS power constraints and with weight vector \( \beta = 1 \). Moreover, the sum capacity with a per BS power constraint is plotted as the absolute upper bound of the scenario.

A large number of randomly chosen initialisations increases the probability for finding a solution close to the global optimum for each channel realisation. It can be seen from Fig. 25 that the impact of the initial beamformer configuration is relatively small and the gain achieved from randomly drawing several initial points saturates rapidly to a fixed value. A different approach was also explored where the initial transmit beamformers were obtained by applying the orthogonal-triangular QR decomposition to the set of dominant right singular vectors of user channels. The singular vectors were ordered in a descending order according to their singular values. The unitary vectors from the resulting Q matrix were used as initial beamformers. There is no randomness involved in the QR decomposition approach since the beamformers depend only on the given channel realisations. The resulting ergodic rates are repeated along the horizontal axis of Fig. 25 for comparison with the random initialisation. As shown in Fig. 25, the QR based initialisation method produces a very efficient starting point, e.g. more than two random initialisations are required to produce a higher ergodic sum rate.

Figs. 26(a) and 26(b) illustrate the ergodic mutual information for different
Fig 25. Impact of beamformer initialisation on ergodic sum rate of \( \{N_T, N_R, N_C, K, M\} = \{2, 2, 1, 2, 2\} \) system with 0 dB RX power imbalance at 0 dB and 10 dB single link SNR.

power imbalance values, and for 0 dB and 10 dB single link SNRs, respectively. The ergodic 2-user sum rate is depicted for the proposed algorithm with QR based initialisation and the corresponding ZF method (Section 6.2) with different power constraints. The single user case (\( K = 1 \)) with and without SHO is also included for comparison.

The sum capacity bounds shown in Figs. 25 and 26 are not generally achievable with a linear transmission strategy. However, the proposed weighted sum rate maximisation algorithm achieves more than 90 percent of the sum capacity with per BS power constraints and single sum power constraints. Note that the resulting ergodic sum rates for the proposed algorithm can still be slightly improved if a few random initialisations for each channel realisation are allowed as seen in Fig. 25.

The zero forcing solution, labelled as ‘ZF’, is depicted for two scenarios: the fully loaded case \( \{m_1, m_2\} = \{2, 2\} \), labelled as ‘FL’, and the partially loaded case, labelled as ‘PL’, where the best allocation of \( m_k \) among possible
Fig 26. Ergodic sum rate of \( \{N_T, N_{R_k}, N_C, K, M\} = \{2, 2, 1, 2, 2\} \) system.
combinations \( \{m_1, m_2\} = \{\{2, 2\}, \{2, 1\}, \{1, 2\}, \{1, 1\}\} \) is selected for each channel realisation. It is seen from the figures that the zero forcing with full spatial load (‘FL’) performs rather badly, especially in the low SNR range. Even with a large imbalance value (−20 dB) both users are intended to be served with two streams. This obviously reduces the achievable rate even below the single link capacity. The zero forcing with partial loading performs reasonably well even at low SNR and approaches the weighted sum rate maximisation algorithm (Algorithm 6) at high SNR. Again, the heuristic power loading solution performs almost as well as the optimal method (91). The ZF method with heuristic power loading and with partial spatial loading is used in the link and system level studies in Sections 6.4.2 and 6.4.4 due to its simplicity and good performance. It also enables decoupling the data streams, and hence, allows for efficient implementation of the bit and power loading algorithms in practical systems.

Maximisation of minimum SINR per stream

The ergodic sum rate for 2-branch SHO achieved using the SINR balancing criterion with a per BS power constraint (Section 6.1.1) is studied in a case where two or four SHO users, labelled as \( K = 2 \) and \( K = 4 \), respectively, are served simultaneously by two base stations in a flat fading scenario at each time instant. Only a single data stream is assigned to each user in this case \( (m_k = 1 \ \forall \ k) \).

Furthermore, the assumption is that the users have identical large scale fading coefficients for simplicity, i.e. \( a_{S_k(1),k} = a_{S_i(1),i} \) and \( a_{S_k(2),k} = a_{S_i(2),i} \ \forall \ i, k \in U \).

Figs. 27(a) and 27(b) illustrate the ergodic mutual information for different power imbalance values \( \alpha = a_{S_k(2),k}^2/a_{S_k(1),k}^2 \) and for 0 dB and 20 dB single link SNRs, respectively. The ergodic sum of individual user rates with different power constraints is depicted for the joint TX-RX optimisation algorithm (Section 6.1.1, labelled as ‘lin. SINR balancing’) and the zero forcing method (Section 6.2.2, labelled as ‘ZF SINR balancing’). Equal weighting of data streams \( \beta_s = \beta_{k,i} = 1 \ \forall \ s, i, k \) is used for both algorithms. The single user capacity \( (K = 1) \) without SHO is also included for comparison. Moreover, the sum capacity for four users with a per BS power constraint is plotted as the absolute upper bound of the scenario.

