Title

A Comparison of Hybrid Beamforming and Digital Beamforming With Low-Resolution ADCs for Multiple Users and Imperfect CSI

Journal

IEEE Journal of Selected Topics in Signal Processing, 12(3)

ISSN

1932-4553

Authors

Roth, Kilian
Pirzadeh, Hessam
Swindlehurst, A Lee
et al.

Publication Date

2018

DOI

10.1109/jstsp.2018.2813973

A Comparison of Hybrid Beamforming and Digital Beamforming with Low-Resolution ADCs for Multiple Users and Imperfect CSI

Kilian Roth, *Member, IEEE,* Hessam Pirzadeh, *Member, IEEE,* A. Lee Swindlehurst *Fellow, IEEE,* Josef A. Nossek, *Life Fellow, IEEE*

***Abstract*—For 5G it will be important to leverage the available millimeter wave spectrum. To achieve an approximately omni-directional coverage with a similar effective antenna aperture compared to state of the art cellular systems, an antenna array is required at both the mobile and basestation. Due to the large bandwidth and inefficient amplifiers available in CMOS for mmWave, the analog front-end of the receiver with a large number of antennas becomes especially power hungry. Two main solutions exist to reduce the power consumption: hybrid beam forming and digital beam forming with low resolution Analog to Digital Converters (ADCs). In this work we compare the spectral and energy efficiency of both systems under practical system constraints. We consider the effects of channel estimation, transmitter impairments and multiple simultaneous users. Our power consumption model considers components reported in literature at 60 GHz. In contrast to many other works we also consider the correlation of the quantization error, and generalize the modeling of it to non-uniform quantizers and different quantizers at each antenna. The result shows that as the SNR gets larger the ADC resolution achieving the optimal energy efficiency gets also larger. The energy efficiency peaks for 5 bit resolution at high SNR, since due to other limiting factors the achievable rate almost saturates at this resolution. We also show that in the multi-user scenario digital beamforming is in any case more energy efficient than hybrid beamforming. In addition we show that if different ADC resolutions are used we can achieve any desired trade-offs between power consumption and rate close to those achieved with only one ADC resolution.**

***Index Terms*—Wireless communication, millimeter Wave, low resolution Analog-to-Digital-Converter (ADC), hybrid beamforming**

I. Introduction

The use of the available bandwidth in the frequency range of 6 to 100 GHz is considered to be an essential part of the next generation mobile broadband standard 5G [1]. Due to the propagation condition 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 spectrum and the availability of consumer grade systems at mmWave frequencies has 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 [1], [2]. The support of a large number of antennas at the mobile and base station 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 an omnidirectional coverage.

A. Related Work

Current Long Term Evolution (LTE) systems have a limited amount of antennas at the base and mobile stations. Since the bandwidth is relatively narrow, the power consumption of a receiver Radio Frequency (RF) chain with a high resolution ADC at each antenna is still feasible. For future mmWave mobile broadband systems, a much larger bandwidth [3] and a much large number of antennas are being considered [1]. The survey in [4] shows that ADCs with a high sampling frequency and a standard number of effective bits of resolution (6-10) consume a considerable amount of power. Consequently, the power consumption of the ADC can be considered as the bottleneck of the receiver [5].

The use of a large antenna array combined with a large bandwidth is a huge challenge for the hardware implementation; essentially the power consumption will limit the design space. At the moment, analog/hybrid beamforming is considered as a possible solution to reduce the power consumption. Analog or hybrid beamforming systems strongly depend on the calibration of the analog components. Another major disadvantage is the large overhead associated with the alignment of the Tx and Rx beams of the base and mobile station. Specifically, if high gain is needed, the beamwidth is small and thus the

K. Roth is with Next Generation and Standards, Intel Deutschland GmbH, Neubiberg 85579, Germany (email: kilian.roth@intel.com)

K. Roth and J. A. Nossek are with the Department of Electrical and Computer Engineering, Technical University Munich, Munich 80290, Germany (email: kilian.roth@tum.de; josef.a.nossek@tum.de)

J. A. Nossek is also with Department of Teleinformatics Engineering, Federal University of Ceara, 60020-180 Fortaleza, Brazil

H. Pirzadeh and A. L. Swindlehurst are with the Center for Pervasive Communications and Computing, University of California, Irvine, CA 92697 USA. (e-mail: hpirzade@uci.edu; swindle@uci.edu).

A. L. Swindlehurst is also with Institute for Advanced Study, Technical University of Munich, Munchen 80333, Germany

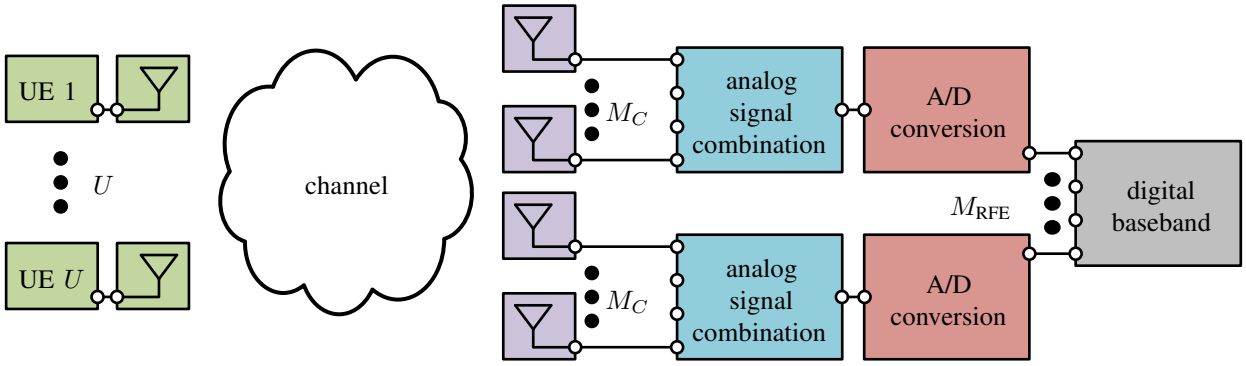

Fig. 1. System Model with U UEs with 1 antenna and M_C antennas at each of the M_{RFE} RF chains at the basestations. Number of receive Antennas M_R is equal to $M_C \times M_{\text{RFE}}$

acquisition and constant alignment of the optimal beams in a dynamic environment is very challenging [6], [7], [8].

The idea of hybrid beamforming is based on the concept of phased array antennas commonly used in radar applications [9]. Due to the reduced power consumption, it is also seen as a possible solution for mmWave mobile broadband communication[10]. If the phased array approach is combined with digital beamforming, the phased array approach might also be feasible for non-static or quasi-static scenarios. In [11], it was shown that considering the inefficiency of mmWave amplifiers and the high insertion loss of RF phase shifters, it is better to perform the phase shifting in the baseband. The power consumption associated with both cases is comparable, as long as the number of antennas per RF-chain remains relatively small.

Another option to reduce the power consumption while keeping the number of antennas constant is to reduce the power consumption of the ADCs by reducing their resolution. This can also be combined with hybrid beamforming. Some of these evaluations consider only the extreme case of 1-bit quantization [8], [12], [5], [13]. In [14], [15] the Analog/Digital (A/D) conversion is modeled as a linear stochastic process. Low resolution A/D conversion combined with Orthogonal Frequency Domain Multiplexing (OFDM) in an uplink scenario are considered in [16], [17].

In [18], [19] hybrid beamforming with low resolution A/D conversion was considered. The energy efficiency / spectral efficiency trade-off of fully-connected hybrid and digital beamforming with low resolution ADCs is assessed in [19]. But in contrast as shown in the system diagram in Fig. 1, we consider a hybrid beamforming system that has exclusive antennas per RF-chain (aka. sub-array hybrid beamforming). In this work we concentrated on effects of the hardware constraints at the receiver, thus we assumed the transmitter to be ideal. In [19], a fully-connected hybrid beamforming system is used, which has a large additional overhead associated with an increased number of phase shifters and larger power combiners. Also in this case additional amplifiers to compensate for the insertion-loss of the RF phase shifters and combiners are required. In [20], analog beamforming is compared with digital beamforming in terms of power efficiency.

