Fakultät für Elektrotechnik und Informationstechnik

Energy Efficient mmWave Receiver Designs

Kilian Peter Anton Roth

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

Doktors der Ingenieurwissenschaften

genehmigten Dissertation.

Vorsitzender: Prof. Dr. sc. techn. Andreas Herkersdorf
Prüfende der Dissertation:

1. Prof. Dr. techn. Dr. h. c. Josef A. Nossek
2. Prof. Dr. A. Lee Swindlehurst
3. Prof. Dr.-Ing. Wolfgang Utschick

Die Dissertation wurde am 10.06.2021 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 25.11.2021 angenommen.

Acknowledgments

I really appreciated the work of Dr. Cecilia Carbonelli, Michael Färber, and Prof. Josef A. Nossek for organizing the setup that lead to me having an PhD student position at Intel Germany. I also want to thank Prof. Josef A. Nossek for being available for technical discussions as well as for all the extensive revisions leading to a substantial improvement of this document. I am very grateful for the freedom to explore different technical directions I got during my work for Michael Färber at Intel. I would like to express my gratitude to Prof. Dr. A. Lee Swindlehurst and Univ.-Prof. Dr.-Ing. Wolfgang Utschick for agreeing to be my second and third examiner.

I want to thank my colleagues Amine Mezghani, Hela Jedda, Jawad Munir, Bernhard Lehmeyer, Qing Bai, Francisco Javier Garcia Gomez, Rana Ali Amjad, Oliver De Candido, and Hans Brunner as well as all the students at NWS and MSV student rooms for giving me a glimpse into the research atmosphere at the university. I also wanted to thank my colleagues at Intel in Germany Miao Honglei, Markus Mück, Miltiades Filippou, Ingolf Karls, Marta Martinez Tarradell, Dario Sabella, Yang Yang, Camila Priale Olivares, and Hyung-Nam Choi for the pleasant office atmosphere as well as always being available for my pestering technical questions. I am very grateful for the technical insights I gained from my Russian Intel colleagues Alexey Vladimirovich Khoryaev, Sergey Vyacheslavovich Panteleev, and Mikhail Sergeevich Shilov as well as their support for my work on different aspects of V2X communication.

As special thanks go to Leonardo Gomes Baltar for working with me first in the university and afterwards in the team at Intel for always keeping me motivated, especially during the last part of finalizing this work. I am especially grateful for sharing the office for a long time with Mustafa Emara. Without his never-ending enthusiasm as well as support I would have likely not finished this work.

I also want to thank my family including Y, B, I, L, L, I, Q, M, T, and P for their constant support and helpful distraction.

Contents

1. Introduction

1.1 Motivation

The use of the available spectrum in the frequency range of 6 to 100 GHz is considered to be an essential part of the mobile broadband standard Fifth Generation mobile communicaiton system (5G) [18]. Due to the propagation conditions at these frequencies, this technology is especially attractive for high data rate, shorter range wireless communication. This frequency range is referred to as millimeter Wave (mmWave), even though it contains the lower centimeter wave range. In recent years, the availability of spectrum and consumer grade systems at mmWave frequencies led to a huge increase in academic and industrial research. However, to fully leverage the spectrum while being power-efficient, the Baseband (BB) and Radio Front-End (RFE) capabilities must be drastically changed from current state-of-the-art cellular devices.

The use of high carrier frequencies above 6 GHz will go hand in hand with the implementation of large antenna arrays [18], [11]. The support of a large number of antennas at the mobile and basestation requires a new RFE design. To attain a similar link budget, the effective antenna aperture of a mmWave system must be comparable to current systems operating at carrier frequencies below 6 GHz. Therefore, an antenna array at both the base and mobile station is unavoidable. Since the antenna gain and thus the directivity increases with the aperture, an antenna array is the only solution to achieve a high effective aperture while maintaining omnidirectional coverage.

Different receiver implementation with an antenna array are possible. The following subsection is introducing these front-end solutions.

1.1.1 Receive Beamforming Architectures

This subsection introduces the different receiver architectures envisioned by academia and industry for future mmWave mobile broadband systems. The architectures considered in this work are Digital Beamforming (DBF), subarray and Hybrid Beamforming (HBF), fully-connected HBF, and subarray HBF using a Butler matrix. The advantages and disadvantages of the different solutions are highlighted in the respective subsection and discussed relative to each other in the subsequent subsection. As in this work we concentrate on receiver aspects we only show receiver architectures. Most of the arguments for different architectures are also valid for transmitters with the same architecture.

1.1.1.1 Digital Beamforming

Current mobile broadband systems based on Third Generation Partnership Project (3GPP) Long Term Evolution (LTE) are mainly based on DBF, similar to the simplified block diagram shown

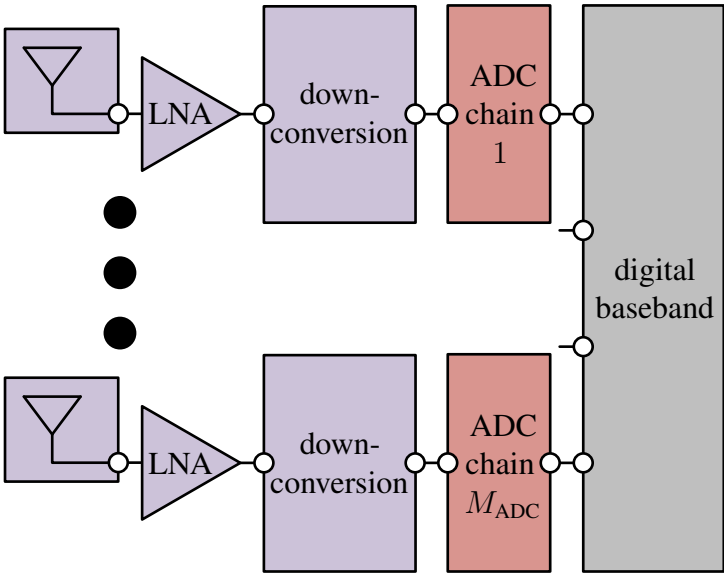

Fig. 1.1. Digital Beamforming (DBF) with M antennas.

in Fig. 1.1. In this architecture the signal from each antenna is available in the digital domain. This has the main advantage that all Degrees of Freedom (DOF) are available for the reception of multiple spatial Multiple Input Multiple Output (MIMO) streams, as well signals from multiple users (MultiUser - Multiple Input Multiple Output (MU-MIMO)). Since only digital processing is required the system is also flexible to adapt to a fast-changing radio environment. However, in contrast to mobile broadband systems operating below a carrier frequency of 6 GHz, the bandwidth available for a mmWave system is more than an order of magnitude larger. As shown in [63] if we would use Analog-to-Digital-Converters (ADCs) with a similar resolution as for current LTE system the power consumption would be prohibiting large. Thus, we only use this architecture in combination with low resolution Analog/Digital (A/D) conversion in this work.

1.1.1.2 Hybrid Beamforming

The idea of analog/hybrid beamforming originated from the phased arrays approach often used for radar applications [49]. The main idea for analog beamforming is that the signal from a group of antennas is combined in the Radio Frequency (RF) or in the analog BB. To align with the channel or the receive angle the signal can be phase shifted and combined to form a virtual antenna with a preferred direction adapted to the signal that should be received [75]. Instead of the phase shifting and combining it is also possible to generate different beams with a Butler matrix as described in [92, 31]. We speak of analog beamforming if a system only has access to one digital signal, as the multi-antenna receive signal is processed purely in the analog domain. In the case there are multiple digital signals available we speak of hybrid beamforming as a combination of analog and digital processing is used to process the multi-antenna signal. The first standard assuming analog beamforming targeting consumer devices was defined in 2012 with WiGig (802.11ad) [37]. In the following two paragraphs we introduce two different flavors of hybrid beamforming with phase shifters. It is important to mention that the phase shifter network could be replaced by a Butler matrix. As shown in [45, 22] it is not for every system the best option to perform the phase shifting in the RF. It is also possible to phase shift the Local Oscillator (LO) or even perform the phase shifting and combining at an Intermediate Frequency (IF) or the analog BB before the A/D con-

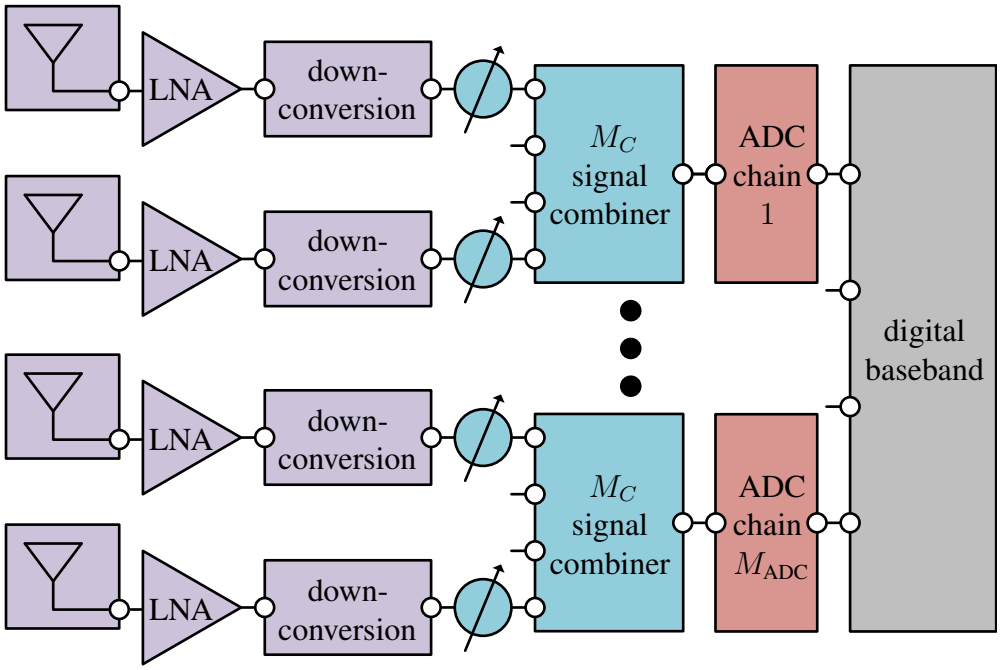

Fig. 1.2. Subarray Hybrid Beamforming (HBF) with M antennas, M_{ADC} ADC chains and M_C antennas per subarray.

version. How such systems would look like is also explained in the following paragraphs.

Subarray Hybrid Beamforming

As shown in Fig. 1.2 this system is composed of multiple phased array antenna subarrays. In this case the signal from each antenna is phase shifted and subsequently combined with the signal from the other antennas in the same subarray. Afterwards, this combined signal from each subarrays is converted into the digital baseband. The transmitted signal from one or multiple transmitting nodes is then estimated via standard digital signal processing techniques. Many proof of concepts demonstrate that such an transceiver architecture is possible [36, 43]. It is also possible to replace the phase shifters and power combiners with a Butler matrix. In this case one of the outputs of the Butler matrix is selected and converted to a digital baseband signal, separately for each subarray. One main disadvantages of this system is that each subarrays needs to find the optimal configuration of the phase shifters or the output of the Butler matrix by a trial and error procedure.

Fully-connected Hybrid Beamforming

Since in the case of subarray hybrid beamforming only the array gain of the subarray is available for each of the digital signals, some work is investigating a system where the antenna gain of the combined array is available at each of the M_{ADC} ADC chains. As shown in Fig. 1.3 the signal at each antenna is split into M_{ADC} signals. Afterwards, each of the signals are phase shifted and combined with a signal from each of the other antennas. The resulting signals are then converted into a digital signal. It is important to note that in this case the power combiners have M inputs instead of M_C as in the case of subarray hybrid beamforming. This system realized with a Butler matrix instead of the signal splitters, phase shifters and combiners would just consist of one big Butler matrix using all antennas and M_{ADC} ADC chains connected to M_{ADC} outputs of the Butler matrix. As power combiners can even theoretically not be implemented without loss this architecture would need addition amplifiers to compensated for the loss inside the power combiners. For a practical system the implementation is even more challenging. Since, in addition to this aspect the

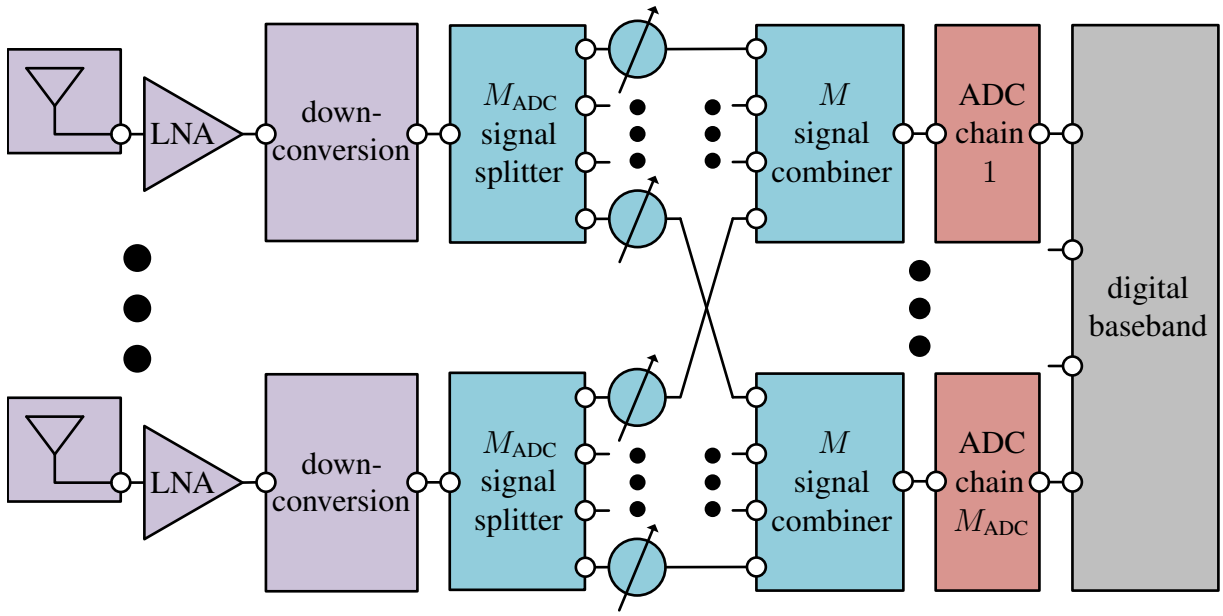

Fig. 1.3. Fully-connected Hybrid Beamforming (HBF) with M antennas and M_{ADC} ADC chains.

design cannot be decomposed into smaller elements containing only one subarray so far it was only considered for theoretical evaluations and there is no attempt to implement it with actual hardware [10, 8, 75, 59]. We will concentrate on the subarray hybrid beamforming, but as many theoretical works focus on fully-connected hybrid beamforming we compare some of our solutions with this approach.

Butler Matrix based subarray Hybrid Beamforming

Another way to realize the analog combining is a Butler matrix first described in [20]. The corresponding receiver is shown in Fig. 1.4. For these systems after the Low Noise Amplifier (LNA) the signal from the M_C antennas belonging to one subarray are combined via a Butler matrix. The Butler matrix consists of hybrid couplers and non nonadjustable phase shifters forming M_C outputs. The following ADC chain is using the selected output via a Multiplexer (MUX). In general, the properties of this approach are similar to subarray HBF, with the only difference being how the different beams are realized.

1.1.1.3 Advantages and Disadvantages of the Different Architectures

One major disadvantage of the hybrid beamforming systems is that as the system output can only be observed after the combining of the signals from one subarray, a trial and error procedure is necessary to test different configurations and afterwards select the best of them. This trial and error procedure is called beam training or beam alignment and can be implemented in different ways, but can always be considered to require a significant amount of time [10, 26, 42]. The first standardized system that included procedures for beam training and constant alignment is 802.11ad (aka WiGig) [37]. Beam management procedures are also be included in 3GPP New Radio (NR) (aka 5G) [5, 6]. We consider that even in a static scenario it has a large overhead to align the beams for a mmWave hybrid or analog beamforming system. The situation becomes even more complicated if we take user mobility and in general a changing environment into account. The work in [14] shows this overhead compared to a digital system. As many beam management procedures assume reciprocity there are stringent requirement on the calibration of the RF front-

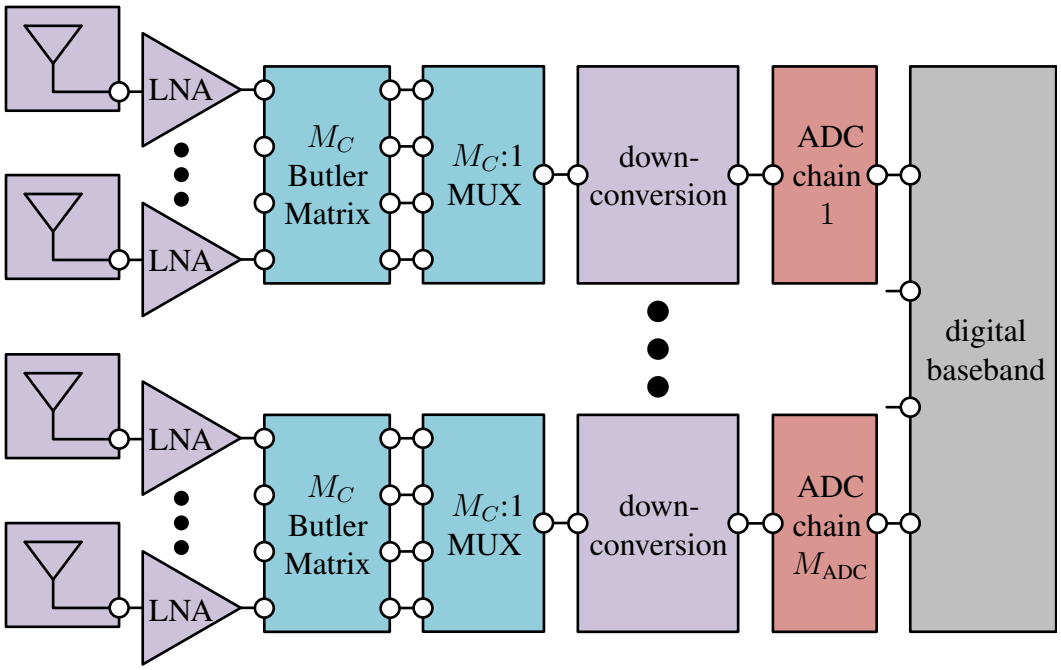

Fig. 1.4. Butler Matrix based subarray Hybrid Beamforming (HBF) with M antennas, M_{ADC} ADC chains and M_C antennas per subarray.

end. To enable optimal beam training the beams need to be designed in advance and calibrated during manufacturing. As again only the combination of the signal from different antennas are observable, external components are necessary to calibrate the transceiver. The major drawback of the digital beamforming system relative to the analog/hybrid ones is the power consumption. At the start of the investigation if mmWave systems might be viable for consumer electronic, the power consumption was believed to be excessive. The combination of these aspects motivates the comparison of these systems especially regarding energy efficiency.

1.2 Contribution and Organization

The work mainly focuses on comparing hybrid and digital beamforming systems as well as showing that the results predicted from information theoretic analysis of digital beamforming systems with low resolution A/D conversion are achievable with standard signal processing techniques. First models for all essential parts of the systems are developed. Based on these models different receiver front-end configurations are compared. We show that current mmWave mobile communication systems can be used with low resolution ADCs without or only with small changes in the specification.

The thesis is divided into two main parts. The first part “Signal Processing Aspects” is covering various signal processing tools for digital beamforming systems with low resolution ADCs and hybrid beamforming system. Essentially, it functions as a toolbox for the evaluations and comparison in the following second part “Spectral and Energy Efficiency Comparison of Analog, Hybrid and Digital Beamforming”, The following items give an overview of the contributions in each chapter.

- **Signal Processing Aspects**
  - *Additive Quantization Noise Model*:

As many past evaluations ignored the effect that the quantization error on multiple antennas are correlated, we also want to investigate this effect. To enable the modeling, we develop a numerical integration based technique to relate the input/output correlation of a quantizer based on the assumption that the input signal is Gaussian distributed. This approach does take the possibility of having ADCs with different resolution at each antenna into account. This work was part of the papers [76, 81].

- *Beam Design for Hybrid or Analog Beamforming*:
  In this section we develop a technique to synthesize a beam according to user defined criteria based on techniques originally developed for radar technology. We are taking all possible hybrid beamforming configurations into account and it is even possible to generate beams that are optimized for multiuser scenarios. A part of the work in this section has also been published in [78].
- *Calculation of Beam Training Parameters*:
  The beam training parameters sequence length and beam training interval are evaluated here. Given a performance quality criterion, the overhead of the sequence used for channel quality assessment is minimized. Assuming low mobility and channel parameters based on past measurements, the interval between consecutive beam training instantiations is assessed.
- *Receiver Front-End Power Consumption Model*:
  As we also wanted to evaluate the energy efficiency, it is necessary to develop a power consumption model. This model is taking into account hardware reported in literature targeting a carrier frequency of 60 GHz using Complementary Metal-Oxide-Semiconductor (CMOS). It does include digital beamforming, hybrid beamforming based on phase shifters and hybrid beamforming based on a Butler matrix in combination with low resolution ADCs. This section was part of the work [77].
- *Channel Estimation*:
  The channel estimation performance of a system with 3GPP NR Type I CP-OFDM DMRS is evaluated. A closed form expression for the channel estimation error is derived. This evaluation was part of the papers [81, 80, 82].

- **Spectral and Energy Efficiency Comparison of Analog, Hybrid and Digital Beamforming**
  - *Energy and Spectral Efficiency Comparison for Single Users Scenarios*:
    In this section the expressions for the energy and spectral efficiency in a multipath propagation scenario are derived and compared for digital and hybrid beamforming with low resolution ADCs. In addition, the impact of imperfect Automatic Gain Control (AGC) adaptation, as well as the modeling of the quantization error correlation is shown. Most of the material in this subsection appeared in [77].
  - *Rate Regions for Hybrid Beamforming and Digital Beamforming with Low Resolution ADCs*:
    This section shows the two user achievable rate region for digital beamforming with low resolution and hybrid beamforming. Especially, in the low Signal-to-Noise-Ratio (SNR) regime the digital system is outperforming the hybrid system. This subsection is based on the article [79].
  - *Energy and Spectral Efficiency Comparison for Multiple Users and Imperfect Channel State Information at the Receiver (CSIR)*:
    The work in this part of the thesis is comparing the energy and spectral efficiency in a multiuser scenario. The following modeling parameters are included in the evaluation: trans-

mitter impairments, receiver channel estimation error, multipath propagation, low resolution ADCs and beamforming. The results show that over the whole range of SNRs typically encountered in a practical scenario the low resolution ADC digital beamforming system is more energy efficient than then hybrid beamforming one. Part of the work in this section also appeared in [81].

- *Robust massive MIMO Equalization for mmWave systems with Low Resolution ADCs*:
  As for massive MIMO system the large number of antennas could be a bottleneck for signal processing. We wanted to clarify that standard signal processing techniques are sufficient to handle the additional challenges arising when considering low resolution ADCs. In addition, we also adapt a MIMO equalization scheme, that has a complexity below Minimum Mean Square Error (MMSE) equalization, to the special case of low resolution and show a slightly better performance than linear schemes. The work in this section was also published in [80].
- *Are the data rates predicted by the analytic analysis of receivers with Low Resolution ADCs achievable?*:
  In this part of the thesis we compare the results of our information theoretic based evaluation to a 3GPP NR link level simulation, both considering low resolution A/D conversion. The results show that the predicted rates are achievable, given the usual gap from a theoretical evaluation to a practical system. The results of this section also appeared in [82].
- *Energy and Spectral Efficiency Comparison Including Beam Training Overhead*:
  In this section we include additional aspects of the beam training overhead into the preceding evaluation. We also added a hybrid beamforming system based on a Butler matrix. For these cases still the low resolution ADCs digital beamforming system is in the practical relevant per antenna SNR up to 0 dB more energy efficient. In the high SNR regime, the Butler matrix based hybrid beamforming system gets more energy efficient.

2. Signal Processing Aspects and Estimation Tools of Hybrid Beamforming and Digital Beamforming with low Resolution ADCs

2.1 Additive Quantization Noise Model

In this paragraph we develop the model we use to characterize the performance of systems including the quantization based on the Bussgang theorem [19]. We use the Additive Quantization Noise Model (AQNM) to model the quantization effects of the ADCs. In contrast to other works we also include the effects of spatially non-white quantization errors and imperfect gain settings of the Variable Gain Amplifier (VGA) by the AGC. We decided to go for the exact expression in contrast other works in this area ignoring this effect [57, 66, 10], or only using the closed form expression available for 1 bit quantization [53]. We start by introducing a general signal model for the quantization and the Bussgang theorem to model the quantization error as additive. Since this requires the calculation of the correlation of the signal after quantization we utilize Price's theorem [70]. This work was part of the papers [76, 81].

The model in this section is covering all aspects of the analog and mixed signal circuits shown in the block diagram Fig. 2.1. For this modeling we only consider the case that the AGC already converged to a stable state. As shown in [91] due to the requirements of the WiGig standard (802.11ad) this can be achieved in less than 1.2 μs. In a practical receiver the variable gain is applied in various amplification stages of a receiver, possibly starting even in the LNA (see [91] as an example). As for the purpose of this work where we only model the equivalent baseband, it is sufficient to model all gain stages combined as an equivalent VGA. Fig. 2.1 only shows the model for either the I or the Q component of one analog baseband signal. As we only assume proper Gaussian signals the same gain is applied for I and Q component of each signal. It is also important that it is possible for different analog signals to have a different gain of the VGA. This might result in different phase shifts of the signal and therefore would change the equivalent channel that will be observed in the digital baseband. However, since the channel estimation in the digital domain would not be able to distinguish the effect of this phase shift from the effect of different channels at each antenna, it will not have an impact on the overall performance, and can thus also be ignored in the modeling. The last point that is important to mention is that for the purpose of the modeling we assume the input signal to the ADC to have unit variance and zero mean, for the case the AGC perfectly adapts the input signal to the available dynamic range of the ADC. This also does not have an impact on the performance as the signal at this point already contains a signal and a noise component, which are scaled together. Thus, this scaling does not change the SNR of the analog baseband signals.

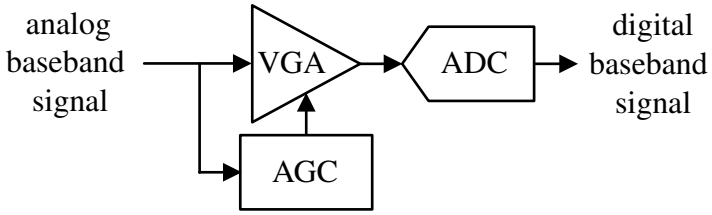

Fig. 2.1. Analog and mixed signal components modeled by the AQNM model.

2.1.1 Decomposition of the Quantization Error

In general, the Bussgang theorem [19] states that if y is a real valued, zero mean Gaussian random variable than if $g(y)$ is any non-linear function

$$r = g(y), \tag{2.1}$$

the cross correlation of y and r differs to the autocorrelation of y only by a scalar S

$$\mathbb{E}[yr] = S\mathbb{E}[yy]. \tag{2.2}$$

In the whole paragraph we always assume the input signal of the quantizer have unit variance, for the case that the AGC perfectly adapts the signal power to the dynamic range of the ADC. Since the quantizers operate on the real and imaginary part separately, this means we scale the real and imaginary part of each element in $\boldsymbol{y}$ to have unit variance. As in a communication system only the relative power of the noise and the signal matter, rescaling of y to simplify the notation does not change the SNR and therefore does not influence the performance of the system. For a vector based received signal $\boldsymbol{y}$, it is necessary to rescale each component individually. As evident from [88] a mean does not transmit any information. Thus, we can safely assume that the receive signal from any communication system would have zero mean. For the thermal noise we can also assume it to have zero mean. Thus, in general for a communication system the requirement of the input of the quantizer having zero mean is satisfied and thus the Bussgang theorem is applicable if the received signal is Gaussian.

Now this can be converted to the case that we now have a circular symmetric, unit variance, zero mean, complex Gaussian vector input $\boldsymbol{y}$ we get the following expression

$$\boldsymbol{r} = \boldsymbol{D}\boldsymbol{r}' = \boldsymbol{D}Q(\boldsymbol{y}), \tag{2.3}$$

where $Q(\cdot)$ the quantization operator, separately operating on each element and the real and imaginary part of $\boldsymbol{y}$. The real, diagonal matrix $\boldsymbol{D}$ is scaling each quantized signal individually. As it scales the signal plus noise together it does not change the SNR and has no influence on the performance of the communication system, but greatly simplifies the calculation in some special cases. The auxiliary variable $\boldsymbol{r}'$ is the result directly after the quantization. Essentially, the operation of the quantization function $Q(\cdot)$ can in general be defined as:

$$r' = Q(y) = v_j \,\forall\, y \;\in]s_{j-1}, s_j]\,. \tag{2.4}$$

Here v_j is the value representing the jth quantization bin with the input interval (position of the steps) $]s_{j-1}\ s_j]$. This is illustrated for a 2 bit uniform quantizer in Fig. 2.2. To cover a real valued input the left limit of the first interval s_0 and the right limit of the last interval s_{N_b} are equal to $-\infty$ and ∞ respectively, where $N_b = 2^b$ is the total number of quantization bins. It is important to keep

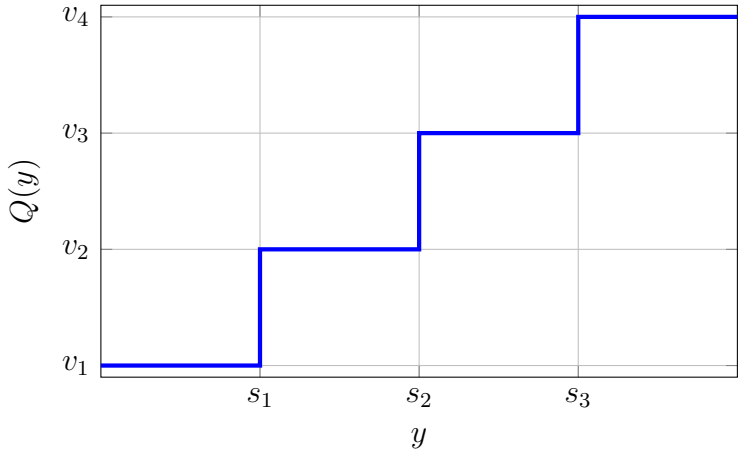

Fig. 2.2. Quantization step function example for a 2 bit uniform quantizer.

in mind that we do not assume the quantizer to be the same for every element of the vector, only for the real and imaginary part of each element.

In case of a vector input the Bussgang theorem transforms to

$$\mathbb{E}\left[\boldsymbol{r}'\boldsymbol{y}^H\right] = \boldsymbol{S}\mathbb{E}\left[\boldsymbol{y}\boldsymbol{y}^H\right]. \tag{2.5}$$

Using this expression, we can decompose the signal after the quantization in the following way:

$$\boldsymbol{r}' = \boldsymbol{S}\boldsymbol{y} + \boldsymbol{e}, \tag{2.6}$$

where $\boldsymbol{e}$ is the quantization distortion uncorrelated with $\boldsymbol{y}$. Following the definition of the Bussgang theorem we can calculate $\boldsymbol{S}$ in the following way:

$$\boldsymbol{S} = \boldsymbol{R}_{\boldsymbol{r}'\boldsymbol{y}}\boldsymbol{R}_{\boldsymbol{y}\boldsymbol{y}}^{-1} = \mathbb{E}\left[\boldsymbol{r}'\boldsymbol{y}^H\right]\left(\mathbb{E}\left[\boldsymbol{y}\boldsymbol{y}^H\right]\right)^{-1} = \boldsymbol{D}^{-1}\boldsymbol{R}_{\boldsymbol{r}\boldsymbol{y}}\boldsymbol{R}_{\boldsymbol{y}\boldsymbol{y}}^{-1} = \boldsymbol{D}^{-1}\mathbb{E}\left[\boldsymbol{r}\boldsymbol{y}^H\right]\left(\mathbb{E}\left[\boldsymbol{y}\boldsymbol{y}^H\right]\right)^{-1}, \tag{2.7}$$

where the second part follows from the definition in (2.3), leading to $\boldsymbol{y}$ and $\boldsymbol{e}$ being uncorrelated:

$$\mathbb{E}\left[\boldsymbol{y}\boldsymbol{e}^H\right] = \boldsymbol{0}. \tag{2.8}$$

It is also necessary to calculate the covariance matrix of the quantization distortion $\boldsymbol{e}$. Since $\boldsymbol{e}$ and $\boldsymbol{y}$ are both zero mean the covariance matrix of $\boldsymbol{r}$ is defined as:

$$\boldsymbol{R}_{\boldsymbol{r}\boldsymbol{r}} = \boldsymbol{D}\left(\boldsymbol{S}\boldsymbol{R}_{\boldsymbol{y}\boldsymbol{y}}\boldsymbol{S}^H + \boldsymbol{R}_{\boldsymbol{e}\boldsymbol{e}}\right)\boldsymbol{D}. \tag{2.9}$$

Solving this equation for $\boldsymbol{R}_{\boldsymbol{e}\boldsymbol{e}}$ and plugging in the definition for $\boldsymbol{S}$ in (2.7) leads to:

$$\boldsymbol{R}_{\boldsymbol{e}\boldsymbol{e}} = \boldsymbol{D}^{-1}\left(\boldsymbol{R}_{\boldsymbol{r}\boldsymbol{r}} - \boldsymbol{R}_{\boldsymbol{r}\boldsymbol{y}}\boldsymbol{R}_{\boldsymbol{y}\boldsymbol{y}}^{-1}\boldsymbol{R}_{\boldsymbol{r}\boldsymbol{y}}^H\right)\boldsymbol{D}^{-1}. \tag{2.10}$$

To calculate the matrices $\boldsymbol{S}$, $\boldsymbol{R}_{\boldsymbol{e}\boldsymbol{e}}$ we need to calculate $\boldsymbol{R}_{\boldsymbol{r}\boldsymbol{y}}$, $\boldsymbol{R}_{\boldsymbol{r}\boldsymbol{r}}$, $\boldsymbol{R}_{\boldsymbol{y}\boldsymbol{y}}$ and $\boldsymbol{D}$. In the following we show how to calculate these matrices.

It is also possible to decompose each element of $\boldsymbol{r}$ in the following way:

$$\Re\{[\boldsymbol{r}]_m\} = [\boldsymbol{D}]_{m,m}\Re\{[\boldsymbol{r}']_m\} = \Re\{[\boldsymbol{y}]_m\} + \Re\{[\boldsymbol{q}]_m\}\ \forall m \in \{1, \cdots, M\}, \tag{2.11}$$

where $\boldsymbol{q}$ is the quantization error, different from $\boldsymbol{e}$. The scaling is $[\boldsymbol{D}]_{m,m}$ is introduced to ensure that the mean square error $\mathbb{E}[\Re\{[\boldsymbol{q}]_m\}^2]$ of the quantization error is minimized. As the quantizer

is assumed to be the same for the real and imaginary part and the input signal is circular symmetric, the same scaling can be applied for the real and imaginary part. From the definition of the quantization error:

$$\begin{gathered}\epsilon_m = \mathbb{E}[\Re\{[\boldsymbol{q}]_m\}^2] = \mathbb{E}[([\boldsymbol{D}]_{m,m}\Re\{[\boldsymbol{r}']_m\} - \Re\{[\boldsymbol{y}]_m\})^2] = \\ \mathbb{E}[\Re\{[\boldsymbol{y}]_m\}^2] + [\boldsymbol{D}]_{m,m}^2\mathbb{E}[\Re\{[\boldsymbol{r}']_m\}^2] - 2[\boldsymbol{D}]_{m,m}\mathbb{E}[\Re\{[\boldsymbol{r}']_m\}\Re\{[\boldsymbol{y}]_m\}],\end{gathered} \tag{2.12}$$

it is obvious that the $[\boldsymbol{D}]_{m,m}$ minimizes this quadratic equation is defined as:

$$[\boldsymbol{D}]_{m,m} = \frac{\mathbb{E}[\Re\{[\boldsymbol{r}']_m\}\Re\{[\boldsymbol{y}]_m\}]}{\mathbb{E}[\Re\{[\boldsymbol{r}']_m\}^2]}. \tag{2.13}$$

In addition, using this optimal scaling results in

$$\mathbb{E}[\Re\{[\boldsymbol{r}]_m\}\Re\{[\boldsymbol{q}]_m\} = [\boldsymbol{D}]_{m,m}\mathbb{E}[\Re\{[\boldsymbol{r}']_m\}\Re\{[\boldsymbol{q}]_m\} = 0 \; \forall m \in \{1, \cdots, M\}. \tag{2.14}$$

Since $[\boldsymbol{y}]_m$ is unit variance Gaussian with zero mean these terms can be calculated in the following way:

$$\mathbb{E}[\Re\{[\boldsymbol{r}']_m\}^2] = \sum_{j=1}^{N_b} (v_j)^2 \left(\Phi\left(s_j\right) - \Phi\left(s_{j-1}\right)\right), \tag{2.15}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the Gaussian distribution with unit variance and zero mean, the s_j are the positions of the quantization steps and v_j is the output value of the quantizer for the input region j between s_{j-1} and s_j. The first and last step are set to be $s_0 = -\infty$ and $s_{N_b} = \infty$. The calculation for the correlation portion leads to the closed form solution

$$\mathbb{E}[\Re\{[\boldsymbol{r}']_m\}\Re\{[\boldsymbol{y}]_m\}] = \frac{-1}{\sqrt{2\pi}} \sum_{j=1}^{N_b} v_j \left(e^{\frac{-\left(s_j\right)^2}{2}} - e^{\frac{-\left(s_{j-1}\right)^2}{2}} \right). \tag{2.16}$$

For calculation of $\boldsymbol{R_{ry}}$ we start by calculating the mean square quantization error ϵ_m as defined in (2.12). All these terms in (2.12) can be calculated from the results in (2.15) and (2.16) and the fact that the variance of $\mathbb{E}[\Re\{[\boldsymbol{y}]_m\}^2]$ is also known. As all $[\boldsymbol{y}]_m$ are circular symmetric complex Gaussian ϵ_m is the same for the real and imaginary part. Combining (2.12) with (2.14) and the circular symmetry of all random variables leads to

$$\mathbb{E}[[\boldsymbol{y}]_m[\boldsymbol{q}]_m^*] = -2\epsilon_m, \tag{2.17}$$

where we used the fact that the real and imaginary part are independent. As shown in [52] as long as (2.14) holds we can calculate $\mathbb{E}[[\boldsymbol{y}]_m[\boldsymbol{q}]_l^*] \; \forall m \neq l$ as

$$\mathbb{E}[[\boldsymbol{y}]_m[\boldsymbol{q}]_l^*] = -2\epsilon_m\mathbb{E}[[\boldsymbol{y}]_m[\boldsymbol{y}]_l^*], \tag{2.18}$$

where again we used the circular symmetry. Combining the results (2.17) and (2.18) we get:

$$\boldsymbol{R_{yq}} = -\boldsymbol{E}\boldsymbol{R_{yy}}, \tag{2.19}$$

where $\boldsymbol{E}$ is a diagonal matrix with the entries $[\boldsymbol{E}]_{m,m} = 2\epsilon_m$. Using these equations, we can calculate $\boldsymbol{R_{ry}}$ as:

$$\boldsymbol{R_{yr}} = \boldsymbol{R_{yy}} + \boldsymbol{R_{yq}} = (\boldsymbol{I} - \boldsymbol{E})\boldsymbol{R_{yy}}. \tag{2.20}$$

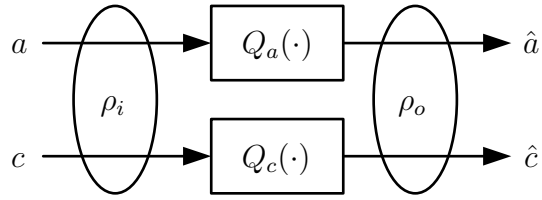

Fig. 2.3. Relationship between input and output correlation coefficients ρ_i and ρ_o.

For the calculation of $\boldsymbol{R_{rr}}$ we first show how to calculate the correlation after the quantizer. To calculate the output covariance matrix of the signal from the covariance matrix before or vice versa it is necessary to perform the calculation per element of the input or output vector. Since for the calculation of rates these calculations need to be executed many times we also show the approach we have taken to generate a table of the required numeric integration, thus dramatically reducing the computational complexity. In the last part of this section we also show how the results of our calculation compares to Monte-Carlo based estimation of the covariance matrix.

Fig. 2.3 shows the relationship of the two real valued, zero mean Gaussian random variable a and c. The inputs of the two quantization function $Q_a(\cdot)$ and $Q_c(\cdot)$ are correlated with ρ_i and the output $\hat{a}$ and $\hat{c}$ with ρ_o. The input and output correlation coefficients ρ_i and ρ_o are defined as:

$$\rho_i = \frac{\mathbb{E}[ac]}{\sigma_a \sigma_c} \text{ and } \rho_o = \frac{\mathbb{E}[\hat{a}\hat{c}]}{\sigma_{\hat{a}} \sigma_{\hat{c}}}, \tag{2.21}$$

where σ_a and σ_c are the standard deviation of the input variables a and c and $\sigma_{\hat{a}}$ and $\sigma_{\hat{c}}$ are the standard deviation of the output variables $\hat{a}$ and $\hat{c}$.

The equation proved for Gaussian random variables in [70] is:

$$\frac{\partial^k \rho_o}{\partial \rho_i^k} = \frac{\sigma_a \sigma_c}{\sigma_{\hat{a}} \sigma_{\hat{c}}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} Q_a^{(k)}(a) Q_c^{(k)}(c) f_{a,c}(a,c)\, da\, dc, \tag{2.22}$$

where $Q_a^{(k)}$ and $Q_c^{(k)}$ are the kth derivative of the non-linear processing functions. $f_{ac}(a,c)$ is the joint probability density function of the variables a and c. The additional normalization factor before the integral needs to be added as in contrast to [70] we use the common definition of a correlation factor. This equation is only valid if the variables a and c are jointly Gaussian. Since we are interested in a simple relationship, we use the first order partial derivative, which means k is equal to 1. In addition, we choose Q_a and Q_c to be the step functions of quantizers with N_a and N_c quantization levels. Since the quantization functions $Q_a(\cdot)$ and $Q_c(\cdot)$ are step functions, the derivative is a sum of Dirac impulses $\delta(x)$ at the position of the steps with the area equal to the step-size:

$$Q_a^{(1)}(x) = \sum_{l=1}^{N_a-1} \left(v_{l+1}^a - v_l^a\right) \delta(x - s_l^a)\ ,\ Q_c^{(1)}(x) = \sum_{j=1}^{N_c-1} \left(v_{j+1}^c - v_j^c\right) \delta(x - s_j^c), \tag{2.23}$$

where the symbols v_l^a, v_j^c, s_l^a and s_j^c represent the value (representative) of the quantization bin and the positions of the step. The difference between the representatives $\left(v_{l+1}^a - v_l^a\right)$ and $\left(v_{j+1}^c - v_j^c\right)$ are the heights of the steps.

The joint probability density function $f_{a,c}(a,c)$ of a and c with variance σ_a^2 and σ_c^2 is defined as:

$$f_{a,c}(a,c) = \frac{1}{2\pi\sigma_a\sigma_c\sqrt{1-\rho_i^2}} \exp\left(-\frac{1}{2(1-\rho_i^2)}\left[\frac{a^2}{\sigma_a^2} + \frac{c^2}{\sigma_c^2} - \frac{2\rho_i ac}{\sigma_a\sigma_c}\right]\right). \tag{2.24}$$

Because the functions $Q_a^{(1)}(x)$ and $Q_c^{(1)}(x)$ are only non-zero at the position of the Dirac impulses, the double integral in (2.22) is reduced to the double sum:

$$\frac{\partial \rho_o}{\partial \rho_i} = \frac{\sigma_a \sigma_c}{\sigma_{\hat{a}} \sigma_{\hat{c}}} \sum_{l=1}^{N_a-1} \left(v_{l+1}^a - v_l^a\right) \sum_{j=1}^{N_c-1} \left(v_{j+1}^c - v_j^c\right) f_{ac}(s_l^a, s_j^c). \tag{2.25}$$

With the initial condition $\rho_i = 0$ at $\rho_o = 0$, the output correlation coefficient ρ_o for a specific input correlation coefficient ρ_i is:

$$\rho_o = \frac{\sigma_a \sigma_c}{\sigma_{\hat{a}} \sigma_{\hat{c}}} \int_0^{\rho_i} \sum_{l=1}^{N_a-1} \left(v_{l+1}^a - v_l^a\right) \sum_{j=1}^{N_c-1} \left(v_{j+1}^c - v_j^c\right) f_{ac}(s_l^a, s_j^c)\, d\rho_i'. \tag{2.26}$$

This equation has in general no closed form solution and must be evaluated numerically. But if the signals a and c are quantized with a 1 bit uniform quantizer with its only step at 0, the previous formula is reduced to

$$\begin{aligned} \rho_o &= \frac{\sigma_a \sigma_c}{\sigma_{\hat{a}} \sigma_{\hat{c}}} \frac{\Delta_a \Delta_c}{2\pi \sigma_a \sigma_c} \int_0^{\rho_i} \frac{1}{\sqrt{1-\rho_i'^2}}\, d\rho_i', \\ \rho_o &= \frac{\Delta_a \Delta_c}{2\pi \sigma_{\hat{a}} \sigma_{\hat{c}}} \int_0^{\rho_i} \frac{1}{\sqrt{1-\rho_i'^2}}\, d\rho_i', \end{aligned} \tag{2.27}$$

where Δ_a and Δ_c are the step-sizes of the quantizers. For this equation the closed form solution is:

$$\rho_o = \frac{\Delta_a \Delta_c}{2\pi \sigma_{\hat{a}} \sigma_{\hat{c}}} \sin^{-1}(\rho_i). \tag{2.28}$$

Assuming that the input variables a and c are zero mean, unit variance Gaussian distributed and a uniform quantizer with minimum variance according to [51] is used, the step-size of quantizer is equal to $\Delta_a = \Delta_c = 2\sigma_{\hat{a}} = 2\sigma_{\hat{a}}$. Therefore, the formula is reduced to:

$$\rho_o = \frac{2}{\pi} \sin^{-1}(\rho_i), \tag{2.29}$$

for this special case. In the following we are going to verify the derived relationship by a simulation. We also show the evaluation for the cases that can only be evaluated numerically.

For the evaluation in Fig. 2.4 the minimum distortion uniform quantizers for Gaussian input symbols according to [51] are used. Fig. 2.4 shows that the more quantization bits we use the more the function converges to a straight line with slope 1, which means that the correlation at the input is the same as the correlation at the output. If we use a minimum distortion uniform quantizer, the correlation coefficient perturbation due to quantization is not noticeable if a quantizer with more than 3 bit resolution is used. As the curves are relatively smooth, we reduce the computational complexity, by storing some values in a look up table and interpolating with cubic splines in between them.

To calculate the whole matrix $\boldsymbol{R_{rr}}$, we first use the formula described in (2.15) to calculate the diagonal elements. Afterwards, to calculate the off-diagonal elements $\boldsymbol{R_{rr}}$ for each possible combination of real or imaginary part of the ith element of $\boldsymbol{r}$ with the real or imaginary part of the jth element of $\boldsymbol{r}$ unequal to i, the equation in (2.26) is solved. From (2.26) it is obvious that

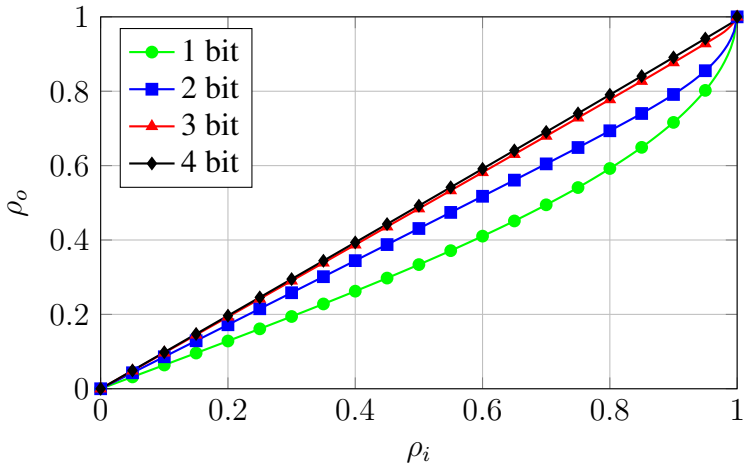

Fig. 2.4. Correlation relationship for different resolution for a uniform quantizer.

except for the normalization regarding the power the formula stays the same, if the quantizer is the same for both input values. Thus, a lookup table for each combination of all possible quantizers is generated. This lookup table is then used to calculate the actual correlation after the quantization. Since this calculation is assuming a real valued signa, we need to calculate it for each combination of real and imaginary part in $\boldsymbol{r}$.

Now we have all the necessary to calculate $\boldsymbol{S}$ and $\boldsymbol{R_{ee}}$.

$$\boldsymbol{S} = \boldsymbol{D}^{-1}\boldsymbol{R_{ry}}\boldsymbol{R_{yy}}^{-1} = \boldsymbol{D}^{-1}\left(\boldsymbol{I}-\boldsymbol{E}\right)\boldsymbol{R_{yy}}\boldsymbol{R_{yy}}^{-1} = \boldsymbol{D}^{-1}\left(\boldsymbol{I}-\boldsymbol{E}\right). \tag{2.30}$$

Therefore, we get for effective quantization distortion covariance $\boldsymbol{R_{ee}}$ the expression:

$$\boldsymbol{R_{ee}} = \boldsymbol{D}^{-1}\left(\boldsymbol{R_{rr}} - (\boldsymbol{I}-\boldsymbol{E})\boldsymbol{R_{yy}}(\boldsymbol{I}-\boldsymbol{E})\right)\boldsymbol{D}^{-1}. \tag{2.31}$$

For the final definition in this section we need to define the transformation matrix of the signal part $\boldsymbol{T}$:

$$\boldsymbol{T} = \boldsymbol{D}\boldsymbol{S} = \boldsymbol{I}-\boldsymbol{E}. \tag{2.32}$$

and the effective covariance matrix $\boldsymbol{R}_{\eta_q\eta_q}$ of the quantization noise $\boldsymbol{\eta}_q$ as:

$$\boldsymbol{R}_{\eta_q\eta_q} = \boldsymbol{D}\boldsymbol{R_{ee}}\boldsymbol{D} = \boldsymbol{R_{rr}} - (\boldsymbol{I}-\boldsymbol{E})\boldsymbol{R_{yy}}(\boldsymbol{I}-\boldsymbol{E}). \tag{2.33}$$

This effective quantization noise $\boldsymbol{\eta}_q$ is use in the analysis of system with quantization in the second part of this work (Chapter 3) and is defined based on the definition of an equivalent quantization model in the following way:

$$\boldsymbol{r} = \boldsymbol{T}\boldsymbol{y} + \boldsymbol{\eta}_q, \tag{2.34}$$

where $\boldsymbol{y}$ and $\boldsymbol{n}$ are the received signal before and after the quantization. To simplify the notation in the next chapter we define the calculations in this section as an operation with the input of the receive covariance matrix and the possibly different quantization functions at each antenna and the output of the transformation matrix and the effective quantization noise covariance matrix in the following way:

$$[\boldsymbol{T}\boldsymbol{R}_{\eta_q\eta_q}] = \mathrm{TF}\left(\boldsymbol{R_{yy}}, Q_1(\cdot), \cdots, Q_M(\cdots)\right), \tag{2.35}$$

where $Q_m(\cdot)$ are the quantization functions for each of the M elements of $\boldsymbol{y}$. It is important to note that in principal they could be different for each element of the vector.