The achievable sum rate with an equal SINR per stream constraint depends
Fig 27. Ergodic sum of user rates of $\{N_T, N_{R_k}, N_C, K, M\} = \{2, 2, 1, 2-4, 2\}$ system with equal SINR per stream constraint.

148
heavily on the imbalance between BS’s. As the imbalance approaches infinity, the case with four scheduled users becomes spatially overloaded. In such a case, the achievable SINR $\gamma_s$ per data stream $s$ with spatial overloading is always sub-unitary, and hence, the maximum achievable rate in the high SNR region approaches four, $4 \log_2(1 + \gamma_s) \leq 4$ [223]. It is then obviously more efficient to serve less users at a time, i.e. to divide the users into subgroups and multiplex them in time domain. As the imbalance decreases, i.e. the users are located closer to the cell edge, the case with higher spatial load becomes more advantageous. In other words, full spatial loading ($K = 4$) becomes more efficient than the time multiplexing of the data streams with partial loading ($K = 2$). The exact crossing point where the sum rate with the full spatial loading exceeds the sum rate with partial loading depends on the SNR region, as seen in Fig. 27.

It can be seen from the figures that the zero forcing with full spatial loading ($K = 4$) performs rather badly in the low SNR range. Also, the achievable sum rate becomes zero as the system becomes spatially overloaded. However, the zero forcing with partial spatial loading ($K = 2$) performs reasonably well with a low SNR as compared to the joint TX-RX optimisation algorithm due to the higher number of spatial degrees of freedom available at both the transmitter and the receiver. With a high SNR and with a small imbalance between BS’s, the performance of the zero forcing method approaches the joint TX-RX optimisation algorithm.

Note that the upper bound is not achievable with a linear transmission strategy, since the sum capacity achieving schemes require the use of the nonlinear dirty paper coding based precoder [54]. However, the proposed linear joint TX-RX optimisation algorithm with an equal SINR per data stream constraint is able to achieve more than 80 percent of the sum capacity, with a small imbalance value between BS’s.

Maximisation of weighted common user rate

The SINR balancing algorithm can be used to provide equal rates per user when $m_k = 1 \ \forall \ k$, since the user rate balancing problem (79) is reduced to (65). On the other hand, an equal SINR per stream constraint could result in a highly sub-optimal solution when the number of available spatial (and/or frequency) dimensions is larger than the number of scheduled users. In such a case, the
proposed balanced rate maximisation solution in Algorithm 8 always provides a higher user rate.

The impact of beamformer initialisation on the performance of the linear weighted common rate maximisation algorithm is studied. Similarly to the weighted sum rate maximisation in Algorithm 6, different initial beamformer configurations \( \{v^{(0)}_1, \ldots, v^{(0)}_S\} \) used in Algorithm 8 may end up in different locally optimal solutions, due to the non-convexity of the original weighted common rate maximisation problem (79). The behaviour of Algorithm 8 in a 2-user channel is compared to Algorithm 6 at a 10 dB single link SNR and with a -3 dB RX power imbalance in Fig. 28. Similarly to Fig. 24, the rate pairs corresponding to different weight vectors and the rate regions are plotted for a single random channel realisation per user and with a per BS power constraint.

Fig 28. Broadcast capacity region and rate region with linear processing for \( \{N_T, N_R, N_C, K, M\} = \{2, 2, 1, 2, 2\} \) system with per BS power constraint.

The example shown in Fig. 28 demonstrates that the achievable rate region
boundaries, which are plotted as convex hulls of all the achievable rate pairs, are identical for both rate maximisation algorithms with linear processing. However, all the rate pairs with a weighted common rate constraint that deviate from the convex hull cannot be claimed as local optima, unlike in the weighted sum rate case. This is due to the different objectives of the two optimisation criteria. Again, the linear part of the convex hull can be achieved only via time sharing. This disfavours the common rate maximisation case especially in the region where the achievable points are inferior to the convex hull, since the user rates are constrained to follow the ratio defined by the weights $\beta_k$. A somewhat higher common rate can still be achieved by allowing time sharing between two rate pairs in both ends of the time sharing region. For instance, a common rate about 0.4 bits/s/Hz higher could be achieved with time sharing in the example shown in Fig. 28. However, finding the optimal pair of weight vectors and the corresponding beamformers for time sharing would require a full search on the rate region.

