6

M4 — Multi-Mode Massive MIMO

Nils L. Johannsen, Peter A. Hoeher

Information and Coding Theory, Kiel University

Nikolai Peitzmeier, Dirk Manteuffel

Institute of Microwave and Wireless Systems, Leibniz University Hannover

CONTENTS

6.1 Introduction

Next generation wireless communication networks will play a crucial role regarding the requirements and possibilities of future society and industry. A tremendous growth in both required data-rates as well as employable number of devices and hence reliability is expected. To solve the difficulties resulting from a large number of users, massive MIMO systems are of interest. These

employ a much larger number of antenna elements and ports at the base station (BS) than are provided in the user equipment.

Traditional wireless communication networks rely on the bandwidth provided in the frequency bands below 10 GHz. To overcome the drawback of limited available bandwidth in these commonly used frequencies, ongoing research focus is on the millimeter wave (mmWave, 30-300 GHz), sub-mmWave (300-1000 GHz) or even THz regions [239, 240]. Because of the large free-space attenuation in these bands, line-of-sight (LOS) beamforming is mandatory. On the one hand, the required space can be reduced by using higher frequencies. On the other hand, using these frequencies requires new hardware being capable of handling the resulting short wavelengths.

Opposed to this approach, focussing on the frequency range below 10 GHz allows employment of mature technology [241, 242]. Consequently, the size of the surface of the desired array increases. To establish targeted data-rates of 100 Gbps and beyond, an ultra-wideband (UWB) communication system is studied in our contribution. The availability of a 2.5 GHz frequency band is assumed, like the frequency range from 6 GHz to 8.5 GHz as offered by our prototype antenna demonstrator.

To be able to reduce the surface of an array providing a certain degree of diversity and enabling UWB communication, multi-mode antennas can be employed. Beginning at the mobile station (MS), the baseplate of the chassis can be thought of as a planar radiator. If its shape can be kept rectangular and does not need to be altered, the antenna can be integrated in the mobile device without an additional radiator. As an advantage, more than one port can be made available, as will be shown subsequently. At the BS, multiple radiators can be operated in parallel in order to boost the number of ports even further. The shape of the radiators can be changed such that the properties in terms of achievable ports and performance are optimized.

Massive MIMO systems use a large number of antenna ports at the BS side to make use of the statistical channel properties. The individual properties average out. This is referred to as channel hardening. Channel hardening allows simple linear precoding techniques to provide good and close to optimal performance [243, 244, 245]. Another aspect is the channel gain, which increases with the number of available ports.

To further increase the achievable data-rates, the transmit signal power is commonly increased. In most wireless communication standards, the equivalent isotropic radiated power (EIRP) is limited. However, the impact of an EIRP constraint on the optimization of massive MIMO systems has rarely been investigated.

The goal of the Multi-Mode Massive MIMO (M4) project is to confirm the ability of multi-mode antennas for UWB massive MIMO systems. A system proposal is targeted. Different challenges shall be determined and addressed. A spectral efficiency of 40 bps/Hz is required in order to achieve a sum data-rate of at least 100 Gbps in the frequency regime between 6 and 8.5 GHz.

Novel contributions include the following aspects:

- Derivation of a theoretical upper bound of orthogonal ports which are realizable on a single radiator.

- Modal analysis of defined antenna structures and design concept for multi-mode antennas.

- Performing power control and precoding for EIRP-limited MIMO systems.

- Single-element beamforming, using the properties of multi-mode antenna radiation patterns.

The chapter is organized as follows: Starting with the theory of characteristic modes, which predicts an upper bound on the number of available orthogonal modes, a design towards both multi-mode antennas suitable for base stations and mobile terminals is proposed. The theory is extended to achieve the theoretical maximum number of ports given a certain antenna structure. Under the assumption of the proposed antenna characteristics, different precoding techniques are investigated. Since the maximum transmit power is limited with respect to either radiation or input power constraints, their impact on the achievable sum-rate using multi-mode antennas is investigated. Different beamforming strategies are discussed. Mode diversity can be used to perform single-element beamforming, which allows steering of the radiation pattern by using just one physical radiator.

6.2 Fundamentals

In this section, fundamentals regarding the M4-Project are provided and principles of multi-mode antennas are introduced.

6.2.1 Application of Characteristic Modes to MIMO Antenna Design

In order to realize MIMO antennas both at the base station and the mobile terminal, the concept of multi-mode antennas based on the theory of characteristic modes is employed. The characteristic modes of an arbitrary perfectly electrically conducting (PEC) object (antenna) form a basis set of characteristic surface current densities and corresponding characteristic electromagnetic fields for the expansion of the total radiated electromagnetic fields of the antenna [246]. The ν-th characteristic surface current density $\mathbf{J}_\nu$ is defined by the following generalized eigenvalue problem [247]:

$$X\left(\mathbf{J}_\nu\left(\mathbf{r}'\right)\right) = \lambda_\nu R\left(\mathbf{J}_\nu\left(\mathbf{r}'\right)\right), \tag{6.1}$$

where $\mathbf{r}' = (x', y', z')^T$ is a point on the antenna surface and λ_ν denotes the corresponding eigenvalue. The linear operators R and X are the real and

imaginary part, respectively, of the complex impedance operator $Z = R + jX$ derived from the electric field integral equation (EFIE) and the electric field boundary condition for perfect electric conductors.

The total surface current density $\mathbf{J}$ on an antenna can be decomposed into a weighted sum of the characteristic surface current densities [247]:

$$\mathbf{J}\left(\mathbf{r}'\right) = \sum_{\nu} \alpha_{\nu} \mathbf{J}_{\nu}\left(\mathbf{r}'\right) = \sum_{\nu} \frac{1}{2} \frac{\langle \mathbf{J}_{\nu}\left(\mathbf{r}'\right), \mathbf{E}_{\text{inc}}\left(\mathbf{r}'\right)\rangle}{1 + j\lambda_{\nu}} \mathbf{J}_{\nu}\left(\mathbf{r}'\right), \qquad (6.2)$$

where α_{ν} is called the modal weighting coefficient. Its numerator is called the modal excitation coefficient and it is defined as the scalar product

$$\left\langle \mathbf{J}_{\nu}\left(\mathbf{r}'\right), \mathbf{E}_{\text{inc}}\left(\mathbf{r}'\right)\right\rangle = \iint_{S'} \mathbf{J}_{\nu}\left(\mathbf{r}'\right) \cdot \mathbf{E}_{\text{inc}}\left(\mathbf{r}'\right) \mathrm{d}S', \qquad (6.3)$$

where the integration is taken over the antenna surface S'. $\mathbf{E}_{\text{inc}}$ denotes the incident electric field on the antenna surface, which may e.g. be impressed by an antenna port. In this notation, the characteristic surface current densities are normalized to radiate unit power (6.4a).

The operators R and X in (6.1) are real-valued symmetric operators. Thus, the eigenvalues and the characteristic surface current densities are real-valued and fulfill the following orthogonality relationships with respect to the impedance operator [247]:

$$\frac{1}{2}\left\langle \mathbf{J}_{\mu}\left(\mathbf{r}'\right), R\mathbf{J}_{\nu}\left(\mathbf{r}'\right)\right\rangle = \delta_{\mu\nu}, \qquad (6.4\text{a})$$

$$\frac{1}{2}\left\langle \mathbf{J}_{\mu}\left(\mathbf{r}'\right), X\mathbf{J}_{\nu}\left(\mathbf{r}'\right)\right\rangle = \lambda_{\nu}\delta_{\mu\nu}, \qquad (6.4\text{b})$$

$$\frac{1}{2}\left\langle \mathbf{J}_{\mu}\left(\mathbf{r}'\right), Z\mathbf{J}_{\nu}\left(\mathbf{r}'\right)\right\rangle = \left(1 + j\lambda_{\nu}\right)\delta_{\mu\nu}, \qquad (6.4\text{c})$$

where $\delta_{\mu\nu}$ is the Kronecker delta. From these orthogonality properties, it can be deduced that the corresponding characteristic far fields $\mathbf{E}_{\mu}$ and $\mathbf{E}_{\nu}$ are orthogonal to each other [248]:

$$\frac{1}{2Z_0} \oiint_{S} \mathbf{E}_{\mu}\left(\mathbf{r}\right) \cdot \mathbf{E}_{\nu}^{*}\left(\mathbf{r}\right) \mathrm{d}S = \delta_{\mu\nu}, \qquad (6.5)$$

where Z_0 is the wave impedance of free space, * denotes the complex conjugate, $\mathbf{r}$ is a point in the far field ($\|\mathbf{r}\| \to \infty$), and the integration is taken over a closed surface S in the far field encompassing the antenna.

The aforementioned orthogonality properties of the characteristic modes make them most suitable for MIMO applications. If the antenna is designed in such a way that its ports excite mutually exclusive sets of characteristic modes, the antenna ports are uncorrelated due to (6.5). This is checked by

evaluating the envelope correlation coefficients ECC_{uv} calculated from the total radiated far fields of the u-th and v-th antenna port [248]:

$$\text{ECC}_{uv} = \frac{\frac{1}{2Z_0} \oiint_S \mathbf{E}_u\left(\mathbf{r}\right) \cdot \mathbf{E}_v^*\left(\mathbf{r}\right) \mathrm{d}S}{\sqrt{P_{\text{rad},u}}\sqrt{P_{\text{rad},v}}}, \tag{6.6}$$

where $P_{\text{rad},u;v}$ denotes the radiated power of the u-th and v-th antenna port, respectively. The total radiated far fields of the u-th and v-th antenna port can each be expanded into a weighted sum of the characteristic far fields in the same manner as in (6.2) [247]:

$$\mathbf{E}_{u;v}\left(\mathbf{r}\right) = \sum_\nu \alpha_{\nu,u;v}\mathbf{E}_\nu\left(\mathbf{r}\right) = \sum_\nu \frac{1}{2}\frac{\left\langle \mathbf{J}_\nu\left(\mathbf{r}'\right), \mathbf{E}_{\text{inc},u;v}\left(\mathbf{r}'\right)\right\rangle}{1 + j\lambda_\nu}\mathbf{E}_\nu\left(\mathbf{r}\right). \tag{6.7}$$

Based on this, it is shown in [249] that the envelope correlation coefficients can be expressed in terms of the modal weighting coefficients alone:

$$\text{ECC}_{uv} = \frac{\sum_\nu \alpha_{\nu,u}\alpha_{\nu,v}^*}{\sqrt{P_{\text{rad},u}}\sqrt{P_{\text{rad},v}}} = \sum_\nu b_{\nu,u}b_{\nu,v}^*, \tag{6.8}$$

where $b_{\nu,u;v}$ are the normalized modal weighting coefficients of the u-th and v-th antenna port, respectively, as defined in [248]:

$$b_{\nu,u;v} = \frac{\alpha_{\nu,u;v}}{\sqrt{P_{\text{rad},u;v}}}, \tag{6.9}$$

which can also be interpreted as the correlation coefficients between the total radiated far field excited by a port and the modal far fields. Equation (6.8) demonstrates the fact that antenna ports are uncorrelated if they excite mutually exclusive sets of characteristic modes, i.e. $b_{n,u} = 0$ if $b_{n,v} \neq 0$ and vice versa. The aim of multi-mode antenna design for MIMO is thus to excite mutually exclusive sets of characteristic modes.

According to (6.2) and (6.7), two criteria have to be fulfilled in order to make a mode contribute significantly to the total excited current density and the total radiated far field of an antenna port:

1. Eigenvalue close to zero and

2. large modal excitation coefficient.

An eigenvalue close to zero is a necessary condition such that the corresponding characteristic mode may potentially contribute significantly. This potential is commonly called the modal significance MS_ν and it is defined as [250]

$$\text{MS}_\nu = \left|\frac{1}{1 + j\lambda_\nu}\right| = \frac{1}{\sqrt{1 + \lambda_\nu^2}}. \tag{6.10}$$

A characteristic mode is considered significant if its modal significance is

greater than $1/\sqrt{2}$ ($-1 \leqslant \lambda_\nu \leqslant 1$). The modal significance depends solely on the antenna geometry. Thus, it has to be designed in such a way that a sufficient number of significant characteristic modes is available for MIMO operation.

In contrast, the modal excitation coefficient as introduced in (6.3) is controlled by the incident electric field, i.e. the antenna port. As the excitation coefficient is a measure of the similarity between the incident electric field and the characteristic surface current density, a proper inspection of the current density to be excited gives insight into how to define a suitable port configuration. Typically, two types of excitations are distinguished: Inductive coupling is performed at local maxima of the current density by introducing slots or loops (inductive coupling elements), whereas capacitive coupling is performed at local minima of the current density (local maxima of the electric field) by introducing additional elements to the antenna (capacitive coupling elements) [251].

The basic work flow of multi-mode antenna design is illustrated in Fig. 6.1 by means of a rectangular PEC plate with dimensions of one wavelength by half a wavelength ($\lambda \times \lambda/2$) (Fig. 6.1(a)). The modal analysis reveals that a rectangular plate of this electrical size offers four significant characteristic modes (Fig. 6.1(b)). The corresponding characteristic surface current densities are shown in Fig. 6.1(c), where the principal current directions are denoted by arrows. A suitable port configuration is derived from the current densities and realized using inductive coupling elements [251] (Fig. 6.1(d)). It is noteworthy that one port consists of several feed points which need to be connected by means of a feed network [253]. This special port configuration leads to orthogonal far field patterns (Fig. 6.1(e)), which may be checked by computing the envelope correlation coefficients (6.6). This initial example demonstrates the basic work flow of multi-mode antenna design. The challenges that arise will be discussed in more detail in Section 6.3 and a systematic design procedure will be presented.

The characteristic mode analysis, which is mandatory for multi-mode antenna design, is commonly performed numerically by means of the method of moments (MoM) [254]. The MoM-impedance matrix is used to solve the generalized eigenvalue problem (6.1). For the characteristic mode computations throughout this chapter, an in-house MoM-code implemented in MATLAB is employed [255]. It is based on a modified version of the code presented in [256], utilizes built-in MATLAB functions for the eigenvalue problem, and offers advanced eigenvalue tracking algorithms [257].

6.2.2   Baseband Channel Modeling

A common equivalent discrete-time baseband system model is described by (see for example [258])

$$\boldsymbol{y} = \boldsymbol{DHPs} + \boldsymbol{n}, \tag{6.11}$$

(a) Rectangular PEC plate as basic geometry.

(b) Modal significances.

Mode 1   Mode 2   Mode 3   Mode 4

(c) Characteristic surface current densities of significant modes with principal current directions denoted by arrows.