2.2 Beam Design for Hybrid or Analog Beamforming

For future mmWave mobile communication systems, the use of analog/hybrid beamforming is envisioned to be an important aspect. The synthesis of beams is a key technology to enable the best possible operation during beam search and data transmission. The method for synthesizing beams developed in this work is based on previous work in radar technology considering phased array antennas. With this technique, it is possible to generate a desired beam of any shape with the constraints of the desired target transceiver antenna frontend. It is not constrained to a certain antenna array geometry, and can handle 1D, 2D and even 3D antenna array geometries, e.g. cylindrical arrays. The numerical examples show that the method can synthesize beams by considering a user defined trade-off between gain, transition width and passband ripples. Since this beam synthesis method is computational complex, it is only suitable for offline calculation during the design or calibration of a device.

To utilize the full potential of the system, it is essential that the beams of Transmitter (Tx) and Receiver (Rx) are aligned. Therefore, a trial and error procedure is used to align the beams of Tx and Rx [37, 41]. This beam search procedure does either utilize beams of different width with additional feedback or many beams of the same width with only one feedback stage [68]. In both cases the beams with specific width, maximum gain and flatness need to be designed.

Based on requirements on the beam shape, this work formulates an optimization problem similar to [86, 62]. Afterwards the optimization problem is solved numerically. This work includes the specific constraints of hybrid beamforming and low resolution phase shifters. In [68], the authors approximate a digital beamforming vector by a hybrid one. We generate our beam by approximating a desired beam instead.

For this evaluation we consider both subarray and fully-connected hybrid beamforming described in Subsection 1.1.1.2. Since the necessary calculations are slightly different all variables used in the context of subarray hybrid beamforming have a superscript s and the ones for fully-connected hybrid beamforming have superscript f. A part of the work in this section has also been published in [78].

2.2.1 Optimum Beam Synthesis

In the following we will develop a strategy to synthesize arbitrary beams based on the formulation of an optimization problem. Furthermore, we show how constraints can be used to model the restrictions of different systems. The array factor $A(\boldsymbol{u}, \boldsymbol{b})$ of an antenna array [13, page 294] is defined as

$$A(\boldsymbol{u}, \boldsymbol{b}) = \boldsymbol{b}^T \boldsymbol{p}(\boldsymbol{u}) \ , \ [\boldsymbol{p}(\boldsymbol{u})]_m = e^{j\frac{2\pi}{\lambda} d_m(\boldsymbol{u})}, \tag{2.36}$$

where $\boldsymbol{b}$ is the beamforming vector, $\boldsymbol{u}$ is the spatial direction combining the azimuth and elevation angle. The scalar $d_m(\boldsymbol{u})$ is the distance from the location of antenna element m to the plane defined by the normal vector $\boldsymbol{u}$ and a reference point. A common choice for the reference point is the position of the first antenna, in this case $d_1(\boldsymbol{u}) = 0$. An example with two antennas is illustrated in Fig. 2.5.

The objective of synthesizing an arbitrary beam pattern can be formulated as a weighted L^p norm between the desired pattern $D(\boldsymbol{u})$ and the absolute value of the actual array factor $|A(\boldsymbol{u}, \boldsymbol{b})|$

$$f(\boldsymbol{b}) = \left(\int W^p(\boldsymbol{u}) \, ||A(\boldsymbol{u}, \boldsymbol{b})| - D(\boldsymbol{u})|^p \, d\boldsymbol{u} \right)^{\frac{1}{p}}, \tag{2.37}$$

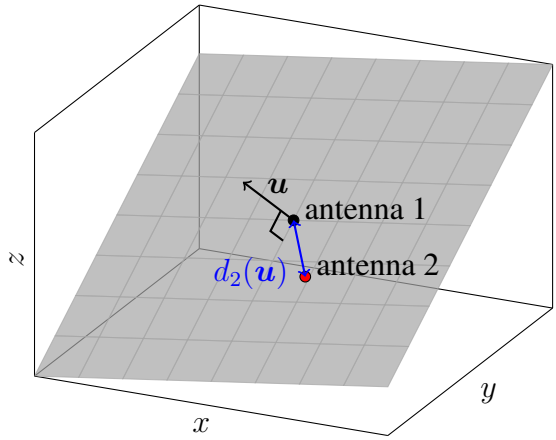

Fig. 2.5. Illustration of the definition of $\boldsymbol{u}$ and $d_m(\boldsymbol{u})$, with antenna 1 being used as the reference point.

where $W(\boldsymbol{u})$ is the weighting. $D(\boldsymbol{u})$ represents the desired absolute value of the array factor in a certain direction $\boldsymbol{u}$. Thus, $D(\boldsymbol{u})$ is always real and greater or equal than zero. This objective function itself is convex over its domain, but the constraints on $\boldsymbol{b}$ shown in the following subsections lead to a non-convex optimization problem. In the case of finite resolution phase shifters, it is even an mixed integer problem. This problem formulation ignores the phase of the array factor, since we require only the magnitude to be of a specific shape. By only optimizing over the array factor we do not take the pattern of the antennas into account. As described in [86] to account for an antenna pattern it is only necessary to divide $D(\boldsymbol{u})$ and $W(\boldsymbol{u})$ by the pattern of the antenna elements.

For many Non-Linear Programing (NLP) solvers it is of advantage to analytically calculate the gradient of the objective function with respect to its parameters. For the shown objective function, the gradient is calculated as (derivation is described in Annex A):

$$\begin{aligned}\nabla f(\boldsymbol{b}) = (f(\boldsymbol{b}))^{1-p} \int W^p(\boldsymbol{u}) \, ||A(\boldsymbol{u},\boldsymbol{b})| - D(\boldsymbol{u})|^{p-1} \\ \operatorname{sgn}\left(|A(\boldsymbol{u},\boldsymbol{b})| - D(\boldsymbol{u})\right) \nabla |A(\boldsymbol{u},\boldsymbol{b})| \, d\boldsymbol{u}.\end{aligned} \tag{2.38}$$

The gradient $\nabla |A(\boldsymbol{u},\boldsymbol{b})|$ depends as shown in Annex A on $\nabla A(\boldsymbol{u},\boldsymbol{b})$. Thus, it also depends on the actual parameters of the array and is therefore different for each of the array types.

To compare with other works in this area most of the following section is considering the transmitter beamforming. However, we also show how the problem formulation can be applied to receiver beamforming. We consider two different hybrid beamforming designs, which are the systems currently considered in literature [68, 75]. In the first case, all M antennas are divided into groups of size M_C. Each subgroup consists of one Digital-to-Analog-Converter (DAC) chain, an M_C signal splitter followed by a phase shifter and a Power Amplifier (PA) at each antenna (see Fig. 2.6). In total there are M_{DAC} DAC chains. The vector $\boldsymbol{\alpha}^s$ contain the different relative power of the signal of different DAC chain. In many cases in order to balance the mainlobe and sidelobe performance small variations in this power can improve the overall performance. Therefore, these parameters are used as part of the optimization. This restricts the beamforming vector $\boldsymbol{b}$ to have the form

$$\boldsymbol{b} = \boldsymbol{W}^s \boldsymbol{\alpha}^s = \begin{bmatrix} \boldsymbol{w}_1^s & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{w}_2^s & \ddots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \boldsymbol{w}_{M_{\text{DAC}}}^s \end{bmatrix} \begin{bmatrix} \alpha_1^s \\ \alpha_2^s \\ \vdots \\ \alpha_{M_{\text{DAC}}}^s \end{bmatrix}, \tag{2.39}$$

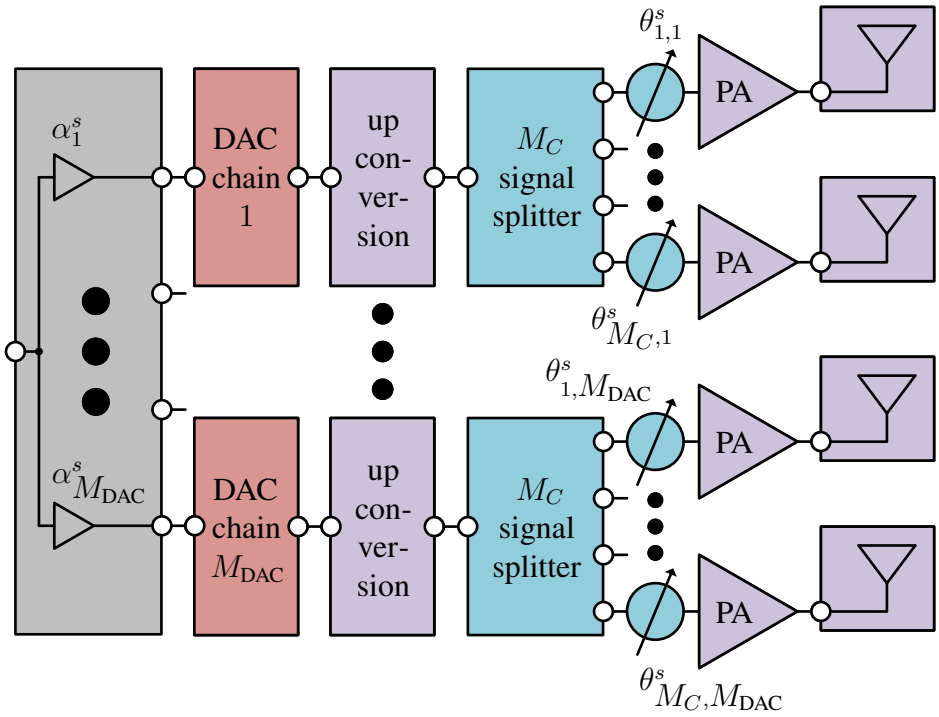

Fig. 2.6. Block diagram of a subarray hybrid beamforming transmitter.

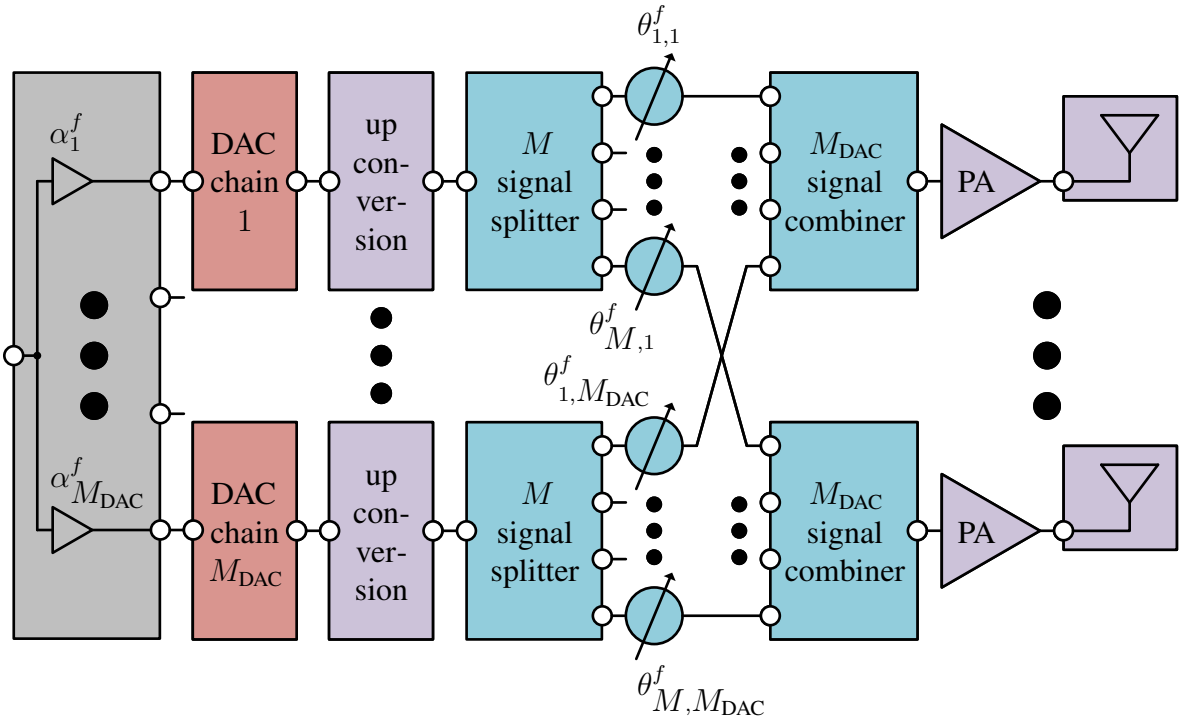

Fig. 2.7. Block diagram of a fully-connected hybrid beamforming transmitter.

where $\boldsymbol{\alpha}^s \in \mathbb{R}^{M_{\text{DAC}}}$ and the vectors $\boldsymbol{w}_i^s$ models the analog phase shifting of group i and therefore have the form

$$\boldsymbol{w}_i^s = \begin{bmatrix} e^{j\theta_{1,i}^s} & e^{j\theta_{2,i}^s} & \cdots & e^{j\theta_{M_C,i}^s} \end{bmatrix}^T. \tag{2.40}$$

The corresponding gradient of the array factor $A(\boldsymbol{u}, \boldsymbol{b})$ with respect to the degrees of freedom $\boldsymbol{\alpha}^s$ and $\boldsymbol{\theta}^s$ is calculated as

$$\begin{bmatrix} \nabla_{\boldsymbol{\alpha}^s} \\ \nabla_{\boldsymbol{\theta}^s} \end{bmatrix} A(\boldsymbol{u}, \boldsymbol{b}) = \begin{bmatrix} (\boldsymbol{W}^s)^T \boldsymbol{p}(\boldsymbol{u}) \\ j(\boldsymbol{\alpha}^s \otimes \mathbf{1}_{M_C}) \circ \boldsymbol{p}(\boldsymbol{u}) \circ \boldsymbol{r}^s \end{bmatrix}, \tag{2.41}$$

where the vector $\boldsymbol{r}^s \in \mathbb{C}^M$ is the column vector with all vectors $\boldsymbol{w}_i^s$ stacked on top of one another

$$\boldsymbol{r}^s = \begin{bmatrix} (\boldsymbol{w}_1^s)^T & (\boldsymbol{w}_2^s)^T & \cdots & (\boldsymbol{w}_{M_{\text{DAC}}}^s)^T \end{bmatrix}^T. \tag{2.42}$$

In this case j is the imaginary unit. The vector $\boldsymbol{\theta}^s \in \mathbb{R}^M$ contains the phase shift at each antenna:

$$\boldsymbol{\theta}^s = \begin{bmatrix} \theta_{1,1}^s & \theta_{2,1}^s & \cdots & \theta_{M_C,1}^s & \theta_{1,2}^s & \cdots & \theta_{M_C,2}^s & \cdots & \theta_{M_C,M_{\text{DAC}}}^s \end{bmatrix}^T. \tag{2.43}$$

In the second case, each of the DAC chain is connected to an M signal splitter followed by a phase shifter for each antenna. At each antenna, the phase shifted signal from each DAC chain is combined and then amplified by a PA followed by the antenna transmission. The corresponding transmitter font-end is shown in Fig. 2.6 and referred to as fully-connected hybrid beamforming. With this system architecture the beamforming vector $\boldsymbol{b}$ can be decomposed into

$$\begin{aligned} \boldsymbol{b} = \boldsymbol{W}^f \boldsymbol{\alpha}^f &= \begin{bmatrix} \boldsymbol{w}_1^f & \boldsymbol{w}_2^f & \cdots & \boldsymbol{w}_{M_{\text{DAC}}}^f \end{bmatrix} \boldsymbol{\alpha}^f \\ &= \begin{bmatrix} e^{j\theta_{1,1}^f} & e^{j\theta_{1,2}^f} & \cdots & e^{j\theta_{1,M_{\text{DAC}}}^f} \\ e^{j\theta_{2,1}^f} & e^{j\theta_{2,2}^f} & \cdots & e^{j\theta_{2,M_{\text{DAC}}}^f} \\ \vdots & \vdots & \ddots & \vdots \\ e^{j\theta_{M,1}^f} & e^{j\theta_{M,2}^f} & \cdots & e^{j\theta_{M,M_{\text{DAC}}}^f} \end{bmatrix} \begin{bmatrix} \alpha_1^f \\ \alpha_2^f \\ \vdots \\ \alpha_{M_{\text{DAC}}}^f \end{bmatrix}, \end{aligned} \tag{2.44}$$

with $\boldsymbol{\alpha}^f \in \mathbb{R}^{M_{\text{DAC}}}$. The gradient of the array factor with respect to the parameters $\boldsymbol{\alpha}^f$ and $\boldsymbol{\theta}^f$ is then:

$$\begin{bmatrix} \nabla_{\boldsymbol{\alpha}^f} \\ \nabla_{\boldsymbol{\theta}^f} \end{bmatrix} A(\boldsymbol{u}, \boldsymbol{b}) = \begin{bmatrix} (\boldsymbol{W}^f)^T \boldsymbol{p}(\boldsymbol{u}) \\ j(\boldsymbol{\alpha}^f \otimes \boldsymbol{p}(\boldsymbol{u})) \circ \boldsymbol{r}^f \end{bmatrix}, \tag{2.45}$$

with $\boldsymbol{r}^f \in \mathbb{C}^{M \cdot M_{\text{DAC}}}$ being the concatenation of all phase shifting vectors $\boldsymbol{w}_i^f$

$$\boldsymbol{r}^f = \begin{bmatrix} (\boldsymbol{w}_1^f)^T & (\boldsymbol{w}_2^f)^T & \cdots & (\boldsymbol{w}_{M_{\text{DAC}}}^f)^T \end{bmatrix}^T. \tag{2.46}$$

The vector $\boldsymbol{\theta}^f \in \mathbb{R}^{M \cdot M_{\text{DAC}}}$ contains the phase shift of each antenna DAC chain combination

$$\boldsymbol{\theta}^f = \begin{bmatrix} \theta_{1,1}^f & \theta_{2,1}^f & \cdots & \theta_{M,1}^f & \theta_{1,2}^f & \cdots & \theta_{M,2}^f & \cdots & \theta_{M-1,M_{\text{DAC}}}^f & \theta_{M,M_{\text{DAC}}}^f \end{bmatrix}^T. \tag{2.47}$$

To limit the maximum output power of the PAs, we need to include the following constraints

$$|[\boldsymbol{b}]_m| \leq 1 \; \forall m. \tag{2.48}$$

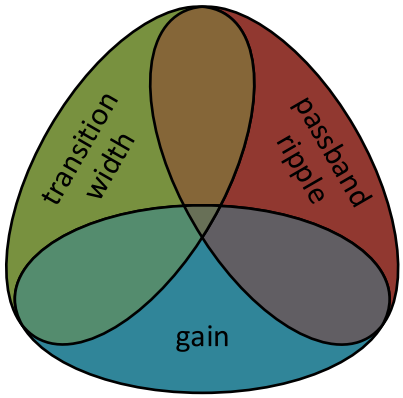

Fig. 2.8. Illustration of the trade-off associated with the beam pattern synthesis.

It is important to keep in mind that this restriction is after the hybrid beamforming, therefore, it is a nonlinear constraint restricting output-power of the PA. Another way to bound the output power is a sum power constraint of the form

$$||\boldsymbol{b}||_2^2 \leq 1. \tag{2.49}$$

If we design beams for a receiver there is no power constraint on the beamforming vector $\boldsymbol{b}$, but rather only the gain difference of α_i^s in the digital domain have an influence. Here it is important to keep in mind that an additional average gain does not increase the SNR, as the noise would be amplified as well.

It is also possible that the resolution of the phase shifters is limited. This means that the values of $\theta_{m,i}^s$ are from a finite set of possibilities

$$\theta_{m,i}^s = -\pi + k_{m,i}\frac{2\pi}{K} \;\; \forall m, i \text{ and } k_{m,i} \in \{0,\ 1,\ \cdots,\ K-1\}, \tag{2.50}$$

where K is the number of possible phases. A possible phase shift in the digital domain need to be taken into account. In the case without finite phase resolution phase shifters, this is redundant with the analog phase shift and thus not taken into account. Therefore, in addition to the scaling $\boldsymbol{\alpha}^f$ or $\boldsymbol{\alpha}^s$, we need to take a phase shift $\boldsymbol{\xi}^f$ or $\boldsymbol{\xi}^s$ into account. For the case of subarray hybrid beamforming with limited resolution RF phase shifters the beamforming vector $\boldsymbol{b}$ takes the form

$$\boldsymbol{b} = \boldsymbol{W}^s \left(\boldsymbol{\alpha}^s \circ \boldsymbol{\xi}^s\right), \tag{2.51}$$

where $\boldsymbol{\xi}^s$ are the digital phase shifts defined as

$$\boldsymbol{\xi}^s = [e^{j\xi_1^s}, e^{j\xi_2^s}, \cdots, e^{j\xi_{M_{\text{DAC}}}^s}]^T. \tag{2.52}$$

The formulation for the fully-connected case does also contain addition phase shifts in the same fashion as the subarray hybrid beamforming system

$$\boldsymbol{b} = \boldsymbol{W}^f \left(\boldsymbol{\alpha}^f \circ \boldsymbol{\xi}^f\right). \tag{2.53}$$

Combining the objective function with the constraints associated with the hardware capabilities lead to the following optimization problem

$$\begin{aligned} &\min f(\boldsymbol{b}) \\ &\text{s.t. } \boldsymbol{g}(\boldsymbol{b}) \leq \boldsymbol{0}, \end{aligned} \tag{2.54}$$

where $\boldsymbol{g}(\boldsymbol{b})$ are the combination of all constraints, that model the desired hardware capabilities. In addition, to the constraints introduced in the preceding paragraph, it is possible that there are other additional constraints on the beamforming capabilities of the system. The first case in the next subsection is considering subarray hybrid beamforming with individual power constraint. For this case the problem would be a specific example of (2.54) described by the following problem formulation:

$$\begin{aligned} &\min f(\boldsymbol{b}) \\ &\text{s.t. } \boldsymbol{b} = \boldsymbol{W}^s \boldsymbol{\alpha}^s, \\ &|[\boldsymbol{b}]_m| \leq 1 \; \forall m \end{aligned} \tag{2.55}$$

It is important to mention that the beam synthesis procedure is similar to digital filter design, therefore, we us the terminology of digital filter design. The weighting $W(\boldsymbol{u})$, the desired pattern $D(\boldsymbol{u})$ and the choice of p in $f(\boldsymbol{b})$, determine which point in the trade-off of the parameters gain, passband ripple and transition width is the target. This trade-off is illustrated in Fig. 2.8.

The case of U beams being used at the same time can be formulated as:

$$\begin{aligned} &\min \sum_{u=1}^{U} f(\boldsymbol{b}_u) \\ &\text{s.t. } \boldsymbol{g}(\boldsymbol{b}_1, \cdots, \boldsymbol{b}_U) \leq \boldsymbol{0}. \end{aligned} \tag{2.56}$$

Essentially, we sum up the objective function for all different beams. The case of a subarray hybrid beamforming transmitter serving two users is illustrated in Fig. 2.9. The case of fully-connected hybrid beamforming is not illustrated. This case would consist of the digital processing block (grey box) in Fig. 2.9 replacing the corresponding digital processing in Fig. 2.7. Since analog beamforming $\boldsymbol{W}$ needs to be the same for each beam u we get the following expression for each beamforming vector

$$\boldsymbol{b}_u = \boldsymbol{W}^{su} \left(\boldsymbol{\alpha}^{su} \circ \boldsymbol{\xi}^{su} \right). \tag{2.57}$$

It is important to note that since for this case we have multiple beams it is important to also consider a phase shift $\boldsymbol{\xi}^s$ in the digital domain for different beams. If we instead would use a fully-connected hybrid beamforming system the individual beamforming vectors have the following expression:

$$\boldsymbol{b}_u = \boldsymbol{W}^{fu} \left(\boldsymbol{\alpha}^{fu} \circ \boldsymbol{\xi}^{fu} \right). \tag{2.58}$$

By adjusting the weighting for each beam $W_u(\boldsymbol{u})$ separately, it is possible to give different priorities to the optimization of different beams. As each beam is having a different shape also the desired beam pattern $D_u(\boldsymbol{u})$ is defined per beam. It is also important that in this case the power constraints need to consider the sum of the power for each beam. The individual per element power constraints have the expression

$$\left| [\sum_{u=1}^{U} \boldsymbol{b}_u]_m \right| \leq 1 \; \forall m. \tag{2.59}$$

In the same fashion as in the single beam case we can also formulate a sum power constraint

$$\left|\left| \sum_{u=1}^{U} \boldsymbol{b}_u \right|\right|_2^2 \leq 1. \tag{2.60}$$

The finite resolution constraint is formulated the same way as in the single beam case, because the analog beamforming is shared among all beams.

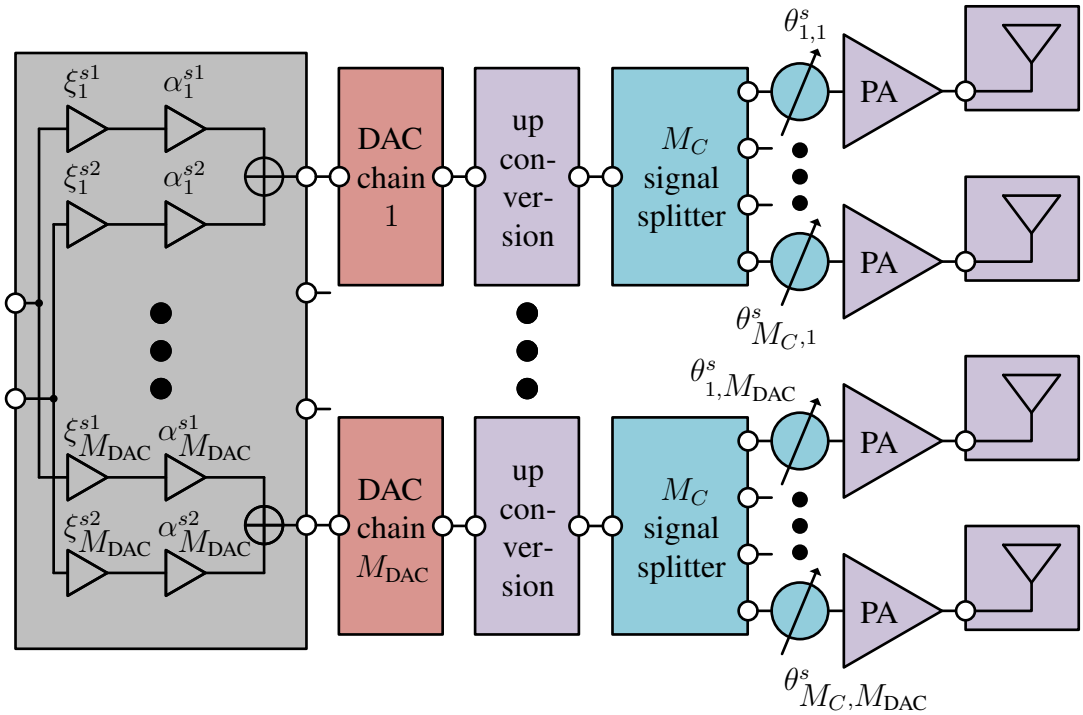

Fig. 2.9. Block diagram of a subarray hybrid beamforming transmitter serving two users simultaneously.

2.2.2 Numerical Results Beam Synthesis

To compare the designed beams, we need to first define some metrics to quantify the difference between them. Some of these metrics are similar to the ones defined in [25]. The first one is the *average gain* in the desired direction. Directly connected to the average gain is the *maximum ripple* of the array factor in the desired directions. For more reliable results, the transition region is excluded from the search of the maximum ripple. A very important criterion to evaluate the performance of a beam for initial access is the *overlap of adjacent beams* of the same width. Here we evaluate the area at which the gain difference between two beams is less than 5 dB, relative to the total area of one beam. The last measure is the *maximum sidelobe* relative to the average gain in the desired directions. In this case the maximum sidelobe is used as this represents the most likely case that a wrong beam is selected during the beam-training. From the formulation of the problem in (2.37) it is obvious that dependent on the selection of p the average sidelobe is weighted more or less relative to the maximum sidelobe. All these measures are illustrated on a beam example in Fig. 2.10. The gain in this figure is defined as $10\log_{10} A(\boldsymbol{u}, \boldsymbol{b})$.

In the following, beams synthesized by the described method are shown. For all systems, the transmitter is equipped with $M_{\text{DAC}} = 4$ DAC chains, connected to 64 Antenna elements, forming an Uniform Linear Array (ULA) with half-wavelength inter-element spacing. Since the antenna array is one dimensional, it is sufficient to look at only one spatial dimension. All plots refer to angle $\psi = \pi \sin(\phi)$, where ϕ is the geometric angle between a line connecting all antennas and the direction of a planar wavefront.

For each system, three beams of width $BW = \pi, \pi/2, \pi/4$ are synthesized. In contrast to the beams in Fig. 2.11 and Fig. 2.12, the beams in Fig. 2.13 and Fig. 2.14 are designed to be used in a multi-beam setup simultaneously. For an ULA, the spatial direction $\boldsymbol{u}$ is fully represented by ψ, therefore $W(\boldsymbol{u})$, $D(\boldsymbol{u})$ and $A(\boldsymbol{u}, \boldsymbol{b})$ depend only on ψ. Since the magnitude of each element of $\boldsymbol{b}$

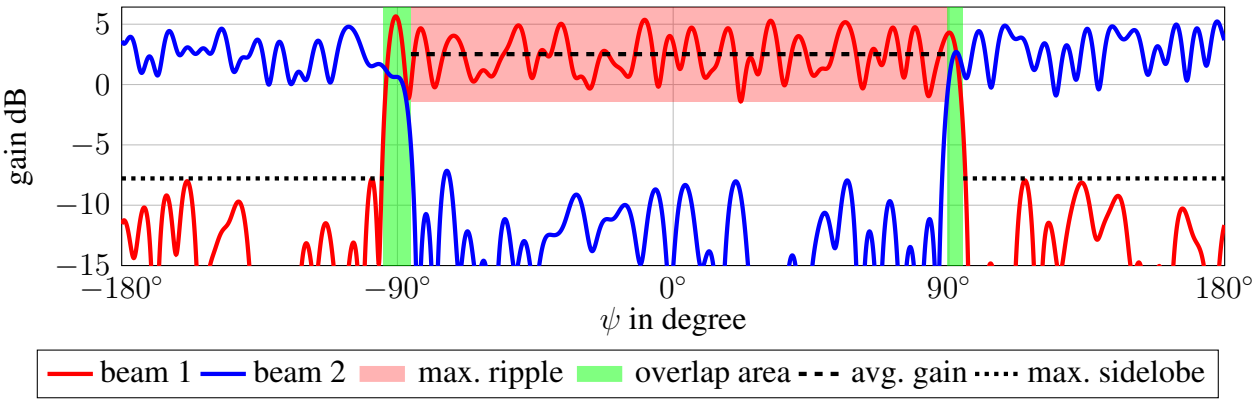

Fig. 2.10. Illustration of the beam comparison metrics.

is less or equal to one, if a perfect flat beam without sidelobes could be constructed, it would have the array-factor $D_{\text{max}} = \sqrt{M 2\pi / BW}$. As also described in [86], such a beam cannot be realized, therefore $D(\psi)$ is equal to βD_{max} at the desired directions and equal to zero, elsewhere. The parameter β ensures the feasibility of a solution.

The weighting of different parts of the beam pattern $W(\psi)$ is uniformly set to 1, except for a small transition region enclosing the desired directions. For all following optimizations we set $p = 4$ in the objective function. This does ensure that the gain in the desired directions as well as the sidelobes have roughly equal ripples. The integral of the objective function over all spatial directions in the objective function is approximated by a finite sum. To ensure a sufficient approximation, the interval is split into 512 elements. As described in [86], the computational complexity can be significantly reduced by reformulating the problem to use Fast Fourier Transformation (FFT)/Inverse Fast Fourier Transformations (IFFTs) to calculate $A(\psi, \boldsymbol{b})$ and the derivatives of the objective function.

For each system, the optimization process was started using several different random initializations. Since the used NLP and Mixed Integer Non-Linear Programing (MINLP) solvers only guarantee to find a local minimum for a non-convex problem, the results were compared and the implementation having to the minimum objective function was selected. Since these solvers are very computational complex and are run for multiple initializations, the overall necessary calculations prohibit an online calculation based on channel measurements. However, for the task of beam training a beam codebook can be offline calculated and stored. The metrics to compare the performance of different beams is shown in Table 2.1 and 2.2 alongside a reference to the respective figures.

The graphs in Fig. 2.11 and 2.12 show the synthesized beams for subarray and fully-connected hybrid beamforming with a per antenna power constraint of one and without resolution constraints on the phase shifters. The gain in the graphs are relative to the average gain given in Table 2.1, which is different for each plot. For (a), (b) and (c) the gain penalty β was selected to be 3 dB, 2 dB and 2 dB, respectively. Compared to the fully-connected case, subarray hybrid beamforming is characterized by more gain ripples and higher sidelobe energy, while having the same transition width.

In Fig. 2.13 and 2.14 fully-connected hybrid beamforming with quantized phase shifters was applied. The beams are designed with the method described in Fig. 2.14. The beams in both figures are optimized to simultaneously transmit both shown beams at each stage (a), (b) and (c). The power constraint for these cases are also different, only the sum power is constrained to be less or

Table 2.1. Comparison of the metrics of the beams designed with individual power constraint.

Beam	avg. gain dB	max. ripple dB	overlap in %	max. sidelobe dB
Fig. 2.11 (a)	18.20	4.00	2.44	-17.40
Fig. 2.11 (b)	21.70	2.89	3.22	-16.20
Fig. 2.11 (c)	26.30	2.76	7.21	-16.30
Fig. 2.12 (a)	18.20	2.04	2.63	-22.60
Fig. 2.12 (b)	22.00	2.10	2.63	-22.80
Fig. 2.12 (c)	24.80	2.35	5.26	-23.30

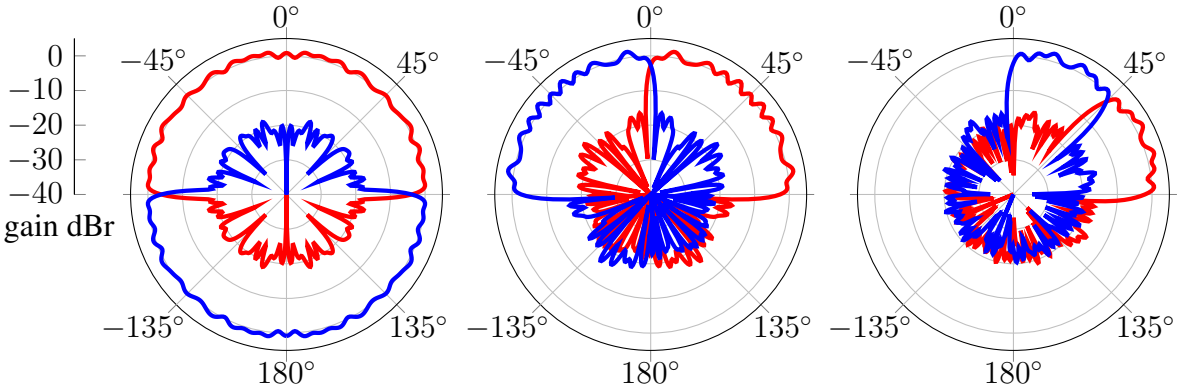

Fig. 2.11. Beams of different width of a subarray hybrid beamforming array.

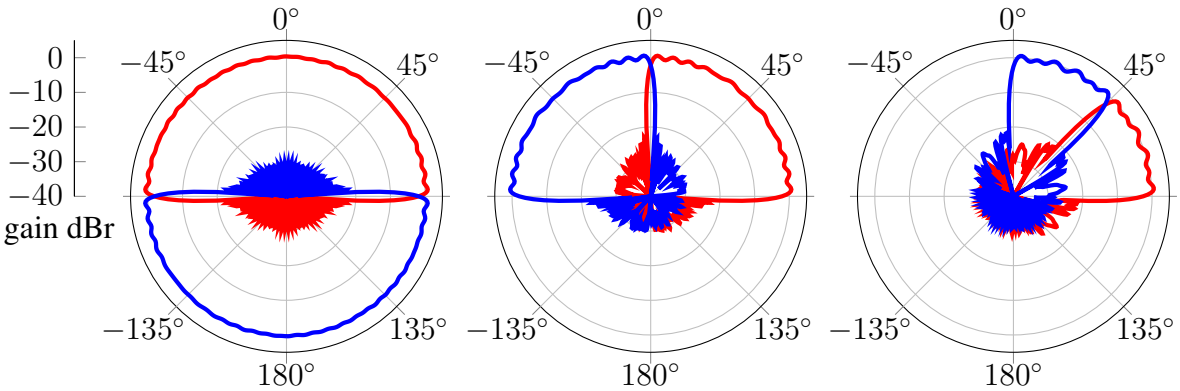

Fig. 2.12. Beams of different width of a fully-connected hybrid beamforming array.

Table 2.2. Comparison of the metrics of the beams designed with sum power constraint.

Beam	avg. gain dB	max. ripple dB	overlap in %	max. sidelobe dB
Fig. 2.13 (a)	2.52	3.90	7.66	-10.30
Fig. 2.13 (b)	5.50	3.01	6.54	-10.10
Fig. 2.13 (c)	8.23	1.47	6.63	-12.70
Fig. 2.14 (a)	2.22	8.82	34.40	-2.16
Fig. 2.14 (b)	5.04	7.25	8.20	-4.04
Fig. 2.14 (c)	8.02	1.49	14.40	-8.97

equal to one.

In Fig. 2.14, and, especially in (a) there are multiple direction where both beams have a similar power. In these directions an estimation of the link quality achieved with both beams is going to be very similar. This can possibly lead to an ambiguous decision and, in turn, to large errors in a multi-stage beam training procedure. On the contrary, the solution evaluated in Fig. 2.13 offers a sharper transition. The sidelobes are also close to uniform, thus enable predictable performance across all directions. The only disadvantage is the larger ripples inside the main beam.

The shortcomings which are observed in Fig. 2.14 are introduced during the generation of $\boldsymbol{b}$. As described in [68] this method approximates a version of $\boldsymbol{b}_d$ generated with the assumption of full digital beamforming. Since for a low number of DAC chains this vector cannot be well approximated, the resulting beam pattern does not correspond well to the desired one. It is also important to mention that there is no one-to-one mapping between the error in approximating $\boldsymbol{b}_d$ and the errors of the corresponding beam. As shown in [68], the method works well if $\boldsymbol{b}_d$ can be well approximated by a larger number of DAC chains.

With the results in Fig. 2.15 we can show that the problem formulation is flexible enough to design two beams that optimally serve two users at the same time. This is highlighting an extreme case with two relatively narrow beams that have a significant difference in the main desired direction. In this case we used the same formulation of the optimization problem, except for the different constraints of subarray and fully-connected hybrid beamforming. This result does showcase that

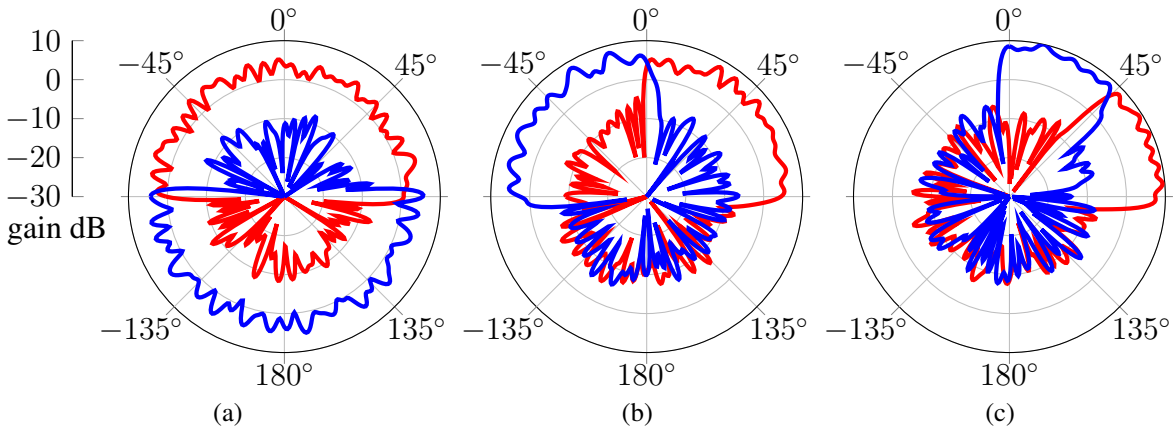

Fig. 2.13. Beams of different width optimized for sidelobe attenuation and with 2 bit quantization of the phase shifters of a fully-connected hybrid beamforming.

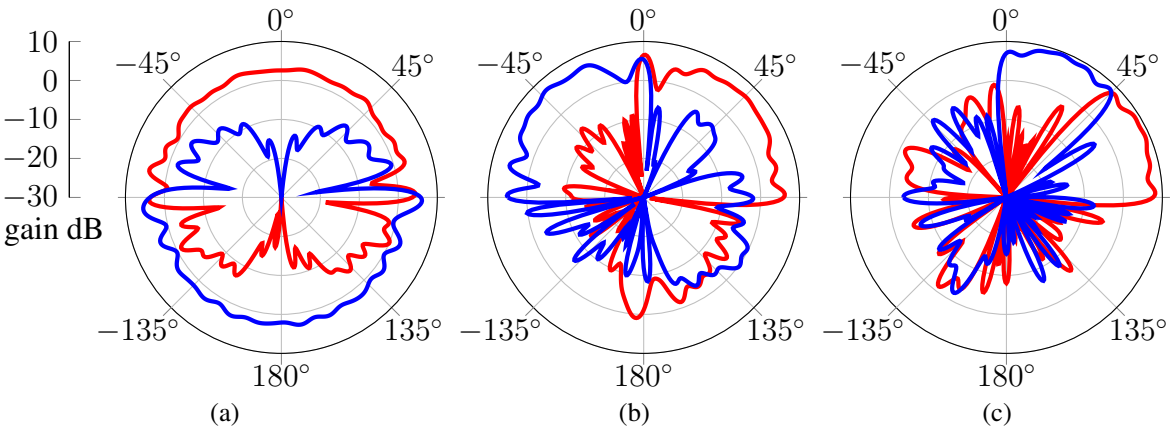

Fig. 2.14. Beams of different width of fully-connected hybrid beamforming array with phase quantization according to [68].

the additional degrees of freedom available for fully-connected hybrid beamforming can be used by the optimization to reduce the sidelobes and smoothen the beam in the desired directions.

2.2.3 Conclusion Beam Design

The developed approach can synthesize any beam-pattern for hybrid beamforming systems. The numerical examples show that a sufficient solution to the underlying optimization problem can be found with high computational complexity. The numeric examples also demonstrated that it is possible to adapt the approach to any type of constraints arising in the context of hybrid beamforming and wireless communication. An interesting extension of this work would be to enable on-the-fly synthesis of beams by reducing the computational complexity of solving the optimization problem.

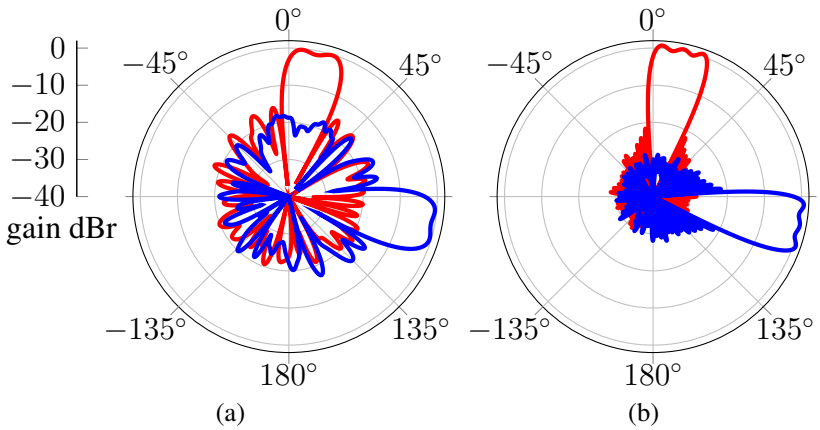

Fig. 2.15. Beams designed to serve multiple users using subarray hybrid beamforming (a) and fully-connected hybrid beamforming (b), otherwise the constraints are the same.

2.3 Calculation of Beam Training Parameters

As described in [3] and [2] a part of the channel quality feedback is based on Reference Signal Received Power (RSRP) measurements. Essentially, for this measurement the average received power of reference signals across the whole bandwidth is used to measure the channel quality. In this part of the work we also want to get a measure of the channel quality given a receiver beamforming configuration. For the following comparison it is also important to know how many reference symbols are necessary to reliably estimate the channel quality.

In this section we first introduce the signal model of the measurement. We continue by calculating the Probability Density Function (PDF) of the given measure. Then we present an algorithm calculating the minimum sequence length to achieve certain quality of the measurement. In the end of the section we also show the relationship between the chosen sequence length and the probability of not selecting the best beam in various scenarios.

2.3.1 Signal Model Beam Training Parameters

Assuming perfect time and frequency synchronization, the received reference signals after analog combining can be described as:

$$y[n] = \boldsymbol{w}^H \boldsymbol{h} x[n] + \boldsymbol{w}^H \boldsymbol{\eta}[n], \tag{2.61}$$

where $y[n]$, $\boldsymbol{h}$, $x[n]$, $\boldsymbol{\eta}[n]$ and $\boldsymbol{w}$ are the received signal, the channel, the transmit signal, the noise and the analog combiner vector, respectively. It is important to mention that in this case $\boldsymbol{w}$ is the analog combiner vector resulting in one digital signal. The details of all possible hybrid and analog beamforming architectures are described in Subsection 1.1.1.2. As the vector $\boldsymbol{w}$ models the analog combing via phase shifters, each entry has the same absolute value. For the derivation in this section we simplify the effective channel $\boldsymbol{w}^H \boldsymbol{h}$ to be a simple scalar value $\alpha e^{j\theta}$. The combination of the noise terms $\boldsymbol{w}^H \boldsymbol{\eta}[n]$ is also combined to form an effective, scalar noise $\eta[n]$. As we assume the entries of $\boldsymbol{\eta}[n]$ to be zero mean circular symmetric complex Gaussian, $\eta[n]$ is also zero mean circular symmetric complex Gaussian with a different variance.

With this we reach the following expression for the received signal assuming perfect time and frequency synchronization:

$$y[n] = \alpha e^{j\theta} x[n] + \eta[n]. \tag{2.62}$$

Since for the estimation of the channel quality the transmitter needs to transmit known reference signals, we know the symbols $x[n]$ inside transmitted reference sequence of length N. We can therefore multiply with $x^*[n]$ to a noisy estimated of the channel:

$$y[n]x^*[n] = \alpha e^{j\theta} x[n]x^*[n] + \eta[n]x^*[n]. \tag{2.63}$$

Without loss of generality we assume that $|x[n]|^2 = 1 \, \forall n \in \{0, \cdots, N-1\}$. Therefore, $\eta[n]x^*[n]$ has the same distribution as $\eta[n]$. To estimate the average receive power of the reference symbols we can combine the N symbols of the transmitted reference sequence the following metric is used:

$$\frac{1}{N^2} \left| \sum_{n=0}^{N-1} y[n]x^*[n] \right|^2 \approx \frac{1}{N^2} \left| \sum_{n=0}^{N-1} \alpha e^{j\theta} + \eta[n] \right|^2 = \frac{1}{N^2} \left| \sum_{n=0}^{N-1} \alpha e^{j\theta} + \eta[n] \right|^2 |e^{-j\theta}|^2, \tag{2.64}$$

where the approximation comes from the fact that $\eta[n]$ is circular symmetric, thus the distribution of $\eta[n]e^{-j\theta}$ is the same as $\eta[n]$. This means that this approximation results in a measure that has an

identical distribution. As for the subsequent evaluation we are only interested in the distribution, the calculated distribution is the same as the one for the original measure given in (2.64). As $|e^{-j\theta}|^2 = 1$, this equation can be further reduced to:

$$\frac{1}{N^2}\left|\sum_{n=0}^{N-1}\alpha+\eta[n]e^{-j\theta}\right|^2 \approx \frac{1}{N^2}\left|\sum_{n=0}^{N-1}\alpha+\eta[n]\right|^2, \tag{2.65}$$

where the approximation follows the same argument as the approximation in the previous equation. Since we want to calculate the channel quality in logarithmic scale, we arrive at the final results of the channel quality measure $\hat{s}$:

$$\hat{s} = 10\log\left(\left|\alpha+\sum_{n=0}^{N-1}\frac{\eta[n]}{N}\right|^2\right). \tag{2.66}$$

2.3.2 Calculation of the PDF of the Channel Quality Measure

To calculate the PDF of the channel quality measure $\hat{s}$ we start by noting that from the definition of $\hat{s}$ in (2.66) we see that the term inside the absolute value is in general a circular symmetric complex Gaussian random variable a with mean μ and variance $2\sigma^2$, which is equivalent to the sum of the real value squared plus the imaginary value squared:

$$|a|^2 = \Re(a)^2 + \Im(a)^2 = a_R^2 + a_I^2. \tag{2.67}$$

Since the real and imaginary part a_R and a_I have non-zero means μ_R and μ_I and in addition their variance σ^2 is also unequal to one, the overall random variable is neither non-central chi-square nor generalized chi-square and we need to derive the PDF ourselves. We take the following steps for deriving the PDF of $\hat{s}$:

- Derive PDF of the real and imaginary part squared
- Calculate the PDF of the combination of the real and imaginary value squared via convolution
- Calculate the PDF including the conversion into a logarithmic scale by the well known change of variable technique

Starting from a real valued Gaussian variable b with mean μ_b and variance σ_b^2 to get the PDF of the random variable $c = g(b) = b^2$ we cannot use the standard formula for a change of random variable as $g(b)$ is not invertible. The Cumulative Distribution Function (CDF) $F_c(c)$ of c be can be defined in terms of the PDF $f_b(b)$ of b in the following way

$$F_c(c) = \int_0^{\sqrt{c}} f_b(b)db + \int_{-\sqrt{c}}^{0} f_b(b)db, \tag{2.68}$$

where we included the aspect that the square of a real valued random variable is always positive, which in turn means that the PDF $f_c(c)$ is zero for values below zero. Since the PDF of a random variable is the derivative of the CDF we calculate $f_c(c)$ as:

$$f_c(c) = \frac{dF_c(c)}{dc} = f_b(\sqrt{c})0.5c^{-0.5} + f_b(-\sqrt{c})0.5c^{-0.5} = \frac{1}{2\sqrt{c}}\left(f_b(\sqrt{c}) + f_b(-\sqrt{c})\right). \tag{2.69}$$

Since b is Gaussian distributed by plugging in the definition of $f_b(b)$ we get the following expression:

$$f_c(c) = \frac{1}{2\sqrt{c}} \left(\frac{1}{\sqrt{2\pi\sigma_b^2}} e^{\frac{-(\sqrt{c}-\mu_b)^2}{2\sigma_b^2}} + \frac{1}{\sqrt{2\pi\sigma_b^2}} e^{\frac{-(-\sqrt{c}-\mu_b)^2}{2\sigma_b^2}} \right). \tag{2.70}$$

This can be simplified to

$$f_c(c) = \frac{1}{\sqrt{8\pi\sigma_b^2 c}} e^{\frac{-\left(c+\mu_b^2\right)}{2\sigma_b^2}} \left(e^{\frac{\sqrt{c}\mu_b}{\sigma_b^2}} + e^{\frac{-\sqrt{c}\mu_b}{\sigma_b^2}} \right) = \frac{1}{\sqrt{2\pi\sigma_b^2 c}} e^{\frac{-\left(c+\mu_b^2\right)}{2\sigma_b^2}} \cosh\left(\frac{\mu_b\sqrt{c}}{\sigma_b^2} \right). \tag{2.71}$$

it needs to be noted that this PDF is only valid for $c \geq 0$. For all $c < 0$ the value of the PDF $f_c(c)$ is equal to zero.