Fig. 29 depicts the ergodic 2-user sum rate as a function of the number of random beamformer initialisations. The simulation setup is identical to that of Fig. 24. The ergodic sum rate is depicted for the user rate balancing algorithm (Algorithm 8 labelled as ‘lin. rate balancing’) and also compared with the linear weighted sum rate maximisation (Algorithm 6). Comparison is carried out with a per BS power constraint and with an equal weight per user and/or stream $\beta_s = \beta_k = 1 \forall s, k$. Moreover, both 2-user sum and symmetric capacity with a per BS power constraint (computed as in Section 6.3) are plotted as the absolute upper bounds. Fig. 29 shows that the rate balancing algorithm achieves more than 80% of the symmetric capacity in the studied scenario despite the linearity constraint. Again, a large number of random initialisations increases the probability of finding a solution close to the global optimum for each channel realisation. The QR based initialisation method introduced in Section 6.4.1 produces a more efficient starting point than a single random initialisation also for the rate balancing algorithm. Fig. 29 also demonstrates that the rate penalty from the additional equal user rate constraint is higher for the linear transceiver than for the capacity achieving scheme, especially with a high SNR. This is due to the fact that the rate pairs achieved with equal user weights $\beta_s = \beta_k = 1 \forall s, k$ and with an equal path loss per user are likely to lie in the
Fig 29. The impact of the beamformer initialisation on sum rate of \( \{N_T, N_{R_k}, N_C, M, K\} = \{2, 2, 1, 2, 2\} \) system with 0 dB RX power imbalance at 0 dB and 10 dB single link SNR.

region where time sharing would provide a higher weighted common rate, as demonstrated in Fig. 28. The difference between the two cases is reduced if some unequal user weights are used.

Figs. 30(a) and 30(b) illustrate the ergodic mutual information for different power imbalance values, and for 0 dB and 20 dB single link SNRs, respectively. The ergodic 2-user sum rate is depicted for the user rate balancing algorithm with QR based initialisation and the corresponding ZF method (Section 6.2.2) with per BS power constraints. Furthermore, the sum rate of the linear SINR balancing algorithm (labelled as ‘lin. SINR balancing’) with a single stream per user \( (m_k = 1) \) is plotted for comparison.

In the given scenario, the SINR balancing solution provides nearly identical results with the user rate balancing algorithm with a low SNR and with a high RX power imbalance. This is due to the fact that the user rate balancing algorithm is also likely to only allocate a single stream per user in such situations.
Fig 30. Ergodic sum of user rates of \( \{N_T, N_{R_k}, N_C, K, M\} = (2, 2, 1, 2, 2) \) system with per BS power constraint.
With a high SNR, however, the user rate balancing algorithm utilises all the available dimensions to maximise the rate per user providing a higher sum rate.

The zero forcing solution, labelled as 'ZF', is depicted for two scenarios: the fully loaded case \( \{m_1, m_2\} = \{2, 2\} \), labelled as 'FL', and the partially loaded case, labelled as 'PL', where the best allocation of \( m_k \) among possible combinations \( \{m_1, m_2\} = \{(2, 2), (2, 1), (1, 2), (1, 1)\} \) is selected for each channel realisation.

The zero forcing with full spatial load gives rather poor performance, especially in the low SNR range. Even with a large RX power imbalance both users are intended to be served with two streams. This reduces the achievable rate down to zero as the system becomes spatially overloaded. The zero forcing with partial loading performs reasonably well even with a low SNR and approaches the user rate balancing algorithm in the high SNR.

Let us now consider an ideal cooperative MIMO–OFDM system \((\Gamma = 1 \text{ and } \gamma_{max} = \infty)\) with \( N_C = 16 \) parallel sub-channels, where four SHO users with equal large scale fading coefficients are served simultaneously by two base stations. The single link SNR is 10 dB and the RX power imbalance is 0 dB. Two GBR users have \( \{r_1, r_2\} = \{6, 4\} \) bits/s/Hz guaranteed throughput requirements and the users have priority weights \( \{\beta_1, \beta_2, \beta_3, \beta_4\} = \{1, 1, 1, 0.5\} \). Fig. 31(a) illustrates the behaviour of Algorithm 10 for a single random channel realisation in such a scenario. It can be seen from Fig. 31(a) that the initial beamformer configuration is not feasible for this particular channel realisation and eight beamforming weight updates for the GBR users are required to reach a feasible starting point. The common rate maximisation phase of Algorithm 10 is then initiated with a feasible beamformer configuration and the user rates converge close to their final values after 30 iterations.

In Fig. 31(b), the behaviour of the algorithm with slightly modified minimum throughput requirements \( \{r_1, r_2\} = \{5, 3\} \) bits/s/Hz is depicted for the same channel realisation. Now the initial beamformer configuration can support the minimum rate requirements, and hence, the feasibility check phase is not required. In this example, the final rate of GBR user 2 is also allowed to exceed the minimum rate requirement since \( \beta_2 r_o > r_2 \).
Fig 31. User rate evolution for a single channel realisation in \( \{ N_T, N_R, M, K, N_C \} = \{ 2, 2, 2, 4, 16 \} \) system at 10 dB single link SNR.
6.4.2 Link level results

In this section, the single user ($|\mathcal{U}| = 1$) link level performance for 2-branch SHO with different power constraints is studied in terms of achievable spectral efficiency. The link level simulation parameters are the same as listed in Table 2 on page 76. Fig. 32 illustrates the achievable spectral efficiency for different power imbalance values and for 0 dB and 12 dB single link SNRs with a fixed $-3$ dB power imbalance. With a small power imbalance the SHO gains can be rather significant. There is very little gain from SHO with a low SNR and with high imbalance between BS’s since the TX power is concentrated on the strongest eigenmode(s) only. This can be observed by looking at Fig. 33 where the achievable spectral efficiency is depicted for different power imbalance values and for 0 dB and 12 dB single link SNRs. However, the weaker BS has in general a larger contribution on the weaker eigenmode(s). Thus, SHO can provide considerable gains with a high SNR even with a large imbalance, as the strongest eigenmode(s) become saturated and more power is poured on the weaker eigenmode(s).