The authors of [21], [22] both analyzed the effect of imperfect channel knowledge on the achievable rate. The channel estimation error is treated as additional noise added to the system. We will use a similar model to include the channel estimation error into our analysis. Since we we have a system involving multiple user with different receive power, we treat the effect of each users separately.

B. Contribution

The contribution of this work can be summarized in the following bulletpoints:

- Achievable rate analysis for digital and hybrid beamforming systems with low resolution ADCs in a multi-user, multipath scenario. In addition the effects of transmitter impairments, channel estimation errors and having a mixed ADC resolutions are considered.
- Analyzing the channel estimation error considering the reference signal patterns already agreed upon for 3GPP NR (aka. 5G).
- Showing the energy efficiency - spectral efficiency trade-off considering the power consumption of the receiver RF front-end.
- Generalizing the Additive Quantization Noise Model (AQNM) to include the effects for quantization error correlation, non-uniform quantization and different ADCs at each antenna.

C. Notation

Throughout the paper we use boldface lower and upper case letters to represent column vectors and matrices. The term $a_{m,l}$ is the element on row m and column l of matrix $\boldsymbol{A}$ and a_m is the mth element of vector $\boldsymbol{a}$. The expressions $\boldsymbol{A}^*$, $\boldsymbol{A}^T$, $\boldsymbol{A}^H$, and $\boldsymbol{A}^{-1}$ represent the complex conjugate, the transpose, the Hermitian, and the inverse of the matrix $\boldsymbol{A}$. The symbol $\boldsymbol{R_{ab}}$ is the correlation matrix of vector $\boldsymbol{a}$ and $\boldsymbol{b}$ defined as $\mathbb{E}[\boldsymbol{ab}^H]$. The Discrete Fourier Transformation (DFT) $\mathcal{F}(\cdot)$ and its inverse $\mathcal{F}^{-1}(\cdot)$ and the Fourier transformation $\mathscr{F}\{\cdot\}$ and its inverse $\mathscr{F}^{-1}\{\cdot\}$ are also used.

II. Signal Model

The system model in Fig. 1 gives a general overview of both investigated systems. For $M_C = 1$ the block analog signal combination is just connecting the input to the output. For

$M_C > 1$ this block contains an analog phase shifter for each signal followed by a power combiner.

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 basestation, 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_R receive antennas. The receive signal $\boldsymbol{y}[n]$ is defined as

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

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 basestation. The transmitter impairments $\boldsymbol{\eta}_u[n]$ are modeled as circular symmetric complex Gaussian noise with zero mean and covariance equal to σ^2_{EVM}. Including the transmit power P_u, this is the classical Error Vector Magnitude (EVM) definition only considering transmitter impairments [23].

Since all noise contributions are Gaussian we can combine them to form a combined noise $\boldsymbol{\eta}'_R[n]$ equal to

$$\boldsymbol{\eta}'_R[n] = \sum_{u=1}^{U} \sqrt{P_u} \sum_{l=0}^{L_u} \boldsymbol{H}_u[l]\boldsymbol{\eta}_u[n-l] + \boldsymbol{\eta}_R[n]. \tag{2}$$

The receive signal is then reduced to

$$\boldsymbol{y}[n] = \sum_{u=1}^{U} \sqrt{P_u} \sum_{l=0}^{L_u} \boldsymbol{H}_u[l]\boldsymbol{x}_u[n-l] + \boldsymbol{\eta}'_R[n]. \tag{3}$$

We restrict the system to have M_C antennas exclusively connected to one RF front-end chain (see Fig. 1). Therefore, the matrix modeling the analog combining at the receiver $\boldsymbol{W}_R$ has the form

$$\boldsymbol{W}_R = \begin{bmatrix} \boldsymbol{w}_R^1 & \boldsymbol{0}_{M_C} & \dots & \boldsymbol{0}_{M_C} \\ \boldsymbol{0}_{M_C} & \boldsymbol{w}_R^2 & \ddots & \boldsymbol{0}_{M_C} \\ \vdots & \ddots & \ddots & \vdots \\ \boldsymbol{0}_{M_C} & \dots & \boldsymbol{0}_{M_C} & \boldsymbol{w}_R^{M_{RFE}} \end{bmatrix} \in \mathbb{C}^{M_R \times M_{RFE}}, \tag{4}$$

where the vector $\boldsymbol{w}_R^i$ is the analog beamforming vector of the ith RF chain. We also restrict our evaluation to the case where each RF chain is connected to the same number of antennas M_C. The vectors $\boldsymbol{w}_R^i$ and $\boldsymbol{0}_{M_C}$ have dimension M_C. The receiver signal $y_C[n]$ after the analog combining $\boldsymbol{y}_C[n]$ is then

$$\boldsymbol{y}_C[n] = \boldsymbol{W}_R^H \boldsymbol{y}[n]. \tag{5}$$

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

For the case of Digital BeamForming (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 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 RF chains, since it takes a non-negligible amount of time to perform the switching. Furthermore, we do not expect the received power to be different on average for different antennas, so allocating the M_h high resolution ADCs to an arbitrary subset of the antennas is a reasonable approach.

A. Channel Model

The measurements in [24] show that for channels at 60 GHz, an exponential Power Delay Profile (PDP) sufficiently approximates a real world scenario

$$\boldsymbol{H}[l] = \frac{1}{\sqrt{M_T}} \alpha(l) \boldsymbol{a}_r(\phi_r(l)) \boldsymbol{a}_t^T(\phi_t(l)). \tag{6}$$

The phase shift between the signal at adjacent antenna elements at the receiver and transmitter $\phi_r(l)$ and $\phi_t(l)$ of path l depend on the angle of arrival $\theta_r(l)$ and departure $\theta_t(l)$

$$\boldsymbol{a}_r^T(\phi_r(l)) = \left[1, e^{j\phi_r(l)}, e^{j2\phi_r(l)}, \cdots, e^{j(M_r-1)\phi_r(l)}\right]. \tag{7}$$

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 circular symmetric Gaussian distributed with zero mean and a variance defined according to

$$v_l = \mathbb{E}\left[|\alpha(l)|^2\right] = e^{-\beta l}. \tag{8}$$

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 elements of the $L \times 1$ vector of variances $\boldsymbol{v}$ are non-zero. We will normalize the variance vector as follows:

$$\boldsymbol{v}_n = \frac{\boldsymbol{v}}{||\boldsymbol{v}||^2}. \tag{9}$$

The Signal to Noise Ratio (SNR) γ_u per user u is defined as as

$$\gamma_u = \frac{P_u \ \mathbb{E}\left[\left|\left|\sum_{l=0}^{L_u} \boldsymbol{H}_u[l]\boldsymbol{x}_u[n-l]\right|\right|_2^2\right]}{\mathbb{E}[||\boldsymbol{\eta}_R[n]||_2^2]}. \tag{10}$$

This formula describes the average SNR at each antenna. It is important to note that the expectation takes the realization of the channel and realizations of $\boldsymbol{x}_i[n]$ into account.

B. Analytic MSE of frequency domain channel estimation with time-frequency interpolation

Assuming perfect synchronization of the timing and carrier frequency, the OFDM receive signal $Y_{(k,\ell,m)}$ of subcarrier k, OFDM symbol ℓ and antenna m can be written as

$$Y_{(k,\ell,m)} = H_{(k,\ell,m)} X_{(k,\ell,m)} + \eta_{(k,\ell,m)}, \tag{11}$$

where we assume that the Channel Impulse Response (CIR) is shorter than the cyclic prefix, and $H_{(k,\ell,m)}$, $X_{(k,\ell,m)}$ and $\eta_{(k,\ell,m)}$ are the channel, transmit signal and white Gaussian

noise of the system, respectively. To include channel estimation errors into the rate analysis, we evaluate the theoretical channel estimation performance. Since frequency domain channel estimation is equivalent transform domain channel estimation in OFDM, we reformulate the theoretical Mean Square Error (MSE) expressions for our system. In [25] the MSE for the reference signal pattern of LTE is calculated. Time-frequency filters are used to interpolate the channel estimate between the position of the reference symbols. The theoretical MSE is identical with the version calculated based on channel realizations. A 2-D time-frequency interpolation method based on a Minimum Mean Square Error (MMSE) criteria as described in [26] is identified as the solution with the best performance.