In the next step we calculate the pdf of $z = |a|^2 = c_R + c_I$, where $c_R = a_R^2$ and $c_I = a_I^2$ both having a PDF as described in (2.71) with the same value of σ but different values μ_R and μ_I. To calculate the PDF of z as the combination of two independent random variable we convolve the PDFs of the both random variables in the following way

$$f_z(z) = \int_{-\infty}^{\infty} f_{c_I}(z - c_R) f_{c_R}(c_R) dc_R. \tag{2.72}$$

As both PDFs $f_{c_R}(c_R)$ and $f_{c_I}(c_I)$ are equal to zero $\forall c_R < 0$ and $\forall c_I < 0$ we change the boundaries of the integration to only include non-zero values

$$f_z(z) = \int_{0}^{z} f_{c_I}(z - c_R) f_{c_R}(c_R) dc_R. \tag{2.73}$$

Plugging in the definition of the PDF in (2.71) into (2.73) and simplifying the expression we get

$$f_z(z) = \frac{1}{2\pi\sigma^2} e^{\frac{-(\mu_R^2+\mu_I^2)}{2\sigma^2}} e^{\frac{-z}{2\sigma^2}} \int_{0}^{z} \frac{1}{\sqrt{(z-c_R)c_R}} \cosh\left(\frac{\mu_I\sqrt{z-c_R}}{\sigma^2} \right) \cosh\left(\frac{\mu_R\sqrt{c_R}}{\sigma^2} \right) dc_R. \tag{2.74}$$

Since the integral has no closed form solution, we need to evaluate it numerically.

In the final step we also need to include the transformation to a logarithmic scale. Since now the function $\hat{s} = g(z) = 10\log(z)$ is invertible we can use the well know formula for transforming random variables. Since

$$\begin{gathered} z = g^{-1}(\hat{s}) = 10^{\frac{\hat{s}}{10}} \\ \text{and } \frac{dg^{-1}(\hat{s})}{d\hat{s}} = \ln(10) 10^{\frac{\hat{s}}{10}-1}, \end{gathered} \tag{2.75}$$

we get the following expression for the PDF of $\hat{s}$

$$\begin{gathered} f_{\hat{s}}(\hat{s}) = \\ \frac{\ln(10) 10^{\frac{\hat{s}}{10}-1}}{2\pi\sigma^2} e^{\frac{-(\mu_R^2+\mu_I^2+10^{\frac{\hat{s}}{10}})}{2\sigma^2}} \\ \int_{0}^{10^{\frac{\hat{s}}{10}}} \frac{1}{\sqrt{(10^{\frac{\hat{s}}{10}}-c_R)c_R}} \cosh\left(\frac{\mu_I\sqrt{10^{\frac{\hat{s}}{10}}-c_R}}{\sigma^2} \right) \cosh\left(\frac{\mu_R\sqrt{c_R}}{\sigma^2} \right) dc_R. \end{gathered} \tag{2.76}$$

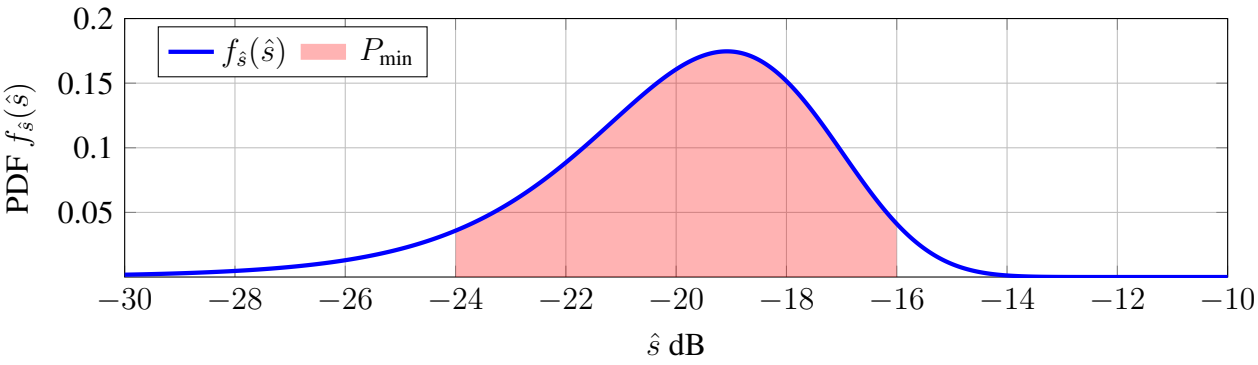

Fig. 2.16. Illustration of the performance criterion on the sequence length design.

2.3.3 Algorithm to Calculate the Minimum Sequence Length

The PDF in (2.76) describes a general case of the situation we have for our channel quality measure. Based on this formula we can now calculate the minimum sequence length. The design criterion is that at a given minimum SNR γ we want the estimated channel power to be within a certain range $\hat{s} \in [10\log(\alpha^2) - d_{\text{max}}, 10\log(\alpha^2) + d_{\text{max}}]$ around the true value $10\log(\alpha^2)$ with minimum probability P_{min}. An example with $\gamma = -20\ dB$, $d_{\text{max}} = 4$ and $P_{\text{min}} = 0.90$ resulting in a minimum length $N = 599$ is illustrated in Fig. 2.16. Since we choose the noise to be unit variance $10\log(\alpha^2) = \gamma$. From the example we can also see that underestimating the channel quality by a large amount is much more probable than overestimating it, even though the estimator itself is unbiased. With these parameters the algorithm is calculating the minimum sequence length achieving this P_{min} in the given range.

The algorithm takes the inputs α, d_{max} and P_{min} to calculate the minimum sequence length achieving P_{min} in the given range. The calculation is divided in two steps: First the order of magnitude of necessary sequence length is calculated. Afterwards, the possible range of sequence lengths is calculated from the order of magnitude. In this range we use interval halving to determine the output sequence length.

To simplify the notation in the algorithm we also need to define the following operators. At first from the definition in (2.66) we can calculate the variance and mean of the resulting circular symmetric Gaussian random variable inside the absolute value operation as:

$$\mu = \mathbb{E}\left[\alpha + \sum_{n=0}^{N-1} \frac{\eta[n]}{N}\right] = \alpha, \text{and } \sigma^2 = \mathbb{E}\left[\left|\sum_{n=0}^{N-1} \frac{\eta[n]}{N}\right|^2\right] = \sum_{n=0}^{N-1} \mathbb{E}\left[\left|\frac{\eta[n]}{N}\right|^2\right] = \frac{1}{N}, \tag{2.77}$$

where we used the fact that the noise $\eta[n]$ is independent across time instances n and has zero mean and unit variance. With these parameters we can express the calculation of the probability to be inside the chosen interval to be:

$$\Pr(\alpha, N, d_{\text{max}}) = \int_{10\log(\alpha^2)-d_{\text{max}}}^{10\log(\alpha^2)+d_{\text{max}}} f_{\hat{s}}(\hat{s}, \alpha, N) d\hat{s}, \tag{2.78}$$

Table 2.3. Minimum sequence length achieving $P_{\text{min}} = 0.90$.

d_{max} \ $10\log(\alpha^2)$	-14	- 12	-10	-8	-6	-4	-2	0
1	2556	1613	1018	642	405	256	162	102
2	633	399	252	159	101	64	40	26
3	276	174	10	70	44	28	18	11
4	151	95	60	38	24	16	10	6
5	92	58	37	24	15	10	6	4

where we get PDF $f_{\hat{s}}(\hat{s}, \alpha, N)$ from the general definition in (2.76) by plugging in the values for the mean and variance:

$$f_{\hat{s}}(\hat{s}, \alpha, N) = \frac{\ln(10) 10^{\frac{\hat{s}}{10}-1} N}{2\pi} e^{\frac{-N(\alpha^2+10^{\frac{\hat{s}}{10}})}{2}} \int_0^{10^{\frac{\hat{s}}{10}}} \frac{1}{\sqrt{(10^{\frac{\hat{s}}{10}} - c_R)c_R}} \cosh\left(N\alpha\sqrt{c_R}\right) dc_R, \quad (2.79)$$

where this formula is (2.76) by setting $\mu_R = \alpha$, $\mu_I = 0$ and $\sigma^2 = \frac{1}{N}$. The double integral to get the probability that the estimate is inside the defined boundaries does also need to be evaluated numerically.

Now we assemble all the mathematical tools to state the algorithm for finding the minimal sequence length $\hat{N}$. The algorithm is described in Algorithm 2.1. As there are three possible parameters we selected a number of possible configurations and show the result in the Table 2.3 and 2.4. From the results we can see that the length increases approximately exponential with the decreasing d_{max} as well as decreases approximately quadratic with increasing of the SNR (here equal to $10\log(\alpha^2)$).

2.3.4 Calculation of the Probability to not Select the best Beam

With the derived expression for the PDF of $\hat{s}$ and given a sequence length we can also calculate the probability to select another beam instead of the optimal one. This is calculated by first setting up the different gains for both beams. In this case α_1 is the gain of the beam with the better performance. This means that the channel gain α_1 is larger than the one of the second best beam α_2. For the calculation we assume that the beam miss-selection is dominated by the beam with second best performance. In terms of the PDF described in (2.79) we can calculate the probability

Table 2.4. Minimum sequence length achieving $P_{\text{min}} = 0.95$.

d_{max} \ $10\log(\alpha^2)$	-14	- 12	-10	-8	-6	-4	-2	0
1	3643	2299	1451	916	578	365	230	146
2	913	577	364	230	145	92	58	37
3	408	257	163	103	64	41	26	17
4	231	146	92	58	37	24	15	10
5	148	94	59	38	24	15	10	6

Algorithm 2.1 Calculation of the minimal sequence length to sufficiently assess the channel quality.

Require: α, d_{max} and P_{min}
$N_{\text{start}} \leftarrow 128$;
$P \leftarrow \Pr(\alpha, N_{\text{start}}, d_{\text{max}})$
if $P \geq P_{\text{min}}$ **then** ▷ Initial length better than requirement
  $N_{\text{high}} \leftarrow N_{\text{start}}$
  $N_{\text{low}} \leftarrow \lfloor N_{\text{high}}/2 \rfloor$
  $P \leftarrow \Pr(\alpha, N_{\text{low}}, d_{\text{max}})$
  while $P \geq P_{\text{min}} \wedge N_{\text{low}} > 1$ **do** ▷ Decrease length till value worse than requirement found
    $N_{\text{high}} \leftarrow N_{\text{low}}$
    $N_{\text{low}} \leftarrow \lfloor N_{\text{high}}/2 \rfloor$
    $P \leftarrow \Pr(\alpha, N_{\text{low}}, d_{\text{max}})$
  end while
else ▷ Initial length worse than requirement
  $N_{\text{low}} \leftarrow N_{\text{start}}$
  $N_{\text{high}} \leftarrow N_{\text{low}} \cdot 2$
  $P \leftarrow \Pr(\alpha, N_{\text{high}}, d_{\text{max}})$
  while $P < P_{\text{min}}$ **do** ▷ Increase length till value better than requirement found
    $N_{\text{low}} \leftarrow N_{\text{high}}$
    $N_{\text{high}} \leftarrow N_{\text{low}} \cdot 2$
    $P \leftarrow \Pr(\alpha, N_{\text{high}}, d_{\text{max}})$
  end while
end if
while $(N_{\text{high}} - N_{\text{low}}) > 1$ **do** ▷ Interval halving till minimum length found
  $N_{\text{mid}} \leftarrow \lfloor (N_{\text{high}} + N_{\text{low}})/2 \rfloor$
  $P \leftarrow \Pr(\alpha, N_{\text{mid}}, d_{\text{max}})$
  if $P \geq P_{\text{min}}$ **then**
    $N_{\text{high}} \leftarrow N_{\text{mid}}$
  else
    $N_{\text{low}} \leftarrow N_{\text{mid}}$
  end if
end while
return N_{high}

of selecting the second best beam P_{miss} as:

$$P_{\text{miss}} = \int_{-\infty}^{\infty} f_{\hat{s}_1}(\hat{s}_1, \alpha_1, N) \left(\int_{\hat{s}_1}^{\infty} f_{\hat{s}_2}(\hat{s}_2, \alpha_2, N) d\hat{s}_2 \right) d\hat{s}_1. \tag{2.80}$$

In this equation the outer integral is integrating over each possible estimated value of the channel quality measure $\hat{s}_1$ of the best beam. The inner integral is calculating the probability that the estimate of the channel quality measure $\hat{s}_2$ of the second beam is larger than the one of the best beam $\hat{s}_1$. As this double integral is composed of sub integrals as described in (2.79) without a closed form solution, we need to again use numeric integration. Since the second best beam has a smaller

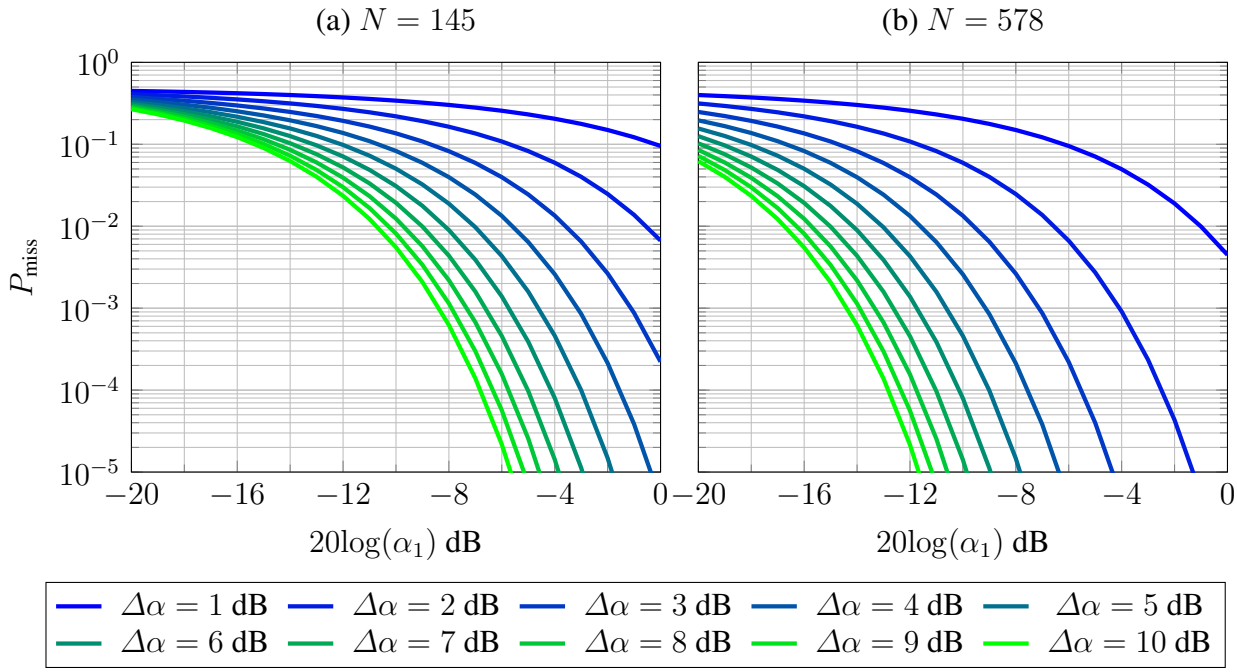

Fig. 2.17. Probability of the selection the wrong beam P_{miss} dependent on the effective channel gain of the best beam α_1 and the power difference between the effective channel gain of both beams $\Delta\alpha$ with sequence length N 145 (b) and 578 (b).

effective channel gain we show the plots in terms of $\Delta\alpha = 20\log(\alpha_1) - 20\log(\alpha_2)$. As the integrals are also asymptotically decaying in both direction from the actual value $\hat{s}_i = 20\log(\alpha_i)$, it is safe to assume that after a distance of 20 dB the values are sufficiently small to have now influence on the overall result.

The probability of selecting the worse beam 2 with α_2 over beam 1 with α_1 is shown considering a sequence length 145 and 578 is shown in Fig. 2.17 (a) and (b). If the power difference between the effective channel of the beams $\Delta\alpha$ is small there is a high chance of selecting the wrong one, but in this case since the power difference is not that large there is only a small penalty in terms of received power if the wrong one is selected. As the channel gain of both beams increased, the estimate of the channel quality also improves, thus P_{miss} decreases. It is also important to mention that if we select the same scenario increasing the sequence length N from 145 to 578 roughly improves the curve by 6 dB.

2.3.5 Approach to Determine the Number of Beams for Beam Training

This subsection was also part of the contribution [77]. In this context is it is also important to know how many beams need to be considered. Based on some simplifying assumptions we determine the number of beams inside the set of all beamforming vectors $\mathbb{B}$ achieving a certain error criterion. In this case we limit each vector in $\mathbb{B}$ to maximize the power coming from a certain angle ϕ_B. Here we assume isotropic minimum scattering antennas. For an ULA with spacing $\lambda/2$, the absolute value of the normalized array factor is defined as [13, page 294]:

$$AF(\phi) = \frac{1}{M_C} \left| \frac{\sin\left(M_C \frac{\pi}{2} \sin\left(\phi - \phi_B\right)\right)}{\sin\left(\frac{\pi}{2} \sin\left(\phi - \phi_B\right)\right)} \right|. \tag{2.81}$$

Table 2.5. Minimum number of beams necessary to achieve mean error $\epsilon = 0.1$.

M_C	minimum number of beams	approximation $4M_C$
2	7	8
4	14	16
8	29	32
16	58	64
32	115	128

That means for actual arriving angle ϕ choosing $\phi_B = \phi$ is optimal. But this would mean that we have an infinite grid of ϕ_B. Assuming a single wavefront arriving at the receiver and an uniformly distributed angle of the arriving signal ϕ, we get the following expression for the average error ϵ:

$$\frac{2}{\Delta}\int_0^{\frac{\Delta}{2}}\left(1-\frac{1}{M_C}\left|\frac{\sin\left(M_C\frac{\pi}{2}\sin\left(x\right)\right)}{\sin\left(\frac{\pi}{2}\sin\left(x\right)\right)}\right|\right)dx \leq \epsilon. \tag{2.82}$$

Setting a maximum allowed ϵ, we can solve the equation for the distance Δ between two angles in the set $\mathbb{B}$:

$$\frac{2}{\Delta}\int_0^{\frac{\Delta}{2}}\left(1-\frac{1}{M_C}\left|\frac{\sin\left(M_C\frac{\pi}{2}\sin\left(x\right)\right)}{\sin\left(\frac{\pi}{2}\sin\left(x\right)\right)}\right|\right)dx - \epsilon = 0. \tag{2.83}$$

The expression in (2.83) can be solved by a bisection based procedure. In this case we select a lower bound Δ_l and an upper bound Δ_u for Δ, these values are chosen in a way to ensure that the value that solves the equation is in between them. Afterwards, (2.83) is solved for Δ_l, Δ_u and $(\Delta_l + \Delta_u)/2$ by numeric integration. Based on the results of this function evaluation we select the bounds for the next iteration of the bisection method. This process is repeated till sufficient accuracy of Δ is reached. If the optimal Δ is known, we get the number of beams by $\lceil\frac{2\pi}{\Delta}\rceil$ as only an integer number of beams is possible. A table for some configurations and the minimum number of elements in $\mathbb{B}$ are shown Table 2.5. It can be observed that for the given parameters, the minimum number of elements can be well approximated by $4M_C$. Therefore, for some of the following evaluations, we select $4M_C$ elements with ϕ_B uniform the range from 0 to 2π to represent the set $\mathbb{B}$.

2.3.6 Conclusion Sequence Length Calculation for Channel Quality Assessment

In the Subsections 2.3.1 to 2.3.5 we have shown how to calculate the necessary length of a sequence to acquire the given channel quality measure given defined constraints on the error probability. We also showed which miss-selection probability is achieved given two example sequence lengths. This result gives us an insight into how much overhead to expect from the channel quality estimation for each beam. In the subsequent chapter we will use the calculated length in conjunction with assumptions about the beam training to get the overall beam training overhead.

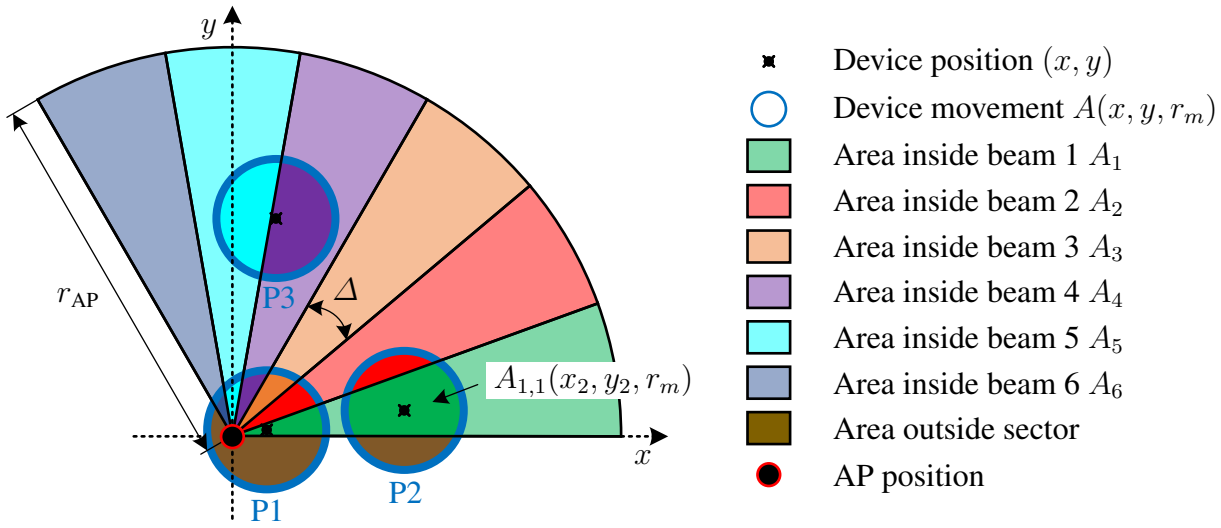

Fig. 2.18. Illustration of the transition of devices at positions P1 at x_1, y_1 to P3 x_3, y_3 from the region covered by one beam to other regions, given a movement radius defined by the maximum speed and beam training time interval. The faded colors show the regions covered by beams 1 to 6 denoted as A_1 to A_6. The other colors show the intersection of the movement at each the example device positions $A(x_i, y_i, r_m)$ with the beam coverage, resulting in the areas $A_{j,l}(x_i, y_i, r_m)$.

2.3.7 Beam Training Time Interval

Another important aspect of beam training is the time interval during which the training needs to be performed. This is dependent on the assumptions of the maximum mobility and the area that needs to be covered. For the evaluation in this work, we consider two simplified scenarios. The first one is only considering Line Of Sight (LOS) connection. The second one models the spatial coherence and therefore, include Non Line Of Sight (NLOS) scenarios. In both cases we assume that the maximum speed of a device would be 30 km/h. Essentially, this speed is modeling a streetcar, bus, or car in a dense urban environment. Essentially, we model a system where mmWave is providing a high data rate connection to the vehicle, and the connectivity for the users inside the vehicle is provided via for example Wifi or another technology.

2.3.7.1 LOS Based Model

In this model we only consider the LOS connection from the device to the Access Point (AP). We also ignore any possible blockage of the LOS path. Assuming the device can move in any direction, the maximum velocity v together with the time between two phases of beam training τ_B define a possible movement radius $r_m = v \cdot \tau_B$. For the AP we assume that the area it should cover is divided into three sectors, each covering $120°$. Each sector is equally divided into areas covered by distinct beams. In example in Fig. 2.18 there are six beams each covering an angle of $\Delta = 20°$. Fig. 2.18 also illustrates three example positions for devices (Device position P1, P2, P3) and the respective potential device movement area $A(x_1, y_1, r_m)$ to $A(x_3, y_3, r_m)$. The area covered by different beams A_i are illustrated with faded colors. The vibrant colors indicate the portion of the area inside the possible movement radius, where a device would transition between beams or leave the coverage area of the sector. These areas can be formed by intersecting the area covered by a beam with the disc describing the movement. In principle we integrate over the area that is inside the current beam for all possible devices starting in this beam. Dividing this by the

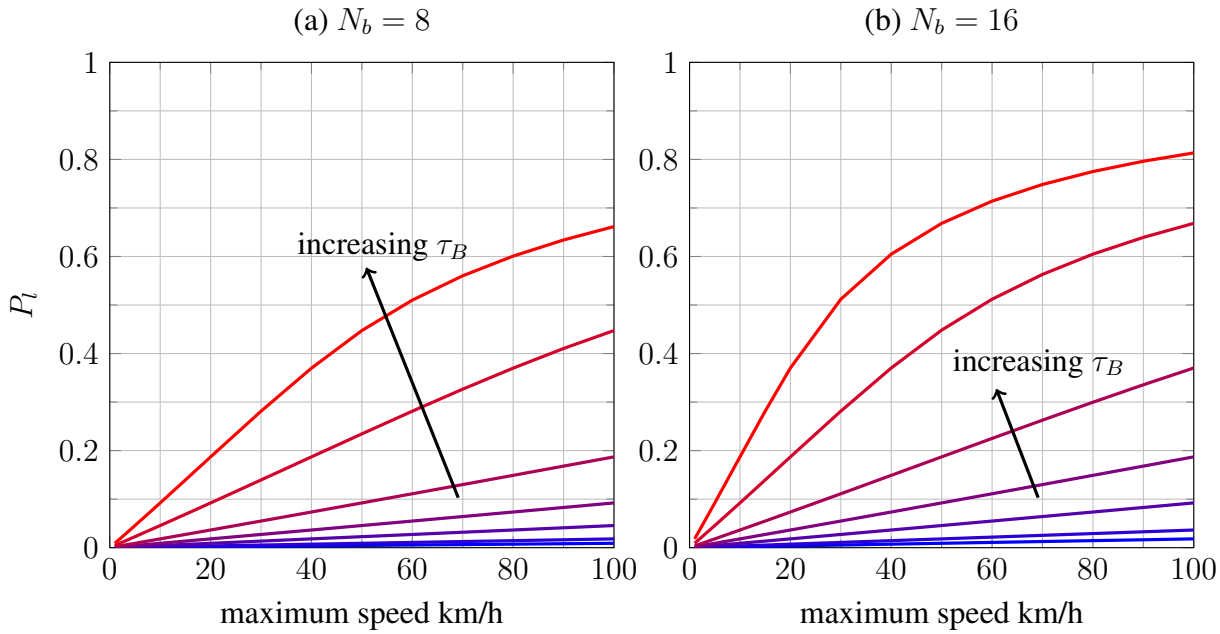

Fig. 2.19. Probability P_l of devices leaving a given coherence area given different beam training intervals $\tau_B \in \{10\text{ ms},\ 20\text{ ms},\ 50\text{ ms},\ 100\text{ ms},\ 200\text{ ms},\ 500\text{ ms},\ 1\text{ s}\}$ and number of beams (a) $N_b = 8$ and (b) $N_b = 16$ for cell radius $r_{\text{AP}} = 50$ m.

area for all possible movements we can get the probability for the device to have the same beam-configuration in two consecutive beam training instantiations. In addition, with this tool it is also possible to generate the transition probability for a Markov-model, where each state is associate with a beam. It needs to be mentioned that this model does not take rotation of the mobile device into account. In fact, we assume that the mobile device can provide sufficient spherical coverage, thus, rotating the device does not change the best beam of the AP.

From the illustration in Fig. 2.18 it is easy to see that all areas $A_{i,j}(x, y, r_m)$ can be formed by intersecting the area describing the device movement $A(x, y, r_m)$ with the beam coverage area A_i. This means assuming that the position x, y is inside the coverage area of beam i we can describe $A_{i,j}(x, y, r_m)$ as:

$$A_{i,j}(x, y, r_m) = A(x, y, r_m) \cap A_j. \tag{2.84}$$

In this case we are mainly interested in the probability that given a movement radius r_m, the beamwidth Δ in radian, and the cell radius r_{AP} we can calculate the probability that we leave coverage of one beam as:

$$P_l = 1 - \frac{1}{O_t} \int_0^{r_{\text{AP}}} \left(\int_0^{\min(\tan(\Delta)\cdot x, \sqrt{r_{\text{AP}}^2 - x^2})} A_{i,i}(x, y, r_m) dy \right) dx, \tag{2.85}$$

with O_t being a normalization factor equal to $O_t = \frac{\Delta}{2\pi} r_{\text{AP}}^2 \pi r_m^2 \pi$. This integral is calculating the full area of the device staying in beam i given that it is already in beam i at the start. As the areas $A_{i,i}(x, y, r_m)$ are dependent on the movement radius r_m and thus implicitly dependent on the time between beam training τ_B we can evaluate this equation for different values of the beam training interval τ_B.

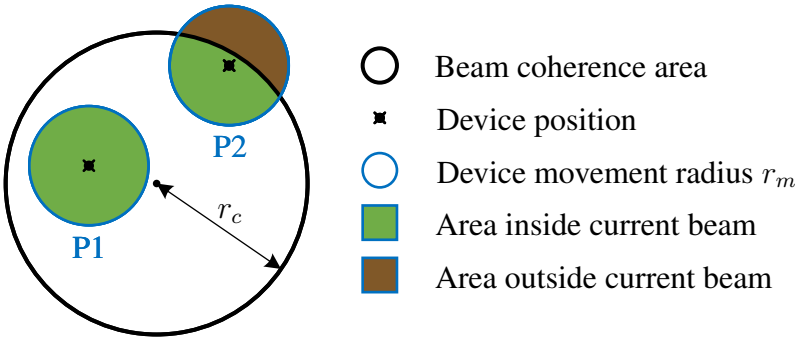

Fig. 2.20. Illustration of the area in which the one beam is optimal including two device positions (P1 and P2) and the respective areas for staying inside the area of this beam or transitioning to another one.

Fig. 2.19 shows the evaluation results for one 120° sector of a cell with radius $r_{\text{AP}} = 50$ m. In the evaluation different number of beams were used. The results show that to have a probability below 10 % of leaving the cell in between two beam training intervals given a speed of 30 km/h, τ_B should be below 100 ms. However, this model does only capture LOS connections. To cover all possible propagation aspects in the next paragraphs a model not restricted to only LOS is developed.

2.3.7.2 Spatial Coherence Based Model

Since the first model was only considering LOS channels, is also interesting to investigate a model not limited by this assumption. For all the aspects in this model it is important to note that the beamforming should only adapt to the large scale parameters of the system. Adapting to small scale

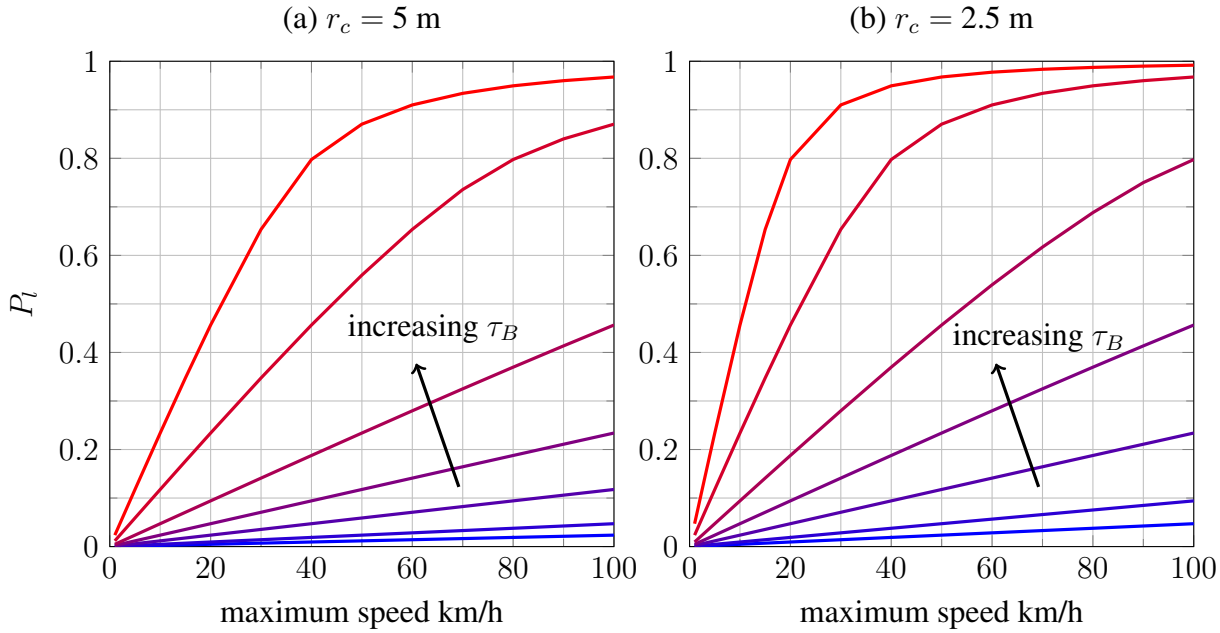

Fig. 2.21. Probability P_l of devices leaving a given coherence area given different beam training intervals $\tau_B \in \{10 \text{ ms}, 20 \text{ ms}, 50 \text{ ms}, 100 \text{ ms}, 200 \text{ ms}, 500 \text{ ms}, 1 \text{ s}\}$ and coherence radius (a) $r_c = 5$ m, and (b) $r_c = 2.5$ m.

parameters would result in a prohibitive beam training overhead, as they too frequently change. The spatial correlation distanced is define in [95] and [40]. Essentially, it describes the spatial distance at which the large scale parameters are highly correlated. In [40] measurements of this parameter at 73 GHz carrier frequency are shown. The resulting correlation distance is in the range of 5 to 10 meters.

In a two dimensional model these aspects can be captured by defining a circle in which the large scale parameters are highly correlated. This results in the system using the same beamforming configuration if the device remains in this area. This model is illustrated in Fig. 2.20. The area of spatially coherent large scale parameters of the channel is modeled as a circle. As the mobile device is currently using the specific beam for this area it must be inside the circle. Now again assuming a certain time interval of the beam training a device could move inside a circle around the starting position. The radius of the circle is defined by the maximum possible speed and the beam training time interval.

The area that a device at a given location x, y would leave the area at which the same beam is used, given its maximum movement radius r_m is define as $A(x, y, r_m)$. From Fig. 2.20 it is evident that for all possible location x, y inside the area served by the same beam, here defined by a circle with radius r_c, this $A(x, y, r_m)$ can be calculated by simple geometric objects like triangles and circles. Combining these for all possible location we get the probability that a device leaves the area:

$$P_l = \frac{2\pi}{O_t} \int_0^{r_c} A(r, 0, r_m) r \, dr, \tag{2.86}$$

where we used the rotational symmetry to simplify the calculation. In this case O_t is a normalization factor equal to $O_t = r_c^2 \pi^2 r_m^2$. The time interval between consecutive beam training periods is defined as τ_B.

In Fig. 2.21 the probability of leaving the area of one beam being the optimal one in between two consecutive beam training intervals is presented. Based on the measured correlation distance of 5 to 10 meters in [40] the coherence radius of r_c was setup to be between 2.5 and 5 meters. As leaving the area for one beam would likely result in a link failure, the probability P_l should not be too large. However, the time between two consecutive beam training intervals τ_B should also not be too large as this would result in a prohibitive large overhead. Thus, judging from the results in Fig. 2.21 we see that a $\tau_B = 50$ ms beam training interval will keep the probability $P_l < 0.1$ for a maximum speed of 30 km/h.

2.4 Receiver Front-End Power Consumption Model

In a 5G millimeter Wave mobile broadband system, it is necessary to utilize large antenna arrays. It is therefore important to compare the power consumption of different receiver architectures. In this section we present a power model for analog/hybrid beamforming and digital beamforming in combination with low resolution ADCs as well as hybrid beamforming using a Butler matrix. This section was part of the contribution [77].

Since the spectrum in the 60 GHz band can be accessed without a license, it got significant attention. Especially the IEEE 802.11ad standard (WiGig) standard operating in this band, increased the transceiver RF hardware Research and Development (R&D) activities. Many chips were reported from industry and academia. Thus, it is safe to assume that the design reached a certain maturity, and performance figures derived from them represent the performance that is possible for

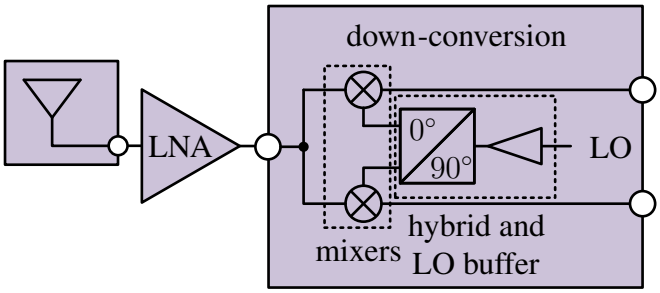

Fig. 2.22. Simplified direct conversion receiver.

a low cost CMOS implementation today. As so far only products with subarray hybrid beamforming are available, we only derive a power model for these type of hybrid beamforming and digital beamforming.

2.4.1 Power Consumption Model for Systems with Digital Beamforming and Hybrid Beamforming based on Phase shifters

According to the discussion in [22], baseband or IF phase shifting in contrast to RF phase shifting is assumed. This has the advantage of increased accuracy, decreased insertion loss, and reduced gain mismatch. In [22], the authors showed that the power consumption for a low number of antennas per ADC chain is equivalent to a system utilizing RF Phase Shifter (PS).

The front-end subsystems shown in Fig. 2.22, 2.23 and 2.24 are the details behind the blocks with the same color shown in the introduction of the different beamforming schemes in Fig. 1.1, 1.2 and 1.3. All systems utilize the same direct conversion receiver (Fig. 2.22) to convert the signal into the analog baseband. For each system, we assume that the LO is shared by the whole system. For the case of analog/hybrid beamforming systems, the analog baseband signals are phase shifted and then combined to generate the input signal of the M_{ADC} ADCs (Fig. 2.23). The A/D conversion consists of a VGA that is amplifying the signal to use the full dynamic range of the ADC (Fig. 2.24). For the special case of 1 bit quantized digital beamforming a VGA is not necessary. It can be replaced by a much simpler Limiting Amplifier (LA). Fig. 2.22 to 2.24 show how the high level blocks in the receiver block diagrams in Fig. 1.1 and 1.2 are subdivided into components. To highlight the connection the same color is used for the corresponding blocks.

The power consumption of each component, including a reference, are shown in Table 2.6. An LO with a power consumption as low as 22.5 mW is reported in [85]. The power consumption of an LNA and a VGA are reported in [87] as 5.4 mW and 2 mW. The 90° hybrid and the LO buffer

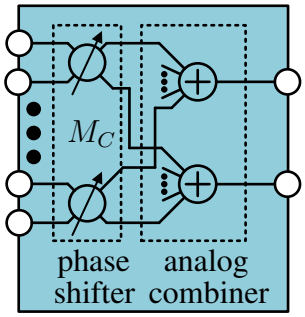

Fig. 2.23. Analog signal combination part of the front-end.

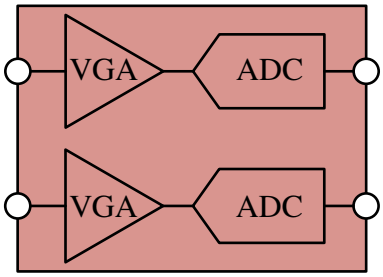

Fig. 2.24. A/D conversion part of the receiver.

reported in [50] have a combined power consumption of 3 mW. The power consumption of the mixer reported in [39] is as low as 0.3 mW. The survey in [63] gives a good overview of state-of-the-art ADCs regarding Effective Number of Bits (ENOB), sampling rate, and power consumption. Taking the predicted curve for the Walden figure of merit in [63] for a sampling frequency of 2.5 GS/s, we get 15 fJ per conversion step. A LA that consumes 0.8 mW is reported in [65]. In the 1 bit quantized system, the LA (aka Schmitt trigger) is already producing a digital signal, therefore the 1 bit ADC can be replaced by a Flip Flop (FF). The power consumption of a FF is negligible compared to the rest of the RF front-end.

With the power consumption of the components, it is possible to compute the power consumption of the overall receiver front-end P_R as:

$$\begin{aligned} P_R = P_{\text{LO}} + M\left(P_{\text{LNA}} + P_{\text{H}} + 2P_{\text{M}}\right) + \\ \text{flag}_C\left(MP_{\text{PS}}\right) + \\ M_{\text{ADC}}\left(\neg\text{flag}_{\text{1bit}}\left(2P_{\text{VGA}} + 2P_{\text{ADC}}\right) + \text{flag}_{\text{1bit}}\left(2P_{\text{LA}}\right)\right), \end{aligned} \tag{2.87}$$

where flag_C is indicating if analog combining is used:

$$\text{flag}_C = \left\{ \begin{array}{ll} 0, & M_{\text{ADC}} = M, M_C = 1 \\ 1, & \text{else} \end{array} \right\}. \tag{2.88}$$

Table 2.6. Components with power consumption.

label	component	power consumption	reference
P_{LO}	LO	22.5 mW	[85]
P_{LNA}	LNA	5.4 mW	[87]
P_{M}	mixer	0.36 mW	[39]
P_{H}	90° hybrid and LO buffer	3 mW	[50]
P_{LA}	LA	0.8 mW	[65]
P_{1}	1 bit ADC	0 mW	
P_{PS}	phase shifter	2 mW	[45, 22]
P_{VGA}	VGA	2 mW	[87]
P_{ADC}	ADC	$15\ \mu\text{W/GHz} \cdot f_s 2^{\text{ENOB}}$	[63] [21] [98]

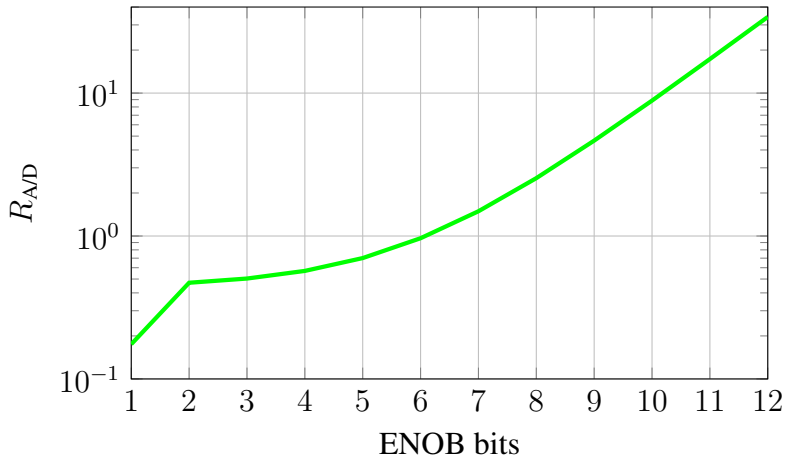

Fig. 2.25. $R_{\text{A/D}}$ for different effective resolution ENOB of the ADC.

The variable $\text{flag}_{\text{1bit}}$ is indicating if 1 or multi bit quantization is used. In the case of 1 bit quantization, the power consumption of the VGA is replaced by the one of the LA and the power consumption of the 1 bit quantizer is neglected. This formula now contains all special cases of digital beamforming ($M_{\text{ADC}} = M$), analog beamforming ($M_{\text{ADC}} = 1$) and hybrid beamforming.

A receiver directly designed for the 1 bit quantization digital beamforming systems is very likely to reduce the power consumption even further. Due to the 1 bit quantization at the end of the analog part of the receiver, the linearity required of the circuits before is greatly reduced. This would enable specialized designs to improve the performance in terms of power consumption, which are not exploited in this work.

Just to give insight into the actual power consumption we show the power consumption of the A/D conversion part relative to the power consumption of one direct conversion part of the receiver:

$$R_{\text{A/D}} = \frac{\neg\text{flag}_{\text{1bit}}\,(2P_{\text{VGA}} + 2P_{\text{ADC}}) + \text{flag}_{\text{1bit}}\,(2P_{\text{LA}})}{P_{\text{LNA}} + P_{\text{H}} + 2P_{\text{M}}}. \tag{2.89}$$

In this case we assume that the LO is shared among all receive antennas. For the given parameters $P_{\text{LNA}}+P_{\text{H}}+2P_{\text{M}}$ is equal to 9.12 mW. The value for $R_{\text{A/D}}$ for different values of the ENOB is shown in Fig. 2.25. Since in the case of 1 bit quantization the A/D conversion stage is very simplified this clearly achieves the lowest power consumption. In the range of 2-6 bit the VGA is consuming more power than the ADC, thus the curve is close to flat in this range. As the power above 6 bit ENOB is dominated by the ADC resolution we see the exponential growth of the relative power consumption.

This power model for the receiver front-end is used in the subsequent chapter to compare the energy efficiency of hybrid and digital beamforming. This model does simplify the situation for both hybrid and digital beamforming but in our opinion, it could help understand the difference in energy efficiency of both systems.

2.4.2 Power Consumption Model for Systems with Hybrid Beamforming based on a Butler Matrix

For the power model considering the front-end based on a Butler matrix we need to modify our previous model. The main difference is in the analog combining subsystem. In this case the analog

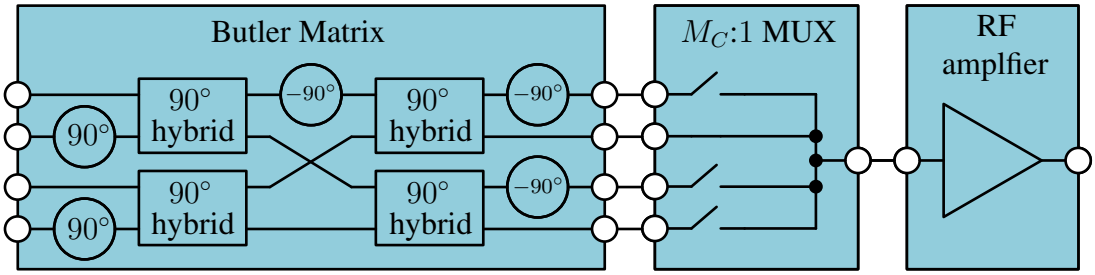

Fig. 2.26. Analog combining for a hybrid beamforming subarray with 4 antennas utilizing a Butler matrix.

combining subsystem consists of the Butler matrix itself, a multiplexer and an additional amplifier to compensate for the power loss in the combining subsystem In the subsequent paragraphs we will also explain why this amplifier is necessary. An example of the analog combining for a subarray with 4 antennas is shown in Fig. 2.26. As described in [20] the Butler matrix aims to replicate the Discrete Fourier Transformation (DFT) across the different antenna elements. This can easily be implemented via a combination of $90°$ hybrids and phase shifters. As show in [12, 7] it even possible to implement a 4x4 or 8x8 Butler matrix on a single layer of a Printed Circuit Board (PCB) utilizing a microstrip implementation. As show in [47, pages 102 - 117] this result can easily be transferred to mmWave. Another aspect is that in contrast to the other systems the down-conversion is after the analog combining. The same down-conversion as shown in Fig. 2.22 is also used for this receiver architecture. The phase shifters at the inputs and outputs of the Butler matrix change the overall transfer function of the 4×4 port to a 4×4 DFT matrix.

After the Butler matrix there are M_C output signals available. Based on a trial and error procedure one of the outputs is selected. How this one is selected will be shown in subsequent sections. The switches of the multiplexer are than configured to select this output of the Butler matrix. These switches are essentially switching between different RF signals. As reported in [69, 46] these switches at 60 GHz have a substantial insertions loss in the range of 2.5 to 3 dB.

The Butler matrix is distributing the input power to all outputs, but we only use the signal from one of the outputs, with an additional loss that according to [47, pages 102 - 117] can be assumed to be in the range of 0.33 dB per $90°$ hybrid. In addition, insertion loss of the phase shifters also need to be taken into account. Combining the insertion loss of all components of the Butler matrix with the insertion loss of the RF switches we get a substantial reduction in power. This needs to be compensated before the down-conversion. Therefore, an additional amplifier is needed. As this amplifier needs to have a similar gain as the LNA in [87] and is also operating at the carrier frequency we assume the power consumption P_{RFA} to be the same.

With the power consumption of the components, it is possible to compute the power consumption of the overall receiver front-end P_R as:

$$\begin{aligned} P_R = P_{\text{LO}} + M P_{\text{LNA}} + M_{\text{ADC}} \left(P_{\text{H}} + 2P_{\text{M}}\right) \\ M_{\text{ADC}} \left(P_{\text{RFA}} + \neg\text{flag}_{\text{1bit}} \left(2P_{\text{VGA}} + 2P_{\text{ADC}}\right) + \text{flag}_{\text{1bit}} \left(2P_{\text{LA}}\right)\right). \end{aligned} \tag{2.90}$$

2.5 Channel Estimation

The evaluation in this subsection is part of the papers [81, 80, 82]. For this evaluation we use the 3GPP NR Orthogonal Frequency Division Multiplexing (OFDM) Type I Demodulation Reference

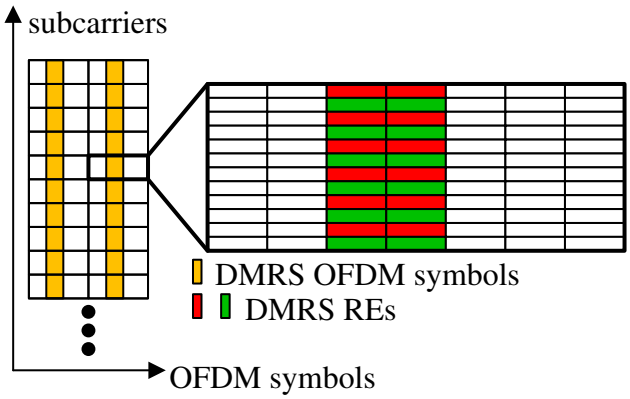

Fig. 2.27. 3GPP NR OFDM type one reference signal pattern for up to 8 users configured to use the 3rd, 4th, 10th and 11th OFDM symbol in a slot. The yellow color shows the position of the DMRS OFDM symbols in a slot. The colors red and green show SC allocated to different groups of DMRS, which are used for different users.