Fig 32. Single user spectral efficiency of $\{N_T, N_{Rx}, N_C, M_k\} = \{2, 2, 64, 2\}$ system with LSO algorithm and with $-3$ dB imbalance between BS’s.
6.4.3 Impact of non-synchronisation between BS antenna heads

In order to perform joint transmission from all the BS antenna heads belonging to the active set, the baseband signals need to have a common carrier phase reference by using some feedback from terminals, for example. The impact of imperfect synchronisation between the BS antenna heads on the achievable gains is addressed with some practical examples.

The impact of imperfect synchronisation between the BS’s is illustrated with an example, where the joint transmission is carried out from two BS’s having a phase mismatch $\phi$ common to all the antenna elements within one BS. In such a case, the estimated channel at the transmitter can be modelled as:

$$\hat{H}_{k,c} = [a_{S_k(1),k}H_{S_k(1),k,c}, a_{S_k(2),k}e^{j\phi}H_{S_k(2),k,c}] .$$  \hspace{1cm} (96)

The transmit pre-coding matrix designed based on the estimated channel is
\[ \hat{M}_{k,c} = \hat{V}_{k,c} \hat{P}_{k,c}^{-\frac{1}{2}}, \] where \( \hat{V}_{k,c} \) contains the first \( m_{k,c} \) columns of \( \hat{V}_{k,c} \) obtained by SVD of \( \hat{H}_{k,c} = \hat{U}_{k,c} \hat{\Lambda}_{k,c} \frac{1}{2} \hat{V}_{k,c}^H \). Now, the SINR \( \gamma_{k,i,c} \) per sub-channel at the receiver after receive filtering (7) can be calculated as [238]

\[ \gamma_{k,i,c} = \frac{1}{\left[ Q_{k,c,i,i} \right]} - 1 \] (97)

where

\[ Q_{k,c} = \left( I + \hat{M}_{k,c}^H \hat{H}_{k,c} R_{k,c}^{-1} \hat{H}_{k,c}^H \hat{M}_{k,c} \right)^{-1}. \] (98)

An example in Fig. 34 illustrates the impact of the phase mismatch between two adjacent BS’s for one random channel realisation \( \{N_T, N_R, N_C, M_k\} = \{2, 2, 64, 2\} \), with different RX power imbalance values \( \alpha = a_{S_k(2),k}^2 / a_{S_k(1),k}^2 \) at \( P_{R_{S_k(1),k}} / N_0 = 10 \text{ dB} \). The SINR values per eigenmode after receive filtering are shown for a single sub-carrier as a function of the phase mismatch. It can be is clearly seen from the figure that the target SINR values for different MCS’s cannot be met at the receiver if there is a phase mismatch between BS’s. However, the phase mismatch can vary between \( \pm 60^\circ \) while the loss in the eigenmode SINR is still less than 1.5 dB. The figure also indicates that there would be no gains from joint processing without phase synchronisation. On the contrary, some additional fading would be introduced (up to 15dB fades in this example).

### 6.4.4 System level evaluation

A realistic multi-cell environment with 57 cells was used for the system level evaluation. Fig. 35 illustrates the simulation scenario. It is assumed for simplicity that the cooperative SHO processing of the transmitted signal is possible between any of the 57 BS’s. Apart from that, the channel model and the simulation environment are the same as in Section 4.3.3. An independent time-continuous fading process is simulated for each MIMO antenna transmitter-receiver pair including both the desired links \( H_{b,k}, b \in S_k \) and the most dominant interference links \( b \notin S_k \) for each terminal \( k \) dropped in the system. Only the center site users’ data is recorded for the user performance statistics. An example distribution of some of the traced user locations is also shown in Fig. 35.
Fig 34. Impact of phase mismatch between BS’s.

Frequency reuse one is assumed. The multiple access scheme is TDMA, where each user is assigned a fixed length transmission slot within a DL TDD frame. The number of DL slots is set to 24, each slot corresponding to 8 OFDM symbols. The same link parameters are used as listed in Table 2 on page 76. Parameter $K$ is adjusted such that different target average DL loads (time slot occupation) are achieved. A simple dynamic transmission slot allocation scheme is used. It is assumed that the BS(s) are aware of each user’s received pilot power levels from all adjacent cells and the user allocations in adjacent cells. A user is allocated to a time slot within the DL TDD frame where the resulting SINR is the highest. The SINR is defined as the ratio between the average RX power from the strongest cell in $S_k$ and the average other-cell interference power. The resulting user allocation tables are maintained fixed during one simulation drop. The maximum TX power in each BS antenna head is fixed at 33 dBm. Only those center site users with an average SINR of more than 0 dB are traced for the statistics while the others are declared to be in the outage (dropped). The user locations are identical between the simulation cases in order to make the results comparable.
Fig 35. System level simulation scenario.