Port 1   Port 2   Port 3   Port 4

(d) Port configuration using inductive coupling elements.
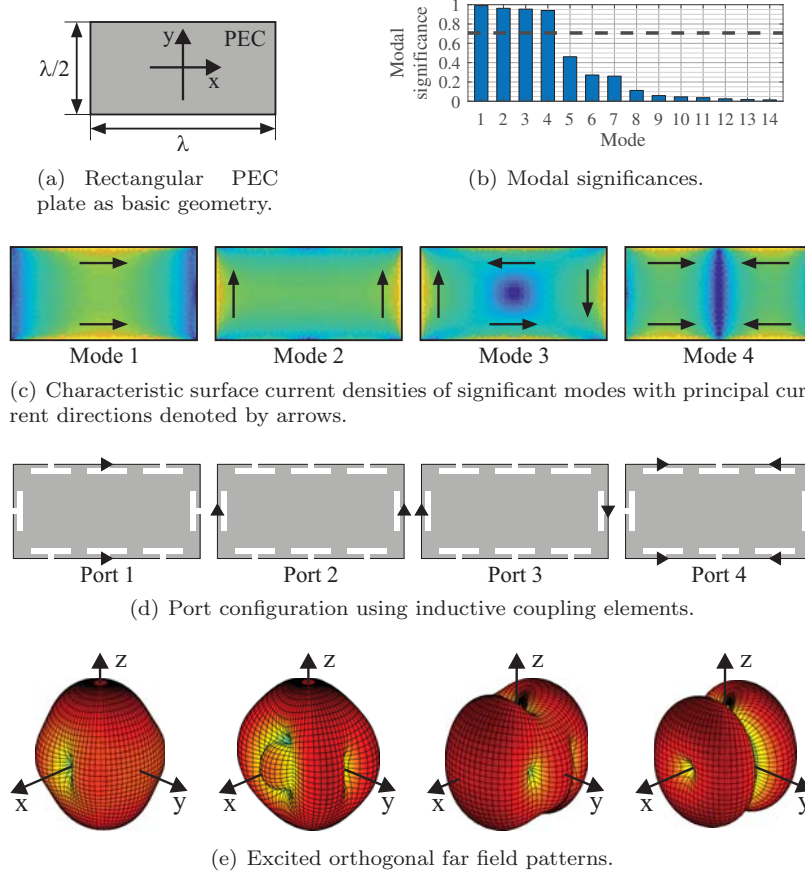

(e) Excited orthogonal far field patterns.

Figure 6.1
Work flow of multi-mode antenna design. From [252] © 2019 IEEE

where $\boldsymbol{y}$ is the observation vector (i.e., the set of received samples), $\boldsymbol{D}$ the decoding matrix, $\boldsymbol{H}$ the channel matrix, $\boldsymbol{P}$ the precoding matrix, $\boldsymbol{s}$ the transmit data vector, and $\boldsymbol{n}$ is an additive white Gaussian noise (AWGN) vector. The transmit vector

$$\boldsymbol{x} = \boldsymbol{P}\boldsymbol{s} \tag{6.12}$$

is formed by the transmit data vector $\boldsymbol{s}$ and the precoding matrix $\boldsymbol{P}$.

WINNER II Channel

The WINNER II channel model [259] has been introduced in order to describe typical channel conditions in several environments in the frequency region of up to 6 GHz. Although the prototype of the multi-mode antenna system under investigation is designed for the frequency region of 6-8.5 GHz, the WINNER

channel model is assumed to behave close to the desired conditions. Concerning system modeling, the multi-mode antennas can be treated as different antennas, each providing a unique radiation pattern, at the same position as the other ports of the same physical radiator. Numerical results are reported subsequently for the non-line-of-sight (NLOS) scenario A1.

Kronecker Channel Model

The Kronecker channel model [260, 261] allows the simulation of correlation in MIMO system models. Commonly, the MIMO channel matrix $\tilde{\boldsymbol{H}}$ is modified such that a desired correlation between antenna ports is reached. The resulting channel can be written as

$$\boldsymbol{H} = \boldsymbol{R}_{\mathrm{rr}}\tilde{\boldsymbol{H}}\boldsymbol{R}_{\mathrm{tt}}. \tag{6.13}$$

$\boldsymbol{R}_{\mathrm{rr}}$ corresponds to the correlation between the antenna ports at the receive or user terminal, while $\boldsymbol{R}_{\mathrm{tt}}$ represents the correlation at the BS antenna array and its elements. Regarding multi-mode antennas, the Kronecker model is of interest to investigate the system performance, if the ports of a radiator cannot be separated due to coupling. The coupling may be due to suboptimal realizations or simpler implementation of feeding points and matching network. To model the coupling between the ports of a single multi-mode radiator, the correlation matrices $\boldsymbol{R}_{\mathrm{rr}}$ and $\boldsymbol{R}_{\mathrm{tt}}$ are set to be block diagonal. The blocks themselves are single-radiator correlation matrices $\boldsymbol{R}_{\mathrm{xx}}$ and $\boldsymbol{R}_{\mathrm{xy}}$, with the block matrices $\boldsymbol{R}_{\mathrm{xy}}$ being zero matrices. These are symmetric, e.g. $\boldsymbol{R}_{\mathrm{xx}} = \boldsymbol{R}_{\mathrm{xx}}^{T}$, where $(.)^{T}$ denotes transpose. The single-radiator correlation matrices are defined by

$$\boldsymbol{R}_{\mathrm{xx}} = \begin{bmatrix} r_{x,x} & r_{1,2} & \cdots & & r_{1,Np} \\ r_{2,1} & r_{x,x} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & & r_{Np-1,Np} \\ r_{Np,1} & \cdots & r_{Np,Np-1} & & r_{x,x} \end{bmatrix}, \tag{6.14}$$

where $r_{i,j}$ corresponds to the correlation coefficient of the ports i and j of the radiator. The matrix entries $r_{i,j}$ are defined by

$$r_{i,j} = \begin{cases} a_{xy}e^{j\phi_{i,j}} & \forall j > i, \\ \sqrt{1 - (N_p - 1)a_{xy}^2} & \forall j = i, \\ r_{j,i} & \forall j < i, \end{cases} \tag{6.15}$$

where a_{xy} results from the given correlation power, namely $a_{xy} = 10^{P_{\mathrm{cor}}/20}$. The correlation power between two ports is given by P_{cor} in decibel. The phase angle $\phi_{i,j}$ is drawn from a uniform distribution in the interval $[0, 2\pi)$. Entries in the principal diagonal are all real-valued, hence the phase angle is zero. Consequently, the lower triangular matrix apart the main diagonal is determined by the symmetry of the resulting single-radiator correlation matrix.

If the contribution of the coupling between neighboring antenna radiators on the data-rate shall be investigated, the form of the matrices $\boldsymbol{R}_{\text{tt}}$ and $\boldsymbol{R}_{\text{rr}}$ is different. Since by definition the correlation between different groups of modes on neighboring radiators is zero, the correlation coefficients need to be defined only to model the mutual coupling between the same group of modes. Each sub-block corresponding to the different antenna radiators $\boldsymbol{R}_{xy}$ then becomes a diagonal matrix. The matrices holding the diagonal of $\boldsymbol{R}_{\text{tt}}$ and $\boldsymbol{R}_{\text{rr}}$, $\boldsymbol{R}_{xx}$, are diagonal matrices according to $a_{xx}\boldsymbol{I}_{Np}$. The identity matrix $\boldsymbol{I}_{Np}$ is of size $N_{\text{p}} \times N_{\text{p}}$, representing the power kept at its origin port. Hence,

$$a_{xx} = \sqrt{1 - \sum_{y \in N_{\text{Ant}} N_{\text{p}} \backslash \{x\}} a_{xy}^2}. \tag{6.16}$$

Like in the aforementioned case for modeling the coupling of a single radiator, the correlation coefficients of the matrices $\boldsymbol{R}_{xy}$ are defined by an arbitrary phase and a given amplitude coefficient. In contrast to (6.15), the matrices $\boldsymbol{R}_{xy}$ are now diagonal. Since the matrices $\boldsymbol{R}_{\text{tt}}$ and $\boldsymbol{R}_{\text{rr}}$ need to be symmetric, $\boldsymbol{R}_{xy} = \boldsymbol{R}_{yx}$.

6.2.3   Definition of System Parameters

Towards system design and optimization, physically and mathematically based system parameters can be used for design and examination. These may include antenna parameters, take power allocation into account, or investigate the system performance in terms of achievable sum-rate. Subsequently, the fundamentals will be defined.

Gain

The gain of an antenna is a radiation performance measure. It defines the capability of any antenna to focus its radiated power to a certain angle. The radiated power at any point in the circumference is compared to a hypothetical ideally isotropic radiating antenna, given the same input power. If the radiation pattern consisting of the transversal electric field components E_ϕ and E_θ of the antenna fulfills

$$1 = \frac{1}{2Z_0} \oint |E_\phi(\phi, \theta)|^2 + |E_\theta(\phi, \theta)|^2 dA, \tag{6.17}$$

the gain pattern of any antenna can be calculated from its electric field components in the angular directions:

$$G(\phi, \theta) = \frac{4\pi}{2Z_0} \left(|E_\phi(\phi, \theta)|^2 + |E_\theta(\phi, \theta)|^2 \right). \tag{6.18}$$

The impedance of free space is given by $Z_0 = 120\pi\,\Omega$. An example is provided in Fig. 6.2, where the gain patterns of the 4-mode prototype antenna presented in [262] are shown [263].
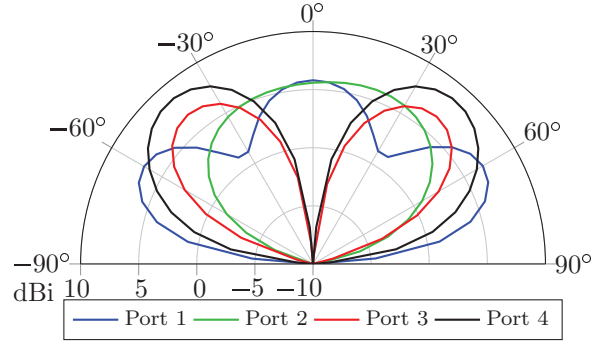
Figure 6.2
Gain patterns in y-z-plane of the 4-port prototype antenna introduced in [262].
From [263] © 2019 IEEE

Element Factor

As the gain pattern of a radiator (also called antenna element) defines the power radiated to any angle, in some cases the power of the sidelobes compared to the mainlobe are of interest. For this purpose, the gain pattern is normalized to the maximum gain, and the resulting pattern is dubbed element factor (EF):

$$EF(\phi, \theta) = \frac{G(\phi, \theta)}{\max\limits_{\phi, \theta}\{G(\phi, \theta)\}}. \tag{6.19}$$

Therefore, the element factor can be interpreted as a measure for the radiation efficiency in the desired direction, given a maximum radiation of one in the direction of the mainlobe.

Gain by Element Factor

Since the mentioned EF on the one hand provides a desired ratio for the comparison of the sidelobes, on the other hand it neglects the information of the size of the lobe. In order to conduct a pattern optimization of an antenna, a new parameter is required, which combines both, gain and element factor [263]. The resulting gain by element factor is defined by

$$GEF(\phi, \theta) = G(\phi, \theta) \cdot EF(\phi, \theta). \tag{6.20}$$

This newly introduced factor is helpful when optimizing the radiation of a single multi-mode antenna element.

Antenna and Array Factor (AAF)

If a set of antenna elements is combined to become an array, on one hand the gain of the elements still contributes to the array performance. On the other hand a major contribution of the array comes from the distribution of elements. The latter is commonly referred to as array factor. A generalization of the array factor is required to include the radiation pattern as well as precoding. To calculate the AAF for multi-mode antennas, the equation has to be generalized further, as is presented in [258]. There, the AAF is defined by

$$
|AAF|^2 = \left| \sum_{n_{\mathrm{T}}} \boldsymbol{x}_{n_{\mathrm{T}}} \cdot F_\theta^{n_{\mathrm{T}}}(\phi, \theta) \cdot e^{j2\pi \boldsymbol{k}(\phi,\theta) \cdot \boldsymbol{r}(n_{\mathrm{T}})} \right|^2 \\
+ \left| \sum_{n_{\mathrm{T}}} \boldsymbol{x}_{n_{\mathrm{T}}} \cdot F_\phi^{n_{\mathrm{T}}}(\phi, \theta) \cdot e^{j2\pi \boldsymbol{k}(\phi,\theta) \cdot \boldsymbol{r}(n_{\mathrm{T}})} \right|^2 .
\tag{6.21}
$$

Recall that the transmit data vector includes the precoding. The angle dependent electric field components of the radiation pattern in ϕ and θ direction are given by $F_\phi^{n_{\mathrm{T}}}(\phi, \theta)$ and $F_\theta^{n_{\mathrm{T}}}(\phi, \theta)$. The wave vector $\boldsymbol{k}(\phi, \theta) = \left[\cos(\phi)\cos(\theta), \sin(\phi)\cos(\theta), \sin(\phi) \right]$ represents the incident angle of any direction of the full sphere, while the position of the current antenna element and port n_{T} is defined by the position vector $\boldsymbol{r}(n_{\mathrm{T}})$ [258].

Equivalent Isotropic Radiated Power (EIRP)

The AAF represents the radiation pattern in a 3D sphere. In contrast to the gain of an antenna, the AAF includes neither the wave impedance nor the power normalization. Hence, for better comparison and fitting the requirements of given radiated power constraints, the EIRP needs to be regarded. The EIRP is defined by [258]

$$
EIRP(\phi, \theta) = \frac{4\pi}{2Z_0} \max_{\phi, \theta} \left\{ AAF(\phi, \theta) \right\}.
\tag{6.22}
$$

Note that when talking about a single multi-mode antenna element, the resulting gain of this element needs to be calculated taking the AAF into account. The multi-mode antenna itself can be regarded as an array, providing N_{p} ports, constructed on a single element. Using the definitions of AAF and EIRP, the parameter gain by element factor rather becomes an EIRP by AAF factor, but still holds the same information.

Sum-Rate

Given the baseband channel model as presented by (6.11), the channel capacity C taking channel matrix and precoding into account can be calculated as (see for example [264])

$$
C = \log_2 \det(\boldsymbol{I}_{N_{\mathrm{R}}} + SNR_0 \ \boldsymbol{H}\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}).
\tag{6.23}
$$

The rate can be reformulated as done in [258], where the contribution of the individual data streams to the data streams can be evaluated:

$$C = \sum_{k=1}^{N_{\mathrm{S}}} \log_2(1 + SINR_k(\boldsymbol{P})). \tag{6.24}$$

The signal-to-interference-plus-noise ratio of the k-th channel is defined by [258]

$$SINR_k(\boldsymbol{P}) = \frac{|\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_k|^2}{1 + \sum_{i \neq k} |\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{p}_i|^2}, \tag{6.25}$$

where $\boldsymbol{h}_k$ and $\boldsymbol{p}_k$ correspond to the k-th row and column of the channel and precoding matrix, respectively. If no interference occurs, e.g. by using a zero forcing (ZF) or block diagonalization (BD) based transmission scheme, the sum-rate of the system can be calculated by neglecting the interference taken into account by the sum in (6.25). Note that in the above equations a decoding matrix is not included.