In contrast, we use a 3-D time-frequency-space filter for smoothing of the estimate in the frequency domain. It is important to note that this technique assumes knowledge of the following statistical channel parameters:

- Doppler shift
- Delay spread
- Signal power of each user
- Noise power
- Spatial correlation

Since in addition we consider a Multi User - Multiple Input Multiple Output (MU-MIMO) scenario we need to ensure that different users have orthogonal reference sequences. In particular, we will assume that the training sequences are orthogonal. We assume that orthogonality is ensured by Frequency Domain Multiplex (FDM) and a cyclic shift of the reference symbols. Therefore, the following calculation is done for each user, and thus no user index is included to simplify the notation.

Assuming a reference symbol is present on subcarrier q and symbol time p we multiply the signal with the known reference signal to obtain the corresponding channel estimate for antenna m

$$\hat{H}_{(p,q,m)} = Y_{(p,q,m)} X^*_{(p,q)} = H_{(p,q,m)} + \eta_{(p,q,m)}, \tag{12}$$

where we assume that $\left|X^*_{(p,q)}\right| = 1$. By combining the channel estimates for all resource elements on K subcarriers, L symbols and M antennas we get

$$\hat{\boldsymbol{h}}_r = \left[\hat{H}_{(1,1,1)}, \hat{H}_{(2,1,1)}, \cdots, \hat{H}_{(K-1,L,M)}, \hat{H}_{(K,L,M)}\right]^T. \tag{13}$$

For all positions where no reference signals were sent the corresponding element of $\hat{\boldsymbol{h}}_r$ is set equal to zero. The set $\mathbb{P}$ contains the indices of the reference symbols in $\hat{\boldsymbol{h}}_r$.

Applying the matrices for interpolation and smoothing in time $\boldsymbol{A}_t$, frequency $\boldsymbol{A}_f$ and space $\boldsymbol{A}_s$ we get the overall estimate of the channel at each position

$$\hat{\boldsymbol{h}} = \boldsymbol{A}_{stf} \hat{\boldsymbol{h}}_r = \left(\boldsymbol{A}_s \otimes \boldsymbol{A}_t \otimes \boldsymbol{A}_f\right) \hat{\boldsymbol{h}}_r. \tag{14}$$

We choose these interpolation matrices separately for each dimension to reduce the complexity. In general to achieve the theoretical optimal performance these interpolation matrices have to be chosen according to the covariance matrix of the channel, which might not be separable. As shown in [26] for the time-frequency case this leads to a minimal performance loss, but with significantly lower complexity. In many cases the covariance is unknown, and one would need to generate the interpolation martrices based on some model for the covariance, whose parameters would also then have to be estimated.

The MSE of the estimate $\hat{\boldsymbol{h}}$ compared to the actual channel $\boldsymbol{h}$ can be calculate as

$$\begin{aligned} &\frac{1}{KLM}\mathbb{E}\left[\left|\left|\hat{\boldsymbol{h}} - \boldsymbol{h}\right|\right|^2\right] = \\ &\frac{1}{KLM}\left(\mathbb{E}\left[\hat{\boldsymbol{h}}^H \hat{\boldsymbol{h}}\right] - 2\Re\left(\mathbb{E}\left[\hat{\boldsymbol{h}}^H \boldsymbol{h}\right]\right) + \mathbb{E}\left[\boldsymbol{h}^H \boldsymbol{h}\right]\right). \end{aligned} \tag{15}$$

We split the term in (15) into three components and calculated them separately.

The third component can be calculated as

$$\mathbb{E}\left[\boldsymbol{h}^H \boldsymbol{h}\right] = \mathrm{tr}(\boldsymbol{R}_{\boldsymbol{hh}}) = \mathrm{tr}(\boldsymbol{R}^s_{\boldsymbol{hh}} \otimes \boldsymbol{R}^t_{\boldsymbol{hh}} \otimes \boldsymbol{R}^f_{\boldsymbol{hh}}). \tag{16}$$

The covariance matrices $\boldsymbol{R}^t_{\boldsymbol{hh}}$, $\boldsymbol{R}^f_{\boldsymbol{hh}}$ and $\boldsymbol{R}^s_{\boldsymbol{hh}}$ are the time, frequency and spatial covariance matrices of the channel. It is important to keep in mind that this separation might not be possible across all domains, dependent on the channel statistics. The channel model chosen in this work allows this separation.

The first component can be calculated as

$$\begin{aligned} \mathbb{E}\left[\hat{\boldsymbol{h}}^H \hat{\boldsymbol{h}}\right] &= \mathrm{tr}\left(\boldsymbol{A}_{stf} \mathbb{E}\left[\hat{\boldsymbol{h}}_r \hat{\boldsymbol{h}}_r^H\right] \boldsymbol{A}_{stf}^H\right), \\ \mathbb{E}\left[\hat{\boldsymbol{h}}_r \hat{\boldsymbol{h}}_r^H\right] &= \sum_{p1\in\mathbb{P}} \sum_{p2\in\mathbb{P}} \left[\boldsymbol{R}_{\boldsymbol{hh}} + \boldsymbol{R}_{\boldsymbol{\eta\eta}}\right]_{p1,p2} \boldsymbol{e}_{p1} \boldsymbol{e}_{p2}^T, \end{aligned} \tag{17}$$

where $\boldsymbol{R}_{\boldsymbol{\eta\eta}}$ is the covariance matrix of the noise across space, time and frequency. The vector $\boldsymbol{e}_p$ is a vector with only zeros, and a one at the pth position. We assume it can be also be separated into the submatrices for space, time and frequency in the same was as the channel:

$$\boldsymbol{R}_{\boldsymbol{\eta\eta}} = \boldsymbol{R}^s_{\boldsymbol{\eta\eta}} \otimes \boldsymbol{R}^t_{\boldsymbol{\eta\eta}} \otimes \boldsymbol{R}^f_{\boldsymbol{\eta\eta}}. \tag{18}$$

The second component of (15) can be calculated in a similar fashion as the previous one

$$\begin{aligned} \mathbb{E}\left[\hat{\boldsymbol{h}}_r^H \boldsymbol{h}\right] &= \mathrm{tr}\left(\mathbb{E}\left[\boldsymbol{h}\hat{\boldsymbol{h}}_r^H\right] \boldsymbol{A}_{tf}^H\right), \\ \mathbb{E}\left[\boldsymbol{h}\hat{\boldsymbol{h}}_r^H\right] &= \sum_{p\in\mathbb{P}} \boldsymbol{R}_{\boldsymbol{hh}} \boldsymbol{e}_p \boldsymbol{e}_p^T, \end{aligned} \tag{19}$$

using that fact that the noise has zero mean.

Plugging (16), (17) and (19) into (15) we get the analytic MSE as

$$\begin{aligned} &\frac{1}{KLM}\mathbb{E}\left[\left|\left|\hat{\boldsymbol{h}} - \boldsymbol{h}\right|\right|^2\right] = \frac{1}{KLM}\Bigg[\\ &\mathrm{tr}\left(\boldsymbol{A}_{stf}\left(\sum_{p1\in\mathbb{P}}\sum_{p2\in\mathbb{P}} \left[\boldsymbol{R}_{\boldsymbol{hh}} + \boldsymbol{R}_{\boldsymbol{\eta\eta}}\right]_{p1,p2} \boldsymbol{e}_{p1}\boldsymbol{e}_{p2}^T\right)\boldsymbol{A}_{stf}^H\right) \\ &- 2\Re\left(\mathrm{tr}\left(\left(\sum_{p\in\mathbb{P}} \boldsymbol{R}_{\boldsymbol{hh}}\boldsymbol{e}_p\boldsymbol{e}_p^T\right)\boldsymbol{A}_{stf}^H\right)\right) \\ &+ \mathrm{tr}\left(\boldsymbol{R}_{\boldsymbol{hh}}\right)\Bigg]. \end{aligned} \tag{20}$$

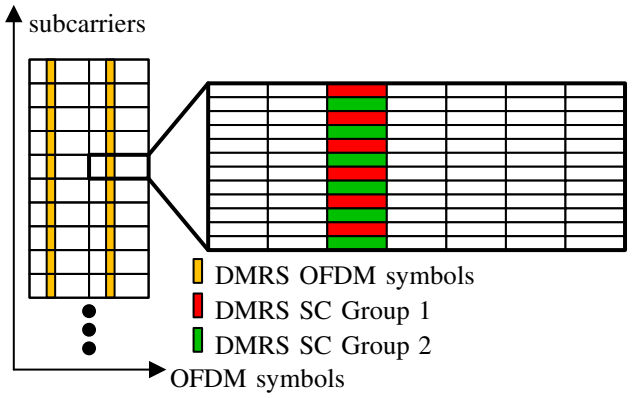

Fig. 2. 3GPP NR OFDM type one reference signal pattern for up to 4 UEs.