Depending on the simulation case, the users that have an identical SHO active set composition can be served within the same time slot using the ZF SDMA method of Algorithm 1 in Section 5.1 with two iterations and the joint user, bit and power allocation algorithm from Section 5.2.4. In order to simplify the simulation setup, the SDMA is not considered for users assigned to a single cell only ($|S_k| = 1$). For single user transmission, the original LSO algorithm from [276, 277] is used. The closed-loop interference non-reciprocity compensation algorithm from Section 4.2, capable of following the time-continuous changes in the interference structure, is applied for each traced user in order to maintain the target quality of service (10% FER) at the receiver. It also compensates for the residual MAI caused by the limited number of iterations in Algorithm 1.

Single user TDMA solution

First, the system level impact of the proposed SHO scheme is studied for the case where only one user can be allocated to a single time slot, i.e., SDMA
between users that have an identical $S_k$ is not allowed. Fig. 36 illustrates the blocking and the total outage (blocking + dropping) probability versus the system load with different SHO parameters. The load is defined such that each user is counted only in the cell he is associated with, i.e. in the cell with the highest received power. Due to the fact that the active SHO connection requires a physical resource (time slot) allocated at each participating BS, the actual load with SHO can be significantly larger than without SHO. Typically, the overhead from SHO varies between 20-50% depending on the parameters used. Thus, the blocking probability can be dramatically increased if the SHO parameters, SHO window and maximum SHO active set size, are too large. On the other hand, the dropping probability is decreased compared to the case without SHO. Thus, the total outage probability with SHO can be even less than without SHO (Fig. 36). It must be noted, that the dropping criteria used in these simulations (0 dB SINR) is rather strict, as for the channel allocation in the beginning of each user drop it is assumed that all the interfering BS’s are transmitting with full power $P_T$. In addition, the SINR after receiver processing can be significantly higher than the pre-processing SINR. Note that the blocking probability also depends on the frame parameters (total number of slots), i.e. the larger the channel pool the lower the call blocking and vice versa.

The CDF of the spectral efficiency per user with different SHO window sizes ($\pm 3$ dB and $\pm 6$ dB) and power constraints is plotted in Fig. 37(a). All the traced users are included in the statistics. The maximum SHO active set size is limited to three in this case. The load used in the simulations in Fig. 37(a) corresponds to 30% time slot occupation without SHO. It is seen from the figure that significant system level gains from cooperative SHO processing are available, especially with a large SHO window. However, the 6 dB SHO window becomes too large when the load increases further from 30% due to increased outage probability (Fig. 36).

The impact of two different power constraints is also compared in Fig. 37(a). A per BS power constraint implies that the power consumption at BS’s may be increased due to SHO overhead. Any BS in $S_k$ can use up to $P_T$ TX power per time slot depending on the users’ received signal strength, imbalance between SHO BS’s, etc. In addition, more inter-cell interference is potentially generated. Therefore, it is interesting to compare this to the case where the
power consumption is maintained the same on average, independent of the SHO parameters used. A shared single BS power constraint does not increase the power consumption at the transmitters even if the SHO overhead is high, since the TX power $P_T$ of a single BS is shared between $M_k$ transmitters. In addition, the inter-cell interference generated from a single BS is reduced with the same ratio. In spite of generating more inter-cell interference, the per BS power constraint results in better overall system performance.

Obviously, the users located in the SHO region may enjoy greatly increased transmission rates, as depicted in Fig. 37(b). The spectral efficiency of those users that are outside of the SHO region is slightly reduced due to the increased inter-cell interference. However, the net performance improvement is clearly positive as seen from Fig. 37(a).
Fig 37. CDF of user spectral efficiency with 30% system load, max. AS size 3 and 3 dB SHO window.
Multiuser SDMA for SHO

Now, SDMA between users with an identical $S_k$ is enabled. The channel allocation algorithm is slightly modified for SDMA. Still, the new user is allocated to the time slot where the resulting SINR is the highest. For each new user $k$, the allocation table is checked for whether there are any earlier allocated users $i$ with the same active set composition, $S_k = S_i$. If the available SINR in the slot of the $i^{th}$ user is larger than or equal to the best free slot, user $k$ is allocated to the same slot with user $i$. The maximum number of users per time slot is limited to $M$, and $MN_T \geq \sum_{k \in U} m_{k,c} \forall c$. The number of beams allocated per user $m_{k,c}$ depends on the loading algorithm (Section 5.2.4). Fig. 38 illustrates the outage probability of the system with different SHO parameters. The probability of finding a user with the same SHO active set composition depends on the system load and on the SHO parameters used. The largest improvement in terms of reduced total outage is achieved with a large SHO window, where the probability for such an occurrence can be up to 40%. SDMA reduces also the inter-cell

![Fig. 38. Blocking and dropping probability as a function of load, SDMA for SHO users.](image)
interference by concentrating the transmissions in fewer time slots. Therefore, it
decreases the dropping probability as well. The overall outage probability is
hence reduced in all SHO cases below the case without SHO.