Power Constraints

When comparing different system proposals, usually the data-rate of the system given a defined transmit power is evaluated. The transmit power of the system can be either defined as a per-user power or sum-power. If the user-power is selected, each column of the precoding matrix, representing a single channel, has the same absolute value. However, commonly the sum-power P_{Tx} is assumed to be set. Then, given a precoding matrix $\boldsymbol{P}$ including power allocation

$$P_{\mathrm{Tx}} \geqslant \|\boldsymbol{P}\|^2 \tag{6.26}$$

needs to be fulfilled.

6.3 Systematic Multi-Mode Antenna Design Based on Symmetry Properties of Characteristic Modes

The challenge that arises in multi-mode antenna design for MIMO applications is to excite mutually exclusive sets of characteristic modes in order to realize uncorrelated antenna ports. In this section, systematic design guidelines are presented which answer the following fundamental questions:

- What antenna geometry should be used?

- How many ports can be realized?

- What is the optimal port configuration?

In order to underline the importance of such guidelines, an introductory example is considered [265].
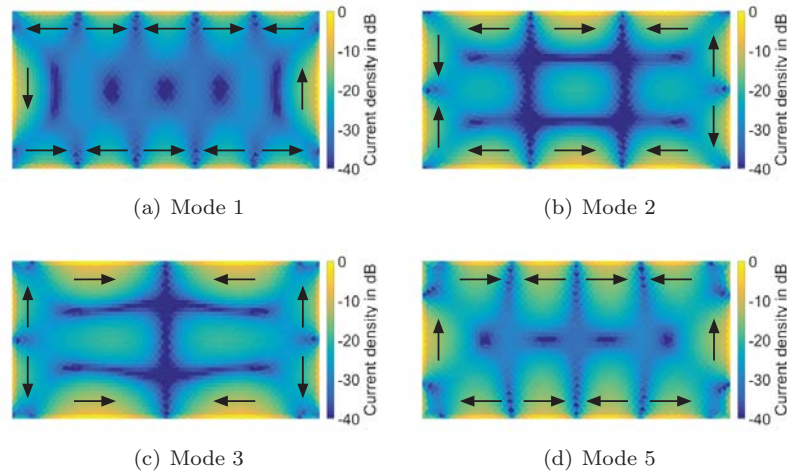
(a) Mode 1

(b) Mode 2

(c) Mode 3

(d) Mode 5

Figure 6.3
Modal analysis of rectangular PEC plate at 7.25 GHz. From [265] © 2018 IEEE
(a)-(d) Normalized surface current densities of four significant characteristic modes.

6.3.1 Modal Analysis of Electrically Large Rectangular Plate

[1]   A rectangular PEC plate of dimensions 120 mm×60 mm (coordinate system as in Fig. 6.1(a)) is used as a simple model for the chassis of a typical modern smartphone (mobile terminal). At the center frequency of the considered frequency range (7.25 GHz), it appears electrically large (approximately $3\lambda \times 3\lambda/2$). As a consequence, the characteristic mode analysis yields 28 significant modes. Apparently, there seems to be a huge potential for creating uncorrelated antenna ports.

However, a closer inspection of the corresponding characteristic surface current densities reveals that the significant modes exhibit rather complex current distributions. Four examples are shown in Fig. 6.3(a)-(d). In particular, the current densities have several local maxima and minima where coupling elements could be placed in order to excite the modes, as explained in Section 6.2.1. The principal difficulty is thus to define a suitable port configuration.

An exemplary port configuration for a smaller rectangular plate has already been introduced in Fig. 6.1. It was derived directly from inspecting the current distributions. It is noted that the four presented characteristic surface current densities (as well as the corresponding antenna ports) have different

---

[1]This section and the figures therein are taken from [265].

Table 6.1
Sorting of significant modes of electrically large rectangular PEC plate according to their symmetry properties. From [265] © 2018 IEEE

Set	Modes	Symmetry	Port
1	3, 8, 14, 18, 22, 24, 27	Rotation about x-, y- and z-axis	1
2	1, 7, 10, 12, 17, 20, 26	Rotation about z-axis	2
3	2, 4, 6, 13, 15, 21, 23	Rotation about x-axis	3
4	5, 9, 11, 16, 19, 25, 28	Rotation about y-axis	4
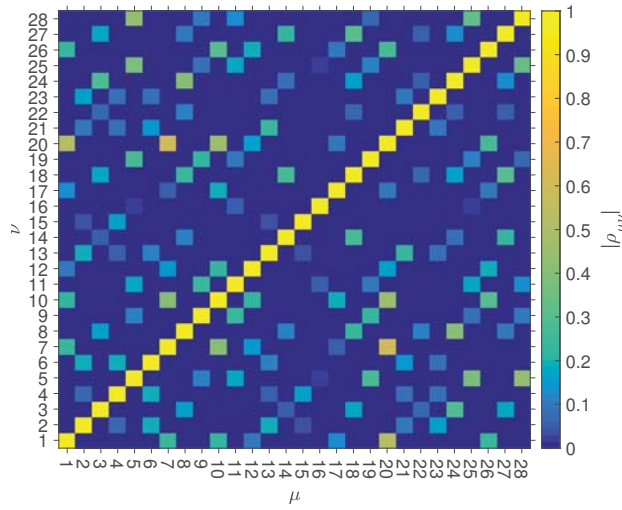
Figure 6.4
Characteristic current correlation of electrically large rectangular PEC plate at 7.25 GHz. From [265] © 2018 IEEE

symmetry properties. Correspondingly, it is possible to sort the characteristic surface current densities of the electrically large rectangular plate into four sets according to their symmetry properties (cf. Fig. 6.3(a)-(d)). This is summarized in Table 6.1. Calculating the correlation coefficients $\rho_{\mu\nu}$ of the characteristic surface current densities according to

$$\rho_{\mu\nu} = \frac{\langle \mathbf{J}_\mu\left(\mathbf{r}'\right), \mathbf{J}_\nu\left(\mathbf{r}'\right)\rangle}{\|\mathbf{J}_\mu\left(\mathbf{r}'\right)\| \, \|\mathbf{J}_\nu\left(\mathbf{r}'\right)\|} \tag{6.27}$$

yields that current densities belonging to different sets, i.e. having different symmetry properties, are orthogonal to each other, whereas those belonging to the same set are correlated to some extent. The characteristic current correlation coefficients are shown in Fig. 6.4.

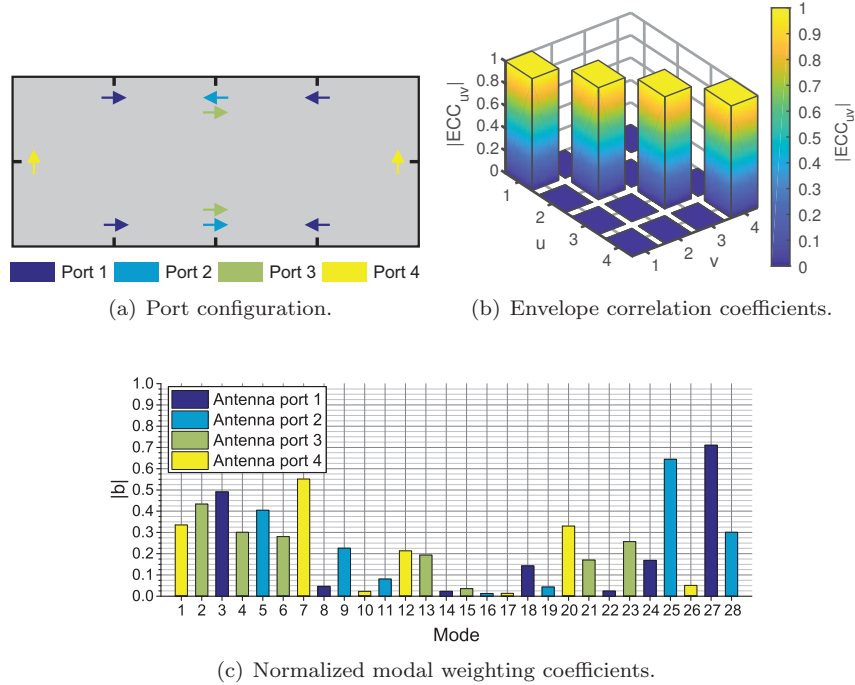From these symmetry properties, and bearing in mind (6.3), it is deduced

(a) Port configuration.

(b) Envelope correlation coefficients.

(c) Normalized modal weighting coefficients.

Figure 6.5
Excitation of electrically large rectangular PEC plate at 7.25 GHz. From [265]
© 2018 IEEE

that the antenna ports should possess the same symmetry properties in order
to be correlated with the current densities of the desired set and orthogonal
to the current densities of the other sets. In other words, the arrangement of
the antenna ports will reflect the symmetry of the characteristic modes which
are to be excited.

Thus, the port configuration as shown in Fig. 6.5(a) is now applied to
the electrically large plate. The feed points, which are modeled as voltage
gap sources in the method of moments [256], are placed at the edges of the
plate and are depicted by the small black strips in Fig. 6.5(a). The colors
denote which feed points have to be driven simultaneously in order to form an
antenna port and the arrows denote the relative phase between feed points.
The resulting symmetry properties of the ports are also listed in Table 6.1.

In order to analyze which characteristic modes are excited by the antenna
ports, the normalized modal weighting coefficients as defined in (6.9) are used.
Their absolute values are plotted in Fig. 6.5(c). Every mode is excited by
only one antenna port or, in other words, every port excites a different set
of significant characteristic modes which means that the antenna ports are
uncorrelated. This is confirmed by Fig. 6.5(b) which depicts the envelope
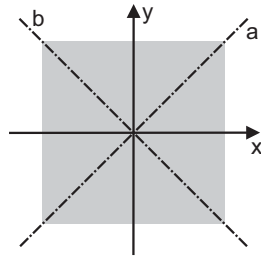
(a) Square plate with coordinate system and symmetry axes.

Symbol	Operation
E	Identity
C_{4z}	Rotation by 90° about the z-axis
C_{4z}^{-1}	Rotation by −90° about the z-axis
C_{2z}	Rotation by 180° about the z-axis
C_{2x}	Rotation by 180° about the x-axis
C_{2y}	Rotation by 180° about the y-axis
C_{2a}	Rotation by 180° about the diagonal a
C_{2b}	Rotation by 180° about the diagonal b

(b) Symmetry operations of group D_4.

Figure 6.6
Symmetry group of square plate. From [249] © 2019 IEEE

correlation coefficients (ECC) calculated according to (6.6). Moreover, the excited sets are exactly the mode sets identified in Table 6.1 and thus all 28 significant characteristic modes are already in use, leaving no further degree of freedom to realize a fifth uncorrelated antenna port.

A three-port prototype realization based on a rectangular plate is presented in [266]. The observations regarding the rectangular plate demonstrate that the correlation of the characteristic currents and the realization of uncorrelated antenna ports are closely related to their symmetry properties. This fundamental relationship is thoroughly analyzed in the following section.

6.3.2 Symmetry Analysis of Characteristic Modes

[2] The fundamental connection between the theory of characteristic modes and the theory of symmetry was established in [267], where it was demonstrated that the characteristic surface current densities are basis functions of the irreducible representations of the symmetry group of the underlying PEC object (antenna). As demonstrated in [249], this connection can be used purposefully for multi-mode antenna design.

The symmetry group of an antenna is the group of geometric transformations (symmetry operations, e.g. rotations) that leave the antenna invariant. As an example, the symmetry group of an infinitely thin square plate as shown in Fig. 6.6(a) is called D_4 (Schoenfliess notation) and it consists of the eight operations listed in Table 6.6(b) [268].

The irreducible representations of the symmetry group describe how the basis functions, i.e. the characteristic surface current densities, transform un-

---

[2]This section and the figures therein are taken from [249].

Table 6.2

Matrix representations of symmetry group D_4. From [249] © 2019 IEEE

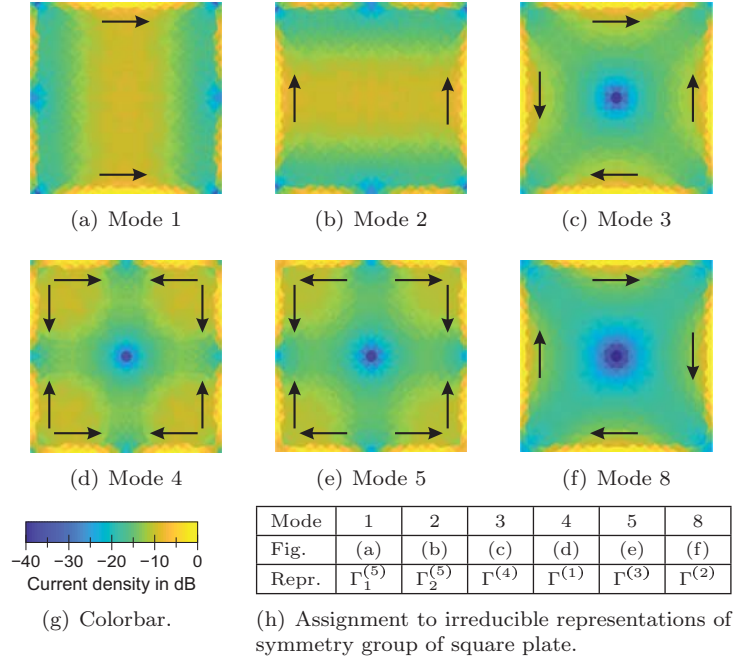| D_4	E	C_{4z}	C_{4z}^{-1}	C_{2z}
$\Gamma^{(1)}$	1	1	1	1
$\Gamma^{(2)}$	1	1	1	1
$\Gamma^{(3)}$	1	−1	−1	1
$\Gamma^{(4)}$	1	−1	−1	1
$\Gamma^{(5)}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$
D_4	C_{2x}	C_{2y}	C_{2a}	C_{2b}
$\Gamma^{(1)}$	1	1	1	1
$\Gamma^{(2)}$	−1	−1	−1	−1
$\Gamma^{(3)}$	1	1	−1	−1
$\Gamma^{(4)}$	−1	−1	1	1
$\Gamma^{(5)}$	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$

der these symmetry operations [269]:

$$P\left(T\right)\mathbf{J}_n^{(p)}\left(\mathbf{r}'\right) = \sum_{m=1}^{d_p} \Gamma_{nm}^{(p)}\left(T\right)\mathbf{J}_m^{(p)}\left(\mathbf{r}'\right), \quad n = 1, 2, \ldots d_p. \tag{6.28}$$

$P\left(T\right)\mathbf{J}_n^{(p)}$ denotes a transformed characteristic surface current density due to the operation T of the symmetry group (where P is called transformation operator). This transformed current density can be expressed as a linear combination of d_p characteristic surface current densities $\mathbf{J}_m^{(p)}$ (hence the term basis functions). The weighting coefficients $\Gamma_{nm}^{(p)}(T)$ are elements of a d_p-dimensional square matrix $\mathbf{\Gamma}^{(p)}(T)$.