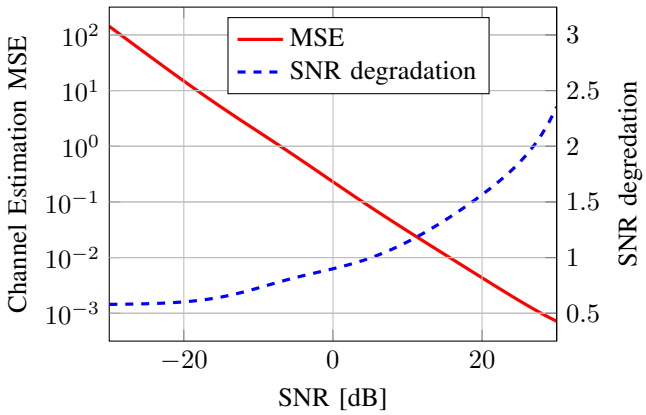

Fig. 3. Channel estimation MSE and resulting SNR degradation dependent on input SNR.

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

$$\frac{1}{KLM}\mathbb{E}\left[\left|\left|\hat{\boldsymbol{h}}-\boldsymbol{h}\right|\right|^2\right]=\frac{1}{KLM}\left[C1-2\Re(C2)+C3\right], \tag{21}$$

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

$$\begin{aligned} C1=&\sum_{p1\in\mathbb{P}}\sum_{p2\in\mathbb{P}}\Big([\boldsymbol{R}^s_{\boldsymbol{hh}}]_{m1,m2}[\boldsymbol{R}^t_{\boldsymbol{hh}}]_{\ell1,\ell2}[\boldsymbol{R}^f_{\boldsymbol{hh}}]_{k1,k2}+\\ &[\boldsymbol{R}^s_{\boldsymbol{\eta\eta}}]_{m1,m2}[\boldsymbol{R}^t_{\boldsymbol{\eta\eta}}]_{\ell1,\ell2}[\boldsymbol{R}^f_{\boldsymbol{\eta\eta}}]_{k1,k2}\Big)\\ &\left([\boldsymbol{A}_s]^H_{m2}[\boldsymbol{A}_s]_{m1}\right)\left([\boldsymbol{A}_t]^H_{\ell2}[\boldsymbol{A}_t]_{\ell1}\right)\left([\boldsymbol{A}_f]^H_{k2}[\boldsymbol{A}_f]_{k1}\right)\\ C2=&\sum_{p\in\mathbb{P}}[\boldsymbol{A}_s^H\boldsymbol{R}^s_{\boldsymbol{hh}}]_{m,m}[\boldsymbol{A}_t^H\boldsymbol{R}^t_{\boldsymbol{hh}}]_{\ell,\ell}[\boldsymbol{A}_f^H\boldsymbol{R}^f_{\boldsymbol{hh}}]_{k,k}\\ C3=&\,\mathrm{tr}\left(\boldsymbol{R}^s_{\boldsymbol{hh}}\right)\mathrm{tr}\left(\boldsymbol{R}^t_{\boldsymbol{hh}}\right)\mathrm{tr}\left(\boldsymbol{R}^f_{\boldsymbol{hh}}\right), \end{aligned} \tag{22}$$

where m, $m1$, $m2$, k, $k1$, $k2$, ℓ, $\ell1$ and $\ell2$ are the space, frequency and time indices corresponding the position of the reference symbols.

The interpolation/spatial smoothing matrices $\boldsymbol{A}_t$ and $\boldsymbol{A}_f$ are chosen according to [26] based on knowledge of the SNR, the delay spread including a model for the PDP and the Doppler spread. Since all these parameters are estimated and afterwards generated according to a model, they will never exactly match the actual PDP and Doppler spread. This introduces a model mismatch that is included in our evaluation.

The time and frequency covariance matrices $\boldsymbol{R}_t$ and $\boldsymbol{R}_f$ can be calculated according to the actual PDP and the Doppler shift including the corresponding model as shown in [25]. Based on the correlation matrix $\boldsymbol{R}_{\text{CIR}}$ of the CIR we can calculate the correlation matrix in the frequency domain $\boldsymbol{R}_f$ as

$$\boldsymbol{R}_f=\boldsymbol{W}\boldsymbol{R}_{\text{CIR}}\boldsymbol{W}^H, \tag{23}$$

where $\boldsymbol{W}$ is the matrix corresponding to a DFT transformation.

In our channel model we assume that the signal arriving at each time instant consists of a single ray. We further assume that the direction of arrival is uniformly distributed and a Uniform Linear Array (ULA) with element spacing of $\lambda/2$ is employed, so that the elements of the spatial correlation matrix can be calculated as:

$$\begin{aligned}[\boldsymbol{R}_s]_{m1,m2}&=E\left[e^{j\pi\sin(\theta)(m1-m2)}\right]\\&=\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{j\pi\sin(\theta)(m1-m2)}d\theta.\end{aligned} \tag{24}$$

This is the definition of the zeroth order Bessel function of the first kind

$$[\boldsymbol{R}_s]_{m1,m2}=\mathrm{J}_0(\pi(m1-m2)). \tag{25}$$

It is important to mention that in the case of hybrid beamforming the spatial correlation after the analog combining is unknown. Since we select the beamforming vectors independently for each RF chain we assume that the resulting channels are spatially uncorrelated. Thus, for this case the spatial correlation matrix $\boldsymbol{R}_s$ is an identity matrix. Based on this calculation we can also generate the spatial interpolation matrix $\boldsymbol{A}_s$ based on the Wiener filter equation as

$$\boldsymbol{A}_s=\boldsymbol{R}_s(\boldsymbol{R}_s+\sigma^2_\eta\boldsymbol{I})^{-1}. \tag{26}$$

Now we have assembled all the necessary mathematical tools to calculate the mean channel estimation error from the given reference signal pattern. A maximum of four User Equipment (UE) are considered. For this system setup it is sufficient to generate reference sequence by cyclic shifting and multiplication with a orthogonal cover code of a Gold sequence sequence as in the future 5G New Radio (NR) standard [27]. As shown in Fig. 2, in contrast to LTE, the DeModulation Reference Signals (DMRS) are located in separate OFDM symbols. As we can see from the figure the different DMRS groups are always allocated to adjacent Sub-Carriers (SCs). For the purpose of calculating the channel estimation mean square error we used the same channels statistics we use later for the rate calculation. Fig. 3 shows the calculated MSE and the corresponding SNR degradation. For the SNR degradation we assume a Single Input Single Output (SISO) system and that the channel estimation error is independent of the actual channel realizations.

TABLE I
COMPONENTS WITH POWER CONSUMPTION.

label	component	power consumption
P_{LO}	LO	22.5 mW
P_{LNA}	LNA	5.4 mW
P_M	mixer	0.3 mW
P_H	90° hybrid and LO buffer	3 mW
P_{LA}	LA	0.8 mW
P_1	1-bit ADC	0 mW
P_{PS}	phase shifter	2 mW
P_{VGA}	VGA	2 mW
P_{ADC}	ADC	$15 \cdot f_s 2^{\text{ENOB}}$ μW/GHz

C. Power Model

For modeling the power of the different RF frontends we use the model described in [28]. This power model is based on components reported in the literature for the WiGig standard (802.11ad) operating in the 60 GHz Industrial, Scientific and Medical (ISM) band. Since this standard was released in 2012 we can safely assume that the designs have reached sufficient maturity to represent low cost, low power power Complementary MetalOxideSemiconductor (CMOS) implementation. Table I shows the power consumption of the different components.