Fig. 37(b) also depicts the spectral efficiency distribution for the case where
SDMA is enabled. Some performance penalty from using the ZF method from
Section 5.1 is caused by its inherent noise amplification property [266]. In
addition, the TX power is shared between the users allocated to the same time
slot. However, the overall reduction is rather small, resulting from the fact
that the fading is independent at each BS antenna site. At the same time less
inter-cell interference is generated. Therefore, the total spectral efficiency with
SDMA (not shown) remains more or less unchanged with the curves shown in
Fig. 37(a).

6.5 Summary and discussion

The joint cooperative processing of the transmitted signal from several MIMO
BS's was considered for users located within an SHO region. The downlink
space-frequency bit and power allocation problem with different BS power
constraints was studied for the considered adaptive MIMO-OFDM system.
The system level gains and trade-offs from cooperative SHO processing were
investigated. The mathematical framework for the SHO based MIMO-OFDM
system was derived and the joint design of linear TX and RX beamformers
in a MIMO multiuser transmission was considered. A general method was
proposed for the linear transceiver design subject to different optimisation
criteria in addition to per BS and/or per antenna power constraints. Solutions
for total power minimisation subject to SINR constraints, maximisation of the
weighted sum rate, finding a maximum weighted common rate achievable for
each scheduled user and maximisation of the minimum weighted SINR per
stream were proposed. In addition, an extension to the frequency selective
downlink channel with OFDM transmission and with additional QoS constraints
was provided. Moreover, efficient resource allocation methods based on zero
forcing transmission were provided. The proposed joint transceiver optimisation
algorithms were compared with corresponding generalised ZF transmission
solutions, as well as with corresponding optimal non-linear transmission methods,
and they were shown to provide very efficient solutions despite the fact that
there is no guarantee of achieving the global optimum due to the non-convexity
of the optimisation problems.

The performance of the proposed heuristic loading method for sum rate
maximisation was shown to be close to the optimal method with per base station
power constraints. Furthermore, the gain from joint processing in SHO quickly
diminishes as the imbalance between the received BS powers increases, especially
with a low SNR. On the other hand, the highest SHO gains are achieved in the
low SNR range with a small power imbalance, where the achievable rates can be
even doubled.

Note that the proposed joint transceiver optimisation algorithms are probably
still too complex for practical implementation. However, most of the gains
from the iterative algorithm are typically achieved from the first few iterations,
depending on the channel realization and the SNR point. Therefore, the algorithm
could be terminated after a few iterations of the main algorithm and would
still provide a relatively good transmit beamformer configuration. Also, an
interesting future research topic is to derive some alternative practical solutions
for the proposed algorithms that rely on the optimality conditions or the dual
formulation of the subproblems, for example.

The impact of the size of the SHO region, overhead from the increased
hardware and physical (time, frequency) resource utilisation, and different
non-reciprocal inter-cell interference distributions due to SHO were evaluated by
system level simulations. A practical user, bit and power allocation method
with different BS power constraints was provided for the proposed cooperative
multiuser MIMO transmission. Although the overhead from the SHO processing
can be significant, it can be mitigated by using zero forcing SDMA for users with
an identical SHO active set composition. In addition, the dropping probability is
decreased, and thus, the total outage probability with SHO can be less than
without SHO depending on the parameters used. The users located in the
SHO region may enjoy greatly increased transmission rates. This translates to
significant overall system level gains from the cooperative SHO processing. The
proposed soft handover scheme can be used to provide more evenly distributed
service over the entire cellular network.

The results also indicate that the joint processing can be even detrimental
for system performance if coarse phase synchronisation between BS’s is not
guaranteed, as some additional fading on the target SNR values is introduced.
7 Conclusion and future work

The scope of this thesis was to develop methods to efficiently convey wireline packets over the wireless last segment between the BS’s and the mobile devices. In particular, issues pertaining to resource management for cooperative broadband wireless systems were examined while considering their special characteristics such as multi-carrier techniques, adaptive radio links and MIMO antenna techniques. Chapter 2 contained the literature review of previous and parallel work which covered several topics related to frequency selective MIMO broadcast channels with multiple users and with multiple transmitters, including channel capacity, optimal and sub-optimal transmitter and receiver design subject to different quality of service criteria, and allocation of available resources over different dimensions. Chapter 3 presented a generic system model for the MIMO-OFDM cellular system with cooperative base station antenna heads.