The matrix $\mathbf{\Gamma}^{(p)}(T)$ is called the representation matrix of the p-th representation for the symmetry operation T. A d_p-dimensional representation is a group homomorphic to the symmetry group that assigns a square matrix of dimension d_p to each element of the symmetry group [270]. For a symmetry group with a finite number of operations such as D_4, there is a finite number of so-called irreducible representations of minimal dimension [269]. The irreducible representations of a finite group are unique up to a similarity transformation. It is thus purposeful to work with these irreducible representations.

The symmetry group of the square plate has five irreducible representations. The representation matrices are given in Table 6.2 [269]. The first four representations $\Gamma^{(1;2;3;4)}$ are one-dimensional ($d_{1;2;3;4} = 1$). The corresponding representation matrices $\mathbf{\Gamma}^{(1;2;3;4)}$ are scalars with values 1 or −1. A single characteristic surface current density forms a basis function for such a represen-

(a) Mode 1   (b) Mode 2   (c) Mode 3

(d) Mode 4   (e) Mode 5   (f) Mode 8

−40 −30 −20 −10 0
Current density in dB

(g) Colorbar.

Mode	1	2	3	4	5	8
Fig.	(a)	(b)	(c)	(d)	(e)	(f)
Repr.	$\Gamma_1^{(5)}$	$\Gamma_2^{(5)}$	$\Gamma^{(4)}$	$\Gamma^{(1)}$	$\Gamma^{(3)}$	$\Gamma^{(2)}$

(h) Assignment to irreducible representations of symmetry group of square plate.

Figure 6.7
Normalized surface current densities of fundamental characteristic modes of square PEC plate with 45 mm edge length at 7.25 GHz. From [249] © 2019 IEEE
Principal current directions denoted by arrows. (a)-(f) Modes 1 to 5 and 8.

tation. The fifth representation $\Gamma^{(5)}$, in contrast, is two-dimensional ($d_5 = 2$) and consists therefore of two-dimensional matrices $\mathbf{\Gamma}^{(5)}$. This has the consequence that a pair of degenerate characteristic surface current densities forms a basis for this representation.

Fig. 6.7 shows six characteristic surface current densities of a square plate with 45 mm edge length evaluated at 7.25 GHz and sorted by their significance. By applying the symmetry operations of the square plate, they can be assigned to the irreducible representations. Mode 4 (Fig. 6.7(d)), for example, is invariant under all symmetry operations. This corresponds to a multiplication of the current density with 1 and it thus forms a basis function of the first representation $\Gamma^{(1)}$ according to (6.28) and Table 6.2. Mode 8 (Fig. 6.7(f)), in contrast, is invariant under all rotations about the z-axis, but is inverted by the rotations about the other axes (multiplication with −1). Therefore, it belongs to the second representation $\Gamma^{(2)}$.

The modes 1 and 2 (Fig. 6.7(a) and Fig. 6.7(b)) are degenerate, i.e. they have the same eigenvalue independent of frequency. Hence, they form a basis for the two-dimensional fifth representation $\Gamma^{(5)}$. Their transformations

are expressed by two-dimensional matrices (Table 6.2). Mode 1 transforms according to the first row of the matrices ($n = 1$) and mode 2 transforms according to the second row ($n = 2$). For instance, a rotation by 180° about the x-axis (C_{2x}) leaves mode 1 invariant, but inverts mode 2. In particular, the current densities transform into each other for some operations. This is e.g. the case for the rotation (of the coordinate system) by 90° about the z-axis (C_{4z}), where mode 1 becomes the inverted mode 2 and mode 2 becomes mode 1. (This is equal to a rotation of the current densities themselves by $-90°$.) The assignment of the six characteristic modes to the irreducible representations is summarized in Table 6.7(h).

The importance of the irreducible representations lies in their orthogonality properties. With respect to the basis functions, the following orthogonality theorem holds [269]:

$$\left\langle \mathbf{J}_m^{(p)}\left(\mathbf{r}'\right), \mathbf{J}_n^{(q)}\left(\mathbf{r}'\right) \right\rangle = 0, \quad p \neq q \quad \vee \quad m \neq n. \tag{6.29}$$

It states that characteristic surface current densities belonging to different representations p and q are orthogonal to each other, and those that form a degenerate set ($m \neq n$) of a multi-dimensional representation are orthogonal to each other. This has the consequence that the six characteristic surface current densities in Fig. 6.7 are orthogonal to each other. Furthermore, all other characteristic surface current densities (higher-order modes) also belong to one of the irreducible representations and are thus correlated to one of the six modes shown. These six modes are called the fundamental modes per representation as they are the first modes to become significant over frequency.

The results of the introductory example in Section 6.3.1 can now be explained. The symmetry group of the rectangular plate is D_2 [268]. It has four one-dimensional representations. Four mutually orthogonal sets of characteristic modes were found (Table 6.1, Fig. 6.4), which exactly correspond to the irreducible representations. In particular, antenna ports were defined that can also be considered as basis functions of the irreducible representations. Therefore, the orthogonality theorem in (6.29) holds, and it explains why the ports only excite a subset of modes and are thus uncorrelated (Fig. 6.5). Moreover, more than four uncorrelated antenna ports cannot be realized as there are only four one-dimensional representations.

Based on the orthogonality theorem for representations, design guidelines for realizing uncorrelated antenna ports on symmetric multi-mode antennas can be deduced. With regard to the questions asked at the beginning of this subsection, the following results are highlighted:

- Symmetric antenna geometries offer the potential to realize perfectly uncorrelated antenna ports.

- The maximum number of uncorrelated antenna ports is dictated by the number and the dimensions of the irreducible representations.

(a) Port configuration.

(b) Envelope correlation coefficients.

(c) Normalized modal weighting coefficients.
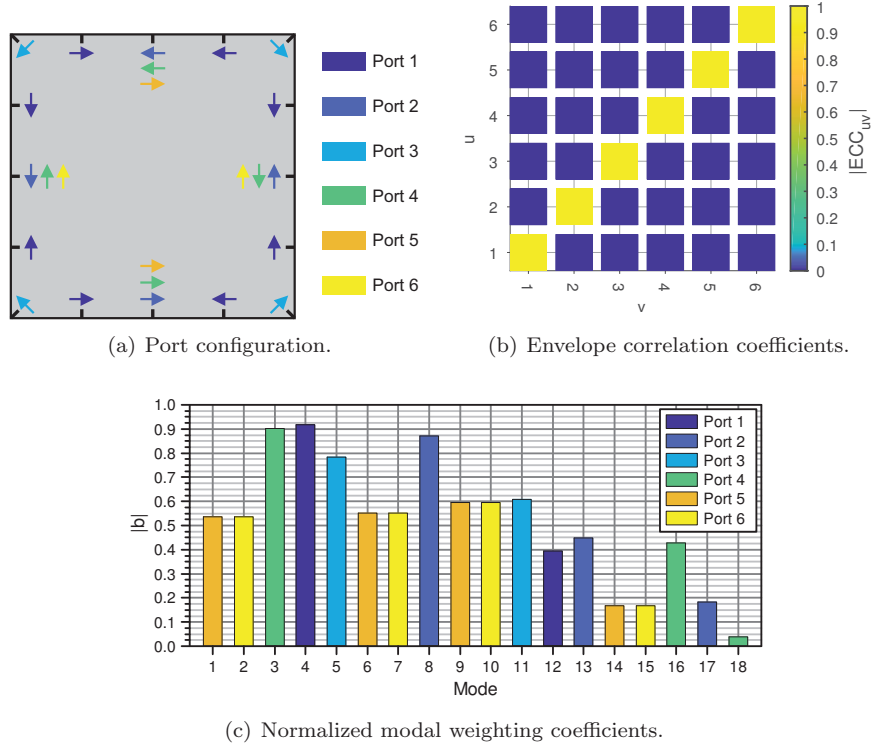
Figure 6.8
Excitation of square PEC plate at 7.25 GHz. From [249] © 2019 IEEE

- Antenna ports have to fulfill the symmetry requirements of the representations.

These guidelines are now applied to the square plate (Fig. 6.6). It has already been shown that the square plate offers six mutually orthogonal sets of characteristic modes since there are four one-dimensional and one two-dimensional representations. It is thus deduced that six uncorrelated antenna ports can be realized. The ports have to be designed as basis functions of the irreducible representations. One possible port configuration using the minimum number of feed points is shown in Fig. 6.8(a). All six ports are designed in such a way that they transform according to Table 6.2.

The inspection of the modal weighting coefficients in Fig. 6.8(c) reveals that every port excites a different set of characteristic modes, as expected. These are the modes that belong to the same representation or the same row of the two-dimensional representation, respectively, as the exciting port. As a result, the ports are uncorrelated, as demonstrated in Fig. 6.8(b).

The design guidelines derived in this section can be applied to arbitrary

Table 6.3
Symmetry group, number of irreducible representations, and number of feasible uncorrelated antenna ports for different symmetric antenna geometries. From [249] © 2019 IEEE

Antenna geometry	Symm. group	No. of irred. repr.			No. of un-corr. ports
		1-D	2-D	3-D	
Isosceles triangular plate	C_2	2	0	0	2
Rectangular plate	D_2	4	0	0	4
Rectangular pyramid	C_{2v}	4	0	0	4
Equilateral triangular plate	D_3	2	1	0	4
Regular triangular pyramid	C_{3v}	2	1	0	4
Square plate	D_4	4	1	0	6
Square pyramid	C_{4v}	4	1	0	6
Regular hexagonal plate	D_6	4	2	0	8
Regular hexagonal pyramid	C_{6v}	4	2	0	8
Rectangular cuboid	D_{2h}	8	0	0	8
Regular triangular prism	D_{3h}	4	2	0	8
Regular tetrahedron	T_d	2	1	2	10
Square cuboid	D_{4h}	8	2	0	12
Regular hexagonal prism	D_{6h}	8	4	0	16
Cube	O_h	4	2	4	20

symmetric antennas. Numerous examples are given in Table 6.3. In particular, a thorough analysis of a hexagonal plate is conducted in [271], yielding a concept for an eight-port multi-mode antenna.

6.3.3 Multi-Mode Antenna Design and Array Arrangement

[3] Based on the guidelines from the previous subsection, a prototype antenna element for arrangement in a massive MIMO array is designed. The square plate is used as the basic geometry. The antenna size is determined by the fact that only the six fundamental modes (Fig. 6.7) need to be significant in the desired frequency range from 6 GHz to 8.5 GHz. Based on a modal analysis, the minimum edge length such that this criterion is fulfilled is found to be 45 mm. In the same manner, the minimum height above the ground plane is found to be 16 mm [272].

The schematic of an ultra-wideband six-port multi-mode antenna element is shown in Fig. 6.9 [272]. Slots are used as excitation elements. Inspired by the feed point positions in Fig. 6.8(a), the slots are placed symmetrically on the antenna element. Thus, although the antenna has been modified, the

---

[3]This section and the figures therein are taken from [272] and [262].
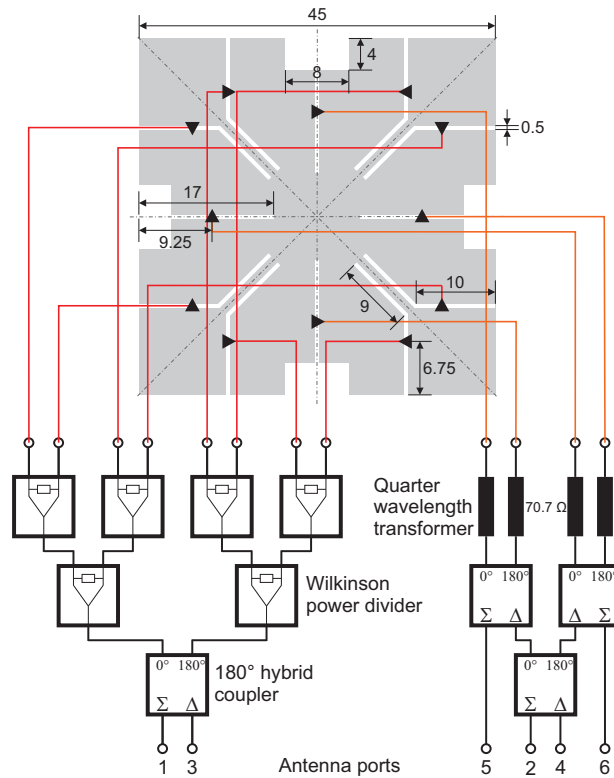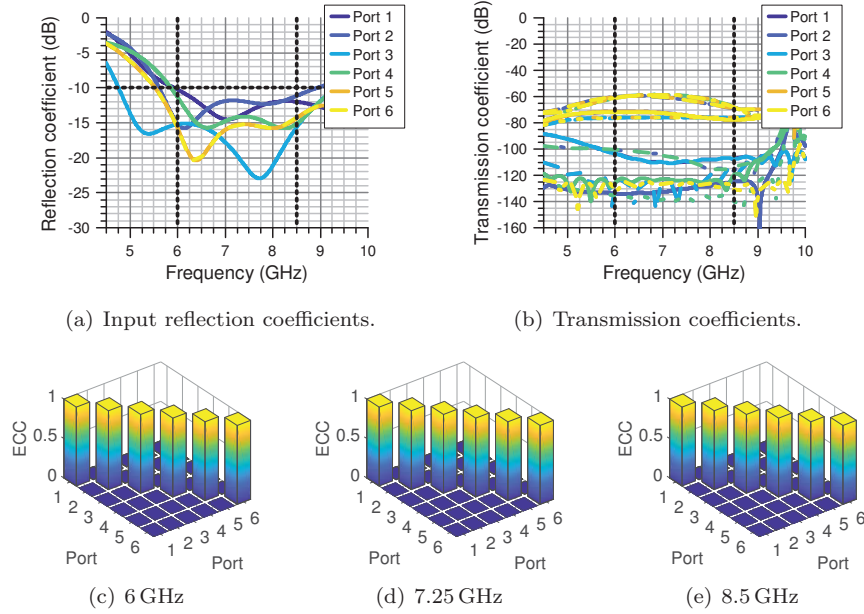
Figure 6.9
Square PEC plate with stepped excitation slots and feed network. The antenna is located 16 mm above a ground plane. Dimensions in mm. From [272] © 2019 IEEE

square symmetry is preserved, such that the orthogonality of the characteristic surface current densities still holds.

The slots are used to control the input impedance of the antenna ports by adjusting their lengths and widths as well as the positions of the feed points within the slots [262]. The feed points within the slots are marked by black arrows in Fig. 6.9 denoting the phase reference. A step in width is introduced for the center slots in order to increase the impedance bandwidth by smoothing the impedance curves.