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

$$\begin{aligned} P_R = P_{LO} + M_R\left(P_{LNA} + P_H + 2P_M\right) + \\ \text{flag}_C\left(M_R P_{PS}\right) + \\ M_h\left(\neg\text{flag}_{\text{1bit}}\left(2P_{VGA} + 2P_{ADC1}\right) + \text{flag}_{\text{1bit}}\left(2P_{LA}\right)\right) + \\ M_t\left(\neg\text{flag}_{\text{1bit}}\left(2P_{VGA} + 2P_{ADC2}\right) + \text{flag}_{\text{1bit}}\left(2P_{LA}\right)\right), \end{aligned} \tag{27}$$

where flag_C indicates if analog combining is used:

$$\text{flag}_C = \begin{cases} 0, & M_{RFE} = M_h + M_t = M_R, M_C = 1 \\ 1, & \text{else} \end{cases}. \tag{28}$$

The variable $\text{flag}_{\text{1bit}}$ indicates if 1 or higher resolution quantization is used. The operator $\neg$ represents a logic negation. In the case of 1-bit quantization, the power consumption of the Variable Gain Amplifier (VGA) is replaced by that of the Limiting Amplifier (LA) and the power consumption of the 1-bit quantizer is negligible compared to the rest of the front-end. This formula now contains all special cases of digital beamforming ($M_{RFE} = M_R$), analog beamforming ($M_R > 0$ and $M_{RFE} = 1$) and hybrid beamforming.

III. RATE EXPRESSION

A. Allocation of RF chains for hybrid beamforming with multiple users

For the following calculations we assume that adjacent antennas are connected to one RF-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) [29], [30].

The overall procedure of 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 for each user i and RF-chain j combination under the assumption that the other users are not present and record the corresponding receive power. Afterwards, the RF-chains are allocated to the users in a resource-fair manner, starting from the RF chain and user with the highest receive power.

As we showed in [28], if the receive antennas form a ULA at each subarray of M_C elements and limiting the beams to receive the signal from only on spacial 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 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:

$$\mathbb{B} = \{\phi_1, \phi_2, \cdots, \phi_{4M_C}\}, \phi_j = -\pi + \frac{j2\pi}{M_C}. \tag{29}$$

Afterwards, for each user u and each sub-array i, all direction are tested, and the one leading to the largest receive power and the corresponding index are stored

$$\begin{aligned} p(j) &= \sum_{l=0}^{L-1} \left|\left|\boldsymbol{w}_j^H \boldsymbol{H}_u^i[l]\right|\right|_2^2 \\ [\boldsymbol{P}]_{u,i} &= \max_j p(j) \\ [\boldsymbol{J}]_{u,i} &= \arg\max_j p(j), \end{aligned} \tag{30}$$

with the vector w_j defined as:

$$\boldsymbol{w}_j = \left[1, e^{\phi_j}, \cdots, e^{(M_C-1)\phi_j}\right]^H. \tag{31}$$

The matrices $\boldsymbol{P}$ and $\boldsymbol{J}$ contain the optimal power and the corresponding direction for all combinations of user u and subarray 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

$$\mathbb{U} = \{1, \cdots, U\}, \quad \mathbb{I} = \{1, \cdots, M_{\text{RFE}}\}. \tag{32}$$

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 the those with higher power also ensures that the rate is optimized. It is important to mention that only selecting the RF-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

Algorithm 1 Selection of the beamforming vectors.

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

this is not desirable we adopted the above procedure. The entire process is summarized in Algorithm 1.

B. Modeling the Quantization

As in [14], [18], we use the Bussgang theorem to decompose the signal after quantization in a signal component and an uncorrelated quantization error $\boldsymbol{e}$:

$$\boldsymbol{r}[n] = Q(\boldsymbol{y}_C[n]) \approx \boldsymbol{F}\boldsymbol{y}_C[n] + \boldsymbol{e}[n], \tag{33}$$

with $\boldsymbol{y}_C[n]$ being the signal after the analog combining at the receiver equal to $\boldsymbol{u}[n]+\boldsymbol{\eta}_r[n]$, where $\boldsymbol{u}[n]$ is the receive signal after the multipath channel. The operation $Q(\cdot)$ represents the quantization, which is performed separately for each element of the vector as well as their real and imaginary parts. This includes the possibility of using ADCs with different resolution at each element.

To include the quantization into the rate analysis we need to calculate $\boldsymbol{F}$ and the covariance matrix $\boldsymbol{R}_{\boldsymbol{ee}}$of $\boldsymbol{e}[n]$. The description in Appendix A shows how to calculate these matrices from the receive covariance matrix $\boldsymbol{R}_{\boldsymbol{y}_C\boldsymbol{y}_C}$ and the quantization functions. For the calculation of the receive covariance matrix we reuse the formulas we derived in [28]. To simplify the notation we use the operands defined in Appendix A

$$\begin{aligned} \boldsymbol{F} &= \text{TF}(Q^1(\cdot), \cdots, Q^{M_{\text{RFE}}}). \\ \boldsymbol{R}_{\boldsymbol{rr}} &= \text{T}\left(\boldsymbol{R}_{\boldsymbol{y}_C\boldsymbol{y}_C}, Q^1(\cdot), \cdots, Q^{M_{\text{RFE}}}\right). \end{aligned} \tag{34}$$

With these results we can calculate the quantization error covariance matrix as

$$\boldsymbol{R}_{\boldsymbol{ee}} = \boldsymbol{R}_{\boldsymbol{rr}} - \boldsymbol{F}\boldsymbol{R}_{\boldsymbol{y}_C\boldsymbol{y}_C}\boldsymbol{F} \tag{35}$$

Now we can calculate the effective channel $\boldsymbol{H}'[l]$ and noise covariance matrix $\boldsymbol{R}_{\boldsymbol{\eta}'\boldsymbol{\eta}'}$of the overall system including the analog combing and the quantization:

$$\boldsymbol{H}'[l] = \boldsymbol{F}\boldsymbol{W}_R^H\boldsymbol{H}[l], \tag{36}$$

and

$$\boldsymbol{R}_{\boldsymbol{\eta}'\boldsymbol{\eta}'} = \boldsymbol{F}\boldsymbol{W}_R^H\boldsymbol{R}_{\boldsymbol{\eta}'_R\boldsymbol{\eta}'_R}\boldsymbol{W}_R\boldsymbol{F}^H + \boldsymbol{R}_{\boldsymbol{ee}}. \tag{37}$$

It is also important to mention that many previous evaluations ([14], [15], [18], [19], [20]) only use a diagonal approximation of the quantization error covariance matrix. As we show in [28], including the off-diagonal elements in the evaluation can have a dramatic impact on the overall performance. Therefore, we generalized our previously derived formulas for the case with different quantization functions to also include the off-diagonal elements in this evaluation.

C. Modeling the Channel Estimation Error

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 receive channel. This is different from the work in [21]. 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 signal receive power goes and thus the rate go to zero. If we look at our simulation of the channel estimation error in Fig. 3 this would be the case for the very low SNR range from -30 to -10 dB. This contradicts the practical observation, that communication at a SNR as low as -10 dB for a SISO system is possible [31]. For a practical massive Multiple Input Multiple Output (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.

The overall covariance matrix of the channel estimation error $\boldsymbol{R}_{\boldsymbol{ww}}$ is defined as a sum of the per user $\boldsymbol{R}_{\boldsymbol{w}_u\boldsymbol{w}_u}$

$$\boldsymbol{R}_{\boldsymbol{w}[f]\boldsymbol{w}[f]} = \sum_{u=1}^{U} \boldsymbol{R}_{\boldsymbol{w}_u[f]\boldsymbol{w}_u[f]}, \tag{38}$$

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

$$\left[\boldsymbol{R}_{\boldsymbol{w}_u[f]\boldsymbol{w}_u[f]}\right]_{m,m} = |[\boldsymbol{h}_u[f]]_m|^2\sigma_u^2. \tag{39}$$

We model each matrix $\boldsymbol{R}_{\boldsymbol{w}_u[f]\boldsymbol{w}_u[f]}$ to be spatially white and thus a diagonal matrix. The values σ_u^2 are determined by calculating the average SNR per antenna per user and then

Algorithm 2 Combined multipath channel from each user $\boldsymbol{H}[l]$, combined transmit impairments EVM co-variance matrix $\boldsymbol{R}_{\eta_T\eta_T}$, combined maximum transmit power constraint $P_{Tx}\boldsymbol{I}$, receiver noise covariance matrix $\boldsymbol{R}_{\eta\eta}$, frequency band from f_1 to f_2, quantization function $Q^m(\cdot)$ separate for each receiver chain m and channel statistics and number of frequency bins N_f.