Chapter 4 focused on the evaluation of adaptive TDD MIMO-OFDM system performance in the presence of non-reciprocal inter-cell interference when the DL interference structure is known only at the receiver. An LMMSE filter was applied at the receiver to suppress the effect of structured inter-cell interference together with a simple and bandwidth efficient closed-loop compensation algorithm. The performance of the proposed framework was analysed and compared to the ideal case where the interference structure per sub-carrier is perfectly known at the transmitter. The proposed closed-loop compensation algorithm combined with interference suppression at the receiver resulted in nearly the same performance as the ideal case in both link and system level studies. The results also demonstrated that in the presence of non-reciprocal inter-cell interference, the quality of service at the receiver cannot be controlled if the transmission parameters are defined based on the reverse link measurements only. Therefore, some feedback to the transmitter is always needed in order to make the cellular MIMO-OFDM system to work properly. Albeit only single user MIMO transmission with TDMA was considered in Chapter 4, the proposed framework is easily extensible to the multiuser MIMO case with inter-cell interference, as was done in Chapter 6. It can also be used to compensate for the impact of channel estimation errors and the imperfect orthogonalization of allocated streams or users, as was done in
Chapter 5.

The implementation of SDMA in the DL of a multi-user MIMO-OFDM system was considered in Chapter 5. Generalised block-ZF processing was used to separate the users in the space domain. A greedy beam and/or user ordering and selection algorithm was proposed for maximising the DL sum rate of the considered system. It was shown that the performance of the block-ZF transmission combined with greedy scheduling approaches the sum rate capacity with a high SNR as the number of users increases. The maximum sum rate or sum spectral efficiency was shown to often be achieved by transmitting to less users/beams than the spatial dimensions available allow, especially in the low SNR region, due to the inherent noise amplification problem of the generalised ZF transmission. An efficient low complexity joint user, bit and power allocation algorithm with low signalling overhead was also proposed, and its performance was shown to be close to the greedy allocation algorithm with Hughes-Hartogs loading. Iterative block-ZF transmission with a single reiteration was shown to provide significant gains to the non-iterative solution, especially if the number of beams assigned per user was less than the number of RX antennas. An LMMSE filter was applied at the receiver to suppress the remaining multi access interference from incomplete orthogonalisation together with a simple compensation algorithm with low rate scalar feedback to the transmitter. The proposed MU MIMO-OFDM system was also shown to be robust against imperfect channel estimation at the transmitter, providing always superior performance to the single user TDMA solution.

BS cooperation with linear transceiver processing in a cooperative cellular system with distributed MIMO antenna heads was studied in Chapter 6. Joint BS processing allowed several users with an identical SHO active composition to be served in the same time-frequency transmission slot by separating their transmissions in the space domain. Different DL space-frequency resource allocation problems with practical per antenna or BS power constraints were studied within the context of the cooperative MIMO-OFDM system. The joint design of the linear TX and RX beamformers in a MIMO multi-user transmission for several optimisation criteria was considered. The studied optimisation criteria included: power minimisation subject to per stream SINR constraints or individual user rate constraints, weighted SINR balancing, weighted sum rate maximisation, and weighted rate balancing. The proposed methods are
able to handle multiple antennas at BS’s and mobile users, and any number of data streams are allowed per scheduled user. The proposed method can easily accommodate supplementary constraints, such as the minimum bit rate per user or upper/lower bounds for the SINR values of individual data streams. It can also handle a fairly general set of practical power constraints for the TX beamformers, e.g. a maximum sum power for any arbitrary subset of the TX antennas.

Unlike the optimal non-linear scheme, the optimisation problems employed in the linear multiuser MIMO transceiver design are generally non-convex. Therefore, the problem of finding the global optimum for most of the optimisation problems is intrinsically non-tractable. An iterative solution was proposed in Chapter 6, where each sub-step can be efficiently solved by using convex optimisation tools. Even though global optimality cannot be guaranteed due to the non-convexity of the original problem, the simulation results demonstrated that the achieved locally optimal solutions are very efficient in several practically relevant scenarios. Furthermore, solutions based on block-ZF transmission were provided for a variety of optimisation criteria. With a ZF constraint, the optimisation problems were simplified to convex power loading problems that could be easily solved using standard optimisation toolboxes. A practical user, bit and power allocation method with different BS power constraints was also provided for the proposed cooperative multiuser MIMO transmission.

Link and system level evaluation was carried out in order to assess the impact of a realistic multi-cell environment on cellular system performance. It was shown that the gain from joint processing in SHO quickly diminishes as the imbalance between the received BS powers increases, especially with a low SNR. On the other hand, the highest SHO gains are achieved in the low SNR range with a small power imbalance, where the achievable rates can be even doubled. The impact of the size of the SHO region, overhead from increased resource utilisation, as well as different non-reciprocal inter-cell interference distributions due to the SHO were evaluated by system level simulations. The overhead from SHO processing was mitigated by using SDMA for users with an identical SHO active set composition. The users located in the SHO region were shown to enjoy greatly increased transmission rates. This was translated to significant overall system level gains from the cooperative SHO processing. The proposed soft handover scheme can be used to provide more evenly distributed service over the
entire cellular network. The results also indicated that joint processing can be even detrimental for system performance if coarse phase synchronisation between the BS’s is not guaranteed. Thus, some feedback must be used to compensate for the RF impairments at the transmitters.