Signals (DMRS) which are based on a sequence that is defined by parts of a length-31 Gold sequence as described in [4]. Throughout this section the property that the auto correlation function of this sequence can be approximated by a Dirac at offset zero is used. There are three techniques used to orthogonalize the reference signal for different users:

- Time domain Cyclic Shift (CS)
- Frequency Division Multiplexing (FDM)
- Code Division Multiplexing (CDM) for the case of more than 4 users

The resource grid in Fig. 2.27 shows a possible allocation of the reference signal Resource Elements (REs) within a slot. Among the three ways to orthogonalize the reference signals sent by different users or for different MIMO streams, FDM and CDM are most commonly used in many communication systems. It is possible to apply CDM in frequency or in time direction. For frequency direction CDM the reference signal is spread into two adjacent Sub-Carriers (SCs). In the case of time direction CDM the reference signal is spread to the same SC in adjacent OFDM symbols. CS is not encountered in many current wireless communication standards except for LTE Uplink (UL), thus more detailed explanation on how it could provide sufficient orthogonality is needed.

A CS represents a circular shift of the time domain signal. The main advantage over frequency direction code division multiplexing is that a timing offset does not have an influence on the orthogonality of the DMRS symbols. This is not the case for frequency direction code division multiplex. Due to the properties of the Fourier transformation, a cyclic shift of half of the symbol in the time domain corresponds to a multiplication of the frequency domain sequence with $\boldsymbol{c} = [-1,\ 1,\ -1,\ 1, \ldots]^T$. This means that the generation of the transmit signal can easily be done in the frequency domain. The reception would consist of a multiplication with the know reference symbol in frequency domain followed by a circular convolution with the frequency domain representation of the time domain windowing. This does correspond to a circular convolution with the known reference sequence followed by the application of a windowing function in time domain.

The reference symbol sequence is generated as Quaternary Phase Shift Keying (QPSK) symbols with the bits defined by a Pseudo Random Binary Sequence (PRBS) sequence. To generate the first estimate of channel at the position of the DMRS, we first transform the signal into the frequency domain. Afterwards, we multiply with the complex conjugate of the known reference

sequence. Since this would correspond to a cyclic convolution of received signal with the known reference sequence $\boldsymbol{s}$ in the time domain this does still include all possible cyclic shifts. Thus, additionally we need to apply a windowing in the time domain to limit the interference among the original sequence and its cyclic shift one. The windowing in the time domain corresponds to a cyclic convolution with the frequency representation of the window in the frequency domain. Combining these observations with the fact that the sequence $\boldsymbol{s}$ should be designed to have a cyclic auto-correlation function with only one strong peak and only small values otherwise, we can conclude that a cyclic shift can sufficiently orthogonalize the signal from different users.

We define the vector $\boldsymbol{s}$ to contain the part of the Gold sequence used for the generation of the DMRS symbols. To randomize the used sequence across different cells the part of the sequence that is taken depends on multiple parameters. As this is known by the receiver and has no influence on a system without inter cell interference this work does not present these details and just assumes known sequence of symbols taken from a Gold sequence. In general, we can describe the reference signal of ith spatial data stream on SC k and OFDM symbol ℓ as

$$a^i_{(k,\ell)} = \alpha_{\text{CS}}(i,k)\alpha_{\text{FDM}}(i,k)\alpha_{\text{CDM}}(i,\ell)[\boldsymbol{s}]_{\lfloor k/2 \rfloor}, \tag{2.91}$$

where $\alpha_{\text{CS}}(i,k)$, $\alpha_{\text{FDM}}(i,k)$ and $\alpha_{\text{CDM}}(i,\ell)$ are the changes of sequence based on CDM, FDM and CS. These parameters are defined in the following way

$$\begin{aligned} \alpha_{\text{CS}}(i,k) &= w^i_{\text{CS}}(\lfloor k/2 \rfloor \bmod 2), \\ \alpha_{\text{FDM}}(i,k) &= \begin{cases} 1, & k \bmod 2 = w^i_{\text{FDM}} \\ 0, & \text{otherwise} \end{cases}, \\ \alpha_{\text{CDM}}(i,\ell) &= \begin{cases} w^i_{\text{CDM}}(0), & \ell = \ell_0 \\ w^i_{\text{CDM}}(1), & \ell = \ell_0 + 1 \end{cases}, \end{aligned} \tag{2.92}$$

where ℓ_0 denotes the OFDM symbols with the DMRS symbols. In NR these are configurable and depend on the system requirements. A table of the parameters w^i_{CS}, w^i_{FDM} and w^i_{CDM} dependent on the spatial data stream i can be found in Table 2.7. It is important to mention that for the case of 1 to 4 MIMO spatial data streams, no CDM based orthogonalization is necessary. Therefore, only one OFDM symbol is used for reference signal transmission and the reference signals for different spatial data streams are orthogonalized via CS and FDM. The Illustration in Fig. 2.28 shows the resulting configuration of the parameters in a specific example.

To limit the computational complexity, we use separate interpolation in the time and frequency dimensions. The work in [35] shows that this leads to only a minor performance degradation compared to joint time-frequency interpolation at a greatly reduced computational complexity. To further reduce the complexity, we also limit the number of SCs that are used for the interpolation in the frequency direction. Another important aspect is that for the channel estimation and equalization techniques to work properly, it must be ensured that the power received from different users is not very different. In a practical system this is ensured via power control.

Assuming perfect synchronization of the timing and carrier frequency, the received OFDM signal $Y_{k,\ell}$ for SC k and OFDM symbol ℓ can be written as

$$Y_{k,\ell} = H_{k,\ell}X_{k,\ell} + \eta_{k,\ell}, \tag{2.93}$$

where we assume that the channel impulse response is shorter than the Cyclic Prefix (CP), and $H_{k,\ell}$, $X_{k,\ell}$ and $\eta_{k,\ell}$ represent the channel, transmit signal and noise, respectively. Note that a part of

Table 2.7. Configuration of NR Type I OFDM reference signals.

spatial data stream i	$[w^i_{\text{CS}}(0),\ w^i_{\text{CS}}(1)]$	w^i_{FDM}	$[w^i_{\text{CDM}}(0),\ w^i_{\text{CDM}}(1)]$
1	$[1,\ 1]$	0	$[1,\ 1]$
2	$[1,-1]$	0	$[1,\ 1]$
3	$[1,\ 1]$	1	$[1,\ 1]$
4	$[1,-1]$	1	$[1,\ 1]$
5	$[1,\ 1]$	0	$[1,-1]$
6	$[1,-1]$	0	$[1,-1]$
7	$[1,\ 1]$	1	$[1,-1]$
8	$[1,-1]$	1	$[1,-1]$

the transmitted signals $X_{k,\ell}$ are known reference signals as described in (2.91). Since the channel estimation procedure is executed separately for each antenna and user, no user or antenna index is used. It is important to note that this technique assumes knowledge of the following statistical channel parameters:

- Doppler spread
- Delay spread
- Received SNR of each user

In a practical system those can be attained via estimation over a longer time period.

Assuming a reference symbol is present on SC k and OFDM symbol ℓ, we multiply the signal with the known reference signal to obtain the corresponding channel estimate for this resource element

$$\hat{H}_{k,\ell} = Y_{k,\ell} X^*_{k,\ell} = H_{k,\ell} + \eta_{k,\ell}, \tag{2.94}$$

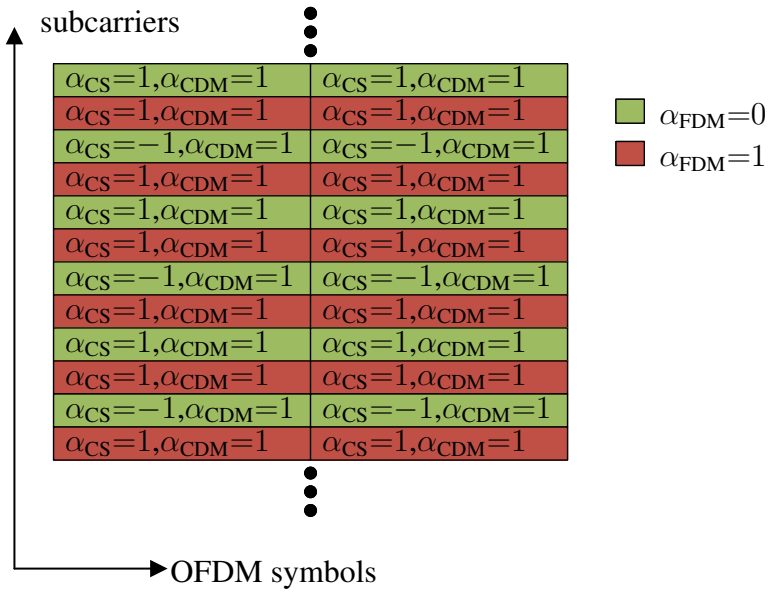

Fig. 2.28. Example of the resulting modifiers α_{CS}, α_{FDM} and α_{CDM} for $i = 2$ (in green) and $i = 3$ (in red) under the assumption that more than 4 spatial data streams are used.

where we assume that $|X_{k,\ell}| = 1$. With this assumption on $X_{k,\ell}$, multiplying the circular symmetric, zero mean, complex Gaussian noise $\eta_{k,\ell}$ with $X^*_{k,\ell}$ does not change the statistics of the noise. By combining the channel estimates for all resource elements on K SCs and L OFDM symbols we get

$$\hat{\boldsymbol{h}}_r = \begin{bmatrix} \hat{H}_{1,1} & \hat{H}_{2,1} & \cdots & \hat{H}_{K,1} & \hat{H}_{1,2} & \cdots & \hat{H}_{K,2} & \cdots & \hat{H}_{K,L} \end{bmatrix}^T. \tag{2.95}$$

For all positions where no reference signals were sent, the corresponding element of $\hat{\boldsymbol{h}}_r$ is set to zero. Applying the matrices for interpolation and smoothing in time $\boldsymbol{A}_t$ and frequency $\boldsymbol{A}_f$ domain, we get the overall estimate of the channel at each position

$$\hat{\boldsymbol{h}} = \boldsymbol{A}_{tf}\hat{\boldsymbol{h}}_r = (\boldsymbol{A}_t \otimes \boldsymbol{A}_f)\,\hat{\boldsymbol{h}}_r, \tag{2.96}$$

where we assume that the interpolation matrix in the time direction on each SC is $\boldsymbol{A}_t$ and the interpolation in frequency direction on each OFDM symbol containing reference signals is $\boldsymbol{A}_f$. If this is not satisfied it is not possible to decompose the overall interpolation matrix in this form. The vector $\boldsymbol{h}$ has the same stacking of SCs and OFDM symbols as $\hat{\boldsymbol{h}}$ and represents the actual channel.

In many cases the covariance of the channel is unknown, and it is necessary to generate the interpolation matrices based on a model for the covariance matrix. The parameters of this model also need to be estimated. It is important to mention that the same interpolation matrix $\boldsymbol{A}_f$ is used for all OFDM symbols containing reference symbols. This means that if the SCs with known reference signals are not the same for different OFDM symbols with reference signals, we need to apply different interpolation matrices $\boldsymbol{A}_f$ for each symbol. In the case of 3GPP NR OFDM Type I DMRS the pattern on each OFDM symbol carrying reference signals is the same. In this work we assume the Doppler spread to be small enough that the channel can be assumed to remain static in one slot of 14 OFDM symbols. Therefore, the time interpolation matrix $\boldsymbol{A}_t$ consists only of an averaging among the OFDM symbols that contain reference symbols.

2.5.1 Calculation of the Interpolation Matrices

The frequency interpolation and smoothing matrix $\boldsymbol{A}_f$ is based on the MMSE solution described in [35], based on the observation that close to no performance loss is observed when separating the time and frequency interpolation of the channel estimation. In this fashion we will first interpolate the channel estimation in all SCs in the OFDM symbols containing DMRS and afterwards interpolate this estimate in time direction.

To generate these matrices based on this method, we need to generate the auto- and cross-correlation of the channel. In order to limit the computational complexity as well as the overfitting to the model the number of SCs with reference signals that are taken into account for the interpolation of the channel estimate at each SC is limited to K_C. For this we first define the set $\mathbb{K}$ containing all combinations of values k, ℓ at SC k and OFDM symbol ℓ where reference signals for a single user are located. We can further define for each SC k the set $\mathbb{K}_k(K_C)$ consisting of the indices of K_C SCs with reference symbols having the minimum distance to k out of $\mathbb{K}$ on the same OFDM symbol. An example for $K_C = 8$ for one value of k is illustrated in Fig. 2.29.

The vectors $\boldsymbol{g} \in \mathbb{C}^{K\times 1}$ is defined as the channel on each SC based on the model that is used for the generation of the interpolation matrix $\boldsymbol{A}_f$. Note that we implicitly assume that the statistics of the channel in a slot stay the same. Based on this definition of $\boldsymbol{g}$ we define the vector $\boldsymbol{g}_k \in \mathbb{C}^{K_C\times 1}$ containing all SCs containing reference signals to be considered for the interpolation of the channel

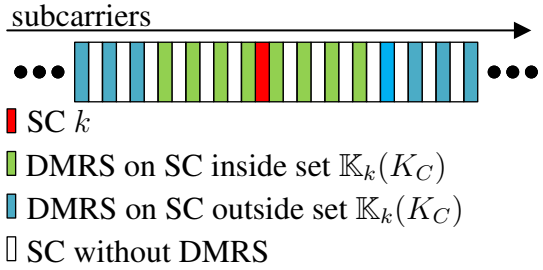

Fig. 2.29. Example of the elements of K_C for a value of k.

at SC k dependent on K_C as:

$$\boldsymbol{g}_k = \sum_{k_1 \in \mathbb{K}_k(K_C)} \boldsymbol{e}_{k_1}^T \boldsymbol{g} \boldsymbol{f}_{i(k_1)}, \tag{2.97}$$

where the vector $\boldsymbol{e}_{k_1} \in \mathbb{C}^{K \times 1}$ is a vector with only zeros, and a one at the k_1th position. The vector $\boldsymbol{f}_{i(k_1)} \in \mathbb{C}^{K_C \times 1}$ is a vector with only zeros, and a one at the $i(k_1)$th position, where $i(k_1)$ is the index of the element k_1 inside the set $\mathbb{K}_k(K_C)$. Based on the definition of these vectors we can define the matrices $\boldsymbol{R}_{\boldsymbol{g}_k \boldsymbol{g}_k} \in \mathbb{C}^{K_C \times K_C}$, $\boldsymbol{R}_{\boldsymbol{\eta}_k \boldsymbol{\eta}_k} \in \mathbb{C}^{K_C \times K_C}$ and the vector $\boldsymbol{r}_{k\boldsymbol{g}_k}^T \in \mathbb{C}^{1 \times K_C}$ as:

$$\begin{aligned} \boldsymbol{R}_{\boldsymbol{g}_k \boldsymbol{g}_k} &= \mathbb{E}\left[\boldsymbol{g}_k \boldsymbol{g}_k^H\right], \\ \boldsymbol{R}_{\boldsymbol{\eta}_k \boldsymbol{\eta}_k} &= \sum_{k_1 \in \mathbb{K}_k(K_C)} \sum_{k_2 \in \mathbb{K}_k(K_C)} \boldsymbol{e}_{k_1}^T \boldsymbol{R}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{mf} \boldsymbol{e}_{k_2} \boldsymbol{f}_{i(k_1)} \boldsymbol{f}_{i(k_2)}^T, \\ \boldsymbol{r}_{k\boldsymbol{g}_k}^T &= \mathbb{E}\left[[\boldsymbol{g}]_k \boldsymbol{g}_k^H\right]. \end{aligned} \tag{2.98}$$

The matrix $\boldsymbol{R}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{mf}$ is the covariance matrix of the noise on all SCs of one OFDM symbol in the model under the assumption that the noise plus interference has zero mean. During the calculation of the interpolation matrix it is possible that frequency selective interference is considered. In this case $\boldsymbol{R}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{mf}$ would not be a scaled identity matrix, but a diagonal matrix.

In addition, for a practical system it is often not possible to directly observe and estimate the covariance matrix of the channel. Consequently, we use a model to generate the covariance matrices. For many real world scenarios, multipath signals that arrive later at the receiver propagate over a longer distance, thus leading to lower energy at the receiver. Since this can be well approximated by an exponential Power Delay Profile (PDP), we use this as a model for the interpolation matrix. Since the elements of $\boldsymbol{R}_{\boldsymbol{g}_k \boldsymbol{g}_k}$ and $\boldsymbol{r}_{k\boldsymbol{g}_k}^T$ represent the cross correlation of different elements of the channel, all elements of these matrices are defined by the cross correlation between the channels on SC k_1 and k_2:

$$\mathbb{E}[g_{k_1} g_{k_2}^*] = \frac{1}{1 - j2\pi\tau_{\text{RMS}}\Delta_f d(k_1, k_2)}, \tag{2.99}$$

where τ_{RMS} is the Root Mean Square (RMS) delay spread as defined in [33] and $d(k_1, k_2)$ the distance in term of SCs between k_1th and k_2th SCs. The term Δ_f is the SC spacing. Another possible assumption that is less frequently used is a rectangular PDP. Since the RMS delay spread usually remains stable over a long period it can be estimated, thus we assume that it is known. In a practical system the delay spread can be estimated.

Based on these definitions we can calculated the kth row of $\boldsymbol{A}_f' \in \mathbb{C}^{K \times K_C}$ denote as $[\boldsymbol{A}_f'^T]_k$

$$[\boldsymbol{A}_f'^T]_k = a_k \left(\boldsymbol{r}_{k\boldsymbol{g}_k}^T \left(\boldsymbol{R}_{\boldsymbol{g}_k \boldsymbol{g}_k} + \boldsymbol{R}_{\boldsymbol{\eta}_k \boldsymbol{\eta}_k}\right)^{-1}\right)^T, \tag{2.100}$$

where the scalar value a_k have the purpose to remove the bias of the resulting interpolation calculated as:

$$a_k = \frac{1}{1 - \boldsymbol{r}_{k\boldsymbol{g}_k}^T \left(\boldsymbol{R}_{\boldsymbol{g}_k\boldsymbol{g}_k} + \boldsymbol{R}_{\boldsymbol{\eta}_k\boldsymbol{\eta}_k}\right)^{-1} \boldsymbol{r}_{k\boldsymbol{g}_k}^*}. \tag{2.101}$$

It needs to be noted that as these statistics used for the calculation of the interpolation matrix will not be the same as the ones of the actual channel it is unclear of removing the bias does improve the system performance or not.

The kth row of $\boldsymbol{A}_f \in \mathbb{C}^{K\times K}$ can then be constructed form the non zero elements as

$$[\boldsymbol{A}_f^T]_k = \boldsymbol{M}_k [\boldsymbol{A}_f'^T]_k, \tag{2.102}$$

where $\boldsymbol{M}_k$ has dimensions $K \times K_C$ is a matrix that is mapping the entries in $[\boldsymbol{A}_f'^T]_k$ to the right position in the $\boldsymbol{h}_\ell$ according to the values in $\mathbb{K}_k(K_C)$, thus having only entries $\in \{0, 1\}$. Note that as the set $\mathbb{K}_k(K_C)$ is different for each SC k the matrix $\boldsymbol{M}_k$ is also different for each k.

For many practical systems these values of $\boldsymbol{A}_f$ can be computed during the system design for multiple SNR and delay spread configurations. Afterwards, based on the estimation of these parameters the correct matrix $\boldsymbol{A}_f$ is selected at runtime.

As written above the examples in this work do not consider the fact that the channel can change in different OFDM symbols of the same transmission block. However, if the coherence time is less than the transmission block this also needs to be considered. In the same way as for the frequency direction the time direction interpolation is based on a MMSE approach as shown in [35]. Following the same arguments as for the frequency direction it is also not possible to directly observe the covariance matrices, thus also a model based approach needs to be chosen. In this case we assume that the Doppler spread was estimated. As this is a parameter that depends on the movement of the transmitter and receiver as well as the environment, it can be assumed to stay constant for multiple transmission blocks.

After the frequency interpolation there is an estimate of the channel available in each SC for each OFDM symbol carrying reference signals. Thus, we can define the set of OFDM symbol indices for these symbols as $\mathbb{L}$. We define the vector $\boldsymbol{g}' \in \mathbb{C}^{L\times 1}$ containing the modeled channel on one SC in all OFDM symbols of a block. Based on the set $\mathbb{L}$ with $|\mathbb{L}| = L_C$, we define a vector containing only the entries at the OFDM symbols with reference signals:

$$\boldsymbol{g}_\ell' = \sum_{\ell\in\mathbb{L}} \boldsymbol{e}_\ell^T \boldsymbol{g}' \boldsymbol{f}_{i(\ell)}, \tag{2.103}$$

where the vector $\boldsymbol{e}_\ell \in \mathbb{C}^{L\times 1}$ is a vector with only zeros, and a one at the ℓth position. The vector $\boldsymbol{f}_{i(\ell)} \in \mathbb{C}^{L_C\times 1}$ is a vector with only zeros, and a one at the $i(\ell)$th position, where $i(\ell)$ is the index of the element ℓ inside the set $\mathbb{L}$.

Based on the definition of these vectors we can define the matrices $\boldsymbol{R}_{\boldsymbol{g}'\boldsymbol{g}_\ell'} \in \mathbb{C}^{L\times L_C}$ and $\boldsymbol{R}_{\boldsymbol{g}_\ell'\boldsymbol{g}_\ell'} \in \mathbb{C}^{L_C\times L_C}$ as:

$$\begin{aligned} \boldsymbol{R}_{\boldsymbol{g}'\boldsymbol{g}_\ell'} &= \mathbb{E}\left[\boldsymbol{g}'\boldsymbol{g}_\ell'^H\right], \\ \boldsymbol{R}_{\boldsymbol{g}_\ell'\boldsymbol{g}_\ell'} &= \mathbb{E}\left[\boldsymbol{g}_\ell'\boldsymbol{g}_\ell'^H\right]. \end{aligned} \tag{2.104}$$

We further define the matrix $\boldsymbol{R}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{mt} \in \mathbb{C}^{L_C\times L_C}$ to be the time direction covariance of the noise for the model used to calculate $\boldsymbol{A}_t$. However, for this noise it is not possible to assume time dependent interference as for the calculation of $\boldsymbol{A}_t$ we assumed that the noise covariance it the same on each OFDM symbol. This means $\boldsymbol{R}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{mt}$ is always a scaled identity matrix that needs to take into account

the remaining noise after the frequency direction interpolation. Not assuming time dependent interference within a slot is a very common case for a synchronized system as for these systems the main source of interference would be inter cell interference. With synchronized slots across cells this interference would change in frequency direction but not in time direction within a slot.

Following the same arguments as for the frequency direction based on the common model of the Jake's spectrum the time direction can be modeled as:

$$\mathbb{E}[g'_{\ell_1} g'^{*}_{\ell_2}] = J_0\left(2\pi f_{\max} T d_t(\ell_1, \ell_2)\right), \tag{2.105}$$

where J_0, $f_{\max}$, T and $d_t(\ell_1, \ell_2)$ are the zero order Bessel function, the maximum Doppler frequency, the length of an OFDM symbol (including the CP) and the difference in OFDM symbols for symbol index ℓ_1 and ℓ_2. The parameter $f_{\max}$ needs to be estimated. As it is dependent only on the relative movement of the transmitter and receiver it can be assumed to stay constant for multiple transmission slots. It is also possible to assume a uniform Doppler spectrum instead of the Jake's one. This results in a sinc function for the correlation.

Combining all these definitions we use the following MMSE based result for $\boldsymbol{A}_t \in \mathbb{C}^{L\times L}$:

$$\boldsymbol{A}_t = \boldsymbol{D}_t \boldsymbol{R}_{\boldsymbol{g}'\boldsymbol{g}'_\ell} \left(\boldsymbol{R}_{\boldsymbol{g}'_\ell \boldsymbol{g}'_\ell} + \boldsymbol{R}^{mt}_{\boldsymbol{\eta}\boldsymbol{\eta}}\right)^{-1} \boldsymbol{M}_t, \tag{2.106}$$

where $\boldsymbol{D}_t \in \mathbb{C}^{L\times L}$ is a diagonal matrix to remove the bias in the estimation for each OFDM symbol with entries defined as:

$$[\boldsymbol{D}_t]_{\ell,\ell} = \frac{1}{1 - \boldsymbol{e}_\ell^T \boldsymbol{R}_{\boldsymbol{g}'\boldsymbol{g}'_\ell} \left(\boldsymbol{R}_{\boldsymbol{g}'_\ell \boldsymbol{g}'_\ell} + \boldsymbol{R}^{mt}_{\boldsymbol{\eta}\boldsymbol{\eta}}\right)^{-1} \boldsymbol{R}^H_{\boldsymbol{g}'\boldsymbol{g}'_\ell} \boldsymbol{e}_\ell}, \tag{2.107}$$

and $\boldsymbol{M}_t$ is the matrix with dimensions $L_C \times L$ that is mapping the entries of $\boldsymbol{g}'_\ell$ to the corresponding entries in $\boldsymbol{g}'$. It needs to be noted that as the actual channel statistics will be different than the ones in the model and it is unclear if removing the bias does improve the system performance or not. As in case of the frequency direction interpolation multiple values for the matrix $\boldsymbol{A}_t$ for different parameters can be precomputed and selected at run time.

2.5.2 Calculation of the Analytic Channel Estimation Error

The Mean Square Error (MSE) of the estimate $\hat{\boldsymbol{h}}$ compared to the actual channel $\boldsymbol{h}$ can be calculated as

$$\begin{gathered}\frac{1}{KL}\mathbb{E}\left[\left|\left|\hat{\boldsymbol{h}} - \boldsymbol{h}\right|\right|_2^2\right] = \\ \frac{1}{KL}\left(\mathbb{E}\left[\hat{\boldsymbol{h}}^H \hat{\boldsymbol{h}}\right] - 2\Re\left\{\mathbb{E}\left[\hat{\boldsymbol{h}}^H \boldsymbol{h}\right]\right\} + \mathbb{E}\left[\boldsymbol{h}^H \boldsymbol{h}\right]\right).\end{gathered} \tag{2.108}$$

We split the term in (2.108) into three components and calculate them separately.

The third component can be calculated as

$$\mathbb{E}\left[\boldsymbol{h}^H \boldsymbol{h}\right] = \mathrm{tr}(\boldsymbol{R}_{\boldsymbol{h}\boldsymbol{h}}) = \mathrm{tr}(\boldsymbol{R}^t_{\boldsymbol{h}\boldsymbol{h}} \otimes \boldsymbol{R}^f_{\boldsymbol{h}\boldsymbol{h}}) = \mathrm{tr}(\boldsymbol{R}^t_{\boldsymbol{h}\boldsymbol{h}})\mathrm{tr}(\boldsymbol{R}^f_{\boldsymbol{h}\boldsymbol{h}}). \tag{2.109}$$

The covariance matrices $\boldsymbol{R}^t_{\boldsymbol{h}\boldsymbol{h}}$ and $\boldsymbol{R}^f_{\boldsymbol{h}\boldsymbol{h}}$ are the time, and frequency covariance matrices of the channel. It is important to keep in mind that this separation might not be possible across all domains, depending on the channel statistics. Note that this also assumes that the channel statistics do not change during the transmission block. The channel model chosen in this work allows this

separation. For many practical channels this is also the case. Based on this assumption the matrices $\boldsymbol{R}^t_{\boldsymbol{hh}}$ and $\boldsymbol{R}^f_{\boldsymbol{hh}}$ are defined as:

$$\begin{aligned}\boldsymbol{R}^t_{\boldsymbol{hh}} &= \mathbb{E}\left[\boldsymbol{h}_t \boldsymbol{h}_t^H\right], \\ \boldsymbol{R}^f_{\boldsymbol{hh}} &= \mathbb{E}\left[\boldsymbol{h}_f \boldsymbol{h}_f^H\right],\end{aligned} \tag{2.110}$$

where $\boldsymbol{h}_t$ is the channel on one SC and $\boldsymbol{h}_f$ is the channel on one OFDM symbol. As we are only interested in the statistics and assume that the statistics are the same on each SC for $\boldsymbol{h}_t$ and on each OFDM symbol for $\boldsymbol{h}_f$, changing which SC and OFDM symbol are used does not change the result.

The first component can be calculated as

$$\begin{aligned}\mathbb{E}\left[\hat{\boldsymbol{h}}^H \hat{\boldsymbol{h}}\right] &= \operatorname{tr}\left(\boldsymbol{A}_{tf} \mathbb{E}\left[\hat{\boldsymbol{h}}_r \hat{\boldsymbol{h}}_r^H\right] \boldsymbol{A}_{tf}^H\right), \\ \mathbb{E}\left[\hat{\boldsymbol{h}}_r \hat{\boldsymbol{h}}_r^H\right] &= \sum_{k_1,\ell_1 \in \mathbb{K}} \sum_{k_2,\ell_2 \in \mathbb{K}} \left[\boldsymbol{R}_{\boldsymbol{hh}} + \boldsymbol{R}_{\boldsymbol{\eta\eta}}\right]_{k_1,k_2} \boldsymbol{e}_{k_1} \boldsymbol{e}_{k_2}^T,\end{aligned} \tag{2.111}$$

where $\boldsymbol{R}_{\boldsymbol{\eta\eta}}$ is the covariance matrix of the noise across time and frequency. For this specific case the notation $k, \ell \in \mathbb{K}$ selects all values of k, ℓ within the set $\mathbb{K}$. It needs to be noted as the same SC could have a known reference signal at different OFDM symbols, some values of k could occur multiple times. We assume it can also be separated into the sub matrices for time and frequency in the same way as the channel:

$$\boldsymbol{R}_{\boldsymbol{\eta\eta}} = \boldsymbol{R}^t_{\boldsymbol{\eta\eta}} \otimes \boldsymbol{R}^f_{\boldsymbol{\eta\eta}}. \tag{2.112}$$

The second component of (2.108) can be calculated in a similar fashion as the previous one

$$\begin{aligned}\mathbb{E}\left[\hat{\boldsymbol{h}}^H \boldsymbol{h}\right] &= \operatorname{tr}\left(\mathbb{E}\left[\boldsymbol{h} \hat{\boldsymbol{h}}_r^H\right] \boldsymbol{A}_{tf}^H\right), \\ \mathbb{E}\left[\boldsymbol{h} \hat{\boldsymbol{h}}_r^H\right] &= \sum_{k,\ell \in \mathbb{K}} \boldsymbol{R}_{\boldsymbol{hh}} \boldsymbol{e}_k \boldsymbol{e}_k^T,\end{aligned} \tag{2.113}$$

using that fact that the noise has zero mean.

Plugging (2.109), (2.111) and (2.113) into (2.108) we get the analytic MSE as

$$\begin{aligned}&\frac{1}{KL}\mathbb{E}\left[\left|\left|\hat{\boldsymbol{h}} - \boldsymbol{h}\right|\right|_2^2\right] = \\ &\frac{1}{KL}\left[\operatorname{tr}\left(\boldsymbol{A}_{tf}\left(\sum_{k_1,\ell_1 \in \mathbb{K}} \sum_{k_2,\ell_2 \in \mathbb{K}} \left[\boldsymbol{R}_{\boldsymbol{hh}} + \boldsymbol{R}_{\boldsymbol{\eta\eta}}\right]_{k_1,k_2} \boldsymbol{e}_{k_1} \boldsymbol{e}_{k_2}^T\right) \boldsymbol{A}_{tf}^H\right)\right. \\ &\quad - 2\Re\left\{\operatorname{tr}\left(\left(\sum_{k,\ell \in \mathbb{K}} \boldsymbol{R}_{\boldsymbol{hh}} \boldsymbol{e}_k \boldsymbol{e}_k^T\right) \boldsymbol{A}_{tf}^H\right)\right\} \\ &\quad \left. + \operatorname{tr}\left(\boldsymbol{R}_{\boldsymbol{hh}}\right)\right].\end{aligned} \tag{2.114}$$

If we can decompose the matrices $\boldsymbol{A}_{tf}$, $\boldsymbol{R}_{\boldsymbol{hh}}$ and $\boldsymbol{R}_{\boldsymbol{\eta\eta}}$ into the Kronecker product of two matrices the computation of the MSE can be simplified to:

$$\frac{1}{KL}\mathbb{E}\left[\left|\left|\hat{\boldsymbol{h}} - \boldsymbol{h}\right|\right|_2^2\right] = \frac{1}{KL}\left[C1 - 2\Re\{C2\} + C3\right], \tag{2.115}$$

with the components $C1$, $C2$ and $C3$ defined as:

$$C1 = \sum_{k_1,\ell_1 \in \mathbb{K}} \sum_{k_2,\ell_2 \in \mathbb{K}} \left([\boldsymbol{R}^t_{\boldsymbol{hh}}]_{\ell_1,\ell_2} [\boldsymbol{R}^f_{\boldsymbol{hh}}]_{k_1,k_2} + [\boldsymbol{R}^t_{\boldsymbol{\eta\eta}}]_{\ell_1,\ell_2} [\boldsymbol{R}^f_{\boldsymbol{\eta\eta}}]_{k_1,k_2} \right) \left([\boldsymbol{A}_t]^H_{\ell_2} [\boldsymbol{A}_t]_{\ell_1} \right) \left([\boldsymbol{A}_f]^H_{k_2} [\boldsymbol{A}_f]_{k_1} \right)$$

$$C2 = \sum_{k,\ell \in \mathbb{K}} [\boldsymbol{A}_t^H \boldsymbol{R}^t_{\boldsymbol{hh}}]_{\ell,\ell} [\boldsymbol{A}_f^H \boldsymbol{R}^f_{\boldsymbol{hh}}]_{k,k}$$

$$C3 = \operatorname{tr}\left(\boldsymbol{R}^t_{\boldsymbol{hh}}\right) \operatorname{tr}\left(\boldsymbol{R}^f_{\boldsymbol{hh}}\right), \tag{2.116}$$

where k, k_1, k_2, ℓ, ℓ_1 and ℓ_2 are the frequency and time indices corresponding the position of the reference symbols defined by the set $\mathbb{K}$.

The time and frequency covariance matrices $\boldsymbol{R}^t_{\boldsymbol{hh}}$ and $\boldsymbol{R}^f_{\boldsymbol{hh}}$ can be calculated according to the actual PDP and the Doppler shift including the corresponding model as shown in [15]. For the theoretical analysis of the channel estimation MSE we therefore need to select a channel model.

Based on the correlation matrix $\boldsymbol{R}_{\text{CIR}}$ of the Channel Impulse Response (CIR) we can calculate the correlation matrix in the frequency domain $\boldsymbol{R}^f_{\boldsymbol{hh}}$ as

$$\boldsymbol{R}^f_{\boldsymbol{hh}} = \boldsymbol{F} \boldsymbol{R}_{\text{CIR}} \boldsymbol{F}^H, \tag{2.117}$$

where $\boldsymbol{F}$ is the matrix corresponding to the DFT transformation. This means $\boldsymbol{R}_{\text{CIR}}$ can also be described as a diagonal matrix that has the elements of the power delay profile with the systems sampling frequency on the diagonal. Note that this is only valid if the system is synchronized in time and the CP is longer than the maximum channel delay.

In this work we assume that the coherence time is sufficiently longer than L. This results in the time covariance matrix $\boldsymbol{R}^t_{\boldsymbol{hh}}$ being a matrix with all entries equal to one.

In some case it might also be interesting to evaluate the channel estimation error individually per element in the time frequency grid. Especially at the boundaries of the allocated resources the interpolation error is going be larger than in the middle of them. To see this position dependency, we derive the per element channel estimation MSE. If we are only interested in the jth element of $\hat{\boldsymbol{h}}$ the channel estimation MSE is defined as:

$$\mathbb{E}\left[\left| \boldsymbol{e}_j^T \hat{\boldsymbol{h}} - \boldsymbol{e}_j^T \boldsymbol{h} \right|^2 \right] = \boldsymbol{e}_j^T \mathbb{E}\left[\hat{\boldsymbol{h}} \hat{\boldsymbol{h}}^H \right] \boldsymbol{e}_j - 2 \boldsymbol{e}_j^T \Re\left\{ \mathbb{E}\left[\boldsymbol{h} \hat{\boldsymbol{h}}^H \right] \right\} \boldsymbol{e}_j + \boldsymbol{e}_j^T \mathbb{E}\left[\boldsymbol{h} \boldsymbol{h}^H \right] \boldsymbol{e}_j. \tag{2.118}$$

If we look at (2.109), (2.111) and (2.113) we see that the calculation of the covariance matrices are the same ones as the ones we need for (2.118). This means we can calculate the MSE as:

$$\begin{aligned} &\mathbb{E}\left[\left| \boldsymbol{e}_j^T \hat{\boldsymbol{h}} - \boldsymbol{e}_j^T \boldsymbol{h} \right|^2 \right] = \\ &\boldsymbol{e}_j^T \boldsymbol{A}_{tf} \left(\sum_{k_1,\ell_1 \in \mathbb{K}} \sum_{k_2,\ell_2 \in \mathbb{K}} [\boldsymbol{R}_{\boldsymbol{hh}} + \boldsymbol{R}_{\boldsymbol{\eta\eta}}]_{k_1,k_2} \boldsymbol{e}_{k_1} \boldsymbol{e}_{k_2}^T \right) \boldsymbol{A}_{tf}^H \boldsymbol{e}_j \\ &- 2\Re\left\{ \boldsymbol{e}_j^T \left(\sum_{k,\ell \in \mathbb{K}} \boldsymbol{R}_{\boldsymbol{hh}} \boldsymbol{e}_k \boldsymbol{e}_k^T \right) \boldsymbol{A}_{tf}^H \boldsymbol{e}_j \right\} \\ &+ \boldsymbol{e}_j^T \boldsymbol{R}_{\boldsymbol{hh}} \boldsymbol{e}_j. \end{aligned} \tag{2.119}$$

Again, if we can decompose the matrices $\boldsymbol{A}_{tf}$, $\boldsymbol{R}_{\boldsymbol{hh}}$ and $\boldsymbol{R}_{\boldsymbol{\eta\eta}}$ into the Kronecker product of two matrices the computation can be simplified to:

$$\mathbb{E}\left[\left|\boldsymbol{e}_j^T\hat{\boldsymbol{h}} - \boldsymbol{e}_j^T\boldsymbol{h}\right|^2\right] = C1 - 2\Re\{C2\} + C3, \tag{2.120}$$

with the components $C1$, $C2$ and $C3$ defined as:

$$\begin{aligned} C1 &= \sum_{k_1,\ell_1\in\mathbb{K}}\sum_{k_2,\ell_2\in\mathbb{K}}\left([\boldsymbol{R}^t_{\boldsymbol{hh}}]_{\ell_1,\ell_2}[\boldsymbol{R}^f_{\boldsymbol{hh}}]_{k_1,k_2} + [\boldsymbol{R}^t_{\boldsymbol{\eta\eta}}]_{\ell_1,\ell_2}[\boldsymbol{R}^f_{\boldsymbol{\eta\eta}}]_{k_1,k_2}\right)[\boldsymbol{A}_t]^*_{\ell_j,\ell_2}[\boldsymbol{A}_t]_{\ell_j,\ell_1}[\boldsymbol{A}_f]^*_{k_j,k_2}[\boldsymbol{A}_f]_{k_j,k_1} \\ C2 &= \sum_{k,\ell\in\mathbb{K}}[\boldsymbol{A}_t]^*_{\ell_j,\ell}[\boldsymbol{R}^t_{\boldsymbol{hh}}]_{\ell_j,\ell}[\boldsymbol{A}_f]^*_{k_j,k}[\boldsymbol{R}^f_{\boldsymbol{hh}}]_{k_j,k} \\ C3 &= [\boldsymbol{R}^t_{\boldsymbol{hh}}]_{\ell_j,\ell_j}[\boldsymbol{R}^f_{\boldsymbol{hh}}]_{k_j,k_j}, \end{aligned} \tag{2.121}$$

where k, k_1, k_2, ℓ, ℓ_1 and ℓ_2 are the frequency and time indices corresponding the position of the reference symbols and ℓ_j and k_j are the time frequency indices corresponding to the selected element j of $\hat{\boldsymbol{h}}$.

As the absolute value of the transmit signal is assumed to be 1 and only one antenna is treated the SNR in dB for the following simulations is defined as

$$\text{SNR} = 10\log\left(\frac{\text{tr}\left(\boldsymbol{R}_{\boldsymbol{hh}}\right)}{\text{tr}\left(\boldsymbol{R}_{\boldsymbol{\eta\eta}}\right)}\right), \tag{2.122}$$

where the expectation of the channel is going over all possible channel realizations of the given time interval, including the changes of the channel inside the interval.

2.5.3 Simulation Results

For the simulations we use the channel model described in [84]. In this case we assumed that the estimation procedures for the delay spread, the Doppler spread and the SNR to be ideal. However, as these parameters do not sufficiently describe the channel given by the channel model this does not result in ideal knowledge of the channel statistics. To generate the following results 1000 power delay profiles were generated with the model described in [84]. Then for each of these power delay profile realization the analytical MSE is calculated. In Fig. 2.30 we show the channel estimation error for different interpolation lengths averaged over 1000 channel realizations. In the low SNR region the additional benefits of averaging leads to the interpolation of length 16 having the best performance. As obvious from (2.100) at low SNR the frequency direction interpolation convergence approximately to an averaging of all K_C SCs with DMRS. Therefore, doubling K_C should result approximately in a gain of 3 dB. This is close to the actual result in our evaluation. In the high SNR regime, the longest interpolation length over-fits the exponential PDP. To achieve a good trade-off between complexity and performance over the whole SNR range we use a maximum interpolation length of eight for the rest of this work.

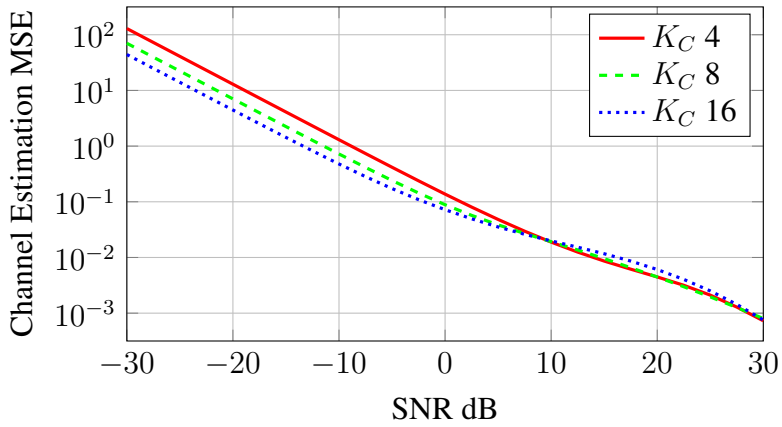

Fig. 2.30. Channel estimation MSE with different maximum interpolation length K_C.

3. Comparison of Analog, Hybrid and Digital Beamforming

3.1 Single User Scenarios

3.1.1 Introduction Single User Scenarios

In this section we compare hybrid and digital beamforming in terms of data rate and energy efficiency in single user scenarios. The general system model is shown in Fig. 3.1. The details of the blocks shown in the system model figure are described in the power model in Section 2.4. This section is structured as follows: First we show the signal model and the rate calculations in relation to the general models introduced in the preceding chapter. Afterwards we show different simulation results in various scenarios. Most of the material in this subsection appeared also in [77].

3.1.2 Signal Model Single User Scenarios

The signal model is shown in Fig. 3.2. The symbols $\boldsymbol{x}[n]$, $\boldsymbol{H}[n]$, $\boldsymbol{\eta}[n]$, and $\boldsymbol{y}[n]$ represent the transmit signal, channel, noise, and receive signal of a system at time n. M_T transmit and M receive antennas are used. Since we assume a channel with multipath propagation the receive signal $\boldsymbol{y}[n]$ is defined as:

$$\boldsymbol{y}[n] = \sum_{l=0}^{L-1} \boldsymbol{H}[l]\boldsymbol{x}[n-l] + \boldsymbol{\eta}[n], \tag{3.1}$$

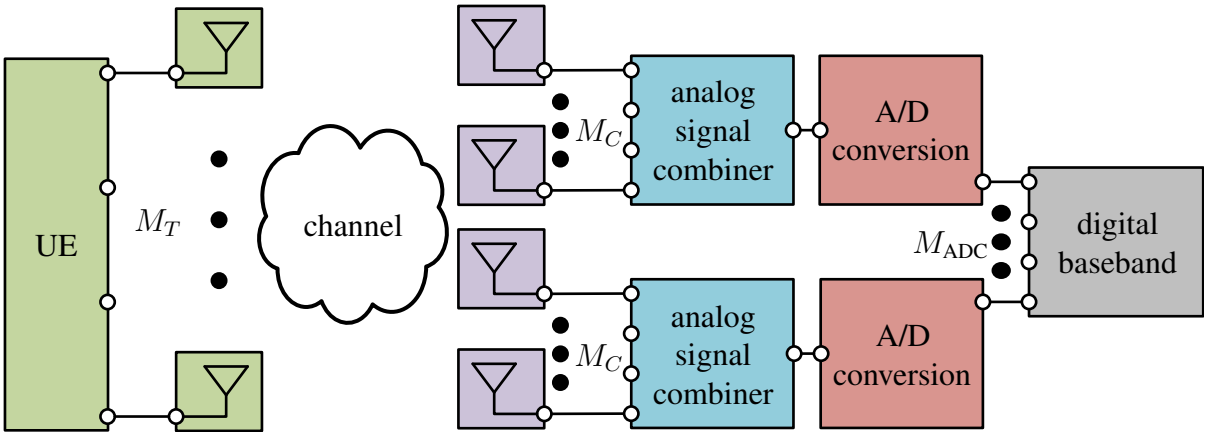

Fig. 3.1. System model with M_T transmit antennas and M_C antennas at each of the M_{ADC} ADC chains. The number of receive antennas M is equal to $M_C \cdot M_{\text{ADC}}$.

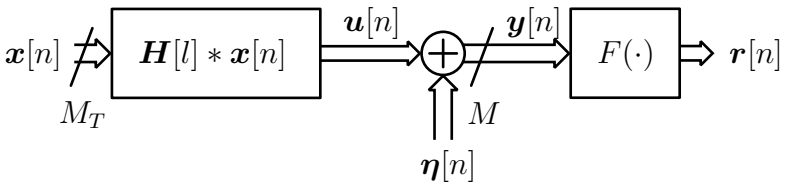

Fig. 3.2. Single user signal model.

where L is the maximum delay of the channel in samples. The operation $F(\cdot)$ is defined as multiplication with the analog receiver beamforming matrix $\boldsymbol{W}_R$ followed by a quantization operation $Q_b(\cdot)$ with resolution of b bits:

$$\boldsymbol{r}[n] = F(\boldsymbol{y}[n]) = Q_b(\boldsymbol{y}_C[n]) = Q_b(\boldsymbol{W}_R^H \boldsymbol{y}[n]). \tag{3.2}$$

We restricted the system to have M_C antennas exclusively connected to one ADC chain (see Fig. 3.1). Therefore, the matrix $\boldsymbol{W}_R$ has the form:

$$\boldsymbol{W}_R = \begin{bmatrix} \boldsymbol{w}_R^1 & \mathbf{0}_{M_C} & \cdots & \mathbf{0}_{M_C} \\ \mathbf{0}_{M_C} & \boldsymbol{w}_R^2 & \ddots & \mathbf{0}_{M_C} \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0}_{M_C} & \cdots & \mathbf{0}_{M_C} & \boldsymbol{w}_R^{M_{\text{ADC}}} \end{bmatrix} \in \mathbb{C}^{M \times M_{\text{ADC}}}, \tag{3.3}$$

where the vector $\boldsymbol{w}_R^i$ is the analog beamforming vector of the ith ADC chain. We also restrict our evaluation to each ADC chain utilizing the same number of antennas M_C. The vectors $\boldsymbol{w}_R^i$ and $\mathbf{0}_{M_C}$ have dimension M_C.

The use of analog beamforming is envisioned in many future mobile broadband systems, especially in the mmWave frequency range ([90, 93]). Since the complete channel matrix cannot be directly observed, one practical solution is scanning different spatial directions (beams) and then select the configuration maximizing the SNR. There are many different possibilities for selecting the optimal beam, e.g. 802.11ad is using a procedure based on exhaustive search [37].

For the evaluation, we assume that the antennas of each ADC chain form an ULA. If a planar wavefront is impinging on the ULA and the spacing of adjacent antennas is $d = \lambda/2$, the receive signal at adjacent antennas is phase shifted by $\psi_i = \pi \sin(\phi_i)$. The angle ϕ_i is the angle of a planar wavefront relative to the antennas of the ULA. This formula assumes that a planar wavefront is impinging at the antenna array, and that the symbol duration is large relative to the maximum delay between two antennas. With the constraint of observing only a single spatial direction, the receive vector $\boldsymbol{w}_R^i$ for an ULA antenna array takes the form:

$$\boldsymbol{w}_R^i = \begin{bmatrix} 1 & e^{j\psi_i} & e^{j2\psi_i} & \cdots & e^{j(M_C-1)\psi_i} \end{bmatrix}^H. \tag{3.4}$$

In the special case of full digital beamforming ($M_C = 1$ and therefore $M_{\text{ADC}} = M$), $\boldsymbol{W}_R$ is equal to the identity matrix $\boldsymbol{I}$ of size $M \times M$.

We use the quantization model developed in Section 2.1. For real world ADCs the difference between representatives of adjacent quantization bins q_j and the size of the quantization bins are uniform. Thus, we limit our evaluation to this set of quantizers. For the evaluation we assume Gaussian transmit signals. Therefore, as the noise is also assumed to be Gaussian the receive signal is also Gaussian distributed. Consequently, we use the step-size to minimize the distortion for

Gaussian signals shown in [51]. Since the actual receive power at each antenna can be different, an AGC needs to adapt a VGA to generate the minimal distortion. To simplify our model, we assume that the AGC is always perfectly adapting to the received power at each ADC chain separately. Since in practice an AGC cannot accomplish this task without error, we will show the impact of an imperfect AGC. We model this by a relative error to the perfect gain value.

For the rest of this section we define the SNR γ as the average per receive antenna SNR in the following way:

$$\gamma = \frac{1}{M} \sum_{m=1}^{M} \frac{\sum_{l=0}^{L-1} \mathbb{E}\left[\left|\left| \boldsymbol{e}_m^T \boldsymbol{H}[l] \boldsymbol{x}[n-l] \right|\right|_2^2 \right]}{\mathbb{E}[||\boldsymbol{e}_m^T \boldsymbol{\eta}[n]||_2^2]}, \tag{3.5}$$

where $\boldsymbol{e}_m$ is a vector of size $\mathbb{C}^{M \times 1}$ with all zeros except for position m which is equal to 1. It is important to note that the expectation takes the realization of the channel and realizations of the transmit signal $\boldsymbol{x}[n]$ into account.