Require: $\boldsymbol{R}_{\eta_T\eta_T}$, $\boldsymbol{R}_{\eta\eta}$, $\boldsymbol{H}[l]$, P_{Tx}, f_1, f_2 and $Q_b(\cdot)$
$\boldsymbol{H}[f] \leftarrow \mathcal{F}(\boldsymbol{H}[l])$
$\boldsymbol{R}_{\boldsymbol{x}[f]\boldsymbol{x}[f]} \leftarrow P_{Tx}\boldsymbol{I}$
$\boldsymbol{R}_{\boldsymbol{yy}} \leftarrow \sum_{f_1}^{f_2} \boldsymbol{H}[f]\left(\boldsymbol{R}_{\boldsymbol{x}(f)\boldsymbol{x}[f]} + \boldsymbol{R}_{\eta_T\eta_T}\right)\boldsymbol{H}^H[f] + \boldsymbol{R}_{\eta\eta}$
$\boldsymbol{R}_{\boldsymbol{rr}} \leftarrow \mathrm{T}(\boldsymbol{R}_{\boldsymbol{yy}}, Q^m(\cdot), \cdots, Q^{M_{RFE}}(\cdot))$
$\boldsymbol{F} \leftarrow \mathrm{TF}(\boldsymbol{R}_{\boldsymbol{yy}}, Q^m(\cdot), \cdots, Q^{M_{RFE}}(\cdot))$
$\boldsymbol{R}_{\boldsymbol{\eta'\eta'}} \leftarrow \boldsymbol{F}\boldsymbol{W}_R^H\boldsymbol{R}_{\boldsymbol{\eta}'_R\boldsymbol{\eta}'_R}\boldsymbol{W}_R\boldsymbol{F}^H + \boldsymbol{R}_{\boldsymbol{rr}} - \boldsymbol{F}\boldsymbol{R}_{\boldsymbol{yy}}\boldsymbol{F}$
$\boldsymbol{H}'[l] \leftarrow \boldsymbol{F}\boldsymbol{H}[l]\ \forall l \in \{0, \ldots, L-1\}$
$\boldsymbol{H}'[f] \leftarrow \mathcal{F}(\boldsymbol{H}'[l])$
$\boldsymbol{R}_{\boldsymbol{w}[f]\boldsymbol{w}[f]} \leftarrow \mathrm{TE}(\boldsymbol{H}[f], \boldsymbol{R}_{\boldsymbol{yy}}, \boldsymbol{R}_{\boldsymbol{\eta\eta}})\ \forall f \in [f_1, f_2]$
$\boldsymbol{R}_{\boldsymbol{\eta'}[f]\boldsymbol{\eta'}[f]} \leftarrow \boldsymbol{R}_{\boldsymbol{\eta'\eta'}} + \boldsymbol{R}_{\boldsymbol{w}[f]\boldsymbol{w}[f]}\ \forall f \in [f_1, f_2]$
$\boldsymbol{A}[f] \leftarrow \boldsymbol{I} + \boldsymbol{R}^{-1}_{\boldsymbol{\eta'}[f]\boldsymbol{\eta'}[f]}(\boldsymbol{H}'[f])\boldsymbol{R}_{\boldsymbol{x}[f]\boldsymbol{x}[f]}(\boldsymbol{H}'[f])^H\ \forall f \in [f_1, f_2]$
$R = \frac{1}{N_f}\sum_{f_1}^{f_2} \log_2\left(\det\left(\boldsymbol{A}[f]\right)\right)$
return R

obtaining the corresponding MSE σ_u^2 from the simulation shown in Fig. 3.

We combine the this calculation into the operator $\mathrm{TE}(\cdot)$

$$\boldsymbol{R}_{\boldsymbol{w}[f]\boldsymbol{w}[f]} = \mathrm{TE}(\boldsymbol{H}[f], \boldsymbol{R}_{\boldsymbol{yy}}, \boldsymbol{R}_{\boldsymbol{\eta\eta}}). \tag{40}$$

D. Combined Rate Expression

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 $\boldsymbol{x}(f)$ is Gaussian
- $\boldsymbol{w}_R^i$ are selected from the derived finite set separately for each antenna group based on an SNR criteria
- Quantization is modeled as additive Gaussian noise with the AQNM model including the off-diagonal elements
- No collaboration among the users

With these simplifications the $\boldsymbol{w}_R^i$ are already defined and we can transform the problem into a frequency domain equation.

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

$$\begin{aligned} R \leq \int_{f1}^{f2} \max_{\boldsymbol{R}_{\boldsymbol{x}(f)\boldsymbol{x}(f)}} & I(\boldsymbol{x}(f), \boldsymbol{r}(f)|\boldsymbol{H}'(f))df \\ \text{s.t. } & \mathbb{E}[||\boldsymbol{x}(f)||_2^2] \leq P_{Tx}\ \forall f \in [f_1, f_2], \end{aligned} \tag{41}$$

where $\boldsymbol{x}(f)$, $\boldsymbol{r}(f)$ and $\boldsymbol{H}'(f)$ represent the input/output signal and equivalent channel of frequency bin f, and $I(\cdot)$ is the mutual information. The frequencies f_1 and f_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 f_1 and f_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:

$$\begin{aligned} & I(\boldsymbol{x}(f), \boldsymbol{r}(f)|\boldsymbol{H}'(f)) = \\ & \log_2\left(\det\left(\boldsymbol{I} + \boldsymbol{R}^{-1}_{\boldsymbol{\eta'\eta'}}\boldsymbol{H}'(f)\boldsymbol{R}_{\boldsymbol{x}(f)\boldsymbol{x}(f)}\boldsymbol{H}'^H(f)\right)\right). \end{aligned} \tag{42}$$

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 $\boldsymbol{H}'(f)$ are dependent on the input covariance matrix $\boldsymbol{R}_{\boldsymbol{x}(f)\boldsymbol{x}(f)}$

The procedure of calculating the sum rate is summarized in algorithm 2.

IV. Simulation Results

Here we describe the chosen evaluation setup and the corresponding results. A basestation with 64 antennas ($M_R = 64$) receives the signal from 4 users ($U = 4$). 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 Hybrid BeamForming (HBF) system, $M_{\mathrm{RFE}} \in \{4, 8, 16, 32\}$ RF 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 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.

A. Average Achievable Rate Results

Fig. 4 (**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. 4 (**A**) we see that at high SNR the rate saturates and there is only minor improvement above a resolution of 5 bits. The reason for the saturation at high SNR are the transmitter impairments and channel estimation error. The same holds true for the HBF case in Fig. 4 (**B**) . But due to the limited degrees of freedom the average achievable rate saturates at a lower value than the DBF case. The results of the DBF mixed case in Fig. 4 (**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 increase.