The center slots are used for ports 2, 4, 5 and 6 (cf. Fig. 6.8(a)), whereas the off-center slots are used for ports 1 and 3. Thus, two feed networks are needed to distribute the input signals of the antenna ports to the corresponding feed points with the correct amplitude and phase relations as required by the irreducible representations. The feed networks used here consist of Wilkinson power dividers and 180° Hybrid couplers with 50 Ω reference impedance [249].

(a) Input reflection coefficients.

(b) Transmission coefficients.

(c) 6 GHz
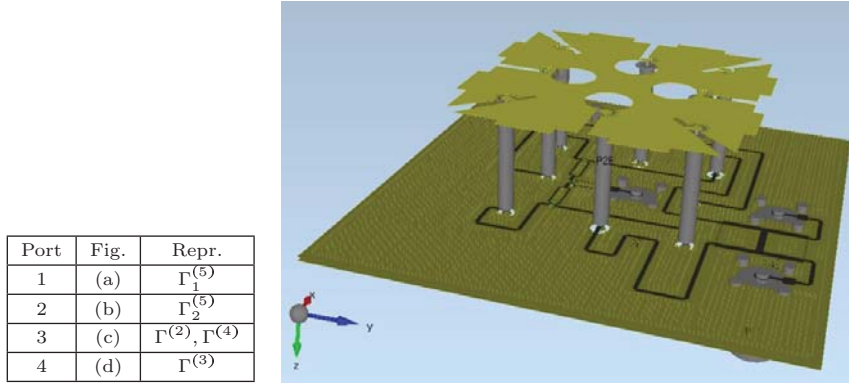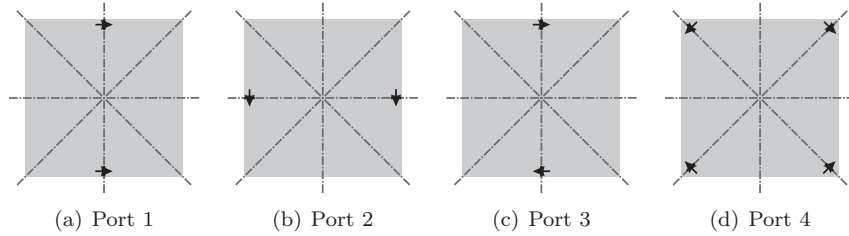

(d) 7.25 GHz

(e) 8.5 GHz

Figure 6.10
Simulation results of square PEC plate with stepped excitation slots and feed
network located 16 mm above ground plane as shown in Fig. 6.9. From [272]
© 2019 IEEE
(c)-(e) Envelope correlation coefficients.

Additionally, quarter wavelength transformers are employed in order to adjust
the impedance levels.

The simulation results of this antenna element are shown in Fig. 6.10. Re-
flection coefficients of less than -10 dB are achieved for all six antenna ports
over the desired ultrawide bandwidth from 6 GHz to 8.5 GHz (Fig. 6.10(a)).
Furthermore, the ports are highly uncoupled (Fig. 6.10(b)) and uncorre-
lated (Fig. 6.10(c)-(e)), which is due to the fact that the symmetry of the
antenna element is preserved and the decoupling of the feed network is as-
sumed to be perfect.

This example demonstrates that an ultra-wideband six-port multi-mode
antenna element can be realized based on the design guidelines derived in the
previous subsection. In the next step, multi-mode antennas are to be imple-
mented in a massive MIMO antenna array prototype (Multi-mode massive
MIMO). Fig. 6.9 shows that the feed network is rather complex if all six an-
tenna ports are to be realized. For the array implementation, the six-port
antenna is reduced to a four-port antenna. Port 1 is not realized. In this
case, port 3 can be designed with only four feeding slots on the diagonals
of the square plate. Furthermore, ports 2 and 4 are merged. This has the

(a) Port 1   (b) Port 2   (c) Port 3   (d) Port 4

Port	Fig.	Repr.
1	(a)	$\Gamma_1^{(5)}$
2	(b)	$\Gamma_2^{(5)}$
3	(c)	$\Gamma^{(2)}, \Gamma^{(4)}$
4	(d)	$\Gamma^{(3)}$

(e) Assignment to irre-
ducible representations of
square plate.

(f) Simulation model with feed network in single-layer
stripline technology.

Figure 6.11
Four-port multi-mode antenna.
(a)-(d) Port configuration.

advantage that the resulting port can be realized by driving port 5 or 6 dif-
ferentially. The new port configuration is shown in Fig. 6.11(a)-(d). As an
additional advantage, the antenna size can be reduced to 40 mm. Moreover,
the feed network is greatly simplified. It consists of only one hybrid coupler
and four Wilkinson power dividers and can be realized on a single stripline
layer. The complete multi-mode antenna model is shown in Fig. 6.11(f) [262].
The connection between the feed network and the antenna element is realized
by means of semi-rigid coaxial cables. Rogers RO4003 is used as substrate
material for both antenna element and feed network.

This antenna element is now arranged in an 11 × 11 planar antenna ar-
ray (Fig. 6.12), resulting in a total of 484 antenna ports. The spacing between
the different antenna elements is chosen such that the mutual coupling to
neighboring elements within the array is less than −20 dB. In addition, the
isolation between neighboring elements within the array is improved by using
choke walls with a height of a quarter wavelength at the center frequency be-
tween the elements (see Fig. 6.12(b)). The choke walls are realized by vertical
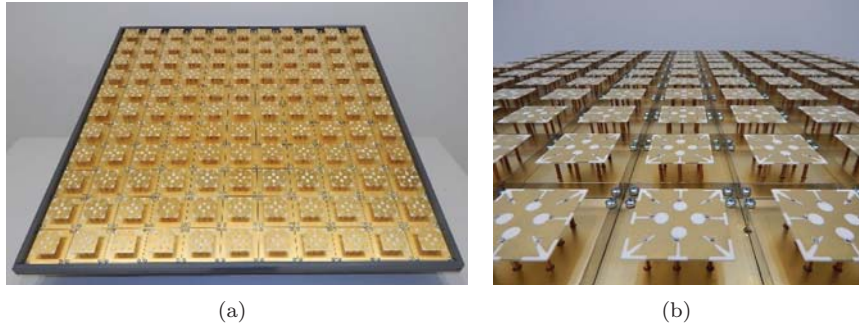
(a) (b)

Figure 6.12
Prototype of an 11 × 11 array of four-port multi-mode antennas. From [262]
© 2016 IEEE

(a) S-parameters of center element.

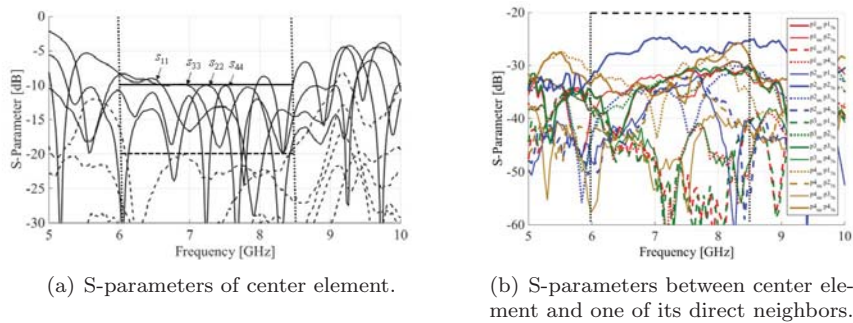(b) S-parameters between center element and one of its direct neighbors.

Figure 6.13
S-parameters of multi-mode massive MIMO array. From [262] © 2016 IEEE

metal strips located between the elements. The separation distance between the elements of the array is 0.58 wavelengths at the center frequency.

The S-parameters of the center element as shown in Fig. 6.13(a) are measured using a four-port network analyzer. It can be observed that the matching of all four antenna ports is less than −10 dB in almost all of the desired frequency band from 6 GHz to 8.5 GHz. The mutual coupling between the antenna ports is less than −20 dB in the entire frequency band. In fact, it is even much better in most parts of the band. Fig. 6.13(b) shows the measured mutual coupling between the center element (element index 66) and one of its direct neighboring elements (element index 76). It can be observed that the mutual coupling between any of the ports is less than −25 dB and therefore fulfills the requirement. Compared to a generic array of crossed dipoles achieving the same decoupling without further measures, the proposed multi-

Table 6.4
Parameters of Multi-Mode Massive MIMO Arrays (λ: Wavelength at center frequency). From [252] © 2019 IEEE

Array	Array of four-port multi-mode antennas	Array of crossed dipoles
Max. inter-element coupling	−25 dB	−25 dB
Number of ports	484	484
Ports per element	4	2
Element size	$0.72\lambda^2$	$0.25\lambda^2$
Inter-element spacing	0.58λ	λ
Total array size	$247\lambda^2$	$544\lambda^2$

mode massive MIMO array offers an overall size reduction of 54 %. Table 6.4 summarizes the important array parameters.

6.4   Baseband Processing System Design

[4] The baseband system discussed in this section is tailored to fit and employ the features provided by the proposed antenna type. Recall that the prototype multi-mode antenna is designed such that transmission is possible in a wide frequency region, as shown in Sec. 6.3. Therefore the system is designed to employ ultra-wideband communications. Because of the broad bandwidth of 2.5 GHz in this concept, a multi-carrier system approach is of interest. As a state-of-the-art and widely applied technique, orthogonal frequency division multiple access (OFDMA) schemes are assumed.

6.4.1   System Architecture of Baseband Signal Processing

In Fig. 6.14 a block diagram of the transmitter under investigation is shown. The different transmit data vectors $\boldsymbol{s}_f$ contain the data symbols of the f-th subcarrier. The digital precoding matrices $\boldsymbol{D}_f$ describes the digital part of the precoding of the f-th subcarrier. The outputs of the digital precoding are connected by N_{RF} radio frequency (RF) chains to an analog precoding matrix $\boldsymbol{A}$. An RF chain includes the digital to analog conversion, filtering and mixing of the signals to the desired frequency band. The analog precoding is assumed to be constant in frequency, hence it contributes equally to any subcarrier.

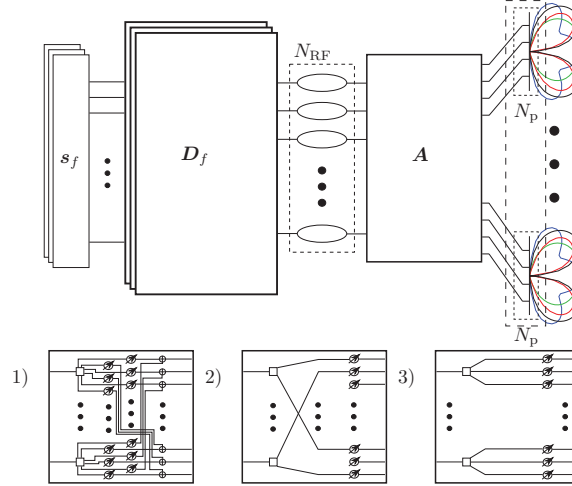[4]This section and the figures therein are taken from [252].

Figure 6.14

Block diagram of the transmitter. From [252] © 2019 IEEE

N_{RF} denotes the number of available RF chains, N_{Ant} is the number of antenna elements, N_{p} corresponds to the number of ports at the multi-mode antenna element. $\boldsymbol{x}_f$ is the data vector of the f-th sub-carrier, $\boldsymbol{D}_{\mathrm{f}}$ the digital precoding matrix of the corresponding sub-carrier. $\boldsymbol{A}$ represents the analog precoding. Inset 1) illustrates the fully-connected approach, where each RF chain is connected to each antenna port by means of a phase shifter. A spatial filtering setup, where a number of ULAs is formed by selecting the same mode is represented in inset 2), while 3) represents single-element beamforming.

Consequently, the transmit vector $\boldsymbol{x}_f$ of subcarrier f can be written as

$$\boldsymbol{x}_f = \boldsymbol{A}\boldsymbol{D}_f\boldsymbol{s}_f. \tag{6.30}$$

6.4.2 Joint Precoding and Power Control

[5] The transmit power P_{S} is obtained by an integration over the whole sphere of the absolute squared AAF. Please note that the transmitted sum-power P_{S}

---

[5]This section and the figures therein are taken from [258].

of the system is smaller or equal to the EIRP:

$$P_{\mathrm{S}} = \frac{1}{2Z_{\mathrm{F0}}} \int\limits_{\phi=0}^{2\pi} \int\limits_{\theta=0}^{\pi} \left(|AAF(\theta,\phi)|^2 \right) \sin(\theta)\,\mathrm{d}\theta\,\mathrm{d}\phi \qquad (6.31)$$

$$\leqslant \frac{1}{2Z_{\mathrm{F0}}} \int\limits_{\phi=0}^{2\pi} \int\limits_{\theta=0}^{\pi} \left(\max_{\theta,\phi} \left\{ |AAF(\theta,\phi)|^2 \right\} \right) \sin(\theta)\,\mathrm{d}\theta\,\mathrm{d}\phi \qquad (6.32)$$

$$= \frac{4\pi}{2Z_{\mathrm{F0}}} \max_{\theta,\phi} \left\{ |AAF(\theta,\phi)|^2 \right\} = \mathrm{EIRP}. \qquad (6.33)$$

This integration can be split-up in the two integrations over the azimuth angle ϕ and the elevation angle θ. Finally, the value is divided by $2Z_{\mathrm{F0}}$. Equation (6.31) implies that for isotropic radiation the EIRP is equal to the sum-power and leads to a theoretical maximum $\mathrm{SNR}_0 = \frac{\mathrm{EIRP}}{N_0}$. For directive radiation of the transmitter, the sum-power of the system will be smaller than the EIRP and therefore only a fraction of SNR_0 is used.

The constraint on the EIRP of the transmitter is given by

$$\frac{4\pi}{2Z_{\mathrm{F0}}} \cdot \max_{\theta,\phi} \left\{ |AAF(\theta,\phi,\boldsymbol{W})|^2 \right\} \leqslant \mathrm{EIRP}_{\mathrm{lim}} \qquad (6.34)$$

and the problem formulation for the maximization of the sum-rate given the constrained EIRP can be written as a conventional minimization problem with one inequality constraint:

$$
\begin{aligned}
&\underset{\boldsymbol{W} \in \mathbb{C}^{N_{\mathrm{T}} \times N_{\mathrm{R}}}}{\text{minimize}} \quad -\sum_{k=1}^{N_{\mathrm{S}}} \log_2 \left(1 + \mathrm{SINR}_k(\boldsymbol{W}) \right) \\
&\text{subject to} \quad \mathrm{EIRP}_{\mathrm{lim}} - \frac{4\pi}{2Z_{\mathrm{F0}}} \max_{\theta,\phi} \left\{ |AAF(\theta,\phi,\boldsymbol{W})|^2 \right\} \geqslant 0.
\end{aligned}
\qquad (6.35)
$$

So far, we have limited the EIRP in spatial domain, separately for each sub-channel. Actually, in several standards the EIRP limit is averaged over time (usually about 1 ms) and frequency (usually several MHz). In this contribution, we assume that each sub-channel has to fulfill the EIRP constraint, which means that the optimization does not include any averaging over time or frequency. It has to be noted that the constraint is a continuous function, which represents an infinite number of constraints for numerical optimization techniques. In the numerical results of this contribution, the characteristic in three-dimensional space is uniformly sampled in θ and ϕ with a spacing of 1°. Therefore, it is possible that the largest radiated peak of the characteristic can lay between two samples. In that case, the actual EIRP could be larger than the determined value by sampling. However, the lack of analytic descriptions of most practical antennas complicates or even prohibits an analytic optimization or a closed-form solution.