For this section different channel models are used to highlight results in different scenarios. These channel models are:

- Finite path model with all paths arriving at the same time
- Finite path model with exponential Power Delay Profile (PDP)

The channel models assume different rays impinging on the receiver antenna array. In the first model, they are assumed to arrive at the receiver antennas at the same time. Under the assumption of an ULA at the transmitter and receiver, a channel consisting of K different rays can be modeled as:

$$\boldsymbol{H} = \frac{1}{\sqrt{K M_T}} \sum_{k=1}^{K} \alpha(k) \boldsymbol{a}_r(\psi_r(k)) \boldsymbol{a}_t^T(\psi_t(k)). \tag{3.6}$$

The vectors $\boldsymbol{a}_r(\psi_r(k))$ and $\boldsymbol{a}_t(\psi_t(k))$ are the array steering vectors at the receiver and transmitter. The phase shift between the signal of adjacent antenna elements $\psi_r(k)$ and $\psi_t(k)$ of path k depend on the angle of arrival $\phi_r(k)$ and departure $\phi_t(k)$.

$$\boldsymbol{a}_r(\psi_r(k)) = \begin{bmatrix} 1 & e^{j\psi_r(k)} & e^{j2\psi_r(k)} & \dots & e^{j(M-1)\psi_r(k)} \end{bmatrix}^T. \tag{3.7}$$

The transmit vectors $\boldsymbol{a}_t^T(\psi_r(k))$ has the same form as $\boldsymbol{a}_r^T(\psi_r(k))$ but different dimension. The complex gains $\alpha(k)$ are circular symmetric Gaussian distributed with zero mean and unit variance. Except for the different normalization factor, this channel model is the same as the one presented in [58]. The difference comes from the fact that the sum power of the transmit signal is constrained to be less or equal to M_T. To set the average per antenna receive power to one we normalize the channel by $\frac{1}{\sqrt{K M_T}}$. The angles of arrival $\phi_r(k)$ and departure $\phi_t(k)$ are uniformly distributed in the range of $-\pi$ to π.

It needs to be mentioned that the channel model above is far from realistic. It is very unlikely that multiple rays with distinct scatterers arrive at the receiver at the same time. We only use this channel model to highlight the effect of not considering multipath propagation in a system with quantization. In addition, it is used to showcase how ignoring the non-diagonal elements of the effective noise covariance influences the system performance.

Since in real world scenario the different rays are reflection of different scatterers, the path of each of these rays from the transmitter to the receiver has a different length. This results in each ray arriving at the receiver at a different time. In a simplified case, it can be expected that the path

that arrives at a later time have a lower power. The measurements in [55] show that for channels at 60 GHz an exponential PDP is sufficiently approximating a real world scenario.

$$\boldsymbol{H}[l] = \frac{1}{\sqrt{M_T}} \alpha[l] \boldsymbol{a}_r(\psi_r[l]) \boldsymbol{a}_t^T(\psi_t[l]). \tag{3.8}$$

Here we assume, that at delay l only one ray arrives at the receiver. Here the complex gain of the ray $\alpha[l]$ is circular symmetric Gaussian distributed with zero mean and a variance defined according to:

$$v_l = \mathbb{E}\left[|\alpha[l]|^2\right] = e^{-\beta l}. \tag{3.9}$$

Note that for this channel model the taps are spaced with the distance of the symbol period. This assumption does simplify the generation of the channel realization but for a real channel this will never be the case. However, this model offers a simplified approach to test the impact of multipath channels. In the following section on multiuser scenarios realistic channel models based on measurements are considered. The parameter β defines how fast the power decays in relation to the sample time. The additional parameters are the maximum channel length in samples L and the number of present channel taps P. This means that for all possible present channel rays $\boldsymbol{v}$ of dimension L, P positions are selected for each channel realization. At all other positions, $\boldsymbol{v}$ is equal to 0. To normalize the average power, the variance vector $\boldsymbol{v}$ is normalized by:

$$\boldsymbol{v}_n = \frac{\boldsymbol{v}}{||\boldsymbol{v}||^2}. \tag{3.10}$$

3.1.3 Rate Calculation Single User

In this subsection achievable rate expressions for different scenarios are derived. The different scenarios are any combination of flat fading channel, multipath channel, hybrid beamforming and digital beamforming with low resolution A/D conversion. In the case of hybrid beamforming, first the analog receive vectors are calculated. Afterwards, the system including the analog combining is treated as an equivalent channel, for which digital beamforming is performed.

3.1.3.1 Selection of the Beamforming Vectors

To mimic the behavior of a spatial scan, we restricted the receive vectors $\boldsymbol{w}_R^i$ of the ith ADC chain to Vandermonde vectors. A practical system would have a set of predefined beamforming configuration that are scanned for every subarray. To obtain the optimal results, all combination of beams needs to be tested by the receiver. This is a combinatorial problem with size growing exponentially with the number of receiver ADC chains. To make the problem feasible, the scan is performed separately for each receiver ADC chain. This problem can be formulated as:

$$\boldsymbol{w}_R^i(\hat{\phi}) = \arg\max_{\boldsymbol{w}_R^i(\phi_B)} \sum_{l=0}^{L-1} \left|\left|\boldsymbol{w}_R^i(\phi_B)^H \boldsymbol{H}^i[l]\right|\right|_2^2 \text{ with } \phi_B \in \mathbb{B}, \tag{3.11}$$

with $\mathbb{B}$ being the set of all spatial direction ϕ_B that are scanned. The number of entries in the set $\mathbb{B}$ are selected to be $4M_C$ according to the quality requirement shown in Subsection 2.3.5. The channel $\boldsymbol{H}^i[l]$ contains the M_C rows of $\boldsymbol{H}[l]$ that belong to the antennas of the ith ADC chain.

This procedure mimics the receive beam training in a practical system as described in [41]. For this case the transmitter is sending a known reference sequence. The receiver tries different receiver

Algorithm 3.1 Selection of the beamforming vectors in single user scenario.

Require: $\boldsymbol{H}[l]$, M_{ADC} and M_C
$\mathbb{B} \leftarrow \{\phi_1,\ \phi_2,\ \cdots,\ \phi_{4M_C}\}$
for $i \leftarrow 1$ to M_{ADC} **do**
  for $j \leftarrow 1$ to $4M_C$ **do**
    $\boldsymbol{w}_{tst} \leftarrow \begin{bmatrix}1 & e^{\phi_j} & \cdots & e^{(M_C-1)\phi_j}\end{bmatrix}^H$
    $p^i(j) \leftarrow \sum_{l=0}^{L-1} \left|\left|\boldsymbol{w}_{tst}^H \boldsymbol{H}^i[l]\right|\right|_2^2$
  end for
  $\hat{j} \leftarrow \underset{j}{\operatorname{argmax}}\, p^i(j)$
  $\boldsymbol{w}_R^i \leftarrow \begin{bmatrix}1 & e^{\phi_{\hat{j}}} & \cdots & e^{(M_C-1)\phi_{\hat{j}}}\end{bmatrix}^H$
end for
return $\boldsymbol{w}_R^i\ \forall i \in \{1,\ \cdots,\ M_{\text{ADC}}\}$

beamforming configurations separately on each subarray i and records the achieved channel quality metric. Afterwards, the configuration resulting in the best channel is selected. In this work such a procedure is emulated by selecting the receive beamforming vector resulting in the highest receive energy, based on the channel knowledge. This procedure avoids lengthy numerical simulation of sequence detection with different configurations but leads to the same beamformer configuration.

After selecting all beamforming vectors $\boldsymbol{w}_R^i$, the overall matrix $\boldsymbol{W}_R$ is constructed. With $\boldsymbol{W}_R$, we can generate the effective channel $\boldsymbol{H}_C[l]$:

$$\boldsymbol{H}_C[l] = \boldsymbol{W}_R^H \boldsymbol{H}[l]. \tag{3.12}$$

and the effective noise covariance matrix $\boldsymbol{R}_{\boldsymbol{\eta}_C\boldsymbol{\eta}_C}$:

$$\boldsymbol{R}_{\boldsymbol{\eta}_C\boldsymbol{\eta}_C} = \boldsymbol{W}_R^H \boldsymbol{R}_{\boldsymbol{\eta}\boldsymbol{\eta}} \boldsymbol{W}_R. \tag{3.13}$$

The effective channel and noise covariance matrix are then input to the digital system with low resolution A/D conversion. Algorithm 3.1 shows the procedure of finding the receiver beamforming vectors $\boldsymbol{w}_R^i$ given the channel and the number of antennas per ADC chain.

3.1.3.2 Calculation of the Receive Covariance Matrix

For the calculations of the parameters of the system including the in Section 2.1 described model for the quantization, it is necessary to calculate the correlation matrix $\boldsymbol{R}_{\boldsymbol{y}_C\boldsymbol{y}_C}$ of the signal after the analog combining. This signal is defined as:

$$\begin{aligned}\boldsymbol{y}_C[n] &= \boldsymbol{W}_R^H \left(\sum_{l=0}^{L-1} \boldsymbol{H}[l]\boldsymbol{x}[n-l] + \boldsymbol{\eta}[n]\right) \\ &= \boldsymbol{W}_R^H \left(\boldsymbol{u}[n] + \boldsymbol{\eta}[n]\right).\end{aligned} \tag{3.14}$$

Since the two random variables $\boldsymbol{x}$ and $\boldsymbol{\eta}$ are independent the covariance matrix decomposes into:

$$\boldsymbol{R}_{\boldsymbol{y}_C\boldsymbol{y}_C} = \boldsymbol{W}_R^H \left(\boldsymbol{R}_{\boldsymbol{u}\boldsymbol{u}} + \boldsymbol{R}_{\boldsymbol{\eta}\boldsymbol{\eta}}\right) \boldsymbol{W}_R. \tag{3.15}$$

Note that in the following derivation we also show that this process is stationary. The remaining matrix that needs to be calculated is $\boldsymbol{R}_{\boldsymbol{uu}}$.

$$\boldsymbol{R}_{\boldsymbol{uu}} = \mathbb{E}\left[\boldsymbol{u}\boldsymbol{u}^H\right]. \tag{3.16}$$

It needs to be added that this expectation is going over the realizations of $\boldsymbol{x}$ only, as for this part we only want to calculate $\boldsymbol{R}_{\boldsymbol{y}_C\boldsymbol{y}_C}$ for one realizations of the channel. To simplify the notation for the derivation we evaluate the elements of the matrix separately:

$$[\boldsymbol{R}_{\boldsymbol{uu}}]_{i,j} = \mathbb{E}\left[[\boldsymbol{u}]_i[\boldsymbol{u}]_j^*\right]. \tag{3.17}$$

This can be expanded based on the definition in (3.14)) to:

$$\mathbb{E}\left[[\boldsymbol{u}]_i[\boldsymbol{u}]_j^*\right] = \mathbb{E}\left[\sum_{l_i=0}^{L-1}\sum_{m_i=1}^{M_T}[\boldsymbol{H}[l_i]]_{i,m_i}[\boldsymbol{x}[n-l_i]]_{m_i}\sum_{l_j=0}^{L-1}\sum_{m_j=1}^{M_T}[\boldsymbol{H}^*[l_j]]_{j,m_j}[\boldsymbol{x}^*[n-l_j]]_{m_j}\right]. \tag{3.18}$$

Which can be rearranged to

$$\mathbb{E}\left[[\boldsymbol{u}]_i[\boldsymbol{u}]_j^*\right] = \sum_{m_i=1}^{M_T}\sum_{m_j=1}^{M_T}\mathbb{E}\left[\sum_{l_i=0}^{L-1}[\boldsymbol{H}[l_i]]_{i,m_i}[\boldsymbol{x}[n-l_i]]_{m_i}\sum_{l_j=0}^{L-1}[\boldsymbol{H}^*[l_j]]_{j,m_j}[\boldsymbol{x}^*[n-l_j]]_{m_j}\right]. \tag{3.19}$$

The linear convolution of the transmit signal with the channel can be replaced by the corresponding frequency domain operation. As we ensure that K is larger than $2 \cdot L - 1$ this linear convolution is the same as the circular convolution of the transmit signal and the channel, if both are extended with zeros to size K. This means that we can express $\sum_{l_i=0}^{L-1}[\boldsymbol{H}[l_i]]_{i,m_i}[\boldsymbol{x}[n-l_i]]_{m_i}$ as:

$$\sum_{l_i=0}^{L-1}[\boldsymbol{H}[l_i]]_{i,m_i}[\boldsymbol{x}[n-l_i]]_{m_i} = \frac{1}{\sqrt{K}}\sum_{k=0}^{K-1}[\boldsymbol{H}[k]]_{i,m_i}[\boldsymbol{x}[k]]_{m_i}e^{j\frac{2\pi}{K}kn}, \tag{3.20}$$

where we define the frequency components of the channel and the transmit signal as $\boldsymbol{H}[k] = \mathcal{F}(\boldsymbol{H}[l])$ and $\boldsymbol{x}[k] = \mathcal{F}(\boldsymbol{x}[n])$. In this case the DFT is operating on all time instances l of each element of the matrix $\boldsymbol{H}[l]$. This means that the elements (3.20) are defined as:

$$\begin{aligned}[\boldsymbol{H}[k]]_{i,m_i} &= \frac{1}{\sqrt{K}}\sum_{n=0}^{K-1}[\boldsymbol{H}'[n]]_{i,m_i}e^{-j\frac{2\pi}{K}kn},\\ [\boldsymbol{x}[k]]_{m_i} &= \frac{1}{\sqrt{K}}\sum_{n=0}^{K-1}[\boldsymbol{x}'[n]]_{m_i}e^{-j\frac{2\pi}{K}kn}.\end{aligned} \tag{3.21}$$

Note that $\boldsymbol{H}'[n]$ is the by $K-L$ zero matrices extended version of $\boldsymbol{H}[l]$ and $\boldsymbol{x}'[n]$ is the by $K-L$ zero vectors extended version or $\boldsymbol{x}[n]$. This can be defined as:

$$\boldsymbol{H}'[n] = \begin{cases}\boldsymbol{H}[n], \text{if } n < L\\ \boldsymbol{0}, \text{otherwise}\end{cases}, \tag{3.22}$$

and

$$\boldsymbol{x}'[n] = \begin{cases} \boldsymbol{x}[n], \text{if } n < L \\ \mathbf{0}, \text{otherwise} \end{cases} . \tag{3.23}$$

Plugging (3.20) into (3.19) we get:

$$\mathbb{E}\left[[\boldsymbol{u}]_i[\boldsymbol{u}]_j^*\right] = \sum_{m_i=1}^{M_T}\sum_{m_j=1}^{M_T}\sum_{k_i=0}^{K-1}\sum_{k_j=0}^{K-1}\mathbb{E}\left[[\boldsymbol{H}[k_i]]_{i,m_i}[\boldsymbol{x}[k_i]]_{m_i}e^{j\frac{2\pi}{K}k_i n}[\boldsymbol{H}^*[k_j]]_{j,m_j}[\boldsymbol{x}^*[k_j]]_{m_j}e^{-j\frac{2\pi}{K}k_j n}\right]. \tag{3.24}$$

We can further combine this to:

$$\mathbb{E}\left[[\boldsymbol{u}]_i[\boldsymbol{u}]_j^*\right] = \sum_{m_i=1}^{M_T}\sum_{m_j=1}^{M_T}\sum_{k_i=0}^{K-1}\sum_{k_j=0}^{K-1}[\boldsymbol{H}[k_i]]_{i,m_i}\mathbb{E}\left[[\boldsymbol{x}[k_i]]_{m_i}[\boldsymbol{x}^*[k_j]]_{m_j}\right][\boldsymbol{H}^*[k_j]]_{j,m_j}e^{(k_i-k_j)j\frac{2\pi}{K}n}. \tag{3.25}$$

Now to calculate the expectation in (3.25) we essentially correlate the frequency domain signal transmitted on transmit antenna m_i on frequency bin k_i with the corresponding signal on transmit antenna m_j on frequency bin k_j. In a linear system the optimal rate can be achieve by optimizing each frequency bin k separately. Thus, we can choose the transmit signal on each frequency bin k to be independent from each other frequency bin. This leads to the expectation $\mathbb{E}\left[[\boldsymbol{x}[k_i]]_{m_i}[\boldsymbol{x}^*[k_j]]_{m_j}\right]$ only being unequal to zero if $k_i = k_j$. Since outside of the transmission bandwidth starting from frequency bin k_1 and ending at frequency bin k_2 the signal is going to be zero, we get the following expression:

$$\mathbb{E}\left[[\boldsymbol{u}]_i[\boldsymbol{u}]_j^*\right] = \sum_{m_i=1}^{M_T}\sum_{m_j=1}^{M_T}\sum_{k=k_1}^{k_2}[\boldsymbol{H}[k]]_{i,m_i}\mathbb{E}\left[[\boldsymbol{x}[k]]_{m_i}[\boldsymbol{x}^*[k]]_{m_j}\right][\boldsymbol{H}^*[k]]_{j,m_j}, \tag{3.26}$$

which shows that the process is stationary. Combining the elements of the first and second summation leads to this expression in matrix vector form:

$$\mathbb{E}\left[[\boldsymbol{u}]_i[\boldsymbol{u}]_j^*\right] = \sum_{k=k_1}^{k_2}[\boldsymbol{H}^T[k]]_i^T\boldsymbol{R}_{\boldsymbol{x}[k]\boldsymbol{x}[k]}[\boldsymbol{H}^H[k]]_j, \tag{3.27}$$

where $[\boldsymbol{H}^T[k]]_i^T$ is the ith row of $\boldsymbol{H}[k]$, $[\boldsymbol{H}^H[k]]_j$ is the jth column of $\boldsymbol{H}^H[k]$ and $\boldsymbol{R}_{\boldsymbol{x}[k]\boldsymbol{x}[k]}$ being defined as $\mathbb{E}\left[\boldsymbol{x}[k]\boldsymbol{x}^H[k]\right]$. Combining all elements of $\boldsymbol{R}_{\boldsymbol{u}\boldsymbol{u}}$ according to the definition in (3.27) leads to:

$$\boldsymbol{R}_{\boldsymbol{u}\boldsymbol{u}} = \sum_{k=k_1}^{k_2}\boldsymbol{H}[k]\boldsymbol{R}_{\boldsymbol{x}[k]\boldsymbol{x}[k]}\boldsymbol{H}^H[k], \tag{3.28}$$

Combining these results we can express the matrix $\boldsymbol{R}_{\boldsymbol{y}_C\boldsymbol{y}_C}$ as:

$$\boldsymbol{R}_{\boldsymbol{y}_C\boldsymbol{y}_C} = \boldsymbol{W}_R^H\left(\sum_{k_1}^{k_2}\boldsymbol{H}[k]\boldsymbol{R}_{\boldsymbol{x}[k]\boldsymbol{x}[k]}\boldsymbol{H}^H[k] + \boldsymbol{R}_{\boldsymbol{\eta}\boldsymbol{\eta}}\right)\boldsymbol{W}_R. \tag{3.29}$$

3.1.3.3 Problem Formulation

For the given signal model, the problem of finding the maximum achievable rate for a multipath channel with full Channel State Information at the Transmitter (CSIT) and CSIR can be formulated as:

$$\begin{aligned} R = B\frac{1}{N} \max_{p(\boldsymbol{x}^N, \boldsymbol{w}_R^i)} & I(\boldsymbol{x}^N, \boldsymbol{r}^N | \boldsymbol{H}[0], \boldsymbol{H}[1], \cdots, \boldsymbol{H}[L-1]) \\ & \text{s.t. } \mathbb{E}[||\boldsymbol{x}||_2^2] \leq P_{Tx} \\ \boldsymbol{w}_R^i = \left[1, e^{j\phi_i}, \cdots, e^{j(M_C-1)\phi_i}\right]^H & \forall i \in \{1, \cdots, M_{\text{ADC}}\}, \end{aligned} \tag{3.30}$$

with $\boldsymbol{x}^N$ and $\boldsymbol{r}^N$ being N input/output samples of the system and B is the bandwidth. Due to the non-linearity of the quantization and the non-trivial problem of finding the optimal beamforming configurations $\boldsymbol{w}_R^i$, we make a number of approximations that make the expression tractable:

- Quantization is modeled as additive Gaussian noise with the Additive Quantization Noise Model (AQNM) model including the off-diagonal elements
- Assume the frequency domain transmit signal $\boldsymbol{x}[k]$ is Gaussian distributed
- For a system with CSIT Singular Value Decomposition (SVD) based precoding is used (the SVD of the effective channel after analog combining)
- $\boldsymbol{w}_R^i$ are selected from the derived finite set separately for each antenna group based on an SNR criterion

With these simplifications the values of $\boldsymbol{w}_R^i$ are already defined. Using the AQNM we get an equivalent linear system with the effective time domain channel $\boldsymbol{H}'[l] = \boldsymbol{T}\boldsymbol{H}[l]$, with $\boldsymbol{T}$ being defined according to the derivation of the AQNM in Section 2.1. For a linear system it is possible to look at equivalent frequency domain system and optimize each frequency bin k separately. As in Subsection 3.1.3.2 we chose a number of frequency bins K that is longer than $2 \cdot L - 1$. This ensure that the linear convolution in the time domain is equivalent to the frequency domain multiplication of the frequency domain representations of the transmit signal and the channel. The equivalent channel on frequency bin k is defined as $\boldsymbol{H}'[k]$. The elements of $\boldsymbol{H}'[k]$ can be calculated via the DFT of $\boldsymbol{H}'[l]$ as described in (3.21) In [32, 60] the achievable rate of a digital beamforming system without quantization, but considering a multipath channel is described. The solution is waterfiling across the frequency bins and the spatial streams. Since for a system with low resolution ADCs the quantization does influence the signal relative to the total power, it is intuitive to use each frequency bin independent of each other. Since the optimization is carried out for each frequency bin k separately, the result only is a lower bound to the joint optimization.

$$\begin{aligned} R \geq B \sum_{k=k_1}^{k_2} \max_{\boldsymbol{R}_{\boldsymbol{x}[k]\boldsymbol{x}[k]}} & I(\boldsymbol{x}[k], \boldsymbol{r}[k] | \boldsymbol{H}'[k]) \\ \text{s.t. } & \mathbb{E}[||\boldsymbol{x}[k]||_2^2] \leq P_{Tx} \; \forall k \in [k_1, k_2], \end{aligned} \tag{3.31}$$

with $\boldsymbol{x}[k]$, $\boldsymbol{r}[k]$ and $\boldsymbol{H}'[k]$ being the input/output signal and equivalent channel of frequency bin k. The frequency bins k_1 and k_2 mark the borders of the band of interest in the equivalent baseband channel. If the sampling rate covers a larger band than the one used by the system, the parameters k_1 and k_2 have to account for the oversampling.

Since all signals are represented by Gaussian random variables, we get the following expression for the mutual information:

$$\begin{aligned} & I(\boldsymbol{x}[k], \boldsymbol{r}[k] | \boldsymbol{H}'[k]) = \\ & \log_2\left(\det\left(\boldsymbol{I} + \boldsymbol{R}_{\boldsymbol{\eta}'\boldsymbol{\eta}'}^{-1} \boldsymbol{H}'[k] \boldsymbol{R}_{\boldsymbol{x}[k]\boldsymbol{x}[k]} \boldsymbol{H}'^H[k]\right)\right). \end{aligned} \tag{3.32}$$

For the non-quantized case the optimal result is the water filling solution. Due to the modeling of the quantization, the effective noise covariance matrix $\boldsymbol{R}_{\boldsymbol{\eta}'\boldsymbol{\eta}'}$ and the effective channel $\boldsymbol{H}'[k]$ are dependent on the input covariance matrix $\boldsymbol{R}_{\boldsymbol{x}[k]\boldsymbol{x}[k]}$.

In a system without quantization, the covariance $\boldsymbol{R}_{\boldsymbol{x}[k]\boldsymbol{x}[k]}$ would be chosen according to the right singular vectors of $\boldsymbol{H}[k]$ to split the channel in orthogonal subchannels [32]. For this system with quantization we also adapt the transmit covariance matrix to the channel without quantization to ensure that the received power before each A/D conversion is approximately equal. Thus, we define $\boldsymbol{R}_{\boldsymbol{x}[k]\boldsymbol{x}[k]}$ to be of the form:

$$\boldsymbol{R}_{\boldsymbol{x}[k]\boldsymbol{x}[k]} = \boldsymbol{V}[k]\boldsymbol{S}[k]\boldsymbol{V}^H[k], \tag{3.33}$$

where $\boldsymbol{V}[k]$ are the eigenvectors of $\boldsymbol{H}^H[k]\boldsymbol{H}[k]$. The diagonal matrix $\boldsymbol{S}[k]$ represents the power allocation to the subchannels. The optimal allocation in a system without quantization follows the water filling solution. Since for a system with quantization at the receiver $\boldsymbol{R}_{\boldsymbol{\eta}'\boldsymbol{\eta}'}$ and $\boldsymbol{H}'[k]$ actually depend on $\boldsymbol{R}_{\boldsymbol{x}[k]\boldsymbol{x}[k]}$, it is difficult to separate the channel into orthogonal subchannels. To make the evaluation tractable, we use the suboptimal precoding vector $\boldsymbol{V}[k]$. For the matrix $\boldsymbol{S}[k]$, we test all different possibilities of allocating equal to power to 1 to S_{max} spatial streams. The number S_{max} is the maximum possible number of spatial streams and is equal to rank$(\boldsymbol{H}[k])$. If we would allow all frequencies to separately allocate the number of streams, we again have a combinatorial problem. Therefore, we check the overall achievable rate for allocating j spatial streams and in the end select one that has the largest achievable rate.

From the calculation of the quantization noise covariance matrix in Section 2.1, we see that the effective noise covariance matrix $\boldsymbol{R}_{\boldsymbol{\eta}'\boldsymbol{\eta}'}$ including the quantization noise is not diagonal. In a system, where the noise covariance matrix is known and independent of the transmit covariance one would simply multiply the receive vector with $\boldsymbol{R}_{\boldsymbol{\eta}'\boldsymbol{\eta}'}^{-\frac{1}{2}}$. This does generate a new system with a different channel $\boldsymbol{R}_{\boldsymbol{\eta}'\boldsymbol{\eta}'}^{-\frac{1}{2}}\boldsymbol{H}'[k]$ and spatial white noise. Afterwards, the water filling solution is applied to the new channel [32]. In general, the achievable rate increases compared to a system with white noise. In a more abstract way, the reason for the improvement is that channels with lower noise power can be used. Dependent on the rank of the channel relative to the number of the receive antennas, the orthogonal subchannels with highest noise power might not be used. In a system, where the channel and the noise depend on the covariance matrix of the transmit signal, it is very difficult to generate precoding and reception matrices that split the channel into orthogonal subchannels. Therefore, with our system, considering the correlation of the quantization noise leads to a decrease in achievable rate. Note that the reason why the channel and noise depend on the covariance matrix of the transmit signal is the AQNM model of the quantization as derived in Section 2.1.

For both calculation of the achievable rate in (3.31) as well as the calculation of the receive signal covariance matrix $\boldsymbol{R}_{\boldsymbol{y}_C\boldsymbol{y}_C}$ in (3.29), it is necessary to add all contribtuons of each frequency bin from k_1 to k_2. We choose the number of frequency bins to make the channel $\boldsymbol{H}[k]$ at each frequency bin sufficiently flat. This leads to a good approximation of the achievable rate. This reduces (3.31) to:

$$\begin{aligned} R \geq B \sum_{k=k_1}^{k_2} \max_{\boldsymbol{R}_{\boldsymbol{x}[k]\boldsymbol{x}[k]}} \log_2 \left(\det\left(\boldsymbol{I} + \boldsymbol{R}_{\boldsymbol{\eta}'\boldsymbol{\eta}'}^{-1}\boldsymbol{H}'[k]\boldsymbol{R}_{\boldsymbol{x}[k]\boldsymbol{x}[k]}\boldsymbol{H}'^H[k]\right)\right) \\ \text{s.t. } \mathbb{E}[||\boldsymbol{x}[k]||_2^2] \leq P_{Tx} \; \forall k \in [k_1, k_2]. \end{aligned} \tag{3.34}$$

Algorithm 3.2 Approximation of the achievable rate of a quantized system with noise covariance matrix $\boldsymbol{R}_{\boldsymbol{\eta\eta}}$, multipath channel $\boldsymbol{H}[l]$ and sum power constraint P_{Tx} and quantization function $Q_b(\cdot)$ with resolution of b bits in the frequency band from k_1 to k_2.

Require: $\boldsymbol{R}_{\boldsymbol{\eta\eta}}$, $\boldsymbol{H}[l]$, P_{Tx}, k_1, k_2 and $Q_b(\cdot)$

$\boldsymbol{H}[k] \leftarrow \mathcal{F}(\boldsymbol{H}[l])$

$S_{max} \leftarrow \min_k \left(\text{rank}\left(\boldsymbol{H}[k]\right)\right)$

for $j \leftarrow 1$ to S_{max} **do**

  $\boldsymbol{S} \leftarrow \boldsymbol{0}$

  $[\boldsymbol{S}]_{i,i} \leftarrow \frac{P_{Tx}}{j}\ \forall i = \{1,\ \cdots,\ j\}$

  $\boldsymbol{R}_{\boldsymbol{x}[k]\boldsymbol{x}[k]} \leftarrow \boldsymbol{V}[k]\boldsymbol{S}\boldsymbol{V}^H[k]\ \forall k \in [k_1, k_2]$ ▷ $\boldsymbol{V}[k]$ is eigenvector of $\boldsymbol{H}^H[k]\boldsymbol{H}[k]$

  $\boldsymbol{R}_{\boldsymbol{yy}} \leftarrow \sum_{k=k_1}^{k_2} \boldsymbol{H}[k]\boldsymbol{R}_{\boldsymbol{x}[k]\boldsymbol{x}[k]}\boldsymbol{H}^H[k] + \boldsymbol{R}_{\boldsymbol{\eta}(t)\boldsymbol{\eta}(t)}$

  $\boldsymbol{R}_{\boldsymbol{rr}} \leftarrow T(\boldsymbol{R}_{\boldsymbol{yy}}, Q_b(\cdot))$

  $[\boldsymbol{T}, \boldsymbol{R}_{\boldsymbol{\eta}_q\boldsymbol{\eta}_q}] \leftarrow \text{TF}\left(\boldsymbol{R}_{\boldsymbol{yy}}, Q_1(\cdot), \cdots, Q_M(\cdots)\right)$

  $\boldsymbol{R}_{\boldsymbol{\eta'\eta'}} \leftarrow \boldsymbol{R}_{\boldsymbol{\eta}_q\boldsymbol{\eta}_q} + \boldsymbol{T}\boldsymbol{R}_{\boldsymbol{\eta\eta}}\boldsymbol{T}^H$

  $\boldsymbol{H'}[l] \leftarrow \boldsymbol{T}\boldsymbol{H}[l]\ \forall l \in \{0,\ \cdots,\ L-1\}$

  $\boldsymbol{H'}[k] \leftarrow \mathcal{F}(\boldsymbol{H'}[l])$

  $\boldsymbol{A}[k] \leftarrow \boldsymbol{I} + \boldsymbol{R}^{-1}_{\boldsymbol{\eta'\eta'}}\boldsymbol{H'}[k]\boldsymbol{R}_{\boldsymbol{x}[k]\boldsymbol{x}[k]}\boldsymbol{H'}^H[k] \forall k \in [k_1, k_2]$

  $R(j) = \sum_{k=k_1}^{k_2} \log_2\left(\det\left(\boldsymbol{A}[k]\right)\right)$

end for

$R_{max} \leftarrow \max_j R(j)$

return BR_{max}

The receive signal correlation matrix can then be calculated as:

$$\boldsymbol{R}_{\boldsymbol{y}_C\boldsymbol{y}_C} = \boldsymbol{W}_R^H \left(\sum_{k=k_1}^{k_2} \boldsymbol{H}[k]\boldsymbol{R}_{\boldsymbol{x}[k]\boldsymbol{x}[k]}\boldsymbol{H}^H[k] + \boldsymbol{R}_{\boldsymbol{\eta\eta}} \right) \boldsymbol{W}_R. \tag{3.35}$$

The channel $\boldsymbol{H}[k]$ or the effective channel $\boldsymbol{H'}[k]$ at the frequency bins k can be calculated from the channel taps $\boldsymbol{H}[l]$ via the DFT $\mathcal{F}(\cdot)$:

$$\boldsymbol{H}[k] = \mathcal{F}(\boldsymbol{H}[l]), \boldsymbol{H'}[k] = \mathcal{F}(\boldsymbol{H'}[l]), \tag{3.36}$$

this follows the same per element DFT as described in (3.21).

We now have all the necessary mathematical tools to approximate the achievable rate of a multipath channel including quantization effects at the receiver. Algorithm 3.2 describes our approximation of the achievable rate for this type of systems.

This approximation is modeling a point to point closed loop spatial multiplexing system. There are many different simple modifications possible to change the modeled system. The following are a non-exhaustive list of examples:

- Systems without CSIT
- Systems with imperfect channel estimation

- Systems with multiple terminals communication with basestation
- Systems with constrained feedback
- Systems with multiple terminals and a basestation

Most of these systems can be modeled by changing the constraints on the transmit covariance matrix $\boldsymbol{R}_{\boldsymbol{x}[k]\boldsymbol{x}[k]}$ and the channel model.

3.1.4 Simulation Results Single User

In this section we evaluate the derived expression for different scenarios. We always include a rate evaluation without quantization. For the system without quantization we apply the water filling solution separate for each frequency bin. For all scenarios the results show the average achievable rate in bps/Hz averaged over 1000 channel realizations.

3.1.4.1 Comparison to Diagonal Approximation

This part of the evaluation compares the difference in performance when considering the non-diagonal elements in the calculation of $\boldsymbol{R}_{\boldsymbol{rr}}$ and therefore $\boldsymbol{R}_{\boldsymbol{\eta'\eta'}}$. The model considering only the diagonal elements of $\boldsymbol{R}_{\boldsymbol{rr}}$ was used in [66] and [57].

For the evaluation a channel of the first channel model is used. Here $K = 7$ separate paths are received at the same time. Different number of transmit and receive antennas are used. From Fig. 3.3 and 3.4 we see that the model considering the off-diagonal elements (ND) has a significant lower performance compared to the model only considering the diagonal elements (D). In fact, for the case of only one transmit antenna ($M_t = 1$) and 1-3 bit A/D conversion, the achievable rate is not maximized at the highest SNR possible, but rather at a finite SNR between 0 and 10 dB. As a reference also a system without limited resolution quantization is shown and labeled as NQ (Not Quantized).

As discussed in Subsection 3.1.3.2, if we compare the results in Fig. 3.3 (a) and (b) to the ones in Fig. 3.4. Considering the off-diagonal elements has only a large influence if the number of receive antennas is larger than the number of transmit antennas. This effect can be explained in the following ways: After spatial whitening, the power distribution of the effective noise is more non-uniform relative to the system that considers only the diagonal component. Since the actual channel and noise covariance matrix depends on the precoding matrix, it is not possible to decompose the channel into orthogonal subchannels with equal SNR. Thus, we cannot avoid using the channel with high noise variance and therefore the overall performance does degrade in the quantization noise limited, high SNR regime. This effect is only dominating the performance in the case of high SNR and very low resolution quantization. The peak in the achievable rate comes from the fact that at a certain SNR the noise provides dithering to randomize this structural performance degradation. At the minimum variance noise, where sufficient dithering is provided, is the peak in the performance. This effect is called statistic resonance and can be found in many non-linear systems [54].

Another important thing to mention is that in a system with multipath propagation and white noise, the covariance matrix $\boldsymbol{R}_{\boldsymbol{yy}}$ of the receive signal is approximated diagonal. This leads to a diagonal matrix $\boldsymbol{R}_{\boldsymbol{rr}}$ and therefore spatial white noise of the quantized system.

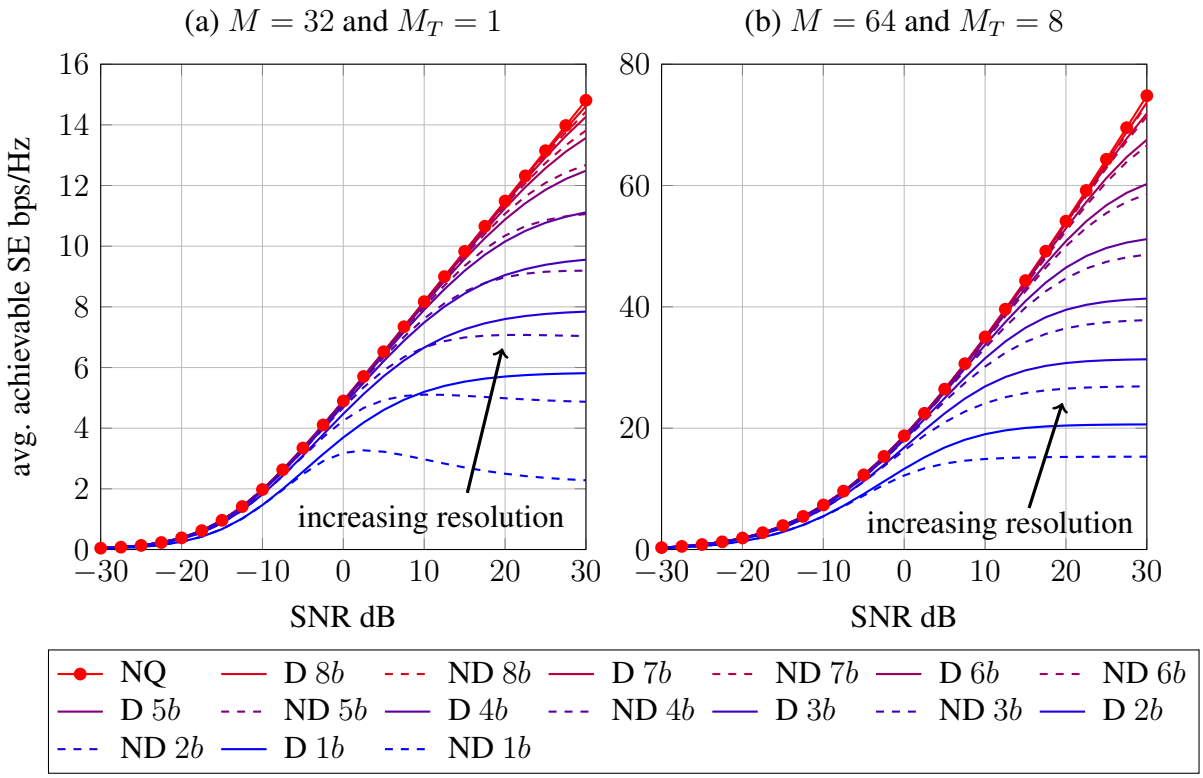

Fig. 3.3. Achievable rate comparison of digital beamforming with different resolution of the ADC for diagonal (D) and non-diagonal (ND) quantization error model and different transmit and receive antenna configuration in (a) and (b).

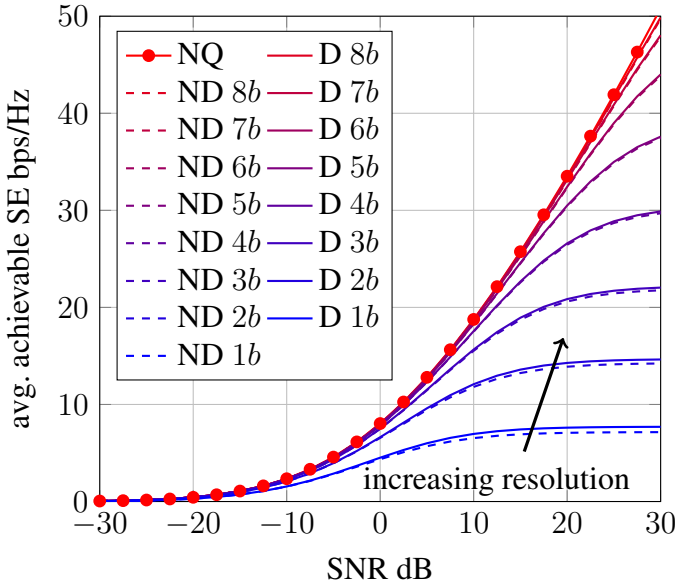

Fig. 3.4. $M = 8$ and $M_T = 8$ digital beamforming with different resolution of the ADC and different quantization noise models.

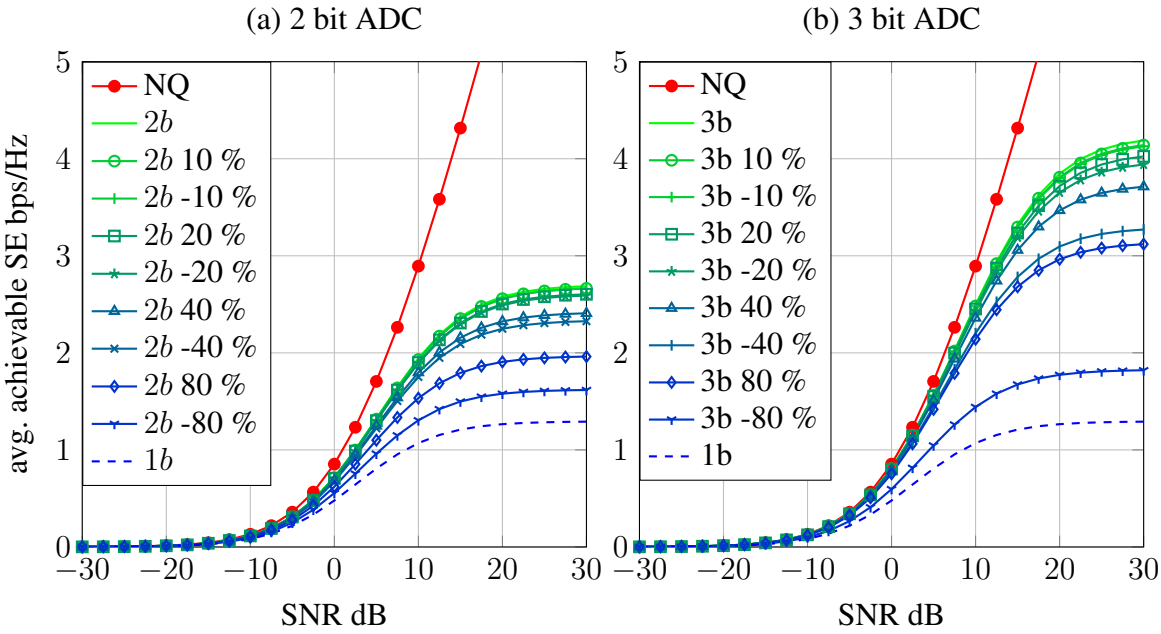

Fig. 3.5. SISO system with imperfect AGC and 2 bit (a) and 3 bit (b) ADC resolution.

3.1.4.2 Influence of AGC Imperfection

In this evaluation, we show the influence of AGC imperfections on the performance. To simplify the evaluation, we choose a Single Input Single Output (SISO) system with the simple multipath model described in the signal model (end of Subsection 3.1.2) with the parameters $L = 32$, $P = 16$ and $\beta = 0.35$. The graphs in Fig. 3.5 (a) and (b) show the average achievable rate with 2 bit resolution and different offset relative to the optimal power at the VGA output. The power after the VGA is defined as

$$\Omega = \Omega_{mq}(1 - \epsilon_{\text{AGC}}), \tag{3.37}$$

where Ω_{mq} and ϵ_{AGC} are the signal variance resulting in the minimal distortion and the AGC error.

The graphs show that an error in the range of -20 % to 20 % has only a minor impact on the performance. But as soon as the error is larger than 20 %, the performance decreases dramatically. Ultimately, the quantization converges to 1 bit quantization and therefore also our achievable rate converges to the one of 1 bit quantization. We can also observe the performance penalty for a larger negative or positive error is different.

Fig. 3.6 shows the average achievable rate at 30 dB SNR and the achievable Signal-to-Quantization-Noise-Ratio (SQNR) given the AGC offset and resolution. These two curves closely resemble each other. From the shape of the curve we can conclude that, if the error of the AGC can be larger than 20 % and our resolution is not too small, it is better to set the reference power for the AGC to a value that is above the actual optimal one. In this situation this would achieve a smaller average performance loss, since the error of estimating the power to larger than the actual one is higher than underestimating it.

3.1.4.3 Downlink Point to Point Scenario

In this subsection, a downlink like scenario is evaluated. A basestation with 64 antennas ($M_T = 64$) is transmitting to a mobile device with 8 antennas ($M = 8$). For the channel model the following parameters are used: $L = 32$, $P = 16$, $\beta = 0.35$. For the hybrid beamforming system

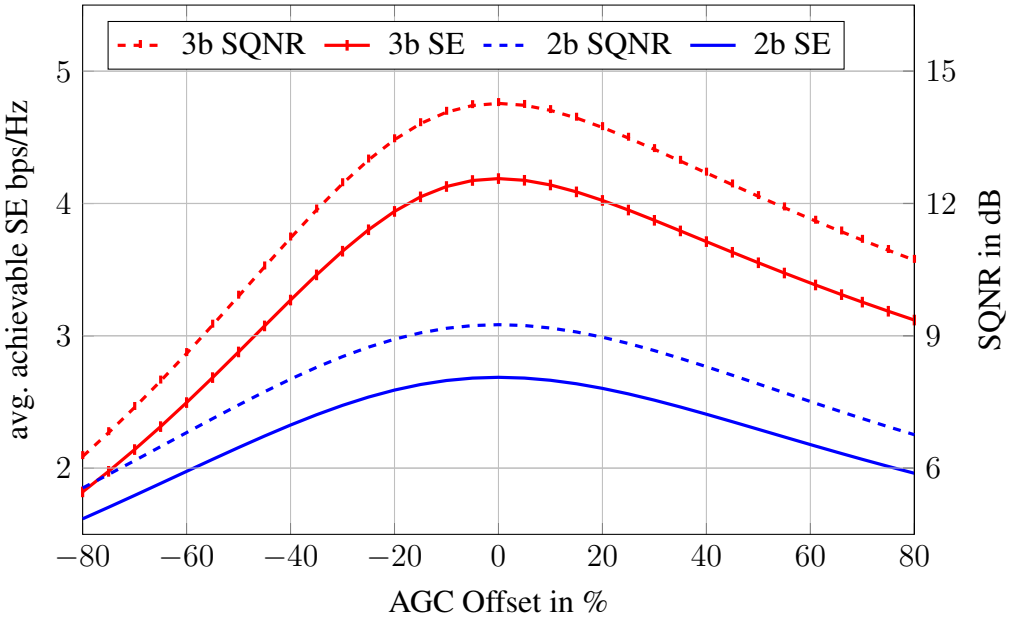

Fig. 3.6. Comparison between high SNR capacity and SQNR of the ADC.

$M_C \in \{2,\ 4,\ 8\}$ and therefore $M_{\text{ADC}} \in \{4,\ 2,\ 1\}$ is used.

Fig. 3.7 shows the average achievable rate for the case of $M_C = 4$ and ADC resolution $b \in \{1, \cdots, 8\}$. The rate curves of the systems including an ADC clearly converge to the ones assuming no quantization, for higher resolution in both cases of hybrid and digital beamforming. Especially, in the low SNR regime (below 0 dB), the performance of the digital beamforming systems with low resolution ADC (1-3 bit) are very close to the performance without quantization. These systems clearly outperform a hybrid beamforming system in this SNR regime. In this evaluation a 4 bit ADC is enough to outperform the hybrid system over the whole SNR range.

Since this system have a different power consumption, we also have to compare the results in terms of energy efficiency. Here we define the Energy Efficiency (EE) as the average achievable

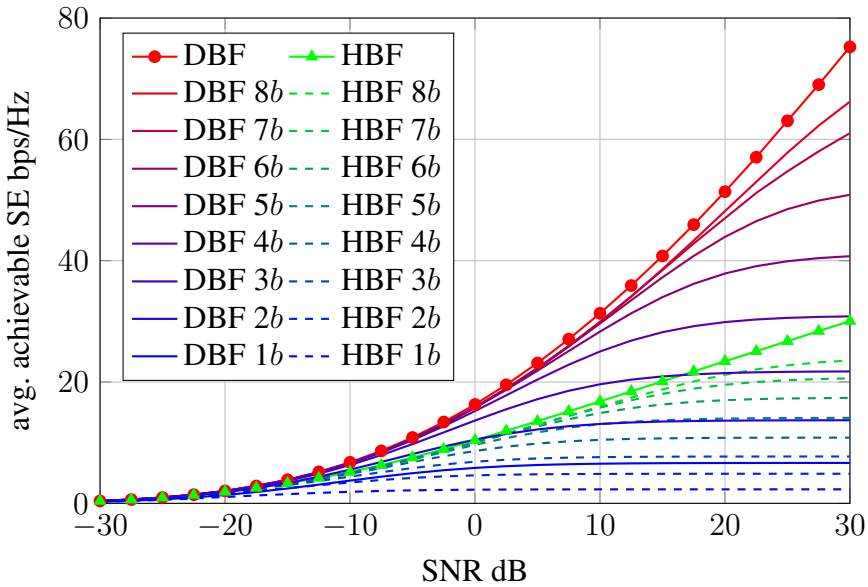

Fig. 3.7. $M = 8$ and $M_T = 64$ $M_C = 4$ different resolution of the ADC b.

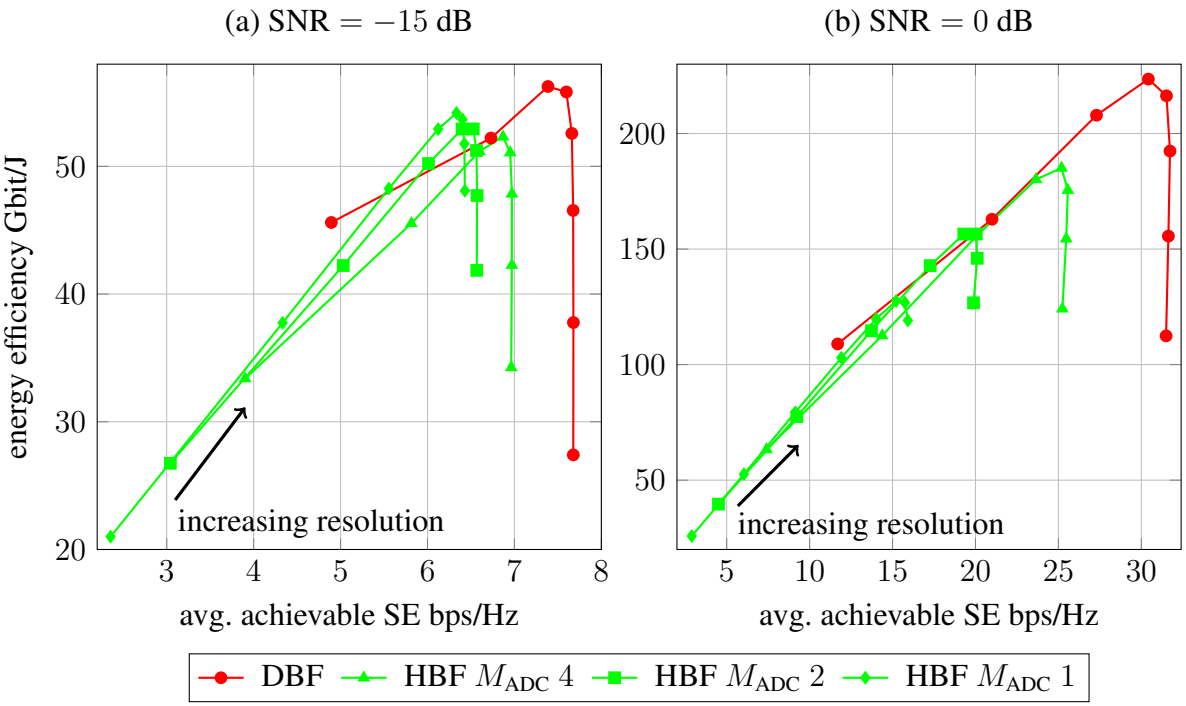

Fig. 3.8. $M = 8$ and $M_T = 64$ different resolution of the ADC b and SNR -15 dB (a) and 0 dB (b).

rate R divided by the power consumption of the RF front-end P_R:

$$\text{EE} = \frac{R}{P_R}. \tag{3.38}$$

Fig. 3.8 (a) and (b) show the energy efficiency over the achievable rate for $M_C \in \{2,\ 4,\ 8\}$ and the resolution of the ADC $b \in \{1,\ \cdots,\ 8\}$ with SNR $\in$ {-15 dB, 0 dB}. For both cases, the digital beamforming achieves a higher data rate and a higher energy efficiency. In the -15 dB SNR case, the difference in energy efficiency is not substantial but in the 0 dB SNR there is a large gap between hybrid and digital beamforming. In the lower SNR case, the energy efficiency peaks at 3 bit ADC resolution. The higher the SNR gets, the larger the ADC resolution that maximizes the energy efficiency. These results show that even when perfect hybrid beamforming without the beam-alignment overhead is considered a digital beamforming system is more energy efficient.