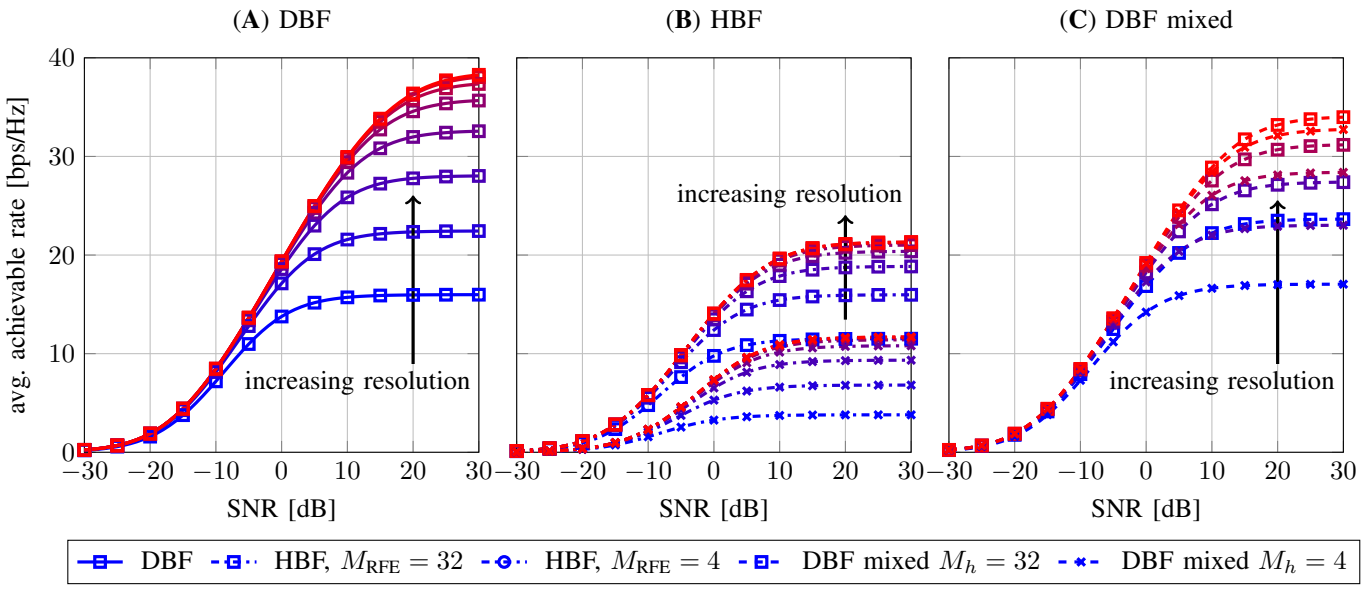

Fig. 4. DBF, HBF and DBF mixed average achievable rate for $M_R = 64$, $U = 4$, $M_{\text{RFE}} \in \{4, 32\}$, $M_h \in \{4, 32\}$ and ADC resolution $b \in \{1, \cdots, 8\}$.

B. Energy Efficiency Results

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

$$\text{energy efficiency} = \frac{R}{P_R}. \tag{43}$$

The scenarios in Fig. 5 (**A**) to (**D**) show the achievable rate and energy efficiency for different SNRs. 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 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. 5 (**A**) to (**C**)) the smaller the improvement of additional RF 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. 5 (**C**) we see that there is little difference between having 8 or 16 RF chains.

As the SNR increases from Fig. 5 (**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. 5 (**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.

V. Conclusion

The evaluations in this paper showed that low resolution ADC digital beamforming systems are more energy efficient and achieves a higher rate than hybrid beamforming systems for multiuser scenario. The reason is that the sub-arrays 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.

Future extensions should consider the following points. For the hybrid beamforming case, the evaluation only shows the result if the beams are already aligned. As shown in [6], beam alignment can require a large overhead. In addition considering what degree of power disparity among the users is possible for different ADC resolutions also provides a interesting scenario to evaluate.

Appendix

In [32] we showed how to calculate the output correlation of a quantized system from the input correlation of a Gaussian signal. However in [28] we assumed, that the same uniform quantizer is used for the signal at each antenna. Here we generalize this result to include non-uniform quantization. We also combine this result with the results in [14] to include the effects of the quantization into the rate calculation. The formula for calculating the quantizer output correlation from the input correlation for a general quantizer and two zero mean Gaussian random variables a and c can be written as

$$\rho_o = \sum_{l=1}^{N_a-1} \sum_{j=1}^{N_c-1} a_l^r c_j^r \int_0^{\rho_i'} f_{ac}(a_l^s, c_j^s, \rho_i)\, d\rho_i, \tag{44}$$

with the joint probability density function f_{ac} defined as

$$f_{ac}(a, c, \rho_i) = \frac{1}{2\pi\sigma_a\sigma_c\sqrt{1-\rho_i^2}} \exp\left(-\frac{1}{2(1-\rho_i^2)}\left[\frac{a^2}{\sigma_a^2} + \frac{c^2}{\sigma_c^2} - \frac{2\rho_i ac}{\sigma_a\sigma_c}\right]\right). \tag{45}$$

The quantizers used for a and c have N_a and N_c quantization levels. The symbols a_l^r, c_l^r, a_l^s and c_j^s value of the quantization bins and the positions of the steps. It is important to mention that we assume that the representatives of the quantization bins and the position of the steps are adapted to the input power. In a practical system this is done by a Automatic Gain Control (AGC) loop. Since we need to perform this transformation for

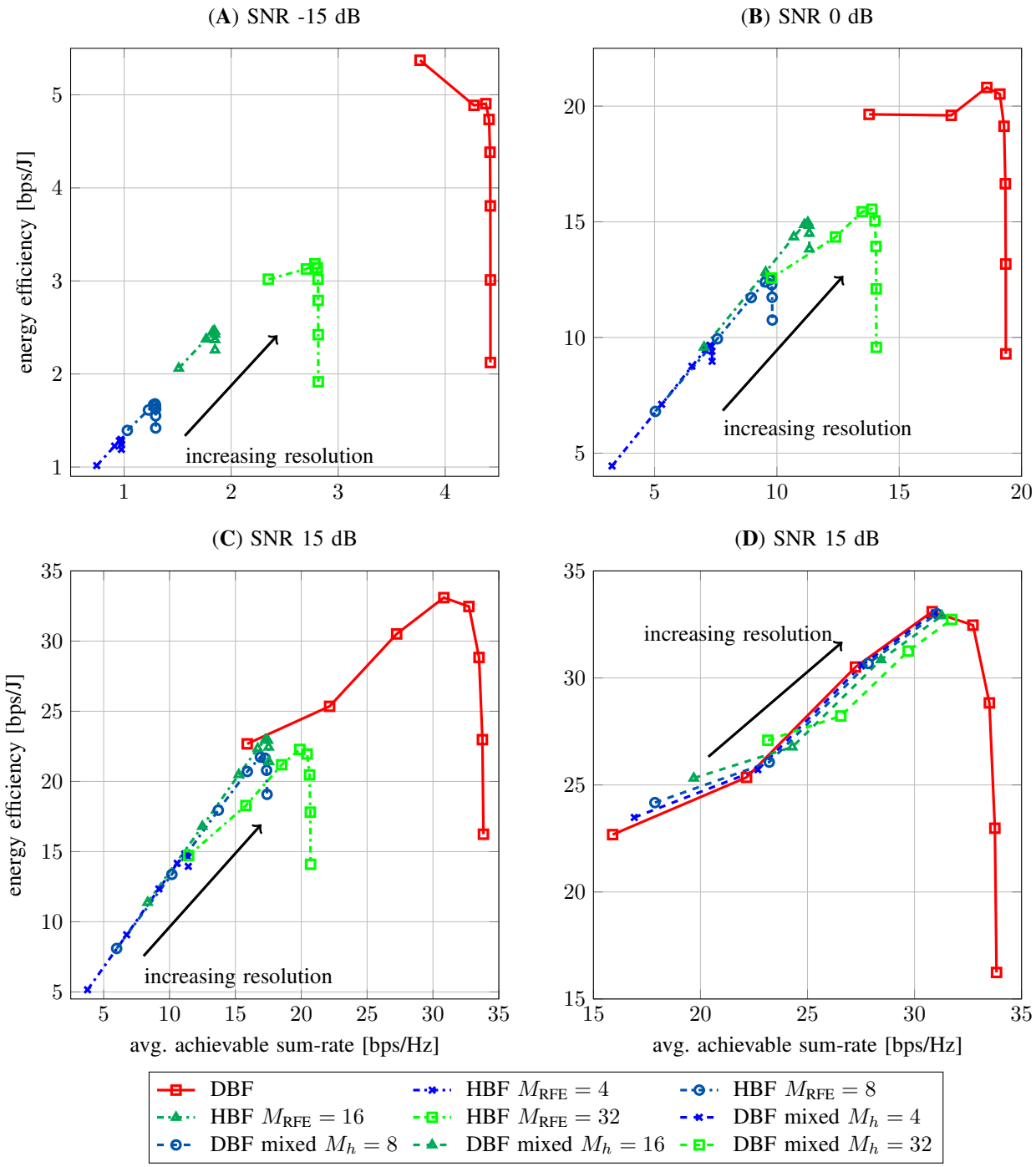

Fig. 5. Spectral and energy efficiency of digital beamforming with/without mixed ADC configuration and hybrid beamforming with $M_R = 64$, $U = 4$, $M_{\text{RFE}} \in \{4, 8, 16, 32\}$, $M_{\text{h}} \in \{4, 8, 16, 32\}$ and ADC resolution $b \in \{1, \cdots, 8\}$, $b_l \in \{1, \cdots, 4\}$ and $b_h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

every antenna pair, calculating the integral for every point is a large overhead. Therefore, we generated a non-uniform grid of input correlation ρ_i' in the range from 0 to 1 and calculate the corresponding output correlation ρ_o via numeric integration. The points in the grid are chosen in way that the change of ρ_o between adjacent points in the grid do not exceed a threshold. Afterwards if we need to calculate the output correlation for a specific input correlation, we use the pre-calculated points and interpolate with cubic splines between them. This approach provides sufficient accuracy with reduced complexity.