Beamforming Strategy Survey

A. Sum-Power-Limited Systems

The given problem formulation is related to sum-power-limited formulations in many multi-user MIMO schemes, where an optimization of the SINRs of the users is considered. When a large fraction of the transmitted power is assigned to one user, the signal of the user will act as a strong interferer for the other users. So the available data-rates improve for one user, while getting worse for other users. To overcome this problem, different strategies have been proposed in the context of a constrained sum-power. One possible strategy specifies minimum SINRs for each of the users as a quality of service (QoS) measure. This results in the so-called "max-min" SINR problem, that maximizes the minimum SINR in the system, while limiting the overall transmit power [273].

When the system data throughput is concerned, another strategy optimizes the power allocation together with matched filter (MF) precoding or zero-forcing (ZF) precoding [274]. These techniques will be referred to as conventional solutions in the rest of this contribution. Without normalization, the formal description of these well known techniques is given by

$$\tilde{\boldsymbol{W}}_{\mathrm{MF}} = \boldsymbol{H}^{\mathsf{H}} \tag{6.36}$$

$$\tilde{\boldsymbol{W}}_{\mathrm{ZF}} = \boldsymbol{H}^{\mathsf{H}} \left(\boldsymbol{H}\,\boldsymbol{H}^{\mathsf{H}} \right)^{-1} . \tag{6.37}$$

Under a constrained sum-power one can write the optimized precoding matrices as

$$\boldsymbol{W}_{\mathrm{MF}} = \boldsymbol{V}_{\mathrm{MF}} \sqrt{\boldsymbol{P}_{\mathrm{MF}}} \tag{6.38}$$

$$\boldsymbol{W}_{\mathrm{ZF}} = \boldsymbol{V}_{\mathrm{ZF}} \sqrt{\boldsymbol{P}_{\mathrm{ZF}}}, \tag{6.39}$$

where $\boldsymbol{V}_{\mathrm{MF}}$ and $\boldsymbol{V}_{\mathrm{ZF}}$ are the normalized versions of $\tilde{\boldsymbol{W}}_{\mathrm{MF}}$ and $\tilde{\boldsymbol{W}}_{\mathrm{ZF}}$ with unit column norms. The matrices $\boldsymbol{P}_{\mathrm{MF}}$ and $\boldsymbol{P}_{\mathrm{ZF}}$ represent the power allocation of the streams. The cost function of (6.35) is reduced in complexity by only considering the power allocation to yield

$$
\begin{aligned}
&\underset{p_1,\ldots,p_{N_{\mathrm{S}}}}{\text{minimize}} \quad -\sum_{k=1}^{N_{\mathrm{S}}} \log_2\left(1 + \mathrm{SINR}_k(\boldsymbol{V}\sqrt{\boldsymbol{P}})\right) \\
&\text{subject to} \quad \sum_{i=1}^{N_{\mathrm{S}}} p_i \leqslant P,
\end{aligned}
\tag{6.40}
$$

where $\boldsymbol{V}$ is either the normalized precoding matrix for MF or ZF, respectively.

The power allocation matrices $\boldsymbol{P}_{\mathrm{MF}}$ and $\boldsymbol{P}_{\mathrm{ZF}}$ are obtained by waterfilling.

The waterfilling algorithm calculates the powers according to

$$\operatorname*{minimize}_{p_1,\ldots,p_{N_S}} \quad -\sum_{i=1}^{N_S} \log_2\left(1 + p_i\alpha_i\right)$$
$$\text{subject to} \quad \sum_{i=1}^{N_S} p_i \leqslant P, \tag{6.41}$$

with the effective channel gains α_i from the respective $\boldsymbol{\alpha}_{\mathrm{MF/ZF}} = \mathrm{diag}\left\{\left|\boldsymbol{HW}_{\mathrm{MF/ZF}}\right|^2\right\}$ and the sum-power constraint P. The power allocation matrix is formed by $\boldsymbol{P} = \mathrm{diag}\{p_i,\ldots,p_{N_S}\}$. The scheme is optimal for ZF and sub-optimal for MF under sum-power constraint.

Recent massive MIMO system proposals have emphasized the usefulness of these rather simple schemes [244, 245, 243]. The reason is that for certain massive MIMO channel models these linear precoding schemes nearly approach the capacity of optimal precoding under a constrained sum-power even without additional power allocation. In order to achieve this optimal bound, nonlinear dirty paper coding (DPC) can be applied [275]. While the capacity can be calculated via iterative waterfilling [276], an actual implementation of the scheme requires a high computational complexity and implementation effort.
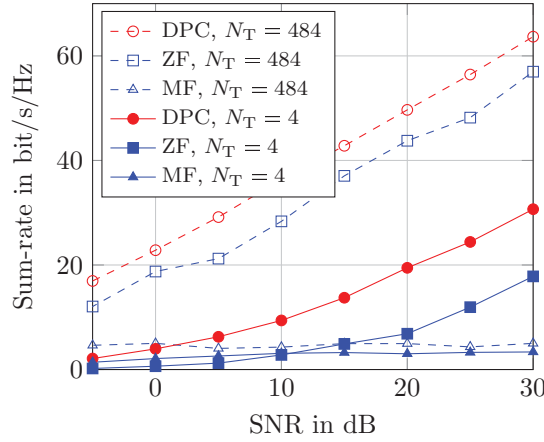
Figure 6.15
Ergodic sum-rate under sum-power constraint for MF, ZF and DPC for the WINNER II A1 NLOS scenario. From [258] © 2017 IEEE
Results are given for multi-mode antennas assuming various base station configurations.

In Fig. 6.15 the ergodic sum-rate for DPC under sum-power constraint is compared to the sum-rate for MF and ZF precoding for various numbers of base station antenna ports. The underlying channel model is a three-dimensional implementation of the WINNER II channel model for an indoor

scenario (A1) under NLOS conditions. The scenario models a rich scattering environment and the three-dimensional approach also includes both polarizations of the scattered waves to be suitable for the multi-mode antennas with their polarization dependent behavior and three-dimensional dimensioning. In Fig. 6.15 the sum-rate for optimized MF and ZF precoding via waterfilling is compared to nonlinear DPC, serving as a benchmark. For a small number of BS ports ($N_T = 4$) and low SNR, MF precoding outperforms ZF precoding, while ZF outperforms MF for higher SNR. For a large number of BS ports ($N_T = 484$), ZF precoding outperforms MF precoding for the considered SNR range.
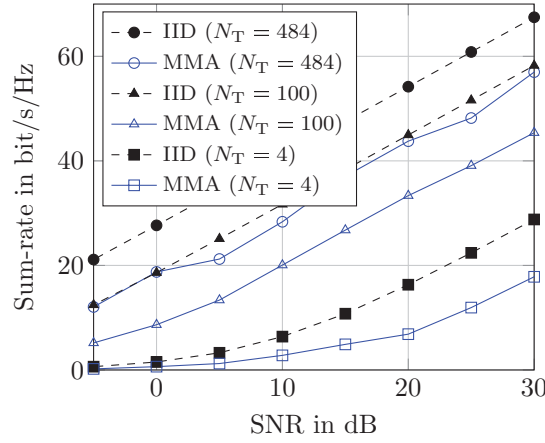
Figure 6.16
Ergodic sum-rate for waterfilled ZF precoding vs. SNR for multi-mode antenna arrays and IID channels. From [258] ⓒ 2017 IEEE
The channel model under investigation is the WINNER II A1 NLOS scenario with one user employing one multi-mode antenna with four receive ports.

In Figs. 6.16 and 6.17 the sum-rate of waterfilled ZF and MF is compared for simulated channels with multi-mode antennas and IID Rayleigh-fading channels. Once again the performance for MF precoding outperforms ZF precoding only for a small number of BS ports in the low SNR regime. Actually, the multi-mode antennas show a similar trend as the IID Rayleigh fading channels for ZF precoding, just with a smaller performance gain per BS size increment. The performance for a channel with multi-mode antennas with four ports at the MS and 484 ports at the BS is very close to the performance for an IID Rayleigh-fading channel with 100×4 elements. However, the performance for MF precoding with multi-mode antennas does not increase significantly with larger BS array sizes. This means that the correlation of the ports of one multi-mode antenna at the MS bounds the performance for MF and we cannot benefit from the massive array at the BS.

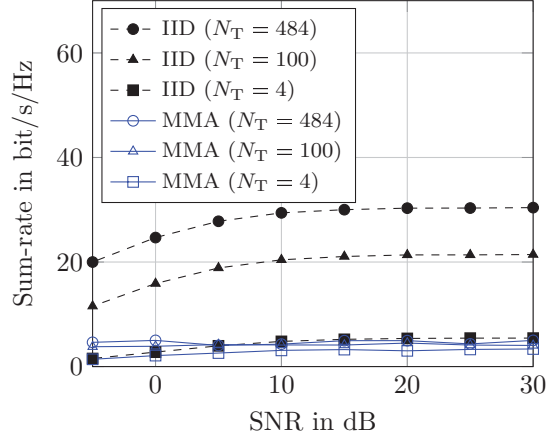The Fig. 6.15–Fig. 6.17 and their discussion show the general potential of

Figure 6.17
Ergodic sum-rate for waterfilled MF precoding vs. SNR for multi-mode antenna arrays and IID channels. From [258] © 2017 IEEE
The channel model under investigation is the WINNER II A1 NLOS scenario with one user employing one multi-mode antenna with four receive ports.

the hardware setup in the context of precoding. The sum-rate performance of ZF under sum-power constraint for example follows the predicted performance of massive MIMO systems qualitatively. Specifically, one can observe that the performance for multi-mode antennas with $N_{\mathrm{T}} = 484$ is almost as good as the performance for IID Rayleigh-fading channels with $N_{\mathrm{T}} = 100$. Therefore, for the conventional sum-power constraint the hardware setup is capable of creating massive MIMO gains.

B. EIRP-Limited Systems

The main difference of the problem formulation in this contribution compared to earlier mentioned publications, which consider the EIRP, is the usage of multi-layer beamforming, i. e., the transmission of multiple data streams in a point-to-point link. In [277] a single data stream is considered, hence the matrix $\boldsymbol{W}$ is a beamforming vector $\boldsymbol{w}$. Given a receive beamformer the authors of [277] claim that the SNR maximization can be reformulated as a second order cone problem (SOCP), which is a convex optimization technique. The problem formulation for this scenario is the maximization of the ℓ_2 norm of the channel matrix times the beamforming vector, which can be transformed to the minimization problem

$$\begin{aligned}
\operatorname*{minimize}_{\boldsymbol{w}\in\mathbb{C}^{N_{\mathrm{T}}}} \quad & -\|\boldsymbol{H}\boldsymbol{w}\|_2^2 \\
\text{subject to} \quad & 1 - \frac{4\pi}{2Z_{\mathrm{F0}}} \max_{\theta,\phi}\left\{|AAF(\theta,\phi,\boldsymbol{w})|^2\right\} \geqslant 0,
\end{aligned} \tag{6.42}$$

where the antenna plus array factor is

$$|AAF(\theta, \phi, \boldsymbol{w})|^2 = \left|\boldsymbol{a}_\vartheta(\theta, \phi)^\mathsf{T}\boldsymbol{w}\right|^2 + \left|\boldsymbol{a}_\varphi(\theta, \phi)^\mathsf{T}\boldsymbol{w}\right|^2 . \qquad (6.43)$$

Further simplifications can be obtained by only considering the constraint in a plane and for one polarization. Solutions to this problem are beyond the scope of this contribution, because the simplification is not suitable for the nature of multi-mode antennas. The considerations in [277] also differ in the sense that transmit and receive beamforming are applied simultaneously. This means that CSI is required at transmitter and receiver side. In [278] the single stream beamforming is optimized jointly over subcarriers of an OFDM system. The problem formulation leads to the maximization of the minimum SNR for the subcarriers, which still can be described as a SOCP. The more general treatment of the input covariance matrices in an EIRP-limited system studied in [279] does not result in full precoding strategies, but only offers conclusions w. r. t. power allocation. Successive work offering sub-optimal vector-perturbation-based beamforming methods again consider a single data stream [280, 281, 282]. The proposed scheme is based on the idea of compressing a known, highly directional beamforming solution to try to utilize the channel in additional spatial directions. Another sub-optimal method for a single data stream is presented in [283] involving Golay sequences. Both of the latter proposals rely on uniform linear arrays to be able to relate the problem formulation to the peak-to-average-power-ratio reduction in OFDM systems.

C. Naive Adaption

When adapting the MF precoding and ZF precoding strategies under a constrained EIRP, the resulting precoding matrices have to be scaled to be compliant with the constraint. The resulting normalized precoding matrices $\boldsymbol{W}'_{\mathrm{MF}}$ and $\boldsymbol{W}'_{\mathrm{ZF}}$ are given as

$$\boldsymbol{W}'_{\mathrm{MF}} = \frac{\boldsymbol{W}_{\mathrm{MF}}}{\sqrt{\mathrm{EIRP}(\boldsymbol{W}_{\mathrm{MF}})}} \qquad (6.44)$$

$$\boldsymbol{W}'_{\mathrm{ZF}} = \frac{\boldsymbol{W}_{\mathrm{ZF}}}{\sqrt{\mathrm{EIRP}(\boldsymbol{W}_{\mathrm{ZF}})}} , \qquad (6.45)$$

where the EIRP in linear scale is computed as a function of the un-scaled versions $\boldsymbol{W}_{\mathrm{MF}}$ and $\boldsymbol{W}_{\mathrm{ZF}}$ as given in (6.38) and (6.39).