3.1.4.4 Uplink Point to Point Scenario

For the configuration of the system, the same parameters as in the Downlink (DL) like setup in the previous subsection is used. The only difference is that in this case the antenna configuration is $M = 64$ and $M_T = 8$.

Fig. 3.9 shows the achievable rate for this case. We observe that the penalty of hybrid beamforming is less severe than in the DL case. The reason is that in this case, the side of the system with less antennas (the mobile device) has no constraints on the front-end which is the exact opposite in the DL. This means that the number of spatial streams is in all cases just limited by the 8 possible streams of the mobile device. Therefore, the penalty of hybrid beamforming is less and the achievable rate of hybrid and digital beamforming rise with the same slope for the case without quantization.

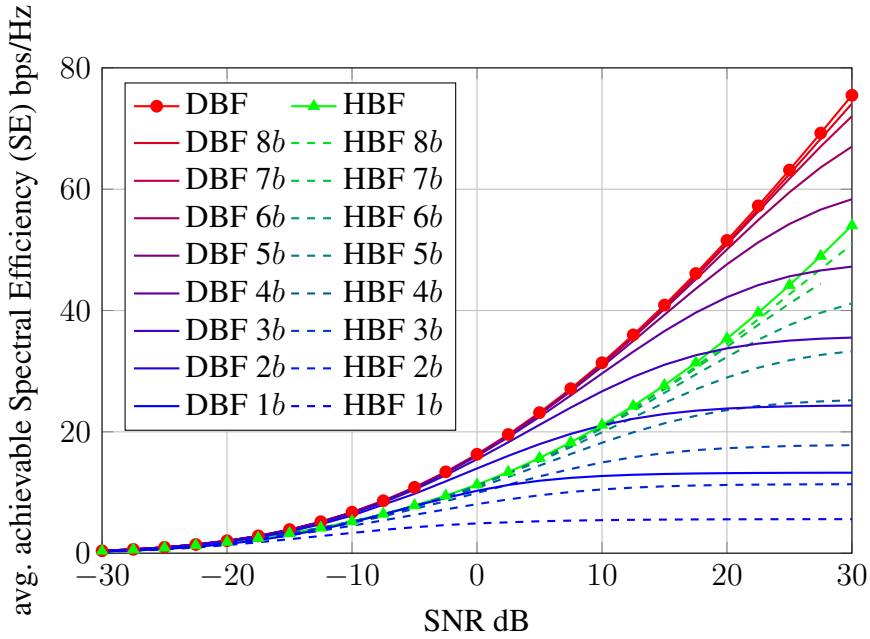

Fig. 3.9. $M = 64$ and $M_T = 8$ $M_C = 4$ different resolution of the ADC b.

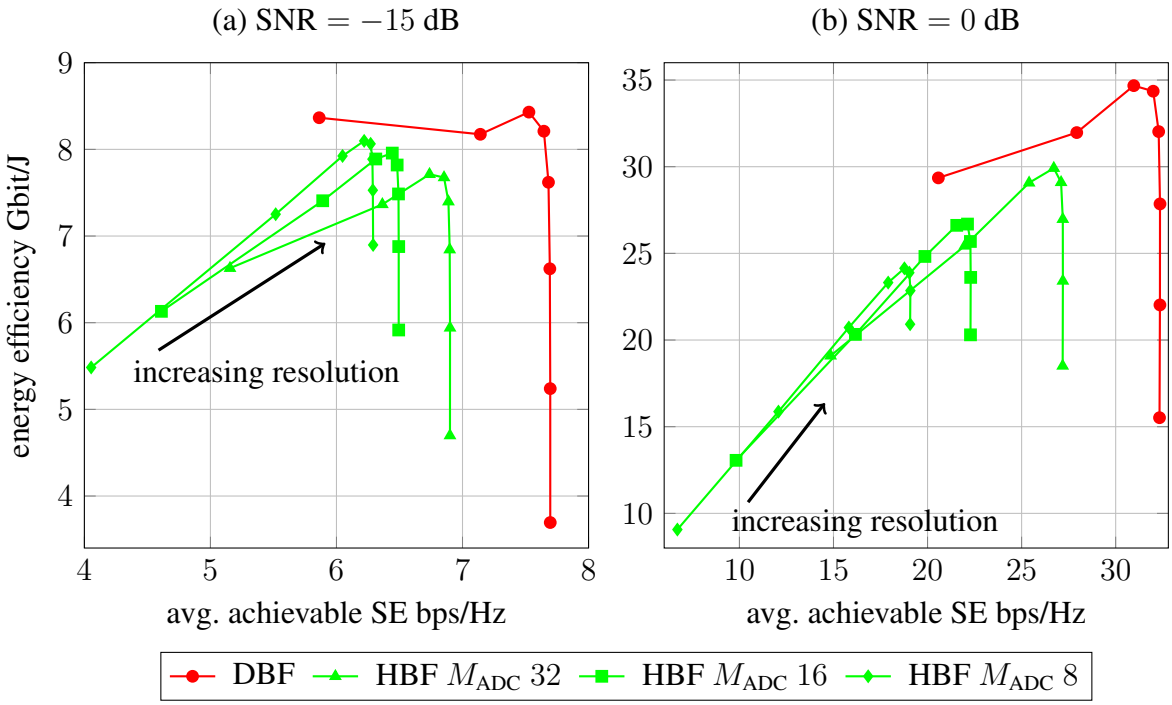

Fig. 3.10. $M = 64$ and $M_T = 8$ different resolution of the ADC b at different SNR -15 dB (a) and 0 dB (b).

In the low to medium SNR, the low resolution ADC digital beamforming systems perform better than the hybrid beamforming one. In the high SNR regime, there is no penalty on the number of possible data streams for the hybrid systems, therefore it performs better in this regime.

Fig. 3.10 (a) and (b) show the energy efficiency over the achievable rate for $M_C \in \{2,\ 4,\ 8\}$ and the resolution of the ADC $b \in \{1,\ \cdots,\ 8\}$ with SNR $\in$ {-15 dB, 0 dB}. As in the DL the digital beamforming system is more energy efficient as well as achieving a higher rate. Due to the small number of antennas, the energy efficiency stays almost constant for 1 to 3 bit ADC resolution. This can be explained with the fact that if the resolution is small the power consumption of the front-end is dominated by the other components, and the fact that we have a large degree of freedom with 64 Antennas and therefore the influence of the quantization noise at each of the antennas is not significant.

3.1.5 Conclusion Single User Performance Comparison

The evaluation in this section showed that low resolution ADC digital beamforming systems are more energy efficient and achieving a higher rate than hybrid beamforming systems for the given scenarios, especially in the low to medium SNR region. We also showed that if the imperfections of the AGC is in the range of -20 % to 20 %, there is no major influence on the performance. The evaluation of including the off-diagonal elements in the quantization error model showed that this could have a substantial impact on the performance with very low resolution ADCs.

3.2 Multiuser Scenarios

3.2.1 Introduction Multiuser Scenarios

The system model in Fig. 3.11 gives a general overview of all possible system configurations in our investigation. For $M_C = 1$ the block analog signal combination just connects the input to the output. For $M_C > 1$ this block contains an analog phase shifter for each signal followed by a power combiner.

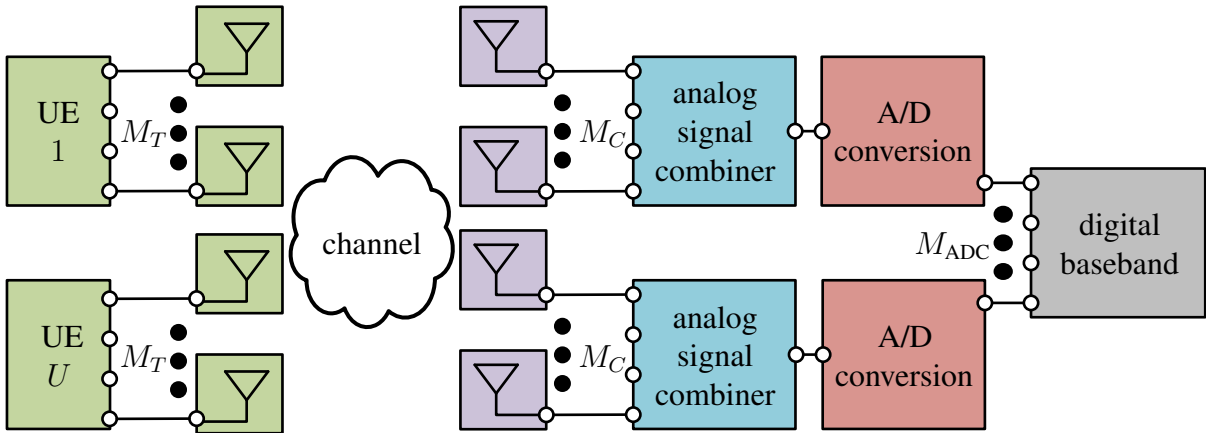

Fig. 3.11. System model with U users with M_T antennas. The receiver has M_C antennas at each of the M_{ADC} ADC chains. The number of receive antennas M is equal to $M_C \cdot M_{\text{ADC}}$.

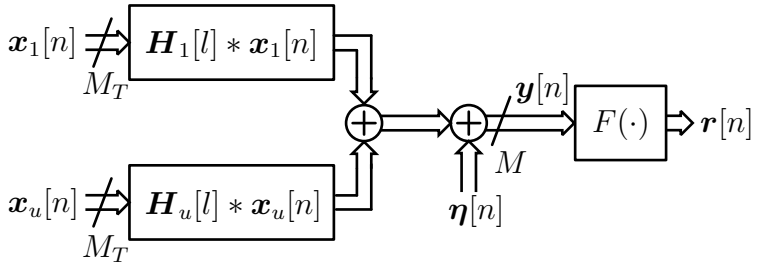

Fig. 3.12. Multiuser signal model.

3.2.2 Signal Model Multiuser Scenarios

The signal model described in this subsection is valid for all subsequent evaluations of data rate and energy efficiency.

The signal model in Fig. 3.12 describes the same system as the block diagram in Fig. 3.11. The symbols $\boldsymbol{x}_u[n]$, $\boldsymbol{\eta}_u[n]$, $\boldsymbol{H}_u[l]$, $\boldsymbol{\eta}_R[n]$, and $\boldsymbol{y}[n]$ represent the complex valued transmit signal of user u, the imperfections of the transmitter of user u, channel from user u to the access point, the noise at the receiver, and the receive signal of the system, respectively. We assume that there are U users with M_T antennas each and a basestation with M receive antennas. The receive signal sample $\boldsymbol{y}[n]$ at time instant n is defined as

$$\boldsymbol{y}[n] = \sum_{u=1}^{U} \sqrt{P_u} \sum_{l=0}^{L_u} \boldsymbol{H}_u[l](\boldsymbol{x}_u[n-l] + \boldsymbol{\eta}_u[n-l]) + \boldsymbol{\eta}_R[n], \tag{3.39}$$

where P_u is the transmit power of user u and L_u is the length of the channel in samples from user u to the access point. The transmitter impairments $\boldsymbol{\eta}_u[n]$ as well as the noise at the receiver $\boldsymbol{\eta}_R[n]$ are modeled as circularly symmetric complex Gaussian noise. The transmitter impairments have zero mean and covariance equal to $\sigma^2_{\text{EVM}}\boldsymbol{I}$ and the receiver noise also has zero mean and a spatially and temporally white covariance that is dependent on the SNR. The samples of $\boldsymbol{\eta}_u[n]$ from different users u and time instances n are independent. Including the transmit power P_u, this is the classical Error Vector Magnitude (EVM) definition only considering transmitter impairments [74]. In this work we only consider the transmitter EVM. For many theoretical evaluations it is not common to include impairments of the transmitter. However, as also shown in [24, 64] the inefficiency of the PAs for mmWave as well as other impairments ([64]), will limit the system performance at a much lower SNR compared to systems operating at lower frequencies. Thus, we think it is necessary to include these effects into our analysis. It is especially important to consider such effects since for a MU-MIMO UL system, the transmit noise components of each user add up at the basestation. Using a general additive model to account for hardware imperfections has been used in prior work and verified experimentally (see e.g., [17, 89, 67, 96]).

Thus, the noise here combines standard thermal noise and non-linear contributions from the whole transmitter hardware, including DAC, PA and the phase noise of the LO. Since all noise contributions in (3.39) are assumed to be Gaussian we can combine them to form a combined noise $\boldsymbol{\eta}'_R[n]$ equal to

$$\boldsymbol{\eta}'_R[n] = \sum_{u=1}^{U} \sqrt{P_u} \sum_{l=0}^{L_u} \boldsymbol{H}_u[l]\boldsymbol{\eta}_u[n-l] + \boldsymbol{\eta}_R[n]. \tag{3.40}$$

The receive signal is then reformulated to

$$\boldsymbol{y}[n] = \sum_{u=1}^{U} \sqrt{P_u} \sum_{l=0}^{L_u} \boldsymbol{H}_u[l]\boldsymbol{x}_u[n-l] + \boldsymbol{\eta}'_R[n]. \tag{3.41}$$

We restrict the system to have M_C antennas exclusively connected to one ADC chain (see Fig. 3.11). Therefore, the matrix $\boldsymbol{W}$ modeling the analog combining at the receiver has the form

$$\boldsymbol{W} = \begin{bmatrix} \boldsymbol{w}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{w}_2 & \ddots & \mathbf{0} \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \boldsymbol{w}_{M_{\text{ADC}}} \end{bmatrix} \in \mathbb{C}^{M \times M_{\text{ADC}}}, \tag{3.42}$$

where the vector $\boldsymbol{w}_i$ is the analog beamforming vector of the ith ADC chain. We also restrict our evaluation to the case where each ADC chain is connected to the same number of antennas M_C. The vectors $\boldsymbol{w}_i$ and $\mathbf{0}$ have dimension M_C. The receiver signal after the analog combining $\boldsymbol{y}_C[n]$ can be calculated as:

$$\boldsymbol{y}_C[n] = \boldsymbol{W}^H \boldsymbol{y}[n]. \tag{3.43}$$

For the case of digital beamforming the matrix $\boldsymbol{W}$ is simply replaced by an identity matrix with the same dimensions.

The receive signal $\boldsymbol{r}[n]$ after the ADC is related to the signal in the analog baseband $\boldsymbol{y}_C[n]$ in the following way

$$\boldsymbol{r}[n] = Q_b(\boldsymbol{y}_C[n]), \tag{3.44}$$

where $Q_b(\cdot)$ is the quantization operator with b bit resolution. The quantization operator $Q_b(\boldsymbol{a})$ treats the Inphase (I) and Quadrature (Q) component of each element of a vector $\boldsymbol{a}$ separately. For a real valued, scalar input a, the output of the operation is defined as:

$$r = Q_b(a) = q^j \; \forall \; a \; \in \left] q_l^{j-1} q_l^j \right]. \tag{3.45}$$

Here q^j is the representative of the jth quantization bin with the input interval $\left] q_l^{j-1} q_l^j \right]$. To cover a real valued input the left limit of the first interval q_l^0 and the right limit of the last interval $q_l^{N_b}$ are equal to $-\infty$ and ∞ respectively. The number of quantization bins N_b is equal to 2^b. For real world ADCs the difference between the representatives of quantization bins q_j and the size of the quantization bins are uniform. Thus, we limit our evaluation to this set of quantizers. For the evaluation we assume Gaussian input signals of the ADC. Consequently, we use the step-size to minimize the distortion for Gaussian signals shown in [51]. In order to ensure minimal distortion of each ADC, an AGC needs to adapt the gain of a VGA for the signal before each ADC separately. To simplify our model, we assume that the AGC always perfectly adapts to the input power. In a practical system a AGC takes some time to adapt to a different signal power. This adaptation time is dependent the actual AGC architecture as we well as system parameters like the bandwidth.

For the case of DBF, we study cases where the ADCs have either uniform resolution or a mixture of different resolutions. In our evaluation, we will restrict our attention to the case of mixed ADC resolution to the following type of scenarios: M_h ADCs with a higher resolution b_h and M_l ADCs with a lower resolution b_l. The channel model assumes the same average receive power at each antenna for each user. This means that the high resolution ADCs can be allocated

to any M_h antennas, and the remaining antennas to the ADCs with lower resolution. In practical scenarios it would be very difficult to adaptively allocate different ADCs to different ADC chains, since it takes a non-negligible amount of time to perform the switching. Furthermore, we do not expect the received power to be on average different at distinct antennas, so allocating the M_h high resolution ADCs to an arbitrary subset of the antennas is a reasonable approach. We will see that the use of mixed resolution ADCs allows one to achieve an arbitrary trade-off of spectral and energy efficiency.

3.2.3 Channel Models Multiuser Scenarios

The measurements in [56] show that for channels at 60 GHz, an exponential PDP sufficiently approximates a real world scenario. The channel associated with the l-th tap of the impulse response is assumed to be expressed as

$$\boldsymbol{H}[l] = \frac{1}{\sqrt{M_T}} \alpha[l] \boldsymbol{a}_r(\psi_r[l]) \boldsymbol{a}_t^T(\psi_t[l]), \tag{3.46}$$

where $\boldsymbol{a}_r(\psi_r[l])$ and $\boldsymbol{a}_t^T(\psi_t[l])$ represent the array response of the receive and transmit arrays, respectively, for arrival angle $\psi_r[l]$ and departure angle $\psi_t[l]$.

For the numeric evaluation, we assume that the antennas of the transmitters and the receiver form an ULA. If a narrowband, planar wavefront is impinging on the ULA and the spacing of adjacent antennas is $d = \pi/\mathrm{k}$, the receive signal at adjacent antennas is phase shifted by $\psi = d\,\mathrm{k}\sin(\phi) = \pi\sin(\phi)$. In this case we use the angular wave number k defined as $2\pi/\lambda$. The angle ϕ is the angle of a planar wavefront relative to the antennas of the ULA. The phase shift between the signal at adjacent antenna elements at the receiver and transmitter $\psi_r[l]$ and $\psi_t[l]$ of path l depend on the angle of arrival $\phi_r[l]$ and departure $\phi_t[l]$ as follows:

$$\boldsymbol{a}_r(\psi_r[l]) = \begin{bmatrix} 1 & e^{j\psi_r[l]} & e^{j2\psi_r[l]} & \dots & e^{j(M-1)\psi_r[l]} \end{bmatrix}^T. \tag{3.47}$$

Here we assume, that at delay l only one ray arrives at the receiver. The complex gain of the ray $\alpha[l]$ is assumed to be circularly symmetric Gaussian distributed with zero mean and a variance defined according to

$$v_l = \mathbb{E}\left[|\alpha[l]|^2\right] = e^{-\beta l}. \tag{3.48}$$

The parameter β defines how fast the power decays in relation to the delay. The other parameters of the model are the maximum channel length in samples L and the number of present channel taps P. This means for any channel realization, only P of the L elements in the vector of variances $\boldsymbol{v}$ are non-zero. We will normalize the variance vector as follows:

$$\boldsymbol{v}_n = \frac{\boldsymbol{v}}{||\boldsymbol{v}||^2}. \tag{3.49}$$

The SNR γ_u per user u is defined as:

$$\gamma_u = \frac{1}{M} \sum_{m=1}^{M} \frac{P_u \sum_{l=0}^{L_u-1} \mathbb{E}\left[\left|\left|\boldsymbol{e}_m^T \boldsymbol{H}_u[l] \boldsymbol{x}_u[n-l]\right|\right|_2^2\right]}{\mathbb{E}[||\boldsymbol{e}_m^T \boldsymbol{\eta}_R[n]||_2^2]}, \tag{3.50}$$

where $\boldsymbol{e}_m$ is a vector of size $\mathbb{C}^{M\times 1}$ with all zeros except for position m which is equal to 1. This formula describes the average SNR at each antenna for user u. It is important to note that the expectation takes the realization of the channel and realizations of $\boldsymbol{x}_u[n]$ into account.

For other parts of the evaluation we use the channel model described in [84]. This model is based on measurements, therefore, it can better reflect real environment. The SNR γ_u per user u is defined in the same way as in the exponential PDP case.

3.2.4 Rate Regions for Hybrid Beamforming and Digital Beamforming with Low Resolution ADCs

This work compares the uplink rate region for systems with hybrid beamforming and digital beamforming with low resolution ADCs, assuming a basestation with a large number of antennas. The evaluation nicely illustrates the strength and weaknesses of the different systems. Most of the results presented for two users can be generalized to a setup with more users. As we will see in subsequent subsection the insights gained from this evaluation can be applied also to system with more than two users. In this subsection we show how to calculate the rate region for the hybrid beamforming and digital beamforming system, using the digital signal processing tools developed in Chapter 2. This subsection is based on the article [79].

3.2.4.1 Calculation of the Rate Regions for Digital Beamforming with Low Resolution ADCs

Given the maximum transmit powers P_1 and P_2, assuming no cooperation among the users, the rate region of the Single Input Multiple Output (SIMO) Gaussian Multiple Access Channel (MAC) is given by [32]

$$\begin{aligned} 0 \leq R_1 &\leq \log_2\left(1+\frac{P_1||\boldsymbol{h}_1||_2^2}{\sigma_\eta^2}\right) \\ 0 \leq R_2 &\leq \log_2\left(1+\frac{P_2||\boldsymbol{h}_2||_2^2}{\sigma_\eta^2}\right) \\ R_1+R_2 &\leq \log_2\left|\boldsymbol{I}+\begin{bmatrix}\boldsymbol{h}_1 & \boldsymbol{h}_2\end{bmatrix}\begin{bmatrix}P_1 & 0\\ 0 & P_2\end{bmatrix}\begin{bmatrix}\boldsymbol{h}_1^H\\ \boldsymbol{h}_2^H\end{bmatrix}\right|, \end{aligned} \tag{3.51}$$

for the case that the transmit covariance matrix is diagonal and not changing. This rate and all following ones are normalized by the available bandwidth. These rate regions are pentagon shaped and assume perfect Channel State Information (CSI) at the receiver (CSIR). The edges of the pentagon at the R_1 and R_2 axis are defined by the single user rates in (3.51) and the points achieving the maximum sum rate by using Successive Interference Cancellation (SIC) in (3.52). By decoding the signal of user one first and user two second, we get the following rates

$$\begin{aligned} R_1^{(1,2)} &= \log_2\left(1+P_1\boldsymbol{h}_1^H\left(\boldsymbol{R}_{\boldsymbol{\eta}\boldsymbol{\eta}}+P_2\boldsymbol{h}_2\boldsymbol{h}_2^H\right)^{-1}\boldsymbol{h}_1\right) \\ R_2^{(1,2)} &= \log_2\left(1+P_2\boldsymbol{h}_2^H\boldsymbol{R}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1}\boldsymbol{h}_2\right) \\ R_1^{(2,1)} &= \log_2\left(1+P_1\boldsymbol{h}_1^H\boldsymbol{R}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1}\boldsymbol{h}_1\right). \end{aligned} \tag{3.52}$$

The superscript in $R_i^{(1,2)}$ show the decoding order. Reversing the order of the decoding in 3.52 we get the last edge of the pentagon.

By assuming the quantization noise is Gaussian distributed the SIMO Gaussian MAC rate region represent a lower bound of the actual achievable rate. Since we assume all signals are Gaussian, the rate expressions are similar to the ones of the case without quantization. Given the transmit powers P_1 and P_2, the quantization function $Q_b(\cdot)$ and perfect CSIR, we can again

construct the two points corresponding to the two possible SIC decoding orders. The calculation formulas are the same as in (3.52). But $\boldsymbol{h}_i$ and $\boldsymbol{R}_{\eta\eta}$ are the ones calculated with help of the model for the quantization developed in Section 2.1.

It is very important to highlight the difference between the case with and without quantization. For the case with quantization we have an effective noise $\boldsymbol{\eta}'$ with a non-diagonal covariance matrix $\boldsymbol{R}_{\eta'\eta'}$

$$\boldsymbol{R}_{\eta'\eta'} = \boldsymbol{R}_{\eta_q\eta_q} + \boldsymbol{T}\boldsymbol{R}_{\eta\eta}\boldsymbol{T}^H \tag{3.53}$$

The transformation matrix $\boldsymbol{T}$ as well as the noise covariance matrix of the quantization noise $\boldsymbol{R}_{\eta_q\eta_q}$ are dependent on the input covariance matrix of the ADCs $\boldsymbol{R}_{\boldsymbol{yy}}$ (Section 2.1). Since $\boldsymbol{R}_{\boldsymbol{yy}}$ is not diagonal also the covariance matrix of the quantization noise $\boldsymbol{R}_{\eta_q\eta_q}$ is not diagonal. In addition, the variance of the quantization noise is always relative to the input variance of the ADC. Therefore, we can conclude that the maximum sum rate is not necessarily achieved by transmitting with the maximum power for both users. Especially for very low resolution, it could also be the case that the interference could provoke the effect of dithering, which would then increase the rate. Thus, the rate with an interfering user could then be even higher than the one without a second user.

Combining these results with time sharing lead to the approximation of the rate region in the case of quantized observation as

$$R_{\text{QMAC}} = \operatorname{co}\left(\bigcup_{\boldsymbol{P}\in\mathbb{P}}\left\{(R_1,R_2):\ \begin{array}{l} 0 \le R_1 \le R_{q1}^{(2,1)} \\ 0 \le R_2 \le R_{q2}^{(1,2)} \\ R_1 + R_2 \le R_{q1}^{(1,2)} + R_{q2}^{(1,2)} \end{array}\right\}\right). \tag{3.54}$$

where $\operatorname{co}(\cdot)$ is the convex hull of all rate regions. The rates $R_{q1}^{(2,1)}$, $R_{q2}^{(1,2)}$ and $R_{q1}^{(1,2)}$ are calculated in the same way as in (3.52) by replacing $\boldsymbol{h}$ with $\boldsymbol{h}' = \boldsymbol{T}\boldsymbol{h}$ and $\boldsymbol{R}_{\eta\eta}$ with $\boldsymbol{R}_{\eta'\eta'}$:

$$\begin{aligned} R_{q1}^{(1,2)} &= \log_2\left(1 + P_1(\boldsymbol{h}')_1^H\left(\boldsymbol{R}_{\boldsymbol{\eta}'\boldsymbol{\eta}'} + P_2\boldsymbol{h}_2'(\boldsymbol{h}')_2^H\right)^{-1}\boldsymbol{h}_1'\right) \\ R_{q2}^{(1,2)} &= \log_2\left(1 + P_2(\boldsymbol{h}')_2^H\boldsymbol{R}_{\boldsymbol{\eta}'\boldsymbol{\eta}'}^{-1}\boldsymbol{h}_2'\right) \\ R_{q1}^{(2,1)} &= \log_2\left(1 + P_1(\boldsymbol{h}')_1^H\boldsymbol{R}_{\boldsymbol{\eta}'\boldsymbol{\eta}'}^{-1}\boldsymbol{h}_1'\right). \end{aligned} \tag{3.55}$$

The set $\mathbb{P}$ contains all possible transmit power combination $\boldsymbol{P} = [P_1\ P_2]$ of user one and user two up to a defined maximum. The rate region for all possible pairs in the set $\boldsymbol{P}$ need to be calculated and then convex hull of all these rate regions give the overall rate region for the case of quantization. As also described in [32] the convex hull is necessary as it is always possible to operate the system in different points for a fraction of the overall time. This leads to all points in a straight line between operations points being achievable by the system.

3.2.4.2 Calculation of the Rate Regions for Hybrid Beamforming

As stated in the signal model, we assume that the antennas belonging to one ADC chain form an ULA. Due the exponential growth even with a medium amount of possible beams and ADC chains, the number of possible beam configuration is very large. In the numerical examples, we use a system with 8 ADC chains and 8 antennas per ADC chain ($M_C = 8$). Using this system configuration and the approximation that $4M_C$ beams need to be check per ADC chain (Subsection 2.3.5), we need to check $32^8 \approx 1.1 \cdot 10^{12}$ possible configurations of the front-end. This is basically

Algorithm 3.3 Setup of the set $\mathbb{W}$ of all possible analog combiner matrices for hybrid beamforming MAC rate regions calculation.

Require: $\hat{\phi}_{(u,i)} \forall u, i$

1: $\mathbb{W} \leftarrow \varnothing$

2: **for** $n \leftarrow 0$ to M_{ADC} **do**

3: $$\mathbb{W} \leftarrow \mathbb{W} \cup \begin{bmatrix} \boldsymbol{w}(\hat{\phi}_{(\pi(1),1)}) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{w}(\hat{\phi}_{(\pi(2),2)}) & \ddots & \mathbf{0} \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \boldsymbol{w}(\hat{\phi}_{(\pi(M_{\text{ADC}}),M_{\text{ADC}})}) \end{bmatrix} \forall \pi \in \mathbb{H}_{M_{\text{ADC}},n}$$

4: **end for**

5: **return** $\mathbb{W}$

infeasible for any practical implementation and our theoretical one. Therefore, we introduce a simplification that reduced the search space.

The overall procedure of selecting the beams is described in the following paragraph in an abstract way. Afterwards, the mathematical details are presented. Therefore, we limit the search for the optimal beamforming configuration in the following way: First, we search the best beam for each user u and ADC chain i combination under the assumption that the other users are not present. Afterwards, all possible combinations of these selected beams are tried in terms of their achievable rate with the maximum transmit power of the users.

Finding the best beam out of a predefined set $\mathbb{B}_{(u,i)}$ for user u on ADC chain i can be formulated as

$$\hat{\phi}_{(u,i)} = \arg \max_{\phi_{(u,i)} \in \mathbb{B}_{(u,i)}} \left| \boldsymbol{w}(\phi_{(u,i)})^H \boldsymbol{h}_u^i \right|_2^2 \text{ with,} \tag{3.56}$$

with $\mathbb{B}_{(u,i)}$ being the set of all spatial directions $\phi_{(u,i)}$ that are scanned. The channel $\boldsymbol{h}_u^i$ contains the M_C elements of $\boldsymbol{h}_u$ corresponding to the antennas of the ith ADC chain. According to the discussion in Subsection 2.3.5 the set $\mathbb{B}_{(u,i)}$ contains $4M_C$ elements. By comparing the maximum receive power, we get the best configuration $\hat{\phi}_{(u,i)}$ for each user u and ADC chain i. The set of all possible $\boldsymbol{W}$ denoted as $\mathbb{W}$ is formed out of all possible combination of $\boldsymbol{w}(\hat{\phi}_{(u,i)})$ assigned to each ADC chain. This process is described in Algorithm 3.3. In this case the set $\mathbb{H}_{M_{\text{ADC}},n}$ contains all possibilities of selecting n ADC chains for user 1. Therefore, it has size $\binom{M_{\text{ADC}}}{n}$. The selection operator $\pi(i)$ is than assigning the different ADC chains to the different users.

Given the beamforming matrix $\boldsymbol{W}$, the system forms an effective channel $\boldsymbol{h}_u^{\text{eff}} = \boldsymbol{W}\boldsymbol{h}_u$ for each user u. This can then be treated as the previously described system without quantization. For each of the beam configuration, a pentagon shaped rate region is formed. To show the full set of possible rates, we form the convex hull of these rate regions for all tested beam configurations.

In the next step, the achievable rate region is calculated as

$$R_{\text{HBFMAC}} = \text{co}\left(\bigcup_{\boldsymbol{W} \in \mathbb{W}} \left\{ (R_1, R_2) : \begin{array}{c} 0 \le R_1 \le R_1^{(2,1)} \\ 0 \le R_2 \le R_2^{(1,2)} \\ R_1 + R_2 \le R_1^{(1,2)} + R_2^{(1,2)} \end{array} \right\} \right). \tag{3.57}$$

For the calculation of the rates (3.52) is used. For each $\boldsymbol{W}$, the pentagon shaped rate region is calculated by replacing $\boldsymbol{h}_u$ with $\boldsymbol{W}\boldsymbol{h}_u$ and $\boldsymbol{R}_{\eta\eta}$ by $\boldsymbol{W}\boldsymbol{R}_{\eta\eta}\boldsymbol{W}^H$.

3.2.4.3 Simulation Results Rate Region

For evaluating the rate region, a system with $M = 64$ receive antennas and 2 users was chosen. For the case of hybrid beamforming, the number of antennas per ADC chain M_C was 8. For evaluation of the comparison of the rate region, two channel types where used. Both represent extreme cases, and the channel of an actual system is likely going to be between those two. The first represents a rich scattering environment, therefore, each element of $\boldsymbol{h}_u$ has independent circular symmetric complex Gaussian entries. In the second case, the channel from each user to the basestation consists only of a single planar wavefront received by the basestation (single ray). The resolution b is varied from 1 to 8 bit for all plots. The graphs Fig. 3.13 show the rate region with different SNR γ_u. In the title of the plots only the SNR assuming the maximum possible transmit power for each user is given.

The results are shown in Fig. 3.13. The results for with/without quantization and hybrid beamforming are labeled as quant, unquant and hybrid BF. In general, the graphs show that the rate region of the quantized observations converges to the one without quantization by increasing b. For the low SNR cases in (a) and (b) even 3 bit resolution is enough to achieve approximately the rate of the system without quantization. For the higher SNR cases (c) and (d) 4 and 5 bit are necessary to achieve a similar performance.

Since the behavior of the quantized system is different to the system without quantization, also the rate points corresponding to transmitting with the highest possible power (quant max $\boldsymbol{P}$) and the rates corresponding to the maximal sum rate (max sum rate) are shown. It is clear that if the system is not limited by the quantization error and is also not asymmetric from the best achievable receive power by both users, it is always best to transmit with the highest possible power (a), (b) and (c). The exception is the case with 1 bit quantization. Due to effect of the correlated effective noise, even the SIMO achievable rate has a peak at about 0 dB SNR and carries over to the SIMO MAC results. From (d), we can deduce that if the users experience clearly different channel conditions and the resolution of the ADC is very low, the system experiences a significant degradation if users would only transmit with the highest possible power. For a practical system, this would call for a scheduling algorithm taking this behavior into account or an accurate power control.

If we now compare the low resolution results to the hybrid beamforming ones, we see that in the low SNR regime (a) and (b) even 1 bit quantization achieves a higher sum rate than hybrid beamforming. The higher the SNR gets (c) and (d), the better hybrid beamforming performs relative to low resolution digital beamforming. Basically, the performance is influenced by the different properties of the two systems. The low resolution ADC digital beamforming system has access to all degrees of freedom, but the quality of the signals is limited. That means in the low SNR regime, where the performance is limited by noise, very little degradation is visible. But dependent on the resolution, the performance saturates in the high SNR regime. For hybrid beamforming, the system does not have access to all degrees of freedom, which especially limits the performance for multiuser scenarios. Therefore, the performance in low SNR is limited by the number of degrees of freedom. In contrast, in the high SNR regime, the performance is less influenced by the limited degrees of freedom. Overall, if the number of users increases, a further degradation of hybrid beamforming relative to the low resolution digital beamforming system is expected, especially in the low SNR regime.

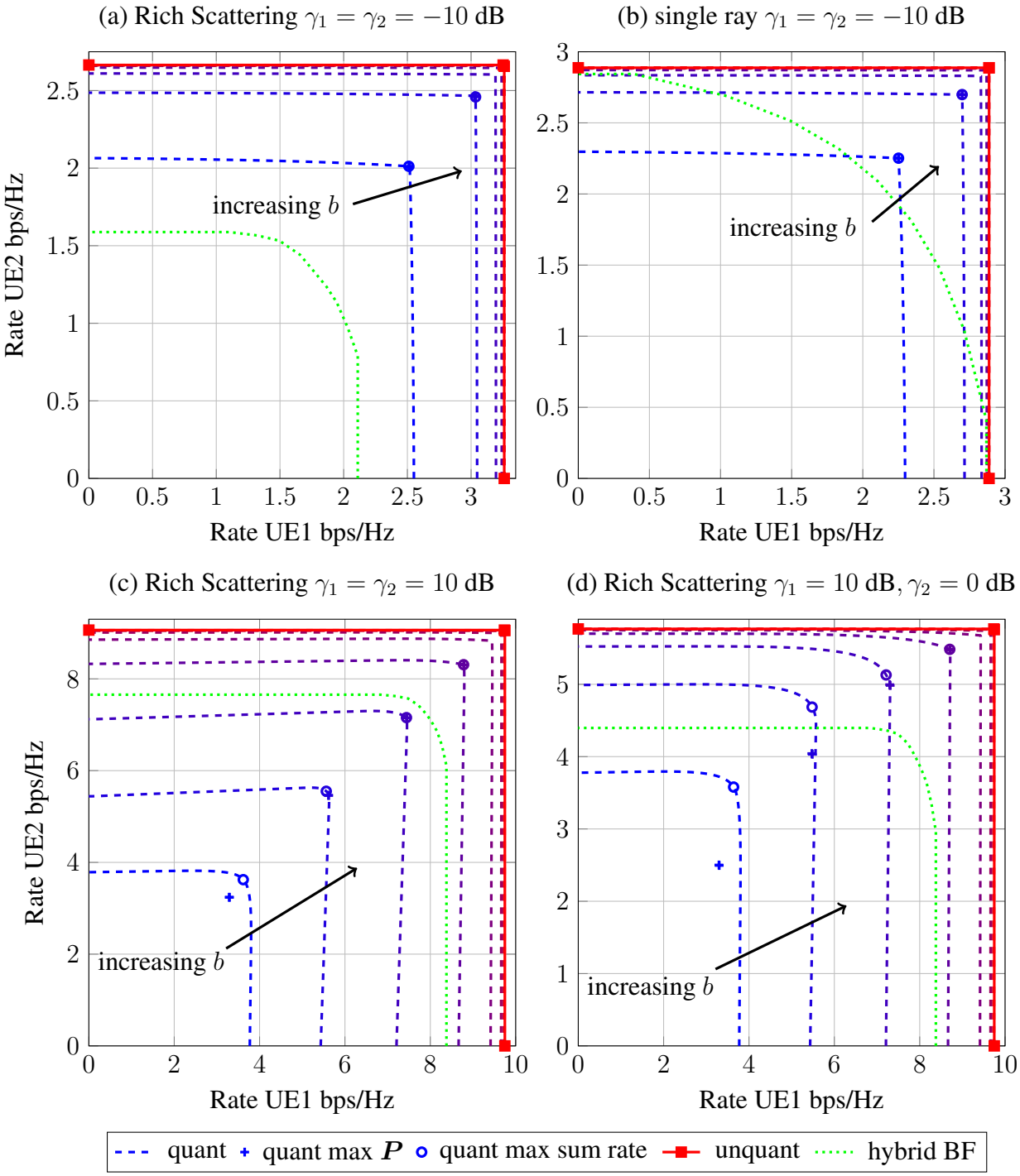

Fig. 3.13. Two user rate region for $M = 64$ hybrid beamfoming and low resolution ADC digital beamforming with different channel conditions. A linear uniform quantizer with resolution b from 1 to 8 bit for I and Q is used.

3.2.4.4 Conclusion Rate Region

In this part of the work, the achievable rate region for low resolution ADC digital beamforming and hybrid beamforming were compared in a massive MIMO scenario. The results show that in the low SNR regime, the low resolution ADC digital beamforming is clearly superior compared to hybrid beamforming. In the high SNR regime, the performance is similar if a higher resolution of about 4 bit is used for digital beamforming. The evaluation also showed that for very low resolution A/D conversion in the range of 1-2 bit and asymmetric maximum receive power of different users, it is essential to have an accurate power control to achieve the maximum sum rate. This clearly limits the applicability of this approach and motivates the use of 3-4 bit resolution to relax the requirements of the power control.

3.2.5 Energy and Spectral Efficiency Comparison for Multiple Users and Imperfect CSIR

The results in this subsection extend the single user analysis in Section 3.1 to a multiuser scenario. In addition, the effects of imperfect CSIR from Section 2.5 and transmitter imperfections from Subsection 3.2.4 are included in the model. The following paragraphs describe the additional modeling that is necessary to extend the results to a multiuser scenario. Part of the work in this section also appeared in [81].

3.2.5.1 Allocation of ADC chains for hybrid beamforming with multiple users

For the following calculations we assume that adjacent antennas are connected to one ADC chain. Finding the optimal configuration of the phase shifters at each antenna to support U users is a non-convex problem, which does not have a trivial solution. Thus, we introduce a number of simplifications that make the problem tractable. At the same time these simplifications are modeling the behavior of practical beamforming systems like WiGig (802.11ad) [37, 41].

The overall procedure for selecting the beams is described in the following paragraph in an abstract way. Afterwards, the mathematical details are presented in the description of the algorithm. We limit the search for the optimal beamforming configuration in the following way: First, we search for the best beam combination for each user u and ADC chain j combination under the assumption that the other users are not present and record the corresponding receive power. Afterwards, the ADC chains are allocated to the users in a resource-fair manner, starting from the ADC chain and user with the highest receive power.

As we showed in Subsection 2.3.5, if the receive antennas form an ULA at each subarray of M_C elements and limiting the beams to receive the signal from only one spatial direction, we achieve 10 % error while having a codebook size of $4M_C$. The first part of the algorithm is thus selecting the best beamforming vectors per User Equipment (UE). Since we assume that all subarrays have the same size M_C we initialize the set of all possible directions $\mathbb{B}$ with $4M_C$ values uniformly spaced from $-\pi$ to π:

$$\mathbb{B} = \{\phi_1,\ \phi_2,\ \cdots,\ \phi_{4M_C}\}, \phi_j = -\pi + \frac{j\pi}{2M_C}. \tag{3.58}$$

Afterwards, for each user u and each subarray i, all directions are tested, and the one leading to the largest receive power and the corresponding index are stored

$$\begin{aligned} p(j,u,i) &= \sum_{l=0}^{L-1} \left|\left| \boldsymbol{w}_j^H \boldsymbol{H}_u^i[l] \right|\right|_2^2, \\ [\boldsymbol{P}]_{u,i} &= \max_j p(j,u,i), \\ [\boldsymbol{J}]_{u,i} &= \arg\max_j p(j,u,i), \end{aligned} \tag{3.59}$$

with the vector w_j defined as:

$$\boldsymbol{w}_j = \begin{bmatrix} 1 & e^{\phi_j} & \cdots & e^{(M_C-1)\phi_j} \end{bmatrix}^H. \tag{3.60}$$

The matrices $\boldsymbol{P}$ and $\boldsymbol{J}$ contain the optimal power and the corresponding direction for all combinations of users u and subarrays j.

The next step is to select which subarray should take which configuration. We at first fill the set $\mathbb{U}$ and $\mathbb{I}$ with all users and subarrays

$$\mathbb{U} = \{1, \cdots, U\}, \;\; \mathbb{I} = \{1, \cdots, M_{\text{ADC}}\}. \tag{3.61}$$

Then we select the subarray user combination leading to the largest receive power and allocate the array steering vector of the selected subarray to this configuration. Since this subarray and user are now allocated, we remove them from the sets $\mathbb{U}$ and $\mathbb{I}$. If the set of remaining users is empty, we reset it to all possible users. This procedure is repeated until all subarrays are allocated. It ensures that the subarrays are distributed among the users under a resource fair constraint. In addition, the selection of those with higher power also ensures that the rate is optimized. It is important to mention that only selecting the ADC chains according to the ones providing the largest receive power, even if considered for all users, would lead to starvation of the users with the worst channels. Since this is not desirable, we adopted the above procedure. The entire process is summarized in Algorithm 3.4. After the analog combining the system is treated in the same way as the full digital system by using the effective channel after the combining.

3.2.5.2 Combined Rate Expression for the Multiuser scenario

After the model for the transmit impairments, the analog combining and the quantization error we have a set of equations that looks fairly similar to a standard MIMO system. We chose to model the channel estimation error as additional noise independent of the receive channel. This is different from the work in [100]. In this work the channel estimation error is also modeled as additional noise. But in addition, the useful signal power is divided between the estimated channel and the channel estimation noise. This has the effect that for cases leading to a large estimation error, the resulting received signal power as well as the rate go to zero. If we look at our simulation of the channel estimation error in Fig. 2.30 this would be the case for the very low SNR range from -30 to -10 dB. This contradicts the practical observation, that communication at SNRs as low as -10 dB for a SISO system is possible [73]. For a practical massive MIMO system this would mean that regardless of the number of antennas it is not possible to be used at low SNR. We therefore think that modeling the channel estimation error as noise is more suitable to reflect the behavior of a practical system. For the rest of this section we assume that each user u uses one transmit antenna.

Algorithm 3.4 Resource fair selection of beamforming vectors in a multiuser scenario.

Require: $\boldsymbol{H}[l]$, U, M_{ADC} and M_C
$\mathbb{B} \leftarrow \{\phi_1,\ \phi_2,\ \cdots,\ \phi_{4M_C}\}$
for $u \leftarrow 1$ to U **do**
  for $i \leftarrow 1$ to M_{ADC} **do**
    for $j \leftarrow 1$ to $4M_C$ **do**
      $\boldsymbol{w}_j \leftarrow \begin{bmatrix}1 & e^{\phi_j} & \cdots & e^{(M_C-1)\phi_j}\end{bmatrix}^H$
      $p(j) \leftarrow \sum_{l=0}^{L-1} \left|\left|\boldsymbol{w}_j^H \boldsymbol{H}_u^i[l]\right|\right|_2^2$
    end for
    $[\boldsymbol{P}]_{u,i} \leftarrow \max_j p(j)$
    $[\boldsymbol{J}]_{u,i} \leftarrow \arg\max_j p(j)$
  end for
end for
$\mathbb{U} \leftarrow \{1,\ \cdots,\ U\}$
$\mathbb{I} \leftarrow \{1,\ \cdots,\ M_{\text{ADC}}\}$
for $i \leftarrow 1$ to M_{ADC} **do**
  $\hat{u}, \hat{i} \leftarrow \arg\max_{u\in\mathbb{U}, i\in\mathbb{I}} [\boldsymbol{P}]_{u,i}$
  $\hat{j} \leftarrow [\boldsymbol{J}]_{\hat{u},\hat{i}}$
  $\boldsymbol{w}_R^{\hat{i}} \leftarrow \begin{bmatrix}1 & e^{\phi_{\hat{j}}} & \cdots & e^{(M_C-1)\phi_{\hat{j}}}\end{bmatrix}^H$
  $\mathbb{I} \leftarrow \mathbb{I} \setminus \hat{i}$
  $\mathbb{U} \leftarrow \mathbb{U} \setminus \hat{u}$
  if $\mathbb{U} = \varnothing$ **then**
    $\mathbb{U} \leftarrow \{1,\ \cdots,\ U\}$
  end if
end for
return $\boldsymbol{w}_R^i\ \forall i = \{1,\ \cdots,\ M_{\text{ADC}}\}$

The overall covariance matrix of the combined channel estimation error covariance matrix $\boldsymbol{R}_{\boldsymbol{w}[k]\boldsymbol{w}[k]}$ on the kth SC is defined as a sum of the per user channel estimation error covariance matrices $\boldsymbol{R}_{\boldsymbol{w}_u[k]\boldsymbol{w}_u}[k]$

$$\boldsymbol{R}_{\boldsymbol{w}[k]\boldsymbol{w}[k]} = \sum_{u=1}^{U} \boldsymbol{R}_{\boldsymbol{w}_u[k]\boldsymbol{w}_u[k]}, \tag{3.62}$$

where the variance of each element of $\boldsymbol{R}_{\boldsymbol{w}_u[k]\boldsymbol{w}_u[k]}$ depends on the channel estimation error σ_u^2 and the actual power of the channel at the corresponding frequency bin k on antenna m:

$$\left[\boldsymbol{R}_{\boldsymbol{w}_u[k]\boldsymbol{w}_u[k]}\right]_{m,m} = |[\boldsymbol{h}_u[k]]_m|^2 \sigma_u^2, \tag{3.63}$$

where $[\boldsymbol{h}_u[k]]_m$ is the channel of user u on the kth SC defined in the same way as for the single user case in (3.21) as:

$$[\boldsymbol{h}[k]]_m = \frac{1}{\sqrt{K}} \sum_{n=0}^{K-1} [\boldsymbol{h}'_u[n]]_m e^{-j\frac{2\pi}{K}kn}. \tag{3.64}$$

Note that $\boldsymbol{h}'_u[n]$ is the by $K - L$ zero matrices extended version of $\boldsymbol{h}_u[l]$ defined as:

$$\boldsymbol{h}'_u[n] = \begin{cases} \boldsymbol{h}_u[n], \text{if } n < L \\ \boldsymbol{0}, \text{otherwise} \end{cases} . \tag{3.65}$$

We model each matrix $\boldsymbol{R}_{\boldsymbol{w}_u[k]\boldsymbol{w}_u[k]}$ to be spatial white and thus a diagonal matrix. The values σ_u^2 are determined by calculating the average SNR per-antenna per user and then obtaining the corresponding MSE σ_u^2 from the simulation shown in Fig. 2.30.

We combine this calculation into the operator TE$(\cdot)$:

$$\boldsymbol{R}_{\boldsymbol{w}[k]\boldsymbol{w}[k]} = \text{TE}(\boldsymbol{H}[k], \boldsymbol{R}_{\boldsymbol{yy}}, \boldsymbol{R}_{\boldsymbol{\eta\eta}}), \tag{3.66}$$

where $\boldsymbol{H}[k]$ is the combined channel from all users to the receiver.