With this technique we can calculate the correlation matrix after the quantization $\boldsymbol{R_{rr}}$ from the correlation matrix before the quantization $\boldsymbol{R_{yy}}$. This procedure consists of the calculation of the diagonal elements of the matrix as

$$[\boldsymbol{R_{rr}}]_{i,i} = (1 - \sigma_{qi}^2)\,[\boldsymbol{R_{yy}}]_{i,i}\,, \tag{46}$$

where σ_{qi}^2 is the variance of the distortion introduced by the quantization. For each off diagonal element we use the formula in Equation (44) for all combinations of real and imaginary

parts to calculate the resulting element in $\boldsymbol{R_{rr}}$. We combine this procedure to form the operator $\mathrm{T}(\cdot)$

$$\boldsymbol{R_{rr}} = \mathrm{T}\left(\boldsymbol{R_{yy}}, Q^1(\cdot), \cdots, Q^{M_{\mathrm{RFE}}}\right). \tag{47}$$

As shown in [14], the matrix $\boldsymbol{F}$ for the Bussgang decomposition is a diagonal matrix. For the case of a different quantizer at each antenna the ith diagonal element is defined as

$$[\boldsymbol{F}]_{i,i} = (1 - \sigma_{qi}). \tag{48}$$

We can combine this operation with an operator TF only dependent on the quantization step functions $Q^i(\cdot)$

$$\boldsymbol{F} = \mathrm{TF}(Q^1(\cdot), \cdots, Q^{M_{\mathrm{RFE}}}). \tag{49}$$

ACKNOWLEDGMENT

This work has been performed in the framework of the Horizon 2020 project ONE5G (ICT-760809) receiving funds from the European Union. The authors would like to acknowledge the contributions of their colleagues in the project, although the views expressed in this contribution are those of the authors and do not necessarily represent the project.

REFERENCES

[1] F. Boccardi et al., "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[2] J. G. Andrews et al., "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[3] "5G white paper," NGNM, Feb. 2015. [Online]. Available: https://www.ngmn.org/uploads/media/NGMN_5G_White_Paper_V1_0.pdf

[4] B. Murmann, "ADC performance survey 1997-2017," 2017. [Online]. Available: http://www.stanford.edu/~murmann/adcsurvey.html

[5] J. Singh, O. Dabeer and U.Madhow, "Communication limits with low precision analog-to-digital conversion at the receiver," *IEEE Trans. Commun.*, vol. 57, no. 12, pp. 3629–3639, Dec. 2009.

[6] C. N. Barati et al., "Directional cell discovery in millimeter wave cellular networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 12, pp. 6664–6678, Dec. 2015.

[7] T. S. Rappaport et al., *Millimeter Wave Wireless Communications*, ser. Prentice Hall Communications Engineering and Emerging Technologies Series from Ted Rappaport. Pearson Education, 2014.

[8] J. Mo and R. W. Heath Jr., "Capacity analysis of one-bit quantized MIMO systems with transmitter channel state information," *IEEE Trans. Signal Process.*, vol. 63, no. 20, pp. 5498–5512, Oct. 2015.

[9] R. J. Mailloux, *Phased Array Antenna Handbook*. Artech House, 2005.

[10] W. Roh et al., "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 106–113, Feb. 2014.

[11] L. Kong, "Energy-efficient 60 GHz phased-array design for multi-Gb/s communication systems," Ph.D. dissertation, EECS Department, University of California, Berkeley, Dec 2014. [Online]. Available: http://www.eecs.berkeley.edu/Pubs/TechRpts/2014/EECS-2014-191.html

[12] A. Mezghani and J. A. Nossek, "On ultra-wideband MIMO systems with 1-bit quantized outputs: performance analysis and input optimization," in *Int. Symp. on Information Theory (ISIT) 2007*, Nice, France, June 2007, pp. 1286–1289.

[13] 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*, vol. 65, no. 15, pp. 4075–4089, Aug 2017.

[14] A. Mezghani and J. A. Nossek, "Capacity lower bound of MIMO channels with output quantization and correlated noise," in *Int. Symp. on Information Theory (ISIT) 2012*, Massachusetts, USA, Jul. 2012.

[15] Q. Bai and J. A. Nossek, "Energy efficiency maximization for 5G multi-antenna receivers," *Transactions on Emerging Telecommunications Technologies*, vol. 26, no. 1, pp. 3–14, Jan. 2015.

[16] C. Studer and G. Durisi, "Quantized massive MU-MIMO-OFDM uplink," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2387–2399, June 2016.

[17] C. Molln et al., "Uplink performance of wideband massive MIMO with one-bit ADCs," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 87–100, Jan. 2017.

[18] J. Mo et al., "Hybrid architectures with few-bit ADC receivers: Achievable rates and energy-rate tradeoffs," *CoRR*, vol. abs/1605.00668, 2016. [Online]. Available: http://arxiv.org/abs/1605.00668

[19] W. bin Abbas, F. Gómez-Cuba and M. Zorzi, "Millimeter wave receiver efficiency: A comprehensive comparison of beamforming schemes with low resolution ADCs," *CoRR*, vol. abs/1607.03725, 2016. [Online]. Available: http://arxiv.org/abs/1607.03725

[20] 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 *Information Theory and Applications Workshop (ITA) 2015*, San Diego, CA, USA, Feb 2015, pp. 191–198.

[21] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading mimo channels with channel estimation error," *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 2203–2214, May 2006.

[22] N. Lee, O. Simeone, and J. Kang, "The effect of imperfect channel knowledge on a mimo system with interference," *IEEE Transactions on Communications*, vol. 60, no. 8, pp. 2221–2229, August 2012.

[23] P. Reynaert and M. Steyaert, "Rf power amplifiers for mobile communications," pp. 184–185, 2006.

[24] "Deliverable D2.1 measurement campaigns and initial channel models for preferred suitable frequency ranges," 2016. [Online]. Available: https://bscw.5g-mmmagic.eu/pub/bscw.cgi/d94832/mmMAGIC_D2-1.pdf

[25] M. Biagini et al., "Time-frequency MSE analysis for pilot aided channel estimation in OFDM systems," in *Int. Conf. on Wireless Communications, Vehicular Technology, Information Theory and Aerospace Electronic Systems (VITAE) 2014*, Aalborg, Denmark, May 2014, pp. 1–5.

[26] P. Hoeher, S. Kaiser and P. Robertson, "Two-dimensional pilot-symbol-aided channel estimation by Wiener filtering," in *Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP) 1997.*, vol. 3, Munich, Bavaria, Germany, Apr. 1997, pp. 1845–1848 vol.3.

[27] 3GPP, "NR; Physical channels and modulation," 3rd Generation Partnership Project (3GPP), TS 38.211, 2017. [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/38211.htm

[28] K. Roth and J. A. Nossek, "Achievable rate and energy efficiency of hybrid and digital beamforming receivers with low resolution ADC," *IEEE Journal of Selected Areas in Communications (IEEE JSAC) Special Issue on Millimeter Wave Communications for Future Mobile Networks (JSACMillimeterWave'2017)*, 2017.

[29] *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*, Std., Dec 2012.

[30] K. Oteri et al., "IEEE 802.11-16/1447r1 further details on multi-stage, multi-resolution beamforming training in 802.11ay," Nov. 2016.

[31] R. Ratasuk, B. Vejlgaard, N. Mangalvedhe, and A. Ghosh, "Nb-iot system for m2m communication," in *2016 IEEE Wireless Communications and Networking Conference*, April 2016, pp. 1–5.

[32] K. Roth et al., "Covariance based signal parameter estimation of coarse quantized signals," in *Int. Conf. on Digital Signal Processing (DSP) 2015*, Singapore, Singapore, July 2015, pp. 19–23.