As an illustrative example, in Fig. 6.18 one can see the radiated pattern of a BS with $N_\mathrm{T} = 24$ antenna ports, while an MS with three ports is considered. Please note that this does not necessarily mean that three streams have to be transmitted. The channel model is a reduced complexity, ray-based channel with two rays with equal power going out of the BS in the directions $\Phi_1 = 0°$ and $\Phi_2 = 45°$. Intuitively, one would expect lobes in the ray directions with equal power, because of the equal powers of the rays. In this example,

however, in Fig. 6.18(a) the pattern created by the MF solution exhibits one powerful lobe in the direction of 0°. If there is a lobe in the other direction, its power is very small. In other words, for the given example the MF beamforming technique tries to focus most available power in one direction. Under a constrained sum-power this might lead to the same performance as forming a beam in the direction of the other ray, or to distribute the available power among both ray directions. For EIRP-constrained systems, however, a single narrow lobe turns out to be sub-optimal, because each lobe has to fulfill the constraint. An additional lobe in the other ray direction will therefore increase the sum-power budget of the transmitter, but does not alter the fact that the other lobe still complies with the EIRP limit. At the same time the instantaneous sum-rate increases because of the additional lobe. The system benefits from the transmission of a second stream over the other lobe. In Fig. 6.18(b) classical phased array beamforming has been applied, where a fixed phase difference between the signals of the elements leads to a specific main radiation direction at a given frequency. In the given setup it is possible to transmit up to three streams. In this example we choose to transmit two streams with the phased array technique to make use of the two spatial rays available. The addition of the second stream is obtained by setting a second beamforming vector. The radiated field for this second beamforming vector simply superimposes the field for the first stream in spatial domain. In Fig. 6.18(b) one can see the superimposed pattern which leads to an increase of the sum-rate compared to the MF technique.

This academic example illustrates that precoding designed for sum-power constrained systems may not be optimal under a constrained EIRP. To quantify the drawback that the EIRP constraint has on the conventional solutions, in Figs. 6.19 and 6.20 the conventional solutions are compared under sum-power and EIRP constraint, respectively. It has to be noted that SNR_0 in the figure is defined by $\text{SNR}_0 = \frac{\text{EIRP}_{\text{lim}}}{N}$ under EIRP constraint and as $\text{SNR}_0 = \frac{P_{\text{S}}}{N}$ under sum-power constraint. Hence, SNR_0 is a ratio of a power measure in three-dimensional space and the noise power in signal space. Therefore, the performance in Figs. 6.19 and 6.20 cannot be compared directly to the performance of the Figs. 6.15–6.17, which are related to the SNR as a pure signal space power ratio. The numerical result is based on the same channel model and scenario as in Fig. 6.15.

Besides the inferior performance for MF precoding that has been observed before, the curves clearly show that the sum-rate under an EIRP constraint is much smaller than under sum-power constraint. This is obvious for the comparison over SNR_0, where the sum-power of the EIRP-limited system is only a fraction of the sum-power that the sum-power limited system can use. The only case, where the sum-power of the EIRP-limited system can be as large as the sum-power of the sum-power limited system, is for isotropic radiation. This case does not occur in practice. Therefore, conventional solutions with a sum-power constraint SNR_0 violate the EIRP constraint SNR_0 every sin-
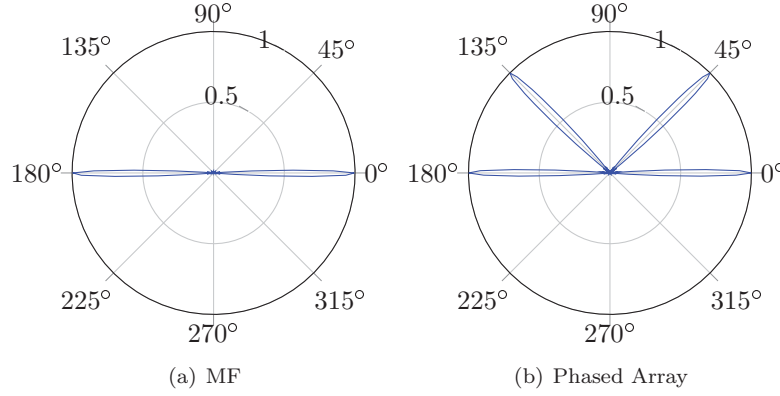
(a) MF

(b) Phased Array

Figure 6.18
Radiated antenna plus array pattern under EIRP constraint, given a two-ray channel model and a BS with $N_\mathrm{T} = 24$ ports for a linear antenna array with omni-directional antennas. From [258] © 2017 IEEE
Case (a) refers to the MF solution ($\boldsymbol{W}_\mathrm{MF}$) and case (b) refers to phased array beamforming, when the two rays are perfectly known and each ray is used by one stream (i. e. each stream forms one lobe). The sum-rate at 30 dB is given by $R_\mathrm{MF} = 19.07$ bps/Hz and $R_\mathrm{phasedarray} = 20.97$ bps/Hz.

gle time. Interestingly, for MF precoding the performance under EIRP-limit catches up with the performance under sum-power limitation.

In the academic example, we showed that the performance could be improved by classical, phased-array based beamforming with separate streams for each spatial ray of the channel. In practice, the implementation of this scheme is not feasible in time-varying environments for several reasons. Firstly, the acquisition or estimation of the angles of departure would be of huge computational complexity and would possibly introduce a large estimation error. Secondly, the number of rays in most practical scenarios is much higher than in the example and the rays are mostly grouped in clusters, what would open the problem of allocation of the streams. Lastly, given the reasons above, the increase of the sum-rate is not guaranteed in any way. However, Figs. 6.18–6.20 motivate an investigation of possible improvements to conventional solutions, when they are applied in systems with a constrained EIRP.

6.4.3 Impact of Inter-Port Coupling

To evaluate the impact of coupling between the ports of multi-mode antenna elements on the sum-rate of the system, the Kronecker model explained in 6.2.2 is used. Hence, the correlation matrices $\boldsymbol{R}_{rr}$ and $\boldsymbol{R}_{tt}$ are set to be block diagonal matrices. By using a WINNER II A1 channel model under NLOS conditions and applying well known zero-forcing precoding with waterfilling
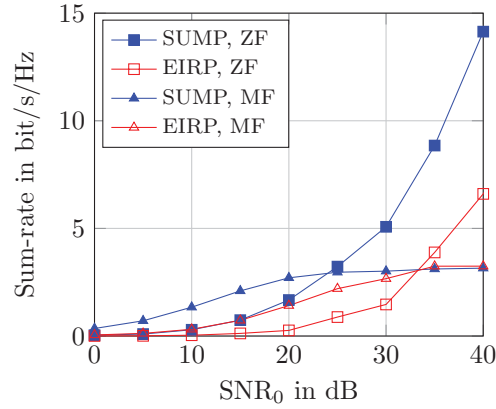
Figure 6.19
Comparison of the system performance in terms of sum-rate between different
constraints for $N_T = 4$. From [258] © 2017 IEEE
SUMP refers to a sum-power constraint and EIRP to an EIRP-limited
system. SNR_0 is defined as $SNR_0 = \frac{EIRP_{lim}}{N}$ under EIRP constraint and as
$SNR_0 = \frac{P_S}{N}$ under sum-power constraint. The numerical data is obtained for
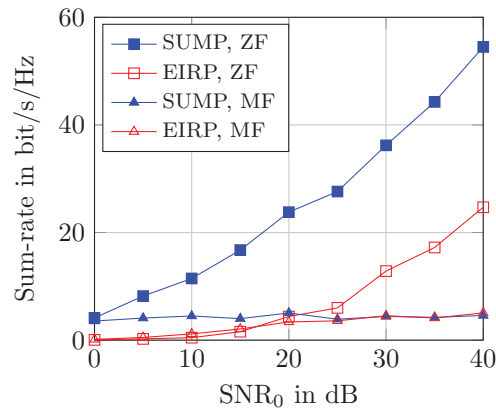the WINNER II A1 NLOS scenario introduced before.

Figure 6.20
Comparison of the system performance in terms of sum-rate between different
constraints for $N_T = 484$. From [258] © 2017 IEEE
SUMP refers to a sum-power constraint and EIRP to an EIRP-limited
system. SNR_0 is defined as $SNR_0 = \frac{EIRP_{lim}}{N}$ under EIRP constraint and as
$SNR_0 = \frac{P_S}{N}$ under sum-power constraint. The numerical data is obtained for
the WINNER II A1 NLOS scenario introduced before.

power allocation, the impact of the correlation can be determined. Note that if the same type of antenna is assumed to be used at both base station and mobile terminal, the correlation is applied twice when employing $\boldsymbol{R}_{rr}$ and $\boldsymbol{R}_{tt}$.
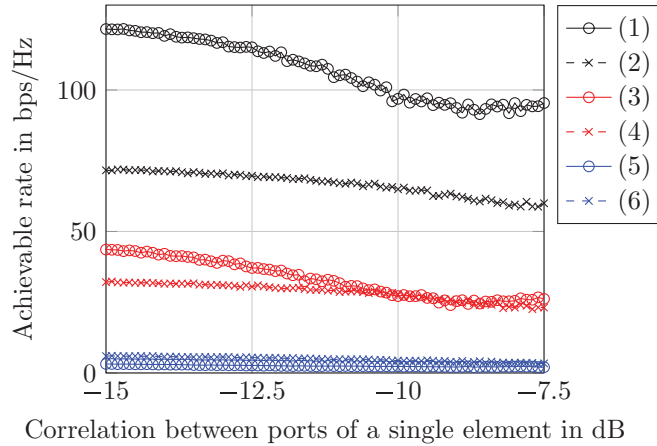

Figure 6.21
The expected achievable data-rate as a function of correlation between the antenna ports of a single element. The assumptions are as follows: The BS has 25 elements, whereas at each receiver one antenna element is employed. A 150x12 (6 ports per element) and 100x8 (4 ports per element) massive MIMO system in downlink scenario is assumed. The channel is generated according to the WINNER II A1 NLOS scenario. ZF precoding is employed for the analysis of the achievable rate. Curve (1): 6 ports, SNR = 20 dB, curve (2): 4 ports, SNR = 20 dB, curve (3): 6 ports, SNR = 10 dB, curve (4): 4 ports, SNR = 10 dB, curve (5): 6 ports, SNR = 0 dB, and curve (6): 4 ports, SNR = 0 dB.

As can be seen from Fig. 6.21, when reaching a mutual coupling greater than -12.5 dB, the achievable rate drops. Since the correlation on both, base station and mobile terminal is set equally, a bonus of three decibels needs to be added. As a result, a mutual coupling of -10 dB is assumed to be sufficient to avoid losses due to mutual coupling.

6.4.4 Beamforming Architectures

[6] As explained in the previous section, diversity and multiplexing can be used either jointly or alternatively. In common baseband signal processing architectures, the corresponding precoding / beamforming is performed mostly in the digital domain. But different system setups are feasible, which are explained subsequently.

---

[6]This section and the figures therein are taken from [252] and [263]

Digital Beamforming

In digital beamforming architectures, each RF chain is connected to an individual antenna. In the special case of digital beamforming, the analog precoding matrix A corresponds to the identity matrix, i.e., the number N_{RF} of RF chains is identical with the number $N_{\mathrm{T}} = N_{\mathrm{Ant,t}} \cdot N_{\mathrm{p,t}}$ of transmit antenna ports. $N_{\mathrm{Ant,t}}$ and $N_{\mathrm{p,t}}$ correspond to the number of antenna elements and ports per element at the base station, respectively. As a consequence, when employing massive MIMO systems, a large number of radio frequency (RF) chains is required. Hence, hardware cost and power consumption are excessive.

Hybrid Beamforming

Besides an identity matrix A (digital beamforming) and a completely filled matrix A (fully-connected analog beamforming), partially filled matrices are also feasible (where each RF chain is connected to a sub-array, either systematically or randomly connected). To reduce the hardware complexity, user-scheduling algorithms can be used. The sum-rate performance of the corresponding digital beamforming is approached by successively increasing the number of scheduled users. Depending on the SNR of the channel, more or less users are activated. Numerous user and stream allocation techniques both for digital beamforming as well as hybrid beamforming have been published. While some reduce a set of defined streams until a maximum sum-rate is reached, others add the most promising data stream until the sum-rate of the system does not increase anymore. However, most hybrid user scheduling algorithms are determined and modified from counterparts used for digital beamforming. Commonly, in hybrid beamforming schemes, the degrees of freedom are less than in digital beamforming. This limitation is due to a limited number of RF chains. An additional limitation is caused by the analog beamforming, since hardware components are not as flexible as software is. Furthermore, the characteristics of the employed hardware need to be known for optimization purposes. This requires some additional channel estimation and measurement effort. Further details on different beamforming techniques can be found in [239].

For reasons of complexity, subsequently focus is on hybrid beamforming. Specifically, the following methods are studied:

1) Fully-connected approach

2) Spatial-filtering approach

3) Single-element beamforming approach.

1) Fully-Connected Approach

In the fully-connected approach, each RF chain is connected to each antenna port by means of steerable phase shifters and matching networks, as presented in inset 1), Fig. 6.14. Hence, the most flexible form of hybrid beamforming is

a fully connected architecture. The elements of matrix $\boldsymbol{A}$ correspond to the phases of the different antenna ports. As this requires additional hardware devices and therefore increases costs while reducing power efficiency, sub-array structures are of interest. In a sub-array architecture, $\boldsymbol{A}$ can be written as a block diagonal matrix. When thinking about multi-mode antennas, two setups based on sub-array precoding are intuitive.

2) Spatial-Filtering Approach

Given an array based on identical multi-mode antenna elements, in the first setup all ports employing the same set of modes are connected to form a sub-array. The spatial characteristics of the structure are used to increase the system performance. This allows simple calculation of the array factor, since simple uniform linear array (ULA) processing can be used if the multi-mode antenna elements are uniformly spaced. Ideally, the array characteristics are employed in order to suppress undesired sidelobes of the chosen group of modes. Therefore, this technique is referred to as spatial filtering.

3) Single-Element Beamforming Approach

The second possibility is to switch between all ports of a single radiation element, subsequently using different characteristics at different positions in the overall sub-array. Thus, mode and space in combination can be used in order to optimize reception and radiation. The single-element beamforming approach requires more sophisticated calculations [258], since the different radiation patterns need to be included. The pattern corresponding to one group of modes of the (sub-) array is only achieved, if only one port of the antenna element is used. The radiation pattern of the port fitting best to the desired direction needs to be chosen. This is referred to as mode-selection.

A more sophisticated method is to combine the individual groups of modes belonging to the ports of a single element. Here, the sub-arrays interfere not only in terms of spatially founded interference, but additionally in terms of mode diversity. The before mentioned sub-arrays holding the same radiation patterns still are employed. The different modes can be combined such that the radiation given a desired direction is optimized. If only one element is used for this kind of beamforming, this is dubbed single-element beamforming subsequently. Each radiating element of a multi-mode antenna array is treated as a sub-array here. By using a priori knowledge of the individual modes, the modes can be combined in a way such that the radiation of a single antenna element is steered towards a single angle. This enables ULA / UPA processing while using the individual antenna elements as a smart antenna. The resulting precoding can be calculated offline and stored as a codebook by means of either a look-up-table or implemented, using phase shifters and switches. One drawback is that additional channel information is required to choose the optimal codebook entry.