At this point we have all the necessary information to calculate the sum rate for the given scenario. We make a number of approximations that make the expression tractable:

- Assume $\boldsymbol{x}[k]$ is Gaussian
- Beamforming vectors $\boldsymbol{w}_R^i$ are selected from the derived finite set separately for each antenna group based on an SNR criterion
- Quantization noise is modeled as additive Gaussian noise with a non-white covariance matrix.
- No collaboration among the users

For constellations of the Quadrature Amplitude Modulation (QAM) family, there exists only a small shaping-gap compared with Gaussian symbols [29]. Since otherwise the overhead for beam training is very large for most theoretical work as well as practical implementations, the vectors are drawn from a predefined set [26, 37]. The assumption of Gaussian quantization noise is not satisfied for very low resolution (1-2) bit in the time domain. However, all rate calculations in this work are in the frequency domain. Due to the central limit theorem [28] the distribution of the quantization noise in the frequency domain converges to Gaussian. We have also verified this in our simulations. For most of the cases in a practical system, users cannot collaborate, because they have different data to transmit and do not know that the others are present. To simplify the evaluation, we limit the transmission from each user to one spatial data stream. Since in this work we concentrate on the effects at the receiver, we do not explicitly model the transmit beamforming at each user. Therefore, the users are modeled to have a single antenna.

The rate analysis is carried out for each frequency bin k separately:

$$\begin{aligned} R \geq B \sum_{k=k_1}^{k_2} \max_{\boldsymbol{R}_{\boldsymbol{x}[k]\boldsymbol{x}[k]}} I(\boldsymbol{x}[k], \boldsymbol{r}[k] | \boldsymbol{H}'[k]) \\ \text{s.t. } \mathbb{E}[||\boldsymbol{x}[k]||_2^2] \leq P_{Tx} \, \forall k \in [k_1, k_2], \end{aligned} \tag{3.67}$$

where $\boldsymbol{x}[k]$, $\boldsymbol{r}[k]$ and $\boldsymbol{H}'[k]$ represent the input/output signal of all users combined and the equivalent channel of frequency bin k, and $I(\cdot)$ is the mutual information. The frequencies k_1 and k_2 mark the borders of the band of interest in the equivalent baseband channel. If the entire band covered by the sampling rate is not available to the system, the parameters k_1 and k_2 must account for the guard band.

Since all signals are represented by Gaussian random variables, we get the following expression for the mutual information:

$$\begin{aligned} I(\boldsymbol{x}[k], \boldsymbol{r}[k] | \boldsymbol{H}'[k]) = \\ \log_2 \left(\det \left(\boldsymbol{I} + \boldsymbol{R}_{\boldsymbol{\eta}'\boldsymbol{\eta}'}^{-1} \boldsymbol{H}'[k] \boldsymbol{R}_{\boldsymbol{x}[k]\boldsymbol{x}[k]} \boldsymbol{H}'^H[k] \right) \right) . \end{aligned} \tag{3.68}$$

Algorithm 3.5 Combined multipath channel from each user $\boldsymbol{H}[l]$, combined transmit impairments EVM co-variance matrix $\boldsymbol{R}_{\boldsymbol{\eta_T\eta_T}}$, combined maximum transmit power constraint $P_{Tx}\boldsymbol{I}$, receiver noise covariance matrix $\boldsymbol{R}_{\boldsymbol{\eta\eta}}$, frequency band from k_1 to k_2, quantization function $Q_m(\cdot)$ separate for each receiver chain m and channel statistics and channel bandwidth B.

Require: $\boldsymbol{R}_{\boldsymbol{\eta_T\eta_T}}$, $\boldsymbol{R}_{\boldsymbol{\eta\eta}}$, $\boldsymbol{H}[l]$, P_{Tx}, k_1, k_2 and $Q_b(\cdot)$

$\boldsymbol{H}[k] \leftarrow \mathcal{F}(\boldsymbol{H}[l])$

$\boldsymbol{R}_{\boldsymbol{x}[k]\boldsymbol{x}[k]} \leftarrow P_{Tx}\boldsymbol{I}\ \forall k \in [k_1, k_2]$

$\boldsymbol{R}_{\boldsymbol{yy}} \leftarrow \sum_{k=k_1}^{k_2} \boldsymbol{H}[k]\left(\boldsymbol{R}_{\boldsymbol{x}[k]\boldsymbol{x}[k]} + \boldsymbol{R}_{\boldsymbol{\eta_T\eta_T}}\right)\boldsymbol{H}^H[k] + \boldsymbol{R}_{\boldsymbol{\eta\eta}}$

$[\boldsymbol{T}, \boldsymbol{R}_{\boldsymbol{\eta}_q\boldsymbol{\eta}_q}] \leftarrow \mathrm{TF}\left(\boldsymbol{R}_{\boldsymbol{yy}}, Q_1(\cdot), \cdots, Q_M(\cdot)\right)$

$\boldsymbol{R}_{\boldsymbol{\eta'\eta'}} \leftarrow \boldsymbol{T}\boldsymbol{W}^H\boldsymbol{R}_{\boldsymbol{\eta}'_R\boldsymbol{\eta}'_R}\boldsymbol{W}\boldsymbol{T}^H + \boldsymbol{R}_{\boldsymbol{\eta}_q\boldsymbol{\eta}_q} + \sum_{k_1}^{k_2} \boldsymbol{H}[k]\boldsymbol{R}_{\boldsymbol{\eta_T\eta_T}}\boldsymbol{H}^H[k]$

$\boldsymbol{H'}[l] \leftarrow \boldsymbol{T}\boldsymbol{H}[l]\ \forall l \in \{0, \cdots, L-1\}$

$\boldsymbol{H'}[k] \leftarrow \mathcal{F}(\boldsymbol{H'}[l])$

$\boldsymbol{R}_{\boldsymbol{w}[k]\boldsymbol{w}[k]} \leftarrow \mathrm{TE}(\boldsymbol{H'}[k], \boldsymbol{T}\boldsymbol{W}^H\boldsymbol{R}_{\boldsymbol{yy}}\boldsymbol{W}\boldsymbol{T}^H, \boldsymbol{R}_{\boldsymbol{\eta'\eta'}})\ \forall k \in [k_1, k_2]$

$\boldsymbol{R}_{\boldsymbol{\eta'}[k]\boldsymbol{\eta'}[k]} \leftarrow \boldsymbol{R}_{\boldsymbol{\eta'\eta'}} + \boldsymbol{R}_{\boldsymbol{w}[k]\boldsymbol{w}[k]}\ \forall k \in [k_1, k_2]$

$\boldsymbol{A}[k] \leftarrow \boldsymbol{I} + \boldsymbol{R}^{-1}_{\boldsymbol{\eta'}[k]\boldsymbol{\eta'}[k]}(\boldsymbol{H'}[k])\boldsymbol{R}_{\boldsymbol{x}[k]\boldsymbol{x}[k]}(\boldsymbol{H'}[k])^H\ \forall k \in [k_1, k_2]$

$R = B\sum_{k=k_1}^{k_2} \log_2\left(\det\left(\boldsymbol{A}[k]\right)\right)$

return R

Due to the transmit noise, the modeling of the quantization and the channel estimation, the effective noise covariance matrix $\boldsymbol{R}_{\boldsymbol{\eta'\eta'}}$ and the effective channel $\boldsymbol{H'}[k]$ depend on the input covariance matrix $\boldsymbol{R}_{\boldsymbol{x}[k]\boldsymbol{x}[k]}$ The procedure for calculating the sum rate is summarized in Algorithm 3.5.

3.2.5.3 Simulation Results Multiuser Performance Comparison

Here we describe the chosen evaluation setup and the corresponding results. A basestation with 64 antennas ($M = 64$) receives the signals from 4 users ($U = 4$) with EVM of -25 dB. For the channel model of each user, identical modeling parameters but different realizations are chosen. We used the following parameters: $L = 128$, $P = 32$, $\beta = 0.5$ for the channel model described in Subsection 3.2.3. For the HBF system, $M_{\mathrm{ADC}} \in \{4,\ 8,\ 16,\ 32\}$ ADC chains are used. For DBF and HBF with uniform quantization we use a resolution of $b \in \{1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8\}$ bits. For the case of DBF with mixed resolution ADCs we used $M_h \in \{4,\ 8,\ 16,\ 32\}$ for the number of ADCs with high resolution. The transmit power for all users is the same. Since on average the channel gain is the same, the powers received from different users is on average similar. Since for the results with uniform quantization we found that the spectral efficiency at high SNR is maximized by an ADC resolution of 5 bits we chose $b_h = 5$. The resolution of the lower resolution ADC is chosen to be $b_l \in \{1,\ 2,\ 3,\ 4\}$ bits.

Fig. 3.15 (a) to (c) show the average achievable rate over 30 channel realizations. The resolution in bits increases from the top to bottom for each group of curves. From the DBF results in Fig. 3.15 (a) we see that at high SNR the rate saturates and there is only minor improvement above a resolution of 5 bits.

To analyze this result analytically we need to look at the maximum SNR of the combined signal from each user and compare this to the quantization noise dependent on the resolution. Based on

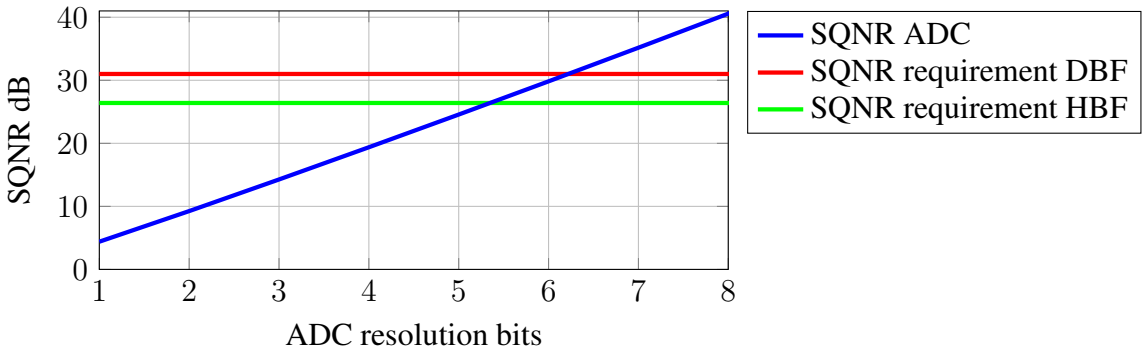

Fig. 3.14. SQNR achieved by an uniform ADC related to the required SQNR for the different system configurations.

the calculations in Section 2.1 the SQNR of an uniform ADC are calculated in Fig. 3.14. The performance is worse than the typical 6 dB per bit, as this calculation in addition includes the effect of clipping.

To calculate the maximum per-antenna SNR we need to first calculate the minimum receive noise power P_n and the maximum combined signal power P_s. In the high SNR regime, our system is limited by the EVM, thus assuming a non-coherent addition of the EVM of each user leads to minimum noise power. In our case with 4 users the minimum noise is therefore equal to

$$P_n = U \cdot 10^{\text{EVM}/10} P_u = 4 \cdot 10^{-25/10} P_u = 1.26 \cdot 10^{-2} P_u, \tag{3.69}$$

where P_u is the receive power from one user. In addition, we assume that scheduling and power control ensures these are the same for each user. For calculating the maximum combined signal power we assume that the signal of each user with power P_u coherently add up as

$$P_s = U^2 P_u = 16.0 P_u. \tag{3.70}$$

Therefore, the maximum SNR in dB is equal to

$$\text{SNR}_{\text{max}} = 10 \log_{10} \left(\frac{U^2 P_u}{U \cdot 10^{\text{EVM}/10} P_u} \right) = 31.0 \text{ dB}. \tag{3.71}$$

Since the quantization noise should be sufficiently smaller than the received noise, we can state that SQNR $>$ SNR$_{\text{max}}$. Based on Fig. 3.14 a resolution of $b \approx 6$ bits is required. This matches our simulation results in Fig. 3.15 (a).

It is important to note that for HBF, due to the analog receive beamforming before the ADC, the signals from all users have different power. In fact, some are largely attenuated since the beamforming is not tailored towards their direction. It is also important to mention that a coherent combining of the signal is also a coherent combining of the transmit noise. Therefore, even though the signal from one user has a larger power at the ADC compared to the DBF case, a smaller dynamic range needs to be covered. This is also obvious from our simulation results in Fig. 3.15 (b).

We can also calculate the minimum resolution at which the ADC is not limiting the performance for HBF, but we need some additional simplifying assumptions. Since each subarray is only adapted to the channel of one user, we assume that this user has a signal gain equal to the subarray size M_C. Since the subarray is not adapted to the other users, we assume that their signal power after the analog combining is equal to P_u. If we now assume that in the worst case the signal from all users coherently add up we get the overall signal power

$$P_s = (M_C + U - 1)^2 P_u = 361 P_u. \tag{3.72}$$

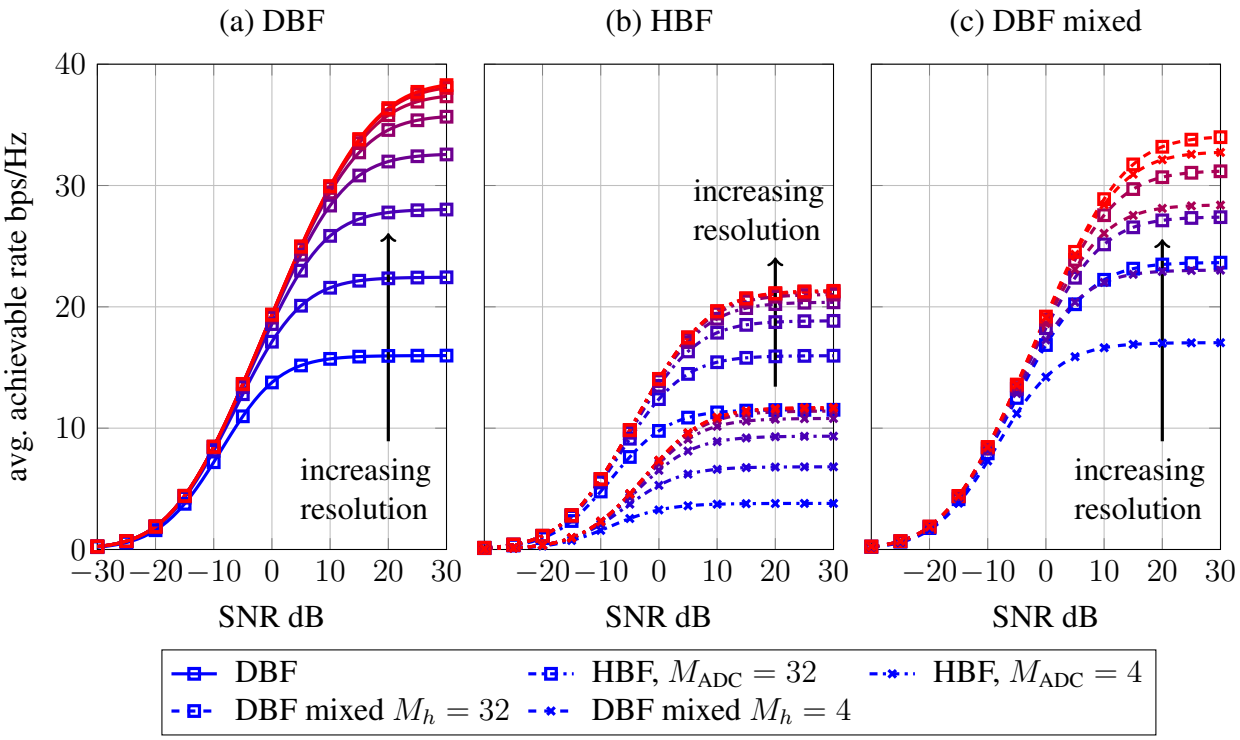

Fig. 3.15. DBF, HBF and DBF mixed average achievable rate for $M = 64$, $U = 4$, $M_{\text{ADC}} \in \{4,\ 32\}$, $M_h \in \{4,\ 32\}$ and ADC resolution $b \in \{1,\ \cdots,\ 8\}$.

Since the major noise component is the transmit noise and the analog combiner is adapted to each user's channel, the transmit noise for this user adds up coherently. Including the non-coherent combination with the transmit noise of the other users the minimum noise in this case is equal to

$$P_n = \left(M_C^2 + U - 1\right) 10^{\text{EVM}/10} * P_u = 0.819 P_u. \tag{3.73}$$

The maximum SNR before the ADC can now be calculated as

$$\begin{aligned} \text{SNR}_{\text{max}} &= 10 \log_{10} \left(\frac{(M_C + U - 1)^2 P_u}{\left(M_C^2 + U - 1\right) 10^{\text{EVM}/10} * P_u} \right) \\ &= 26.4 dB. \end{aligned} \tag{3.74}$$

Again the statement SQNR $>$ SNR_{max} must hold for the system not being limited by the quantization. The resolution b is thus required to be around 5 bits (see Fig. 3.14). This is the same result as in Fig. 3.15 (b) if we consider $M_C = 16$.

There are multiple aspects leading to the effect that DBF outperforms HBF especially in the low SNR regime. First, all possible degrees of freedom are available for each user in the case of DBF. Since in the case of HBF each subarray only uses a phase-shifter configuration optimized for one user, the resulting overall receive beamforming is far from optimal considering the sum of the available users. Since in the low SNR regime, the quantization noise is smaller than the noise in the receiver, the system is not limited by it. This is also evident from the fact that the rate curves in Fig. 3.15 (a) are on top of each other.

The results of the DBF mixed case in Fig. 3.15 (c) show that this approach can offer all possible rates in between the results of having only one ADC resolution, offering all possible values of energy and spectral efficiency around the values for DBF with only one ADC resolution. Combining

the observations of the achievable rate we can predict that the energy efficiency for an ADC resolution above 5 bits will not improve, since the achievable rate only shows limited improvement, while the power consumption of the front-end will dramatically increases.

We define the energy efficiency as the average achievable sum rate R divided by the power consumption of the RF front-end P_R

$$\text{energy efficiency} = \frac{R}{P_R}. \tag{3.75}$$

The power consumption of the different receiver front-ends are calculated according to Section 2.4. The scenarios in Fig. 3.16 (a) to (d) show the achievable rate and energy efficiency at different SNR values. For each curve the ADC resolution increases from the leftmost point of the curve. This point represents 1 bit resolution for all ADCs, or 1 bit resolution for the ones with lower resolution ADCs in the case of mixed-ADC DBF. For all cases we see that the DBF system is more energy efficient compared to HBF. The major reason for this is that the digital system retains all available degrees of freedom. We can see that as the SNR increases (Fig. 3.16 (a) to (c)) the smaller the improvement of additional ADC chains. The explanation for this is that even though we gain more degrees of freedom we still need to divide them among the users. In Fig. 3.16 (c) we see that there is little difference between having 8 or 16 ADC chains.

As the SNR increases from Fig. 3.16 (a) to (c) the optimal resolution in terms of energy efficiency improves. As predicted from the achievable rate curves, above a resolution of 5 bits the energy efficiency decreases for all cases. The results for DBF with mixed configurations in Fig. 3.16 (a) to (d) show that these curves are tightly clustered around the curves for the case with only one resolution. This shows that this approach can achieve all possible different values in the rate - energy efficiency trade-off.

3.2.5.4 Conclusion Multiuser Performance Comparison

The evaluations in this paper showed that low resolution ADC digital beamforming systems are more energy efficient and achieve a higher rate than hybrid beamforming systems for multiuser scenarios. The reason is that the subarrays of hybrid beamforming must focus on a single user. Evaluations with mixed ADC configurations showed that such systems can achieve different achievable rate and energy efficiency values around the ones achieve by a uniform ADC configuration.

3.2.6 Robust massive MIMO Equalization for mmWave systems with low resolution ADCs

Many of the investigated channel estimation and detection schemes require detailed knowledge about the structure of the channel. For example, algorithms like GAMP [58, 23] are very sensitive to the case that their modeling assumptions are not fully valid. Other algorithms like Expectation Maximization (EM) require accurate knowledge about the sparsity or related parameters [58, 23]. In a practical system, this knowledge is hard to obtain. In addition, the systems should be robust regarding cases, where the assumptions leading to a specific algorithm are not fully satisfied.

Many massive MIMO equalization schemes consider only perfect channel estimation [30, 97] without channel coding. We think the propagation of the channel estimation error inside the MU-MIMO equalization is not straightforward. Only for linear methods the influence of the channel estimation error can be investigated theoretically [27]. This motivated us to investigate channel estimation, MIMO equalization in combination with channel coding and low resolution ADCs.

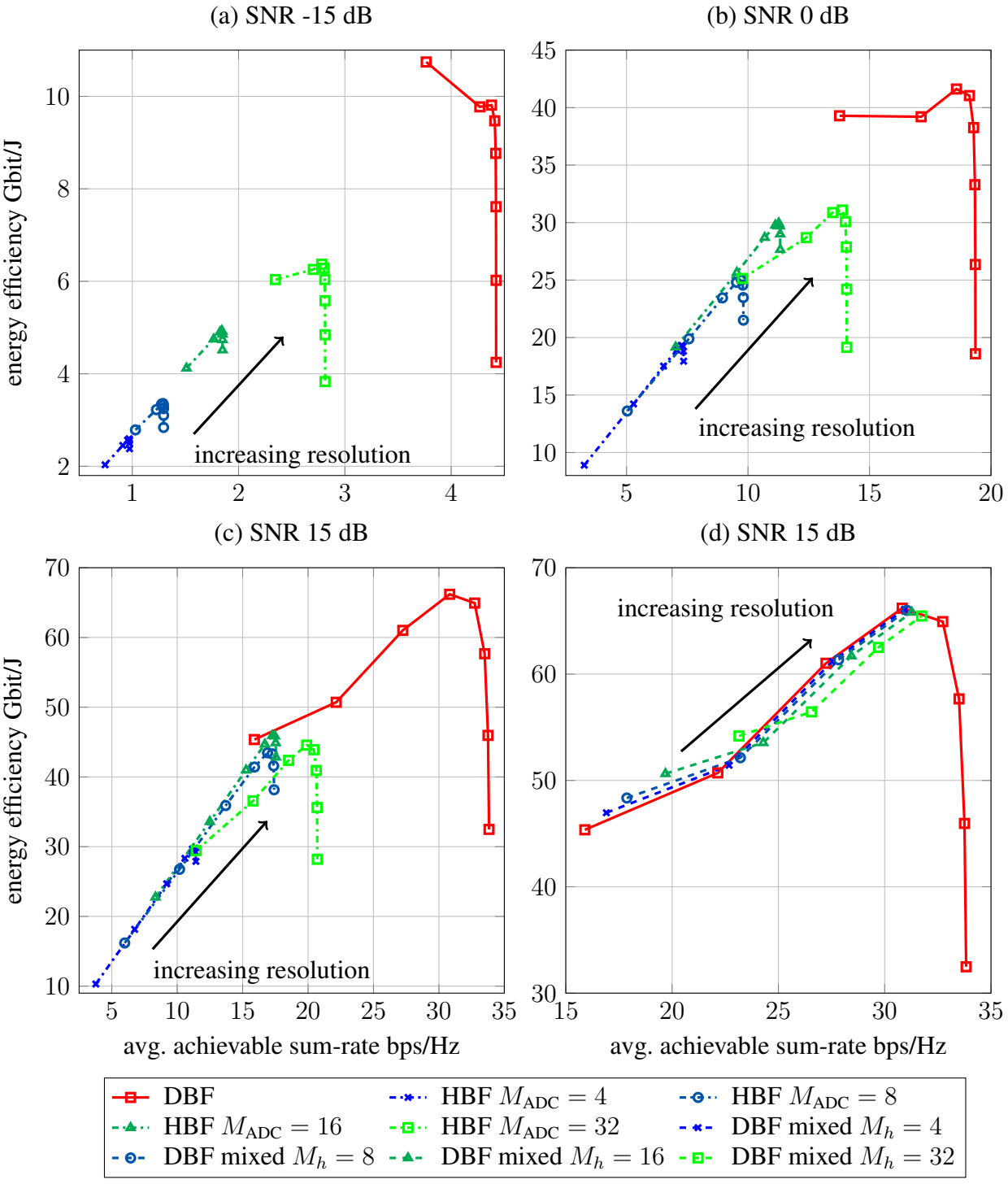

Fig. 3.16. Spectral and energy efficiency of digital beamforming with/without mixed ADC configuration and hybrid beamforming with $M = 64$, $U = 4$, $M_{\text{ADC}} \in \{4,\ 8,\ 16,\ 32\}$, $M_{\text{h}} \in \{4,\ 8,\ 16,\ 32\}$ and ADC resolution $b \in \{1,\ \cdots,\ 8\}$, $b_l \in \{1,\ \cdots,\ 4\}$ and $b_h = 5$ at SNR $\in \{-15\ \text{dB},\ 0\ \text{dB},\ 15\ \text{dB}\}$.

Therefore, we decided to investigate the performance of these systems without limiting assumptions on the statistics of the channel. We also put the focus on linear, low complexity algorithms while considering reference signals developed for NR in 3GPP. Even though we do not consider finite resolution calculation in this work, it is important to mention that massive MIMO is very robust to these effects [34]. This could further simplify the required calculation and lead to an implementation that might be feasible for mobile devices.

In this subsection we introduce the sequential Dichotomous Coordinate Descent (DCD) algorithm for MU-MIMO equalization. We also show a performance and computational complexity based comparison to MMSE MU-MIMO detection using a 3GPP NR signals. The work in this section was also published in [80].

3.2.6.1 DCD MIMO Detection Algorithm

In this section we show how the sequential DCD with bound is derived from a relaxation of the Maximum Likelihood (ML) MIMO equalization. The classical problem of ML detection can be formulated as

$$\hat{\boldsymbol{x}} = \underset{x_n \in \mathbb{X}}{\operatorname{argmin}} \, ||\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}||_2^2 , \tag{3.76}$$

where $\mathbb{X}$ is the set containing all possible transmit symbols. The symbols $\boldsymbol{x}$, $\boldsymbol{H}$, $\boldsymbol{y}$ and $\hat{\boldsymbol{x}}$ represent the transmit symbol, the channel, the receive symbol and the symbol after the detection of the system, respectively. The complexity of this discrete optimization problem grows exponentially with the dimensions of $\boldsymbol{x}$ and the size of the set $\mathbb{X}$. Thus, for higher number of spatial streams as envisioned for massive MIMO it is not feasible to solve this problem.

Fortunately, as the number of receive antennas grows large with respects to the number of simultaneously served users, the MMSE solution to the relaxed optimization problem

$$\hat{\boldsymbol{x}} = \operatorname{argmin} \mathbb{E}\left[||\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}||_2^2\right] , \tag{3.77}$$

approaches the performance of the ML detection [99]. Unfortunately, the close form solution to this problem requires knowledge about the noise covariance matrix. For a system with a large number of antennas this is hard to attain.

Therefore, we choose to relax the ML detection problem in a way that reduces the complexity, but does not make any assumptions on the noise statistics

$$\hat{\boldsymbol{x}} = \underset{\Re(x_n)\in[-B,B],\ \Im(x_n)\in[-B,B]}{\operatorname{argmin}} \, ||\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}||_2^2 . \tag{3.78}$$

The variable B forces the real and imaginary part of each element of the vector $\hat{\boldsymbol{x}}$ to be in the range from $-B$ to B. In the following paragraphs we will show how to solve this optimization problem efficiently and that we do not need to make any assumption on the noise statistics.

This problem can be reformulated into solving the following linear system of equations

$$\boldsymbol{H}^H\boldsymbol{H}\boldsymbol{x} = \boldsymbol{H}^H\boldsymbol{y}. \tag{3.79}$$

Thus, we can utilize a coordinate descend based method to solve this problem. To reduce the complexity we selected the step-size to be of the form 2^{-l}, where l is an integer. This has the advantage that all multiplications with this number can be implemented by a bit shift. These algorithms are called DCD as described in [71, 72] for multiuser detection in a Code Division Multiple Access

Algorithm 3.6 Sequential DCD with bound

Require: $\boldsymbol{A}$, $\boldsymbol{b}$, N, H, B, N_u, M_b
Initialization: $\boldsymbol{x} \leftarrow \boldsymbol{0}$, $\boldsymbol{r} \leftarrow \boldsymbol{b}$, $\alpha \leftarrow H$, $m \leftarrow 0$,
UpdateFlag $\leftarrow$ false, $k \leftarrow 0$
while $m < M_b$ **do**
  for $n \in \{1, \cdots, N\}$ **do**
    if $\alpha/2[\boldsymbol{A}]_{n,n} < |[\boldsymbol{r}]_n|$ **then**
      $t \leftarrow [\boldsymbol{x}]_n + \text{sign}([\boldsymbol{r}]_n)\alpha$
      if $t \leq B$ **then**
        $[\boldsymbol{x}]_n \leftarrow t$
        $\boldsymbol{r} \leftarrow \boldsymbol{r} - \text{sign}([\boldsymbol{r}]_n)\alpha\boldsymbol{a}_n$
        UpdateFlag $\leftarrow$ true, $k \leftarrow k + 1$
      end if
    end if
  end for
  if $k \geq N_u$ **then**
    return $\boldsymbol{x}$, $\boldsymbol{r}$
  end if
  if UpdateFlag **then**
    UpdateFlag $\leftarrow$ false
  else
    $m \leftarrow m + 1$, $\alpha \leftarrow \alpha/2$
  end if
end while
return $\boldsymbol{x}$, $\boldsymbol{r}$

(CDMA) system and is shown in Algorithm 3.6. The parameters H, B, N_u and M_b are the maximum step-size, the upper bound of the detected symbols, the maximum number of updates and the maximum number of step-size divisions by 2. The parameter B should be chosen in a way to accommodate the QAM constellation. For the case that the constellation is not bounded these parameters can be set to a reasonable large value or even infinity. For example, in a DFT-spread-OFDM (DFT-s-OFDM) system the constellation in the frequency domain is not bounded. The value of H should be of the form 2^{-l}, where l is an integer. Since it is not useful to start with a step-size that is larger than the final bound, H should also be smaller than B. A good way to choose H is $H = 2^{\lfloor \log_2(B) \rfloor}$. This ensures that we start with the maximum possible step-size to enable fast convergence. The symbols $\boldsymbol{A}$, $\boldsymbol{b}$ and N define the linear systems of equations

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \tag{3.80}$$

and the size of the vector $\boldsymbol{b}$. The vectors $\boldsymbol{x}$ and $\boldsymbol{r}$ represent the resulting vector and the residual error, which is updated in every step. Since algorithm 3.6 solves only real, linear systems of equations $\boldsymbol{A}$ and $\boldsymbol{b}$ are related to $\boldsymbol{H}$ and $\boldsymbol{y}$ in the following way

$$\boldsymbol{A} = \begin{bmatrix} \Re(\boldsymbol{H}^H\boldsymbol{H}) & -\Im(\boldsymbol{H}^H\boldsymbol{H}) \\ \Im(\boldsymbol{H}^H\boldsymbol{H}) & \Re(\boldsymbol{H}^H\boldsymbol{H}) \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} \Re(\boldsymbol{H}^H\boldsymbol{y}) \\ \Im(\boldsymbol{H}^H\boldsymbol{y}) \end{bmatrix}. \tag{3.81}$$

Since we solve the equivalent real, linear system of equations the value of N is double the number of users/spatial streams to be detected. The resulting value of $\boldsymbol{x}$ is also going to be split between

Table 3.1. Simulation Parameter for DCD based MIMO equalization.

Parameter	Value
Reference Signal	3GPP NR OFDM Type I DMRS
Channel Estimation	2x1D MMSE and ideal
Number of Users	8
Number of receive antennas	64
Channel model	Exponential PDP (no Doppler spread)
SNR definition	Average per user per antenna SNR
Channel code	LTE turbo code with rate 0.9
MIMO detection algorithms	MMSE and DCD-Bound
ADC resolution	2 bit
Modulation format	16 QAM

real and imaginary part in the same way as $\boldsymbol{b}$. Combining the real and imaginary part in $\boldsymbol{x}$ lead to the detected symbol.

3.2.6.2 Simulation Results DCD MIMO Equalization

In this section we compare the performance of sequential DCD to MMSE equalization. Since we consider system with low resolution ADCs and channel estimation error, the noise is in general not white. We also tested additional noise whitening for both systems, but since no performance gain was observed at an addition computational cost, we only show results without noise whitening. The simulation parameter in Table 3.1 shows the most important simulation parameters. It is also important to mention that the bounds for the sequential DCD equalization tightly encloses the QAM constellation.

The uncoded and coded Bit Error Ratio (BER) results are shown in Fig. 3.17 and 3.18. Due to the frequency selective channel and the 2 bit resolution ADCs, the uncoded BER does exhibits an error floor. It is also obvious that error floor of the DCD is lower than the one of the MMSE algorithm w/o ideal channel estimation. The channel estimation is implemented as described in Section 2.5. As we can see from the zoomed in part around 10^{-2} BER, the performance of DCD is slightly more robust to channel estimation errors. Fig. 3.18 shows that these results translate well to a system with channel coding.

Since the DCD algorithm has no multiplications, we need to compare the complexity to MMSE by mapping additions and multiplications to logic operations. The work in [44, 16] offer a mapping

Table 3.2. Complexity per operation.

operation	real additions	real multiplications	logic operations
$\boldsymbol{H}^H\boldsymbol{H}$	8128	8192	19038400
$\boldsymbol{h}^H\boldsymbol{y}$	2032	2048	5521600
$\boldsymbol{H}^H\boldsymbol{H}+\boldsymbol{I}$	16	0	2000
$(\boldsymbol{H}^H\boldsymbol{H}+\boldsymbol{I})^{-1}$	1700	1900	4392500
$(\boldsymbol{H}^H\boldsymbol{H}+\boldsymbol{I})^{-1}\boldsymbol{h}^H\boldsymbol{y}$	240	256	593200
Sequential DCD with bound	2000	0	250000

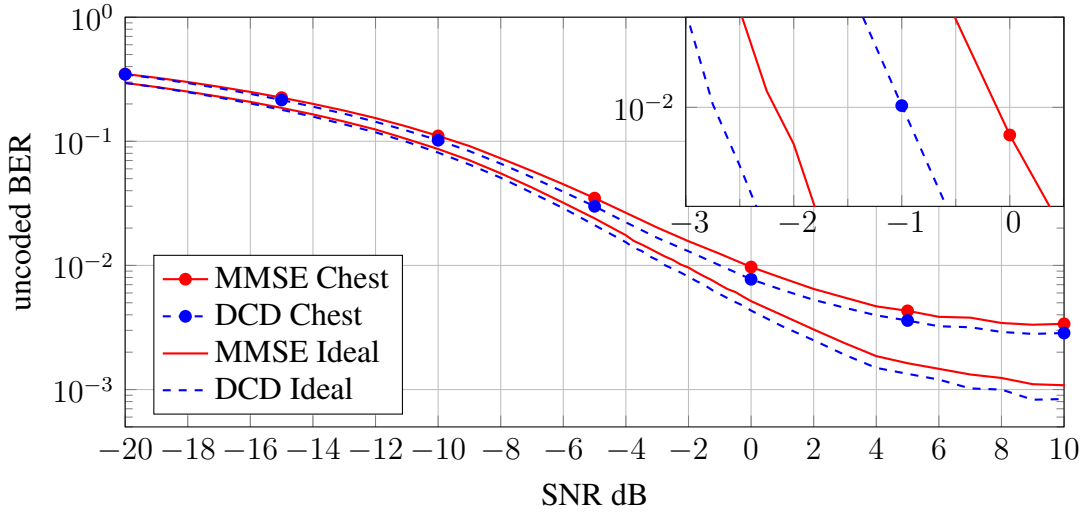

Fig. 3.17. Uncoded BER results.

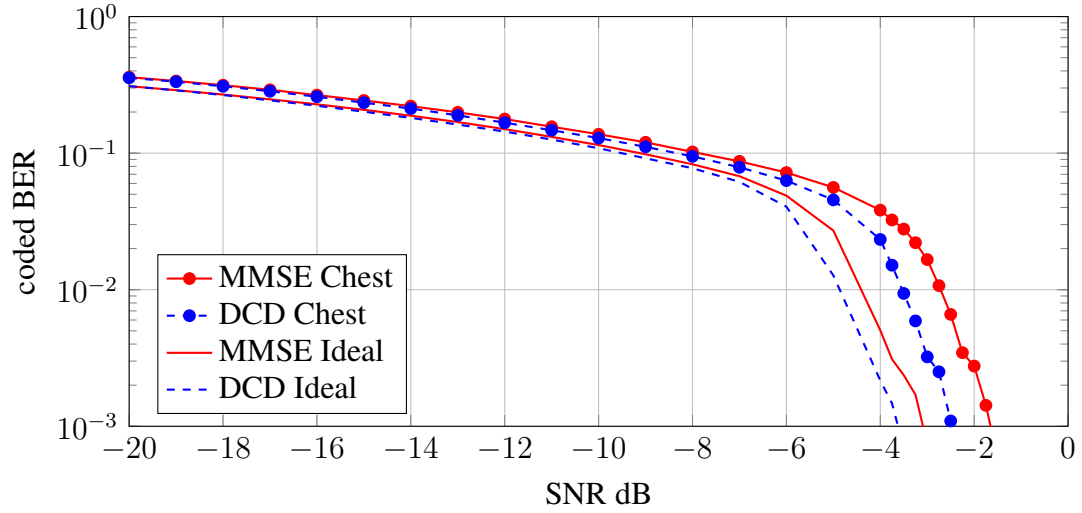

Fig. 3.18. Coded BER with LTE turbo code rate = 0.9.

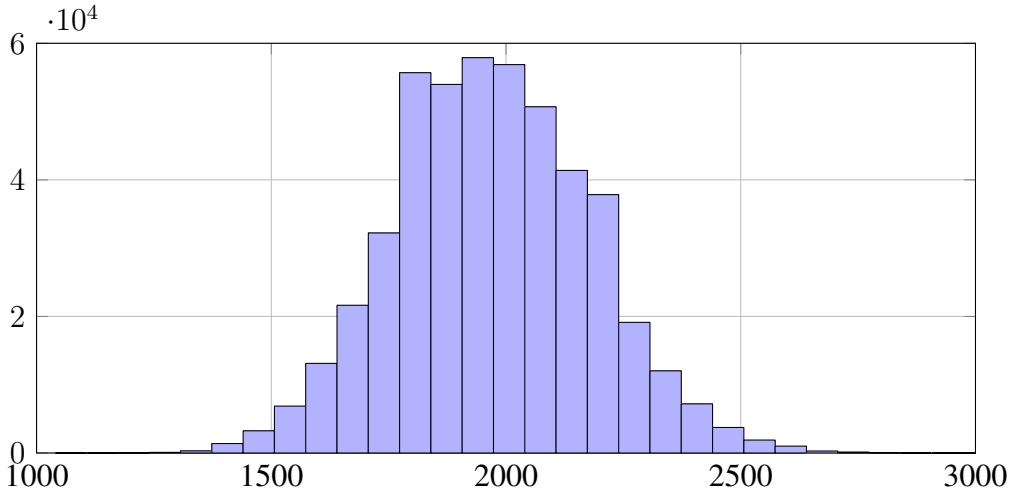

Fig. 3.19. Histogram of the additions require to converge.

Table 3.3. Complexity results calculated for 14 consecutive OFDM symbols.

detection algorithm	scenario 1	scenario 2
MMSE	29547700	109040100
Sequential DCD with bound	24810000	99840800

of real additions and multiplications to logic gates. Assuming 18 bit signed fixed point calculation an addition and a multiplication can be implemented using 125 and 2200 NOT-AND logical gate (NAND) gates with two inputs, respectively. To compare the different algorithms, we compare this number of logic gates as the number of logic operation required to calculate the result.

As the sequential DCD algorithm is an iterative procedure, we need to investigate the convergence of it. The histogram in Fig. 3.19 shows the number of additions used to process one symbol vector for the simulation parameters presented in the preceding paragraph. The comparisons are implemented as a subtraction followed by checking if the sign bit is set or not. Therefore, they are counted to have equal complexity compared to an addition. We would also like to mention that the number of comparisons is low compared to additions used for updating the residual vector $\boldsymbol{r}$. The average complexity is 1972 real additions. For simplicity we use 2000 for the following analysis.

From the Table 3.2 it is easy to see that even just the multiplication with the already inverted matrix is more complex than solving the linear system of equations with the Sequential DCD algorithm with bound. It is also obvious that the complexity is dominated by the computation of the Gram matrix ($\boldsymbol{H}^H\boldsymbol{H}$). It is important to mention that for the calculation of the Gram matrix, we already exploited the symmetry of the resulting matrix to minimize the necessary multiplications and additions. All other computation steps have a much lower complexity. In this investigation, we neglected the necessary complexity for normalization of the signal power and the additional complexity for making the MMSE equalizer unbiased.

To compare the MMSE equalizer to the algorithm developed here, we compare two scenarios. In the first scenario, the matrix computed to generate the MMSE result is calculated separately for each SC (scenario 1). In the second case we assume that the matrix can be reused to detect the symbol in 14 consecutive OFDM symbols on the same SC (scenario 2). There are few common operations to both systems and we assume that this intermediate calculation result can be stored and reused for scenario 2. The complexity for both algorithms is shown in Table 3.3. The overall computational complexity of our approach compared to MMSE is reduced while at the same time the performance is improved. In the first scenario the improvement is about 16 % in the second it is in the range of 10 %.

3.2.6.3 Conclusion DCD MIMO Equalization

Our investigation showed that a bounded DCD MIMO equalization algorithm does outperform a MMSE based equalization. In addition, sequential DCD has a lower computational complexity. But it is important to mention that in contrast to many other papers considering the complexity for massive MIMO we showed that the complexity is dominated by the computation of the Gram matrix and not the matrix inversion. This evaluation shows that it is possible to achieve high data rates with digital beamforming mmWave system with low resolution ADCs by considering algorithms with low complexity.

3.2.7 Are the data rates predicted by the analytic analysis of receivers with Low Resolution ADCs achievable?

Evaluation of digital beamforming systems with low resolution ADCs has mainly concentrated on isolated analysis of either signal processing aspects or achievable rate. There have been analyses considering uplink wideband systems [61], multiple users in the uplink [38] or the effects of imperfect CSI at the receiver [48]. Different signal processing issues such as channel estimation [9] and MIMO equalization [94, 61] have been analyzed in isolation of other receiver aspects.

In this work we want to bridge the gap between these two evaluation methods. The main target is to show that indeed we can approach the attractive data rates promised by analytic methods using standard signal processing techniques. In [83] the authors show a similar comparison for a point-to-point LTE system. Since we concentrate on a system setup that is almost identical with a 3GPP NR system and a SC spacing scaled to 480 kHz, we believe that these results show that such a receiver might be attractive for a practical system. The results of this section also appeared in [82].

3.2.7.1 Rate Evaluation Information Theory vs Link Level Simulation

For the comparison we use almost the same rate calculations as developed in Subsection 3.2.5.2. The only difference is that in this work we included the effects of linear MMSE MU-MIMO equalization into our evaluation. To explain the main principles of this calculation we just highlight the effects taken into account as well as the modeling assumptions. The rate calculation takes the following effects into account:

- Multiple users
- Multiple receive antennas
- Frequency-selective channel
- Low resolution ADCs with quantization error modeled as non-white additive noise with variance dependent on the receive power
- Imperfect CSI at the receiver based on derived expressions for the channel estimation MSE
- Transmitter impairments modeled as an additive EVM
- Perfect timing and carrier frequency synchronization
- MMSE MU-MIMO spatial channel equalization

To make the calculation tractable we used the following assumptions:

- Frequency domain sub-channels are frequency flat
- Transmit signal is Gaussian
- No collaboration among the users

These assumptions are justified below. If the size of the DFT is chosen to be sufficiently large, the resulting sub-channels can be well approximated to be frequency flat. For constellations of the QAM family, there exists only a small shaping-gap compared with Gaussian symbols [29]. The assumption of Gaussian quantization noise is not satisfied for very low resolutions (1-2 bit) in the time domain. However, all rate calculations in this work are made in the frequency domain. Due to the central limit theorem [28] the distribution of the quantization noise in the frequency domain converges to Gaussian. We have also verified this in our simulations. For most of the cases in a practical system, users cannot collaborate, because they have different data to transmit and do not know that the others are present.

For the link-level simulations we take the same effects into account as described in the preceding paragraph for the theoretical evaluation. There are some additional effects that must be accounted for, and we describe them in the following paragraph.

As mentioned above, since the processing is performed in the frequency domain, the quantization noise can be modeled as Gaussian. Thus, in all the link-level processing steps we simply treat the quantization error as additional noise. Based on the channel estimate we also determine the signal power by averaging over the estimated power at each SC. In a similar fashion we estimate the noise power by regenerating the receive signal from the channel estimate and the known reference signal. Afterwards, we subtract this regenerated receive signal from each user to generate noise samples and then average over their power. This estimate is used to align the power of the noise at each antenna.

As in Subsection 3.2.6 we use an unbiased MMSE receiver for the MIMO equalization. We use the channel estimates described in Section 2.5 to generate the MMSE MIMO equalizer. We also use the rate matching channel code and the Cyclic Redundancy Check (CRC) from a standard 3GPP LTE system defined in [1]. To adapt to different channel conditions, we use QPSK, 16-QAM and 64-QAM modulation with different code rates. For each SNR we calculate the Block Error Ratio (BLER) for each of the modulation formats and combine that with the bits transmitted in one block to calculate the transmitted data rate. Then we select the highest data rate separately for each SNR.

3.2.7.2 Simulation Results Information Theory vs Link Level Simulation

All major simulation parameters are shown Table 3.4. These describe a 3GPP NR system with a SC spacing of 480 kHz, but we used an LTE turbo code instead of the codes defined for NR. To enable simulations over a wide SNR range, the modulation and code rate combinations in Table 3.5 are used. At the modulation format switching point we always choose a configuration for each of the two modulation formats to have the same spectral efficiency. This enables us to compare the effects of the different modulation formats. For the calculation of the spectral efficiency we included the CP and channel estimation reference signal overhead in both the theoretical and link-level evaluations. We also only used the occupied bandwidth of $1200 * 32 * 15$ kHz $= 576$ MHz for the spectral efficiency calculation, thus including the guardband required in a practical system. It is important to mention that the oversampling is also taken into account in the theoretical rate calculation, as it changes the quantization noise power for each frequency bin.

The results for the comparison between the rate calculation and the link-level simulations are shown in Fig. 3.20, and we see a gap between the theoretical and link-level simulation results. There are multiple effects that make the results of the link-level simulations worse than the theoretical evaluation:

- Transmit symbols are not Gaussian
- Theoretical evaluation does not take the Inter Symbol Interference (ISI) of channels that are longer than the CP into account
- LTE turbo code performance is not ideal

Considering all these reasons for the performance gap, we believe that these results show that the interesting rates predicted in many theoretical evaluations could translate well to a practical implementation. The performance gap is consistent with that observed in similar evaluations for a point-to-point LTE system in [83]. We note that the low-SNR regime is the more likely operating condition for such systems, and the observed performance difference is much smaller there. For

Table 3.4. Simulation Parameters.

Parameter	Description
Reference Signal	3GPP NR OFDM Type I DMRS
Reference OFDM symbols per slot	2
Channel estimation interpolation	1D MMSE for time and frequency dimensions
Number of users	4
Number of receive antennas	64
Channel model	NYU channel model [84]
SNR definition	Average per user per antenna SNR
Channel code	LTE turbo code
MIMO detection algorithms	MMSE
ADC resolution	1-3 bit
Modulation format	QPSK, 16-QAM, 64-QAM
Number of channel realizations	30
SC spacing	480 kHz
Allocated SCs	1200
OFDM symbols per slot	14
Slots simulated per channel realization	10
EVM	-25 dB

Table 3.5. Link-level simulation modulation and Code Rate (CR) combinations.

Modulation format	Code Rate (CR)	bits per symbol
QPSK	0.1	0.2
QPSK	0.2	0.4
QPSK	0.3	0.6
QPSK	0.4	0.8
QPSK	0.5	1.0
QPSK	0.6	1.2
16-QAM	0.3	1.2
16-QAM	0.4	1.6
16-QAM	0.5	2.0
16-QAM	0.6	2.4
64-QAM	0.4	2.4
64-QAM	0.5	3.0
64-QAM	0.6	3.6
64-QAM	0.7	4.2
64-QAM	0.8	4.8

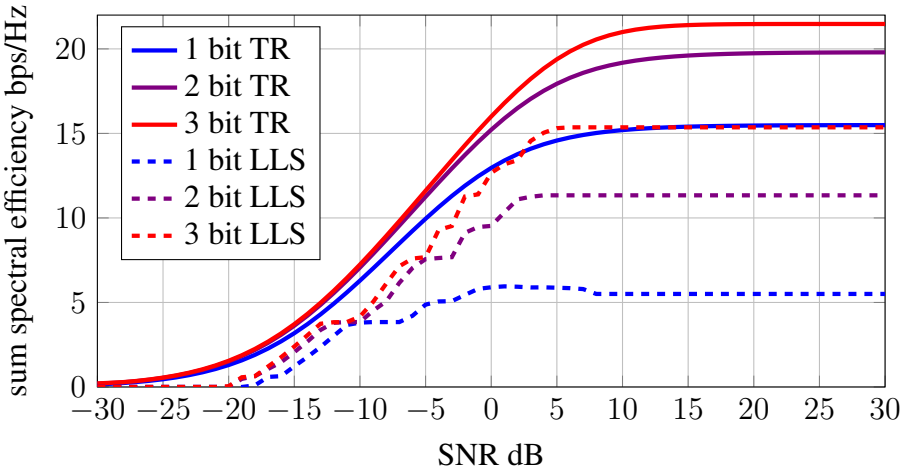

Fig. 3.20. Comparison of spectral efficiency results based on Theoretical Rate (TR) results and Link-Level Simulation (LLS) including reference and CP overhead.

the modulation and coding schemes that result in the same spectral efficiency, there is a performance gap of 5 dB between QPSK and 16-QAM modulation for 1 bit resolution, showing that all amplitude information is lost during the quantization of the time domain symbols. As the systems using ADCs with 2-3 bits of resolution do not remove the amplitude information for each time domain sample, the performance difference for the configuration with the same spectral efficiency is in the range of 0 to 1 dB. We also highlight that the Spectral Efficiency (SE) of 10 bps/Hz results in a combined data rate of about 5.8 Gbit/s.

3.2.7.3 Conclusion Information Theory vs Link Level Simulation

This comparison shows that indeed the data rates predicted for mmWave digital beamforming with low resolution ADCs translates well into link-level simulation results. At low per-antenna SNRs where such systems will likely operate, the performance of the link-level simulation can achieve the same data rate at about 4 dB lower SNR compared to the theoretical prediction. At high SNRs, the additional limitations of the constellation used in this work leads to a maximum data rate that is substantially lower compared to the theoretical prediction.

3.2.8 Energy and Spectral Efficiency Comparison Including Beam Training Overhead

In this this subsection we compare the energy and spectral efficiency results including the beam training overhead. In addition, to the phase array based hybrid beamforming approach included in the preceding subsection we also use Butler matrix based hybrid beamforming in this subsection. For the calculation of the rates we use the formulas derived in Subsection 3.2.5 including the effects of low resolution ADCs and imperfect channel estimation. For this evaluation we also used beams specifically derived for different number of beams per subarray and the subarray size based on the design methodology defined in Section 2.2.