A major challenge for wireless communication systems is how to allocate resources among users across the space (including different cells), frequency and time dimensions and jointly design all the transceivers with different system optimisation objectives. This remains unresolved for a large variety of optimisation criteria, especially when combined with practical modulation and coding schemes. The problem is a difficult non-convex combinatorial problem with integer constraints and finding jointly optimal solutions is most likely intractable. Therefore, efficient sub-optimal solutions are required. Also, an interesting future research topic is to derive some alternative practical solutions, potentially with a lower computational complexity for the proposed algorithms that rely on the optimality conditions or the dual formulation of the sub-problems, for example.

The optimal beamformer design for MIMO transceivers as well as resource allocation with imperfect CSI due to the channel estimation uncertainty at the transmitter are still largely unresolved research problems. The feedback method proposed in this thesis for compensating for the imperfect TX CSI is rather simple because it does not consider any statistics of the CSI imperfection. A refined approach considering an elaborated robust design against imperfect CSI at the transmitter is a promising research line for future work.

An interesting future extension for interference studies is to evaluate the effect of bursty packet traffic and the impact of scheduling decisions on system performance. The fact that a higher initial power adaptation value is needed in the beginning of packet transmissions reduces the gains available from fast power offset adaptation for short packets. In practical cellular networks, the inter-cell interference experienced by a given user may sometimes vary even faster than the channel itself, depending on the fast frame-by-frame channel allocations in the neighbouring cells. Different users may be served in subsequent frames. This affects the precoder design at the neighbouring BS transmitters, which in turn changes the structure of the interference received by a given user. Another future research topic is the further refinement of the feedback scheme. Since the scalar feedback affects only the power (and the MCS) allocated to
different eigenmodes/subcarriers, the beamforming vectors are not affected. Thus, the impact on the structure of the interference experienced by other users is significantly smaller than in the case where the interference covariance matrices are ideally available at the transmitters. Obviously, the feedback may still cause some oscillation in the adaptive process due to the distributed nature of the algorithm. One possible solution for such a problem is to distribute the feedback information between all the affected nodes via a backbone network and to design the feedback strategies jointly across the nodes.

Typically, the parameters of the OFDM transmission, e.g. frame length and CP length, are designed based on the characteristics of the radio environment, such as average or worst case delay and Doppler spreads. Cooperative BS processing may potentially induce increased delay spreads due to the fact that the distributed antenna heads can be placed far apart. An interesting future topic would be to evaluate the impact of the distributed antenna system on the design of OFDM parameters.
References


19. 3rd Generation Partnership Project (3GPP); Technical Specification Group Radio Access Network (2006). Requirements for evolved UTRA (E-UTRA) and evolved UTRAN (E-UTRAN) (TR 25.913 version 7.3.0 (release 7)).


idea whose time has come. IEEE Communications Magazine 28(5): 5–14. Telebit Corporation, USA.


Turkey, vol. 4, 1641–1646.


with adaptive subcarrier, bit, and power allocation. IEEE Journal on Selected Areas in Communications 17(10): 1747–1758.


148. Heath RW & Paulraj AJ (2005) Switching between diversity and multi-


on Wireless Communications 3(6): 2097–2107.


systems under multiuser interference based on worst case noise behavior.
EURASIP J. Wireless Communications and Networking, Special Issue:


systems with MIMO. IEEE Communications Letters 6(6): 242–244.


modulation over wireless channels with non reciprocal interference. Proc.

201. Cover TM & Gamal AAE (1979) Capacity theorems for the relay channel.

202. van der Meulen EC (1977) A survey of multiway channels in information


capacity theorems for relay networks. IEEE Transactions on Information
Theory 51(9): 3037–3063.

I: System description. IEEE Transactions on Communications 51(11):
1927–1938.

II: Implementation aspects and performance analysis. IEEE Transactions

wireless networks: Efficient protocols and outage behavior. IEEE

for exploiting cooperative diversity in wireless networks. IEEE Transactions

outage-optimal opportunistic relaying. IEEE Transactions on Wireless Communications 5.


Communications 2(4): 773–786.


281. 3GPP TSG RAN WG1 (2001) A standardized set of MIMO radio propa-


286. Tapurungsananogorn, Ataphongse (2007) Evaluation of MIMO radio channel characteristics from TDM-switched MIMO channel sounding

287. Elsiä, Ulla (2007) Knowledge discovery method for deriving conditional probabilities from large datasets


291. Lyöri, Veijo (2007) Structural monitoring with fibre-optic sensors using the pulsed time-of-flight method and other measurement techniques


294. Gore, Amol (2008) Exploring the competitive advantage through ERP systems. From implementation to applications in agile networks

Antti Tölli

RESOURCE MANAGEMENT IN COOPERATIVE MIMO-OFDM CELLULAR SYSTEMS