6.4.5   Single-Element Beamforming

Codebook Optimization

[7]

Towards the optimization of the gain (or other criteria) of a single multi-mode antenna element, the complex weighting coefficients c_m and electric field components F_ϕ and F_θ are rewritten in matrix notation as

$$\boldsymbol{F} = \begin{bmatrix} F_{\phi,1} & \cdots & F_{\phi,M} \\ F_{\theta,1} & \cdots & F_{\theta,M} \end{bmatrix}, \quad \boldsymbol{c} = \begin{bmatrix} c_1 & \cdots & c_M \end{bmatrix}^T. \tag{6.46}$$

Note that for better readability the angular arguments ϕ and θ are neglected. Then, the problem structure is defined by maximizing (6.21) by using the matrix notation from (6.46) as

$$\arg\max_{c_m}\big\{G(\phi,\theta)\big\} = \arg\max_{\boldsymbol{c}}\big\{|\boldsymbol{F}\cdot\boldsymbol{c}|^2\big\}. \tag{6.47}$$

Upon designing codebooks, several approaches are feasible. In the case of digital beamforming, each port of the multi-mode antenna is connected to the digital processing unit via a full radio frequency (RF) chain and digital-to-analog conversion. This is referred to as <u>digital mode combining</u> in the remainder. Digital mode combining provides best performance due to maximum flexibility for the weighting coefficients. According to (6.47) and given the constraint $\sum_m |c_m|^2 = 1$, the coefficient vector $\mathbf{c}$ can be written as

$$\mathbf{c} = \frac{1}{\sqrt{C}}\big[a_1\,e^{j\alpha_1}, \ldots, a_m\,e^{j\alpha_m}, \ldots, a_M\,e^{j\alpha_M}\big]^T, \tag{6.48}$$

with $C = \sum\limits_{m=1}^{M} a_i^2, \quad a_i \epsilon \Re^{\geqslant 0}$ being a normalization constant to fit the power constraint. When trying to reduce the hardware effort, hybrid beamforming structures come into play. Different approaches are possible. Starting with a simple structure, only the port providing the best performance in a certain direction can be selected:

$$\mathbf{c} = \big[0, \ldots, 0, 1, 0, \ldots 0\big]^T. \tag{6.49}$$

This strategy has been suggested in [241] and is called <u>mode selection</u> subsequently.

Alternatively, adjustable phase shifters can be used to connect the RF chain with all ports of the antenna element. In this case, all ports are active by definition. The optimized phases can be stored in a codebook, according to

$$\mathbf{c} = \frac{1}{\sqrt{M}}\big[e^{j\alpha_1}, \ldots, e^{j\alpha_m}, \ldots, e^{j\alpha_M}\big]^T. \tag{6.50}$$

---

[7]This section and the figures therein are taken from [263].

Due to pure analog beamforming, this version is called <u>analog mode combining</u>.

A more flexible version is to implement the phases by using fixed phase shifters. Triggered by a processor, selected ports and the desired phases can be chosen. Because of the hybrid nature, we call this technique <u>hybrid mode combining</u>. Consequently,

$$\mathbf{c} = \frac{1}{\sqrt{N_C}} \Big[a_1 \, e^{j\alpha_1}, \dots, a_m \, e^{j\alpha_m}, \dots, a_M \, e^{j\alpha_M} \Big]^T, \qquad (6.51)$$

where the amplitudes before normalization a_m are either zero or one. N_C represents the number of selected ports in the current setup. Since an arbitrary number of ports are selected in order to optimize the overall <u>gain</u> (or other criteria), this case is more flexible and hence an improved performance can be obtained.

Numerical Results

[8] In common system simulations, the antenna radiation pattern is included in the channel model. In the above mentioned scenario, the EIRP is employed in terms of a power allocation and precoding constraint. The resulting precoding matrix depends on the given channel. To allow hardware reduction and previous offline precoding, a codebook precoding is targeted. Independent of the channel the given radiation characteristics of the antenna element can be used to construct a codebook, optimizing the radiation to a certain set of angles. Given by (6.18), all available radiation patterns contribute to the radiation pattern of the multi-mode antenna.

In Fig. 6.22, the individual gains of the four ports are plotted in curves (1)-(4) for illustrative purposes. <u>Mode selection</u> can be achieved by appointing the port providing the largest gain at a certain angle. The corresponding gain is depicted in curve (5). <u>Mode selection</u> as defined in Sec. 6.4.4 serves as a benchmark.

In order to compare the performance of the different optimization criteria, the achieved performances of the codebooks are plotted for each criterion. The achieved gain is depicted in Fig. 6.23. As can be seen from curve (1), the codebook of the optimization performed to maximize the <u>gain</u> per angle performs best in terms of achieved gain. Close to this performance comes the achieved gain of the codebook designed to fit the new criterion <u>gain by element factor</u>, as depicted in curve (2). Finally, the gain of the codebook designed to fit the best <u>element factor</u> is presented in curve (3). As can be seen, the achieved gain is fluctuating. In some points, it is even worse than the gain of the <u>mode selection scheme</u>, given in curve (4).

Fig. 6.24 shows that the optimal performance of the <u>element factor</u>, namely $EF = 1$, is reached for the corresponding codebook in nearly the entire angular region. The element factor is presented by curve (1). A good perfor-

---

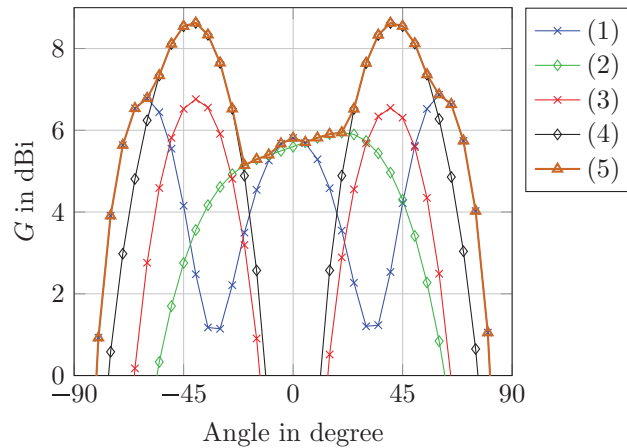[8]This section and the figures therein are taken from [263].

Figure 6.22
Maximum gain in a certain direction which can be derived by switching be-
tween the modes of the multi-mode antenna element. This method is dubbed
mode selection. The modes one to four are represented by (1)-(4). Mode se-
lection is shown by curve (5). From [263].

mance is also reached by using the codebook designed for the gain by element
factor, which is demonstrated in curve (2). The codebook optimized to fulfill
maximum gain requirements is represented by curve (3). It has the weakest
performance in terms of element factor, which results in larger interference
power to neighboring users.

   When comparing the gain by element factors of the three codebooks, as
expected, the performance of codebook optimized to this criterion performs
best, as is shown by curve (1). In the angular region below -30° and above
30°, the codebook for maximum gain achieves a similar performance, given
by curve (2). In contrast, in the central region around 0°, its performance is
roughly 1.5 dB less compared to the achievements of the codebook in curve
(1). Here, the performance of the codebook designed to fit the element factor,
curve (3), is better.

   In Fig. 6.26, the gain achieved by the different mode combining techniques
is shown. Digital mode combining, featured in curve (1), provides the best
results. An exhaustive search considering all possible gain factors and phase
angles of the weighting coefficients **c** is conducted. It enables up to 10 dBi,
which is an improvement of close to 3 dB in some directions. Furthermore, the
gain is improved nearly for the entire angular domain. The gain achieved by
just using phase shifters and switching between contributing ports is shown
in curve (2). Compared to digital mode combining, the loss of hybrid mode
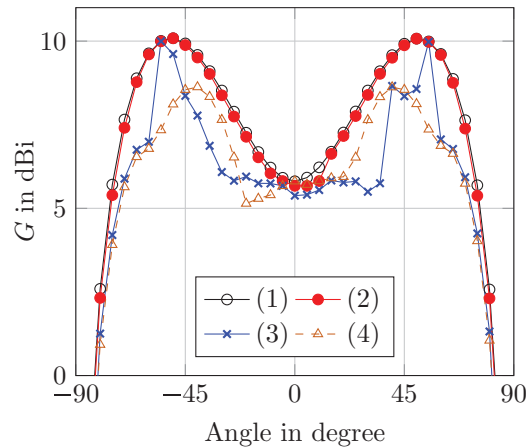combining is fairly small, but hardware complexity is significantly reduced.

Figure 6.23
Achievable gain. From [263] © 2019 IEEE
Performance of the three codebooks for (1) gain, (2) element factor, (3) gain
by element factor and (4) mode selection as reference.

Analog mode combining has the weakest gain, as shown by curve (3). Curve
(4) is the reference plot defined by mode selection.

In order to draw conclusions about the necessity of the individual ports and
combinations, codebooks providing the best performances are investigated. In
Fig. 6.27(a) the incidence for all ports available is shown. As can be observed,
the first port, corresponding to the gain of the blue curve in Fig. 6.22, is em-
ployed at all angles for the two of the three codebooks depending on the gain
of the system. This is remarkable, since it provides the largest individual gain
in only very sparse cases, as compared to the fourth port shown in black. The
radiation pattern corresponding to Port 4 is scheduled in most angles as well.
The codebook designed to fit the element factor employs the patterns corres-
ponding to the Ports 1 and 4 less often. It more frequently enables Port 2 to
reduce the interference, which is employed in the codebook towards gain very
sparsely. Port 3 is generally used most rarely. Regarding the required hard-
ware and precoding effort, the number of scheduled modes per angle segment
is evaluated in Fig. 6.27(b). As can be seen, in nearly all cases two or more
modes are employed in order to achieve the performance shown in Fig. 6.26,
curve (1). To address the defined criteria, in roughly 70% of the angles two
modes are sufficient. Here, the codebook designed to fit the element factor
provides the greatest distribution in the usage of modes. Only in 50% of the
angles exactly two modes are employed. In about 10% of the angles only one
port is sufficient. Looking at the codebook designed to fit the gain by element
factor, always two or more modes are required. In 10% of the angles even all
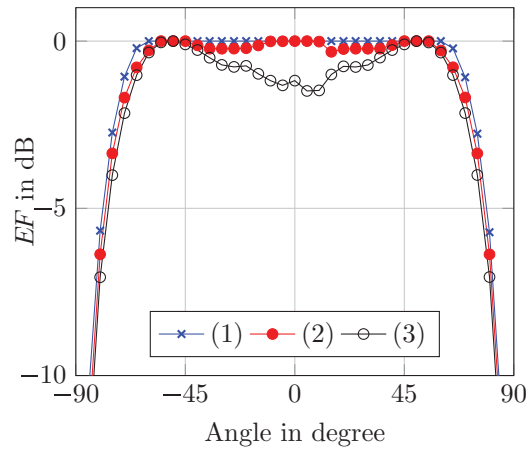four modes are selected.

Figure 6.24
Achievable element factor. From [263] © 2019 IEEE
Performance of the three codebooks for (1) element factor, (2) gain by
element factor and (3) gain.

6.4.6 Acknowledgement

The authors acknowledge first face contributions by Niklas Doose, Kiel University.
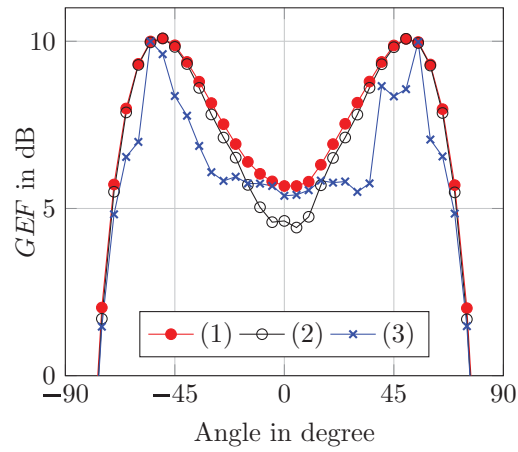
Figure 6.25
Achievable gain by element factor. From [263] © 2019 IEEE
Performance of the three codebooks for (1) gain by element factor, (2) gain, and (3) element factor.
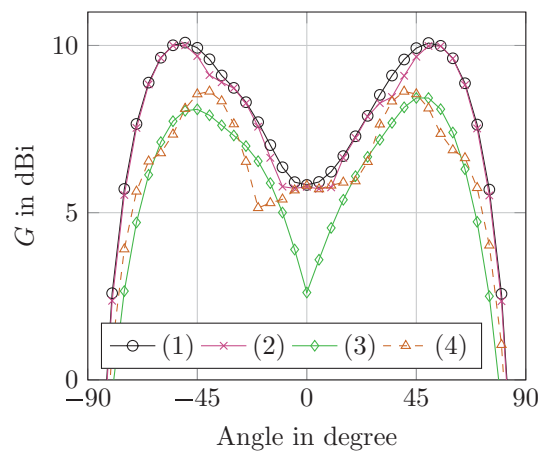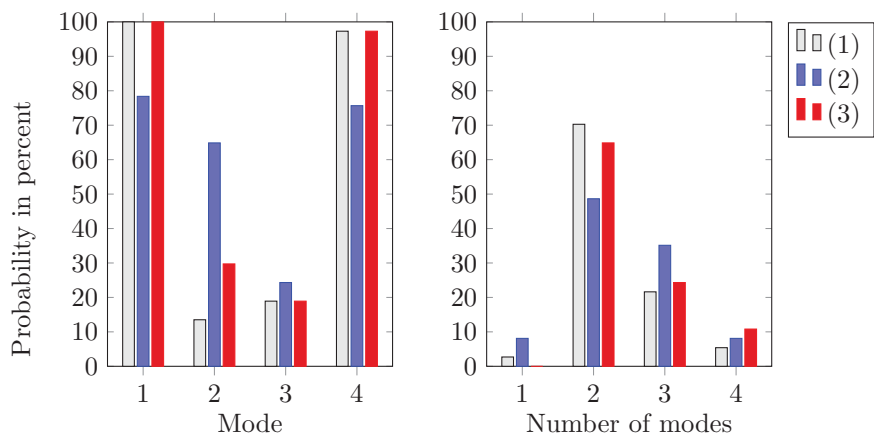

Figure 6.26
Achievable gain by different mode combining techniques. From [263] © 2019 IEEE
(1) Digital mode combining, (2) Hybrid mode combining, (3) Analog mode combining, (4) Mode selection.

(a) Incidence in percent each port is employed.

(b) Incidence in percent a certain number of ports is employed.

Figure 6.27
The bar charts are generated from the digital-mode-combining codebook for 37 equidistant angles of the 180 degree hemisphere. From [263] ⓒ 2019 IEEE The incidences are shown for the three codebooks designed to fit (1) Gain, (2) Element factor, (3) Gain by element factor.