3.2.8.1 Beam Training Overhead Estimation

We use the algorithm in Subsection 2.3.7.2 to determine a suitable sequence length for reliably detecting the channel quality. From the previous results in Subsection 3.2.5 we set the value of the

Table 3.6. Portion of the time available for data transmission given the assumptions on the beam training.

Antennas \ Beams	4	8	16	32
4	98.7 %	97.4 %	94.8 %	89.6 %
8	99.6 %	99.2 %	98.6 %	97.2 %
16	99.9 %	99.7 %	99.6 %	99.2 %

minimum per antenna SNR to -26 dB. Assuming the analog combining with the maximum array gain of $10\log_{10}(M_C)$ the resulting post combining SNR of -20, -17, and -14 dB is used for the M_C equal to 4, 8 and 16. The resulting sequence lengths are 2330, 1168 and 586 samples. Now in addition we need to add the time it takes to switch between beams. We assume that the switching time is similar to the switching time of a RF switch in the range of 100 ns [69], which results in 200 sample for the assumed sampling rate of 2 GSamples/s. We combine this with the beam training interval of 100 ms derived in Subection 2.3.7.2 and assumption that the beams of an average of 16 users are trained every time interval. Based on these calculations we get the amount of time that is available for data transmission in the system in Table 3.6. In this case the beam training is performed separately for each subarray.

3.2.8.2 Designed Beams for Different Subarray Size

We used the approach in Section 2.2 to design different codebooks for the phase shifter based hybrid beamforming architecture. As MU-MIMO should be supported this codebook is used for each subarray independently. The measurements leading to the channel model in [84] showed that even for a NLOS mmWave channel, there is a path or cluster of paths dominating the receive power. Therefore, the beams have combined a uniform coverage, whereas each beam is desired to be flat in its own region. To show the impact of different number of beams and thus different beamwidth beams with different width were designed. For the codebook generation it is assumed that each subarray forms an ULA.

The beams in Fig. 3.21 are designed for different codebook size. As combined a uniform coverage should be achieved, the optimal beam would be the maximum gain in the desired direction without any sidelobes. This is not achievable for an actual beam. The gain in Fig. 3.21 is always relative to the ideal beam of the same size. Using a Butler matrix with $M_C = 8$ the beam codebook in Fig. 3.22 can be achieved.

3.2.8.3 Simulation Parameters and Results Performance Comparison including Butler Matrix and Beam Training Overhead

All configuration parameters of the following simulations are given in Table 3.7. The rate calculations make the same assumption and for the hybrid beamforming use the same beam selection mechanism as in Subsection 3.2.5. For the calculation of the energy efficiency the power model from Section 2.4 is used.

A comparison of some of the receiver configurations is given in Fig. 3.23. As for the full digital case Fig. 3.23 (a) all degrees of freedom are available it achieves the best performance. From the results in Fig. 3.23 (b) we see that the amount of different beams thus their gain has a very limited effect on the overall performance. This is also evident from the comparison of the phase shifter based HBF in Fig. 3.23 (b) to the Butler matrix based one in Fig. 3.23 (c). We can therefore state

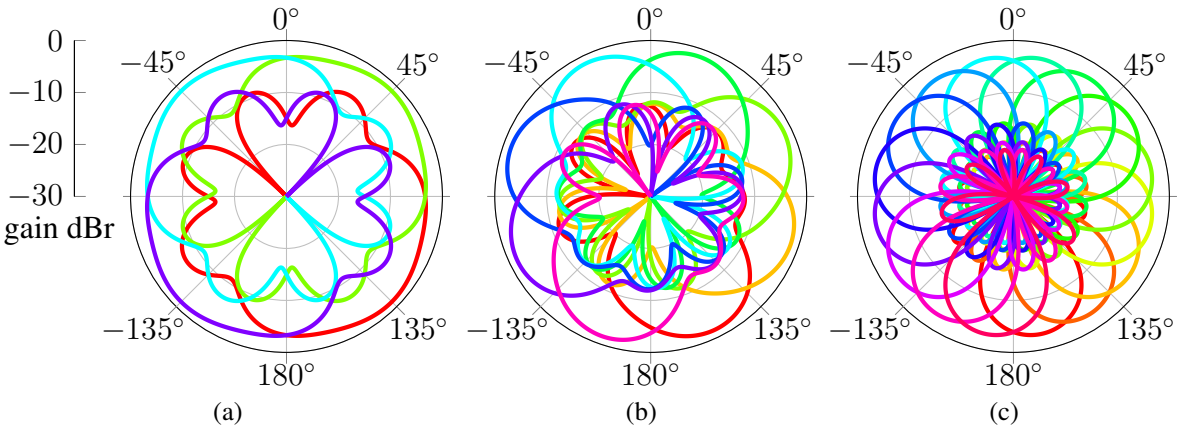

Fig. 3.21. Relative receive power of beam codebooks designed for $M_C = 8$ with (a) 4 beams, (b) 8 beams and (c) 16 beams.

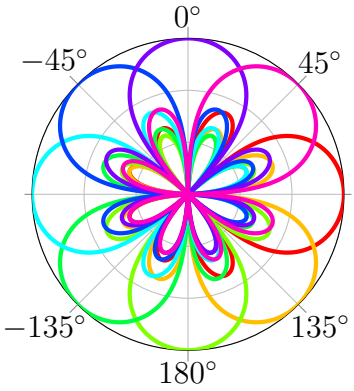

Fig. 3.22. Relative receive power of beam codebooks designed for $M_C = 8$ based on a Butler Matrix.

Table 3.7. Simulation Parameters.

Parameter	Description
Channel estimation error	Based on MMSE interpolation of 3GPP Type I OFDM DMRS (see Section 2.5)
Number of users	4
Number of receive antennas	64
Channel model	NYU channel model [84]
SNR definition	Average per user per antenna SNR
ADC resolution	1-8 bit
Subarray size M_C	4, 8, 16
Number of channel realizations	100
EVM	-25 dB

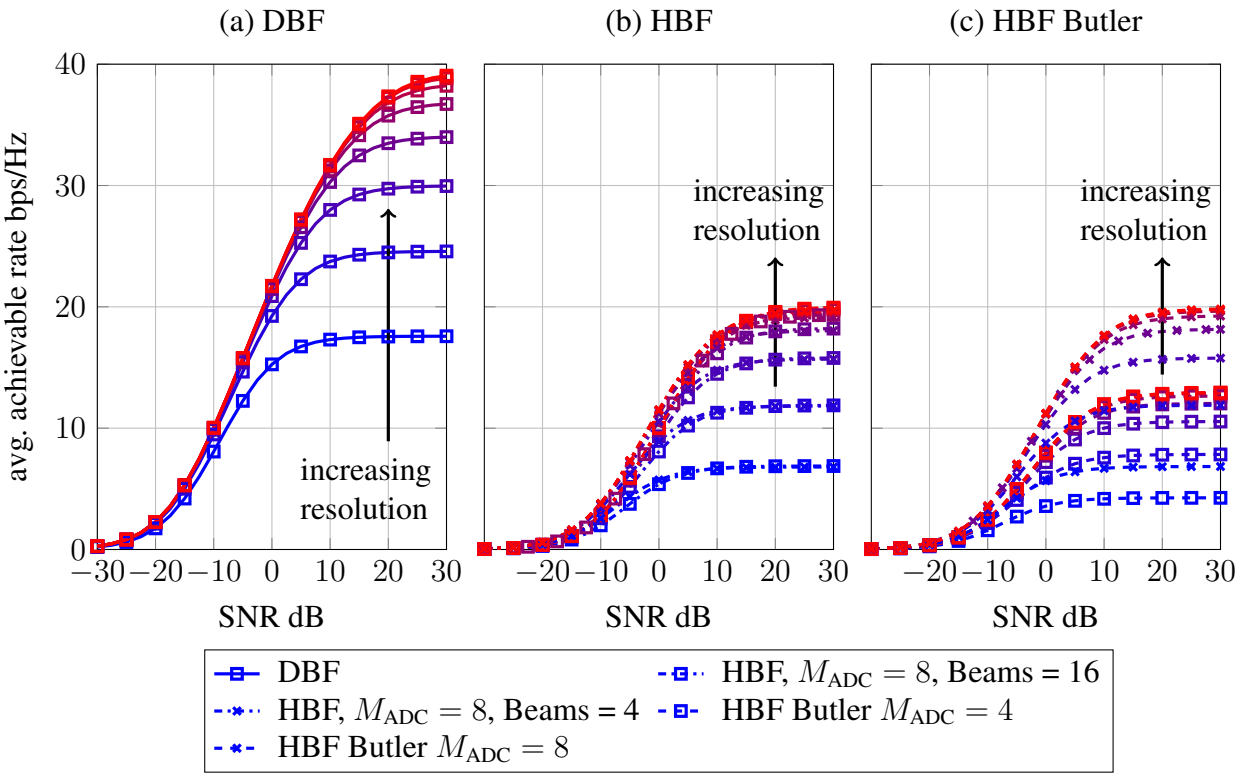

Fig. 3.23. DBF, HBF and HBF Butler average achievable rate for $M = 64$, $U = 4$, $M_{\text{ADC}} \in \{4,\ 8\}$, Beams $\in \{4,\ 16\}$ and ADC resolution $b \in \{1,\ \cdots,\ 8\}$.

that for HBF in a MU-MIMO scenario the size of the codebook as a significantly smaller impact on the performance than the available degrees of freedom.

As the results of the spectral efficiency evaluation reveled, that the impact of the number of beams on the performance is minimal for the evaluation of the energy efficiency only the case with equal number of subarray antennas and beams is used. The results in Fig. 3.24 show that in the case of low per antenna SNR the digital system is substantially more energy as well as spectral efficient. However, starting from 0 dB SNR the energy efficiency of the digital system and the system based on a Butler matrix get comparable. This can also be explained by the fact that in this SNR range the transmitter imperfection become more and more a limiting factor of the system. Since the sum spectral efficiency of multiple users is mainly dependent on the degrees of freedom and not the accuracy of the beams and in addition the receiver of the HBF system based on a Butler matrix consumes less power than the one based on phase shifters, it is more energy efficient.

3.2.8.4 Conclusion Performance Comparison including Butler Matrix and Beam Training Overhead

In this section we showed that considering many different aspects of the communication systems the low resolution digital beamforming system remains an attractive options. The low SNR region is likely to be the operating region of mmWave systems.

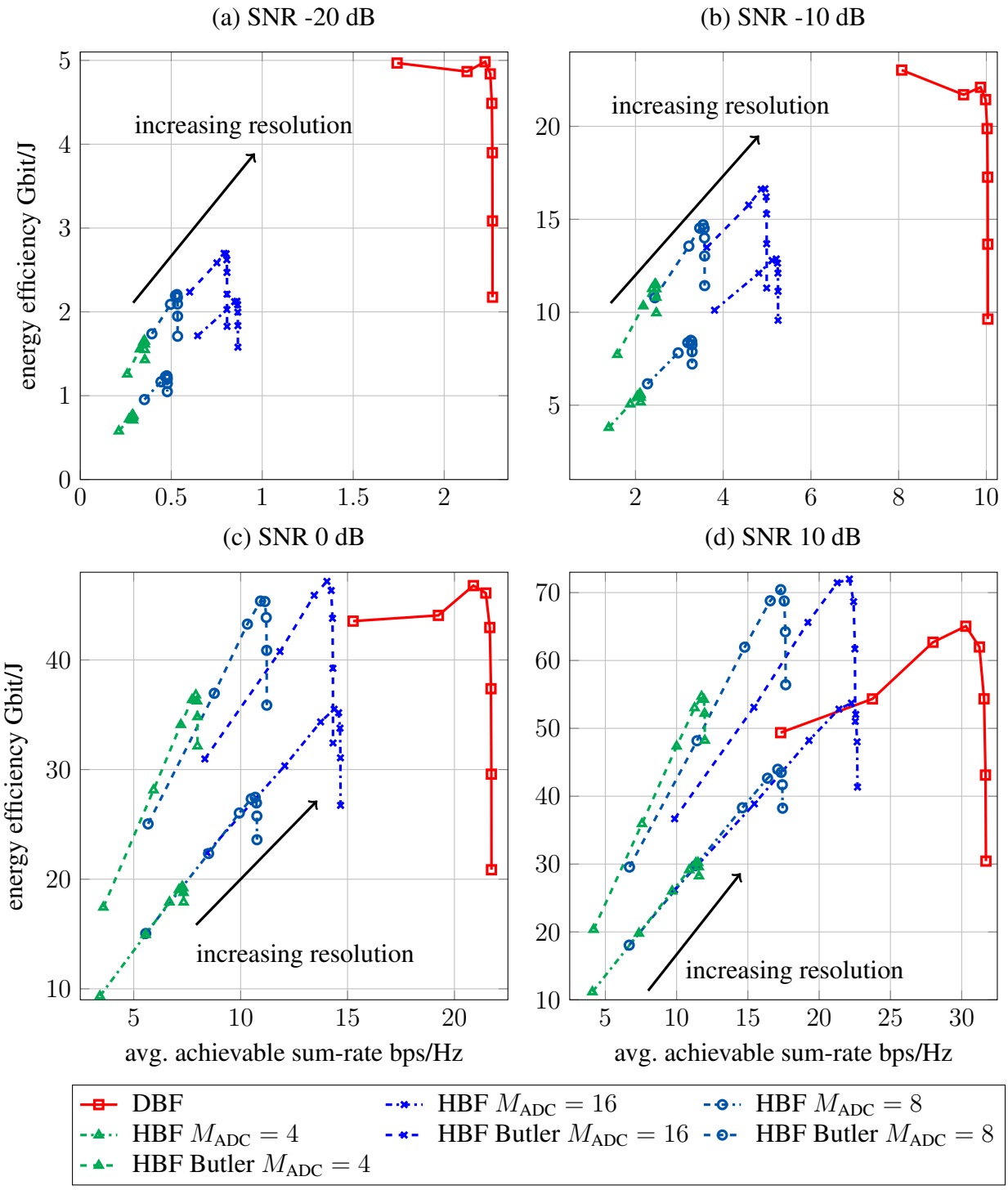

Fig. 3.24. Spectral and energy efficiency of digital beamforming and hybrid beamforming based on phase shifters and a Butler matrix $M = 64$, $U = 4$, $M_{\text{ADC}} \in \{4,\ 8,\ 16\}$ and ADC resolution $b \in \{1,\ \cdots,\ 8\}$ at SNR $\in \{-20\ \text{dB},\ -10\ \text{dB},\ 0\ \text{dB},\ 10\ \text{dB}\}$.

4. Concluding Remarks

In this work we enhanced the models for the quantization error. We showed that including the aspects of correlation has an impact on the performance bounds. We transferred beam design strategies from radar technology to communications technology. The resulting beams were shown to outperform other strategies reported in the literature. To model the beam training, we also showed strategies to design training sequences for the channel quality assessment as well as chose the beam training interval based on available channel measurements. We evaluated the impact of an imperfect AGC adaptation on the system performance. We showed that for low resolution ADCs of 2-3 bit the system is quite robust regarding the imperfections of the AGC.

We showed that in contrast to the currently dominating opinion mmWave systems with digital beamforming can be more energy efficient than hybrid beamforming systems, especially if multiple simultaneous users are considered. It is important to mention that due to the propagation conditions the per antenna SNR regime between -20 and 0 dB is the most likely operation region of future mmWave applications. We also showed the rate regions that are achievable by the different systems. The evaluation of hybrid beamforming systems based on a Butler matrix are more energy efficient compared to hybrid beamforming systems based on adaptive phase shifters. However, compared to the digital system they are less energy efficiency in the practical interesting SNR range.

The full link level evaluation essentially using a standard 3GPP NR system using low resolution ADCs at the receiver showed that it is feasible to use currently defined mmWave systems using OFDM without any change in specification. Our comparison of the link level performance with our theoretical bounds also showed that the theoretical performance sufficiently approximates the performance of a practical system.

Based on the results in this work we think that digital beamforming with low resolution A/D conversion is under the following conditions an attractive system solution: multiple simultaneous users, small range possibly indoor, high carrier frequency. Especially an application like wireless Virtual Reality (VR) might benefit from this approach. This work was limited to receiver aspects, but to implement energy efficient mmWave systems a joint design of the receiver, the transmitter and the corresponding baseband system is necessary.

5. Annex

5.1 Annex A: Derivatives of the Beam Design Objective Function

In this section we derive the cost function of the beam design:

$$f(\boldsymbol{b}) = \left(\int W^p(\boldsymbol{u}) \, ||A(\boldsymbol{u}, \boldsymbol{b})| - D(\boldsymbol{u})|^p \, d\boldsymbol{u} \right)^{\frac{1}{p}}, \tag{5.1}$$

with respect to the design parameters contained in $\boldsymbol{b}$. The outermost part can be described as

$$\nabla f(\boldsymbol{b}) = \frac{1}{p} \left(g_1(\boldsymbol{b})\right)^{\frac{1}{p}-1} \nabla g_1(\boldsymbol{b}) = \frac{1}{p} \left(f(\boldsymbol{b})\right)^{1-p} \nabla g_1(\boldsymbol{b}), \tag{5.2}$$

with

$$g_1(\boldsymbol{b}) = \int W^p(\boldsymbol{u}) \, ||A(\boldsymbol{u}, \boldsymbol{b})| - D(\boldsymbol{u})|^p \, d\boldsymbol{u}. \tag{5.3}$$

The derivation $\nabla g_1(\boldsymbol{b})$ is:

$$\nabla g_1(\boldsymbol{b}) = \int W^p(\boldsymbol{u}) \nabla \left|g_2(\boldsymbol{b})\right|^p d\boldsymbol{u}, \tag{5.4}$$

with $g_2(\boldsymbol{b})$ is equal to $|A(\boldsymbol{u}, \boldsymbol{b})| - D(\boldsymbol{u})$. The gradient of $\nabla \left|g_2(\boldsymbol{b})\right|^p$ is only differentiable if $g_2(\boldsymbol{b})$ is unequal to zero. For this system this means that for this spatial direction $\boldsymbol{u}$ the desired array factor is exactly the same as the actual one. During the iterative, numeric solving of this problem this is very unlikely. However, we mitigated this problem by replacing all possible occurrence of zero by the smallest floating point value. As this is also inside the integral the error due to the mitigation of this problem is minimal.

The result in (5.4) can be further reduced to

$$\nabla g_1(\boldsymbol{b}) = \int W^p(\boldsymbol{u}) p \left|g_2(\boldsymbol{b})\right|^{p-1} \nabla \left|g_2(\boldsymbol{b})\right| d\boldsymbol{u}. \tag{5.5}$$

As state in the preceding paragraph the case $g_2(\boldsymbol{b}) = 0$ can be excluded. Thus, the gradient is defined as:

$$\nabla \left|g_2(\boldsymbol{b})\right| = \nabla \left(\left(g_2(\boldsymbol{b})^2\right)^{\frac{1}{2}} = \frac{1}{2} \left(\left(g_2(\boldsymbol{b})^2\right)^{-\frac{1}{2}} \nabla \left(g_2(\boldsymbol{b})\right)^2 \right. . \tag{5.6}$$

Which is further expanded to

$$\nabla \left|g_2(\boldsymbol{b})\right| = \frac{1}{2} \left|g_2(\boldsymbol{b})\right|^{-1} 2 g_2(\boldsymbol{b}) \nabla g_2(\boldsymbol{b}) = \left|g_2(\boldsymbol{b})\right|^{-1} g_2(\boldsymbol{b}) \nabla g_2(\boldsymbol{b}) = \operatorname{sgn}\left(g_2(\boldsymbol{b})\right) \nabla g_2(\boldsymbol{b}). \tag{5.7}$$

Plugging in the definition of $g_2(\boldsymbol{b})$ this expression is converted to

$$\nabla |g_2(\boldsymbol{b})| = \operatorname{sgn}\left(|A(\boldsymbol{u},\boldsymbol{b})| - D(\boldsymbol{u})\right) \nabla |A(\boldsymbol{u},\boldsymbol{b})| \,. \tag{5.8}$$

For this we again need to ensure that $A(\boldsymbol{u},\boldsymbol{b})$ is not equal to zero. We use the same technique as for $g_2(\boldsymbol{b})$ and replace the values with a very small numeric value for the process of calculating the numeric solution of the problem. As $A(\boldsymbol{u},\boldsymbol{b})$ is complex valued the second part of (5.9) needs to be expanded in the following way:

$$\nabla |A(\boldsymbol{u},\boldsymbol{b})| = \frac{1}{|A(\boldsymbol{u},\boldsymbol{b})|} \Re\left\{A(\boldsymbol{u},\boldsymbol{b})\right\} \nabla \Re\left\{A(\boldsymbol{u},\boldsymbol{b})\right\} + \Im\left\{A(\boldsymbol{u},\boldsymbol{b})\right\} \nabla \Im\left\{A(\boldsymbol{u},\boldsymbol{b})\right\}. \tag{5.9}$$

Combing definitions in (5.2), (5.4) and (5.8) the overall gradient can be expressed as:

$$\begin{aligned} \nabla f(\boldsymbol{b}) = (f(\boldsymbol{b}))^{1-p} \int W^p(\boldsymbol{u}) \, ||A(\boldsymbol{u},\boldsymbol{b})| - D(\boldsymbol{u})|^{p-1} \\ \operatorname{sgn}\left(|A(\boldsymbol{u},\boldsymbol{b})| - D(\boldsymbol{u})\right) \nabla |A(\boldsymbol{u},\boldsymbol{b})| \, d\boldsymbol{u}. \end{aligned} \tag{5.10}$$

Bibliography

[1] 3GPP. Evolved Universal Terrestrial Radio Access (E-UTRA); Multiplexing and channel coding. TS 36.212, 3rd Generation Partnership Project (3GPP), Sep. 2018.

[2] 3GPP. Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation. TS 36.211, 3rd Generation Partnership Project (3GPP), Sep. 2018.

[3] 3GPP. Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer; Measurements. TS 36.214, 3rd Generation Partnership Project (3GPP), Sep. 2018.

[4] 3GPP. NR; Physical channels and modulation. TS 38.211, 3rd Generation Partnership Project (3GPP), Sep. 2018.

[5] 3GPP. NR; Physical layer procedures for control. TS 36.213, 3rd Generation Partnership Project (3GPP), Sep. 2018.

[6] 3GPP. NR; Physical layer procedures for data. TS 36.214, 3rd Generation Partnership Project (3GPP), Sep. 2018.

[7] G. A. Adamidis, I. Vardiambasis, M. Ioannidou, and T. Kapetanakis. Design and Implementation of Single-Layer 4×4 and 8×8 Butler Matrices for Multibeam Antenna Arrays. *International Journal of Antennas and Propagation*, Mar. 2019.

[8] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath. Hybrid Precoding for Millimeter Wave Cellular Systems with Partial Channel Knowledge. In *Information Theory and Applications Workshop (ITA) 2013*, pages 1–5, Feb. 2013.

[9] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath. Channel Estimation and Hybrid Precoding for Millimeter Wave Cellular Systems. *IEEE Journal of Selected Topics in Signal Processing*, 8(5):831–846, Oct. 2014.

[10] A. Alkhateeb, J. Mo, N. Gonzalez-Prelcic, and R. W. Heath. MIMO Precoding and Combining Solutions for Millimeter-Wave Systems. *IEEE Communications Magazine*, 52(12):122–131, Dec. 2014.

[11] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang. What Will 5G Be? *IEEE Journal on Selected Areas in Communications*, 32(6):1065–1082, June 2014.

[12] S. A. Babale, S. K. Abdul Rahim, O. A. Barro, M. Himdi, and M. Khalily. Single Layered 4×4 Butler Matrix Without Phase-Shifters and Crossovers. *IEEE Access*, 6:77289–77298, 2018.

[13] C. A. Balanis. *Antenna Theory: Analysis and Design Third Edition*. John Wiley & Sons, 2005.

[14] C. N. Barati, S. A. Hosseini, S. Rangan, P. Liu, T. Korakis, S. S. Panwar, and T. S. Rappaport. Directional Cell Discovery in Millimeter Wave Cellular Networks. *IEEE Transactions on Wireless Communications*, 14(12):6664–6678, Dec. 2015.

[15] M. Biagini, S. Morosi, F. Argenti, and E. Del Re. Time-Frequency MSE Analysis for Pilot Aided Channel Estimation in OFDM Systems. In *International Conference on Wireless Communications, Vehicular Technology, Information Theory and Aerospace Electronic Systems (VITAE)*, pages 1–5, May 2014.

[16] S. Bilavarn, G. Gogniat, J. . Philippe, and L. Bossuet. Design Space Pruning Through Early Estimations of Area/Delay Tradeoffs for FPGA Implementations. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 25(10):1950–1968, Oct. 2006.

[17] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah. Massive MIMO Systems With Non-Ideal Hardware: Energy Efficiency, Estimation, and Capacity Limits. *IEEE Transactions on Information Theory*, 60(11):7112–7139, Nov. 2014.

[18] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski. Five Disruptive Technology Directions for 5G. *IEEE Communications Magazine*, 52(2):74–80, Feb. 2014.

[19] J. J. Bussgang. Crosscorrelation Functions of Amplitude Distorted Gaussian Signals. Technical Report 216, Massachusetts Institute of Technology, 1952.

[20] J. Butler and R. Lowe. Beam-Forming Matrix Simplifies Design of Electronically Scanned Antenna. *Electronic Design*, 9:170–173, 1961.

[21] C. H. Chan, Y. Zhu, S. W. Sin, S. Pan U, and R. P. Martins. A 5.5mW 6b 5GS/S 4x-Interleaved 3b/cycle SAR ADC in 65nm CMOS. In *IEEE International Solid-State Circuits Conference (ISSCC) 2015*, pages 1–3, Feb. 2015.

[22] J. Chen. *Advanced architectures for efficient mm-Wave CMOS wireless transmitters*. PhD thesis, EECS Department, University of California, Berkeley, May 2014.

[23] J. Choi, J. Mo, and R. W. Heath. Near Maximum-Likelihood Detector and Channel Estimator for Uplink Multiuser Massive MIMO Systems With One-Bit ADCs. *IEEE Transactions on Communications*, 64(5):2005–2018, May 2016.

[24] D. C. Daly, L. C. Fujino, and K. C. Smith. Through the Looking Glass - The 2017 Edition: Trends in Solid-State Circuits from ISSCC. *IEEE Solid-State Circuits Magazine*, 9(1):12–22, 2017.

[25] D. De Donno, J. P. Beltrán, D. Giustiniano, and J. Widmer. Hybrid Analog-Digital Beam Training for mmWave Systems with Low-Resolution RF Phase Shifters. In *IEEE International Conference on Communications Workshops (ICC) 2016*, pages 700–705, May 2016.

[26] D. De Donno, J. Palacios, and J. Widmer. Millimeter-Wave Beam Training Acceleration Through Low-Complexity Hybrid Transceivers. *IEEE Transactions on Wireless Communications*, 16(6):3646–3660, June 2017.

[27] E. Eraslan, B. Daneshrad, and C. Lou. Performance Indicator for MIMO MMSE Receivers in the Presence of Channel Estimation Error. *IEEE Wireless Communications Letters*, 2(2):211–214, Apr. 2013.

[28] W. Feller. The Fundamental Limit Theorems in Probability. *Bull. Amer. Math. Soc.*, 51:800–832, 1945.

[29] G. D. Forney and G. Ungerboeck. Modulation and Coding for Linear Gaussian Channels. *IEEE Transactions on Information Theory*, 44(6):2384–2415, Oct. 1998.

[30] X. Gao, L. Dai, Y. Hu, Z. Wang, and Z. Wang. Matrix Inversion-Less Signal Detection Using SOR Method for Uplink Large-Scale MIMO Systems. In *IEEE Global Communications Conference (GLOBECOM) 2014*, pages 3291–3295, Dec. 2014.

[31] A. Garcia-Rodriguez, V. Venkateswaran, P. Rulikowski, and C. Masouros. Hybrid Analog-Digital Precoding Revisited Under Realistic RF Modeling. *IEEE Wireless Communications Letters*, 5(5):528–531, Oct. 2016.

[32] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath. Capacity Limits of MIMO Channels. *IEEE Journal on Selected Areas in Communications*, 21(5):684–702, June 2003.

[33] Andrea Goldsmith. *Wireless Communications*. Cambridge University Press, 2005.

[34] S. Gunnarsson, M. Bortas, Y. Huang, , L. Van der Perre, and O. Edfors. Lousy Processing Increases Energy Efficiency in Massive MIMO Systems. In *European Conference on Networks and Communications (EuCNC) 2017*, pages 1–5, June 2017.

[35] P. Hoeher, S. Kaiser, and P. Robertson. Two-Dimensional Pilot-Symbol-Aided Channel Estimation by Wiener Filtering. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASP) 1997*, volume 3, pages 1845–1848 vol.3, Apr. 1997.

[36] W. Hong, K. Baek, Y. Geon Kim, Y. Lee, and B. Kim. mmWave Phased-Array with Hemispheric Coverage for 5th Generation Cellular Handsets. In *The 8th European Conference on Antennas and Propagation (EuCAP) 2014*, pages 714–716, Apr. 2014.

[37] IEEE standards. IEEE Standard for Information technology–Telecommunications and information exchange between systems–Local and metropolitan area networks–Specific requirements-Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 3: Enhancements for Very High Throughput in the 60 GHz Band. Technical report, IEEE, Dec. 2012.

[38] S. Jacobsson, G. Durisi, M. Coldrey, U. Gustavsson, and C. Studer. Throughput Analysis of Massive MIMO Uplink With Low-Resolution ADCs. *IEEE Transactions on Wireless Communications*, 16(6):4038–4051, June 2017.

[39] Y. Jin, J. R. Long, and M. Spirito. A 7dB NF 60GHz-Band Millimeter-Wave Transconductance Mixer. In *Radio Frequency Integrated Circuits Symposium (RFIC) 2011*, pages 1–4, June 2011.

[40] S. Ju and T. S. Rappaport. Millimeter-wave Extended NYUSIM Channel Model for Spatial Consistency. In *Global Communications Conf. (GLOBECOM) 2018*, Dec. 2018.

[41] K. Oteri et al. IEEE 802.11-16/1447r1 Further Details on Multi-Stage, Multi-Resolution Beamforming Training in 802.11ay, Nov. 2016.

[42] J. Kim and A. F. Molisch. Fast Millimeter-Wave Beam Training with Receive Beamforming. *Journal of Communications and Networks*, 16(5):512–522, Oct. 2014.

[43] M. Kobayashi, H. Motozuka, T. Urushihara, N. Shirakata, and K. Takinami. IEEE 802.11ad/WiGig based Millimeter-Wave Small Cell Systems with Adjacent Channel Interference Suppression. In *IEEE Conference on Standards for Communications and Networking (CSCN) 2016*, pages 1–5, Oct. 2016.

[44] V. K. Kodavalla. IP Gate Count Estimation Methodology during Micro-Architecture Phase.

[45] L. Kong. *Energy-Efficient 60 GHz Phased-Array Design for Multi-Gb/s Communication Systems*. PhD thesis, EECS Department, University of California, Berkeley, Dec. 2014.

[46] C. Kuo, H. Kuo, H. Chuang, C. Chen, and T. Huang. A High-Isolation 60GHz CMOS Transmit/Receive Switch. In *IEEE Radio Frequency Integrated Circuits Symposium (RFIC) 2011*, pages 1–4, June 2011.

[47] B. Lehmeyer. *Receiver and Transmitter Topologies*. PhD thesis, Technical University Munich, June 2018.

[48] Y. Li, C. Tao, G. Seco-Granados, A. Mezghani, A. L. Swindlehurst, and L. Liu. Channel Estimation and Performance Analysis of One-Bit Massive MIMO Systems. *IEEE Transactions on Signal Processing*, 65(15):4075–4089, Aug. 2017.

[49] R. J. Mailloux. *Phased Array Antenna Handbook*. Artech House, 2005.

[50] C. Marcu. *LO Generation and Distribution for 60 GHz Phased Array Transceivers*. PhD thesis, EECS Department, University of California, Berkeley, Dec. 2011.

[51] J. Max. Quantizing for Minimum Distortion. *IRE Transactions on Information Theory*, 6(1):7–12, Mar. 1960.

[52] A. Mezghani and J. A. Nossek. On Ultra-Wideband MIMO Systems with 1-bit Quantized Outputs: Performance Analysis and Input Optimization. In *IEEE International Symposium on Information Theory (ISIT) 2007*, pages 1286–1289, June 2007.

[53] A. Mezghani and J. A. Nossek. Capacity Lower Bound of MIMO Channels with Output Quantization and Correlated Noise. In *IEEE International Symposium on Information Theory (ISIT) 2012*, Jul. 2012.

[54] S. Mitaim and B. Kosko. Adaptive Stochastic Resonance. *Proc. IEEE*, 86(11):2152–2183, Nov. 1998.

[55] mmMAGIC. Deliverable D2.1 Measurement Campaigns and Initial Channel Models for Preferred Suitable Frequency Ranges, 2016.

[56] mmMAGIC. Deliverable D2.2 Measurement Results and Final mmMAGIC Channel Models, 2017.

[57] J. Mo, A. Alkhateeb, S. Abu-Surra, and R. W. Heath. Hybrid Architectures With Few-Bit ADC Receivers: Achievable Rates and Energy-Rate Tradeoffs. *IEEE Transactions on Wireless Communications*, 16(4):2274–2287, Apr. 2017.

[58] J. Mo, P. Schniter, N. G. Prelcic, and R. W. Heath. Channel Estimation in Millimeter Wave MIMO Systems with One-Bit Quantization. In *Asilomar Conference on Signals, Systems and Computers 2014*, pages 957–961, Nov. 2014.

[59] A. F. Molisch, V. V. Ratnam, S. Han, Z. Li, S. L. H. Nguyen, L. Li, and K. Haneda. Hybrid Beamforming for Massive MIMO: A Survey. *IEEE Communications Magazine*, 55(9):134–141, Sep. 2017.

[60] A. F. Molisch, M. Steinbauer, M. Toeltsch, E. Bonek, and R. S. Thoma. Capacity of MIMO Systems Based on Measured Wireless Channels. *IEEE Journal on Selected Areas in Communications*, 20(3):561–569, Apr. 2002.

[61] C. Mollén, J. Choi, E. G. Larsson, and R. W. Heath. Uplink Performance of Wideband Massive MIMO With One-Bit ADCs. *IEEE Transactions on Wireless Communications*, 16(1):87–100, Jan. 2017.

[62] A. F. Morabito, A. Massa, P. Rocca, and T. Isernia. An Effective Approach to the Synthesis of Phase-Only Reconfigurable Linear Arrays. *IEEE Transactions on Antennas and Propagation*, 60(8):3622–3631, Aug. 2012.

[63] B. Murmann. ADC Performance Survey 1997-2017, 2017.

[64] A. M. Niknejad, S. Thyagarajan, E. Alon, Y. Wang, and C. Hull. A Circuit Designer's Guide to 5G mm-Wave. In *IEEE Custom Integrated Circuits Conference (CICC) 2015*, pages 1–8, Sep. 2015.

[65] A. Oncu, B. B. M. W. Badalawa, and M. Fujishima. 60 GHz-Pulse Detector Based on CMOS Nonlinear Amplifier. In *IEEE Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems (SiRF) 2009*, pages 1–4, Jan. 2009.

[66] O. Orhan, E. Erkip, and S. Rangan. Low Power Analog-to-Digital Conversion in Millimeter Wave Systems: Impact of Resolution and Bandwidth on Performance. In *2015 Information Theory and Applications Workshop (ITA)*, pages 191–198, Feb. 2015.

[67] P. Zetterberg. Experimental Investigation of TDD Reciprocity-Based Zero-Forcing Transmit Precoding. *EURASIP Journal on Advances in Signal Processing*, 2011(1):137541, Dec. 2010.

[68] J. Palacios, D. De Donno, D. Giustiniano, and J. Widmer. Speeding Up mmWave Beam Training through Low-Complexity Hybrid Transceivers. In *IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC) 2016*, pages 1–7, Sep. 2016.

[69] Peregrine Semiconductor. *UltraCMOS SPDT RF Switch, 9 kHz - 60 GHz*, 2016.

[70] R. Price. A Useful Theorem for Nonlinear Devices Having Gaussian Inputs. *IRE Transactions on Information Theory*, 4(2):69–72, June 1958.

[71] Z. Quan, J. Liu, and Y. Zakharov. FPGA Implementation of DCD Based CDMA Multiuser Detector. In *IEEE International Conference on Digital Signal Processing (DSP) 2007*, pages 319–322, Jul. 2007.

[72] Z. Quan, J. Liu, and Y. Zakharov. FPGA Design of Box-constrained MIMO Detector. In *IEEE International Conference on Communications (ICC) 2009*, pages 2866–2870, 2009.

[73] R. Ratasuk, B. Vejlgaard, N. Mangalvedhe, and A. Ghosh. NB-IoT System for M2M Communication. In *Wireless Communications and Networking Conference (WCNC) 2016*, pages 1–5, Apr. 2016.

[74] P. Reynaert and M. Steyaert. RF Power Amplifiers for Mobile Communications, 2006.

[75] W. Roh, J. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar. Millimeter-Wave Beamforming as an Enabling Technology for 5G Cellular Communications: Theoretical Feasibility and Prototype Results. *IEEE Communications Magazine*, 52(2):106–113, Feb. 2014.

[76] K. Roth, J. Munir, A. Mezghani, and J. A. Nossek. Covariance Based Signal Parameter Estimation of Coarse Quantized Signals. In *IEEE International Conference on Digital Signal Processing (DSP) 2015*, pages 19–23, Jul. 2015.

[77] K. Roth and J. A. Nossek. Achievable Rate and Energy Efficiency of Hybrid and Digital Beamforming Receivers With Low Resolution ADC. *IEEE Journal on Selected Areas in Communications*, 35(9):2056–2068, Sep. 2017.

[78] K. Roth and J. A. Nossek. Arbitrary Beam Synthesis of Hybrid Beamforming Systems for Beam Training. *IEEE Wireless Communications Letters*, 6(6):714–717, Dec. 2017.

[79] K. Roth and J. A. Nossek. MAC Rate Region of Low Resolution ADC Digital and Hybrid Beamforming Massive MIMO UL. In *IEEE International Conference on Communications Workshops (ICC Workshops) 2017*, pages 1227–1232, May 2017.

[80] K. Roth and J. A. Nossek. Robust Massive MIMO Equilization for mmWave Systems with Low Resolution ADCs. In *IEEE Wireless Communications and Networking Conference Workshops (WCNCW) 2018*, pages 113–118, Apr. 2018.

[81] K. Roth, H. Pirzadeh, A. L. Swindlehurst, and J. A. Nossek. A Comparison of Hybrid Beamforming and Digital Beamforming With Low-Resolution ADCs for Multiple Users and Imperfect CSI. *IEEE Journal of Selected Topics in Signal Processing*, 12(3):484–498, June 2018.

[82] K. Roth, H. Pirzadeh, A. L. Swindlehurst, and J. A. Nossek. Are the Data Rates Predicted by the Analytic Analysis of Receivers with Low Resolution ADCs Achievable? In *European Conference on Networks and Communications (EuCNC) 2018*, pages 378–9, June 2018.

[83] M. Rupp, S. Schwarz, and M. Taranetz. *The Vienna LTE-Advanced Simulators: Up and Downlink, Link and System Level Simulation*, page 187. Springer Publishing Company, Incorporated, 1st edition, 2016.

[84] M. K. Samimi and T. S. Rappaport. 3-D Millimeter-Wave Statistical Channel Model for 5G Wireless System Design. *IEEE Transactions on Microwave Theory and Techniques*, 64(7):2207–2225, Jul. 2016.

[85] K. Scheir, S. Bronckers, J. Borremans, P. Wambacq, and Y. Rolain. A 52 GHz Phased-Array Receiver Front-End in 90 nm Digital CMOS. *IEEE Journal of Solid-State Circuits*, 43(12):2651–2659, Dec. 2008.

[86] D. P. Scholnik. A Parameterized Pattern-Error Objective for Large-Scale Phase-Only Array Pattern Design. *IEEE Transactions on Antennas and Propagation*, 64(1):89–98, Jan. 2016.

[87] Y. Shang, D. Cai, W. Fei, H. Yu, and J. Ren. An 8mW Ultra Low Power 60GHz Direct-conversion Receiver with 55dB Gain and 4.9dB Noise Figure in 65nm CMOS. In *IEEE International Symposium on Radio-Frequency Integration Technology (RFIT) 2012*, pages 47–49, Nov. 2012.

[88] C. E. Shannon. A Mathematical Theory of Communication. *Bell Systems Technical Journal*, 27:623–656, 1948.

[89] C. Studer, M. Wenk, and A. Burg. MIMO Transmission with Residual Transmit-RF Impairments. In *International ITG Workshop on Smart Antennas (WSA) 2010*, pages 189–196, Feb. 2010.

[90] S. Sun, T. S. Rappaport, R. W. Heath, A. Nix, and S. Rangan. MIMO for Millimeter-Wave Wireless Communications: Beamforming, Spatial Multiplexing, or Both? *IEEE Communications Magazine*, 52(12):110–121, Dec. 2014.

[91] A. Tomkins, A. Poon, E. Juntunen, A. El-Gabaly, G. Temkine, Y. To, C. Farnsworth, A. Tabibiazar, M. Fakharzadeh, S. Jafarlou, A. Abdellatif, H. Tawfik, B. Lynch, M. Tazlauanu, and R. Glibbery. A 60 GHz, 802.11ad/WiGig-Compliant Transceiver for Infrastructure and Mobile Applications in 130 nm SiGe BiCMOS. *IEEE Journal of Solid-State Circuits*, 50(10):2239–2255, Oct. 2015.

[92] C. Tseng, C. Chen, and T. Chu. A Low-Cost 60-GHz Switched-Beam Patch Antenna Array With Butler Matrix Network. *IEEE Antennas and Wireless Propagation Letters*, 7:432–435, 2008.

[93] V. Venkateswaran and A.-J. van der Veen. Analog Beamforming in MIMO Communications With Phase Shift Networks and Online Channel Estimation. *IEEE Transactions Signal Processing*, 58(8):4131–4143, Aug. 2010.

[94] S. Wang, Y. Li, and J. Wang. Multiuser Detection in Massive Spatial Modulation MIMO With Low-Resolution ADCs. *IEEE Transactions on Wireless Communications*, 14(4):2156–2168, Apr. 2015.

[95] Y. Wang, Z. Shi, L. Huang, Z. Yu, and C. Cao. An Extension of Spatial Channel Model with Spatial Consistency. In *IEEE Vehicular Technology Conference (VTC-Fall) 2016*, pages 1–5, Sep. 2016.

[96] M. Wenk. *MIMO-OFDM Testbed: Challenges, Implementations, and Measurement Results*. Series in microelectronics. ETH, 2010.

[97] M. Wu, C. Dick, J. R. Cavallaro, and C. Studer. High-Throughput Data Detection for Massive MU-MIMO-OFDM Using Coordinate Descent. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 63(12):2357–2367, Dec. 2016.

[98] B. Xu, Y. Zhou, and Y. Chiu. A 23mW 24GS/s 6b Time-Interleaved Hybrid Two-Step ADC in 28nm CMOS. In *Symposium on VLSI Circuits (VLSI-Circuits) 2016*, pages 1–2, June 2016.

[99] S. Yang and L. Hanzo. Fifty Years of MIMO Detection: The Road to Large-Scale MIMOs. *IEEE Communications Surveys Tutorials*, 17(4):1941–1988, 2015.

[100] T. Yoo and A. Goldsmith. Capacity and Power Allocation for Fading MIMO Channels with Channel Estimation Error. *IEEE Transactions on Information Theory*, 52(5):2203–2214, May 2006.

Acronyms

3GPP	Third Generation Partnership Project
5G	Fifth Generation mobile communicaiton system
A/D	Analog/Digital
ADC	Analog-to-Digital-Converter
AGC	Automatic Gain Control
AP	Access Point
AQNM	Additive Quantization Noise Model
BB	Baseband
BER	Bit Error Ratio
BLER	Block Error Ratio
CDF	Cumulative Distribution Function
CDM	Code Division Multiplexing
CDMA	Code Division Multiple Access
CIR	Channel Impulse Response
CMOS	Complementary Metal-Oxide-Semiconductor
CP	Cyclic Prefix
CR	Code Rate
CRC	Cyclic Redundancy Check
CS	Cyclic Shift
CSI	Channel State Information
CSIR	Channel State Information at the Receiver
CSIT	Channel State Information at the Transmitter
DAC	Digital-to-Analog-Converter
DBF	Digital Beamforming
DCD	Dichotomous Coordinate Descent
DFT	Discrete Fourier Transformation
DFT-s-OFDM	DFT-spread-OFDM
DL	Downlink
DMRS	Demodulation Reference Signals
DOF	Degrees of Freedom
EE	Energy Efficiency
EM	Expectation Maximization
ENOB	Effective Number of Bits

EVM	Error Vector Magnitude
FDM	Frequency Division Multiplexing
FF	Flip Flop
FFT	Fast Fourier Transformation
HBF	Hybrid Beamforming
I	Inphase
IF	Intermediate Frequency
IFFT	Inverse Fast Fourier Transformation
ISI	Inter Symbol Interference
LA	Limiting Amplifier
LNA	Low Noise Amplifier
LO	Local Oscillator
LOS	Line Of Sight
LTE	Long Term Evolution
MAC	Multiple Access Channel
MIMO	Multiple Input Multiple Output
MINLP	Mixed Integer Non-Linear Programing
ML	Maximum Likelihood
MMSE	Minimum Mean Square Error
mmWave	millimeter Wave
MSE	Mean Square Error
MU-MIMO	MultiUser - Multiple Input Multiple Output
MUX	Multiplexer
NAND	NOT-AND logical gate
NLOS	Non Line Of Sight
NLP	Non-Linear Programing
NR	New Radio
OFDM	Orthogonal Frequency Division Multiplexing
PA	Power Amplifier
PCB	Printed Circuit Board
PDF	Probability Density Function
PDP	Power Delay Profile
PRBS	Pseudo Random Binary Sequence
PS	Phase Shifter
Q	Quadrature
QAM	Quadrature Amplitude Modulation
QPSK	Quaternary Phase Shift Keying
R&D	Research and Development
RE	Resource Element
RF	Radio Frequency
RFE	Radio Front-End
RMS	Root Mean Square
RSRP	Reference Signal Received Power
Rx	Receiver

SC	Sub-Carrier
SE	Spectral Efficiency
SIC	Successive Interference Cancellation
SIMO	Single Input Multiple Output
SISO	Single Input Single Output
SNR	Signal-to-Noise-Ratio
SQNR	Signal-to-Quantization-Noise-Ratio
SVD	Singular Value Decomposition
Tx	Transmitter
UE	User Equipment
UL	Uplink
ULA	Uniform Linear Array
VGA	Variable Gain Amplifier
VR	Virtual Reality
WiGig	IEEE 802.11ad standard

Glossary

$\boldsymbol{A}$	Generic matrix
a	Generic scalar random variable
$\boldsymbol{a}$	Generic column vector
$\boldsymbol{A}^H$	Hermitian of generic matrix $\boldsymbol{A}$
$[\boldsymbol{A}]_i$	Column vector consiting of all elements of the ith column of matrix $\boldsymbol{A}$
$[\boldsymbol{A}]_{i,j}$	Element in ith row and jth column of matrix $\boldsymbol{A}$
$[\boldsymbol{a}]_i$	Element i of column vector $\boldsymbol{a}$
$\boldsymbol{A}^{-1}$	Inverse of generic matrix $\boldsymbol{A}$
$A(\boldsymbol{u}, \boldsymbol{b})$	Antenna Array factor dependent on spatial direction $\boldsymbol{u}$ and beamforming vector $\boldsymbol{b}$
$\boldsymbol{A}^T$	Transpose of generic matrix $\boldsymbol{A}$
$\boldsymbol{b}$	Beamforming vector
$\text{co}(\cdot)$	Convex hull operator
$a(n) * b(n)$	Linear convolution of a(n) and b(n)
$\cosh(x)$	Hyperbolic cosine equal to $\frac{e^{-x}+e^{x}}{2}$
$D(\boldsymbol{u})$	Desired form of the beam dependent on spatial direction $\boldsymbol{u}$
$\det(\boldsymbol{A})$	Determinant of matrix $\boldsymbol{A}$
d_{max}	Maximum allowed distance for the channel quality estimation in dB
$\mathbb{E}$	Expectation operator
$\boldsymbol{e}_i$	Column vector with only one entry equal to 1, the position is defined by the subscript i
$\boldsymbol{\eta}$	Noise vector
$\eta[n]$	Scalar noise at discrete time instant $[n]$
$\boldsymbol{A} \circ \boldsymbol{B}$	Hadamard (element-wise) product of matrix $\boldsymbol{A}$ and $\boldsymbol{B}$
$\boldsymbol{h}$	Channel vector
j	Imaginary unit
$\boldsymbol{A} \otimes \boldsymbol{B}$	Kronecker-product of matrix $\boldsymbol{A}$ and $\boldsymbol{B}$
$\wedge$	Logic AND operation

$\ln$	Base e logarithm
$\log$	Base 10 logarithm
$\vee$	Logic OR operation
M	Number of antennas
M_C	Number of antennas per subarray
M_{ADC}	Number of ADC chains
M_{DAC}	Number of DAC chains
M_T	Number of transmit antennas
μ	Mean of a scalar random variable
N	Sequence or block length
$[n]$	Discrete time instant
$\neg$	Logic negation
$\mathbf{1}_N$	All ones column vector of dimension $N \times 1$
$Q(\cdot)$	Quantization function
$\text{TF}(\cdot)$	System transformation function giving $\boldsymbol{T}$ and $\boldsymbol{R}_{\hat{\boldsymbol{r}}_q \hat{\boldsymbol{r}}_q}$ out of the receive signal covariance matrix $\boldsymbol{R}_{\boldsymbol{yy}}$ and the quantization functions $Q_m(\cdot)$
$\boldsymbol{r}$	Receive signal after quantization
$\boldsymbol{R}_{\hat{\boldsymbol{r}}_q \hat{\boldsymbol{r}}_q}$	Effective quantization error covariance matrix
$\lfloor x \rfloor$	Rounding to nearest integer smaller or equal for each element in x.
$\lceil x \rceil$	Rounding to nearest integer larger or equal for each element in x.
$\hat{s}$	Channel quality measure
σ^2	Variance of a scalar random variable
$\boldsymbol{T}$	Transformation matrix for the received signal relating signal after quantization to before.
γ	Signal-to-Noise-Ratio
$\text{tr}(\boldsymbol{A})$	Trace of matrix $\boldsymbol{A}$
U	Number of users
$\boldsymbol{u}$	Spatial direction combining the azimuth and elevation angle
$\boldsymbol{w}$	Analog combining vector of one subarray at the receiver
$\boldsymbol{W}$	Analog combining matrix at the receiver
$x[n]$	Transmit signal scalar at discrete time instant $[n]$
$\boldsymbol{y}$	Receive signal
$y[n]$	Receive signal scalar at discrete time instant $[n]$
$\mathbf{0}$	All zeros vector or matrix

List of Figures

List of Tables

List of Algorithms