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 for a wideband multipath model. 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 Signal to Noise Ratio (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 mixed ADC resolutions are used we can achieve any desired trade-off between power consumption and rate close to those achieved with only one ADC resolution.**

***Index Terms*—Wireless communication, millimeter Wave, low resolution 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 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 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 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 Analog-to-Digital-Converter (ADC) at each antenna is still feasible. For future mmWave mobile broadband systems, a much larger bandwidth [3] and a much larger 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,

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

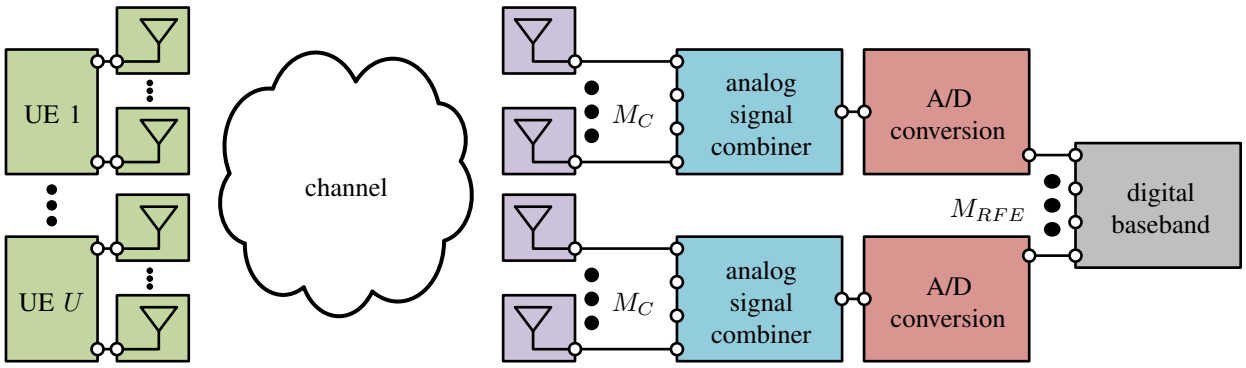

Fig. 1. System Model with U UEs and a basestation with M_C antennas at each of the M_{RFE} RF chains. Number of receive Antennas M_R is equal to $M_C \times M_{\text{RFE}}$.

if high gain is needed, the beamwidth is small and thus the 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, it may 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. Most theoretical analyses of low resolution ADCs have only considered the extreme case of one-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 Division Multiplexing (OFDM) in an uplink scenario is 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 (i.e., sub-array hybrid beamforming). In this work we concentrate on effects of the hardware constraints at the receiver, thus we assume 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] analyzed the effect of imperfect channel knowledge on the achievable rate. The channel estimation error is treated as additional noise in the system. We will use a similar model to include the channel estimation error in our analysis. Since we have a system involving multiple users with different receive power, we treat the effect of each user separately.

B. Contribution

The contribution of this work can be summarized as follows:

- Achievable rate analysis for hybrid beamforming systems and digital systems with low resolution ADCs in a multi-user, multipath, wideband scenario. In addition the effects of transmitter impairments, channel estimation errors and having mixed ADC resolutions are considered.
- Analyzing the channel estimation error considering the reference signal patterns already agreed upon for 3GPP NR (5G).
- Illustrating 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 $\boldsymbol{R}_{\boldsymbol{ab}}$ is the correlation matrix of zero-mean vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ defined as $\mathbb{E}[\boldsymbol{a}\boldsymbol{b}^H]$. The Discrete Fourier Transform (DFT) $\mathcal{F}(\cdot)$ and its inverse $\mathcal{F}^{-1}(\cdot)$ and the Fourier transform $\mathscr{F}\{\cdot\}$ and its inverse $\mathscr{F}^{-1}\{\cdot\}$ are also used. A list of commonly used symbols can be found in Appendix B.

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 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.

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 $\boldsymbol{\eta}_u[n]$ 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 Signal-to-Noise-Ratio (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 [23]. 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], [25] the inefficiency of the Power Amplifiers (PAs) for mmWave as well as other impairments ([25]), 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 Multi User - Multiple Input, Multiple Output (MU-MIMO) Uplink (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., [26], [27], [28], [29]).

Thus, the noise here combines standard thermal noise and non-linear contributions from the whole transmitter hardware, including Digital-to-Analog-Converter (DAC), PA and the phase noise of the Local Oscillators (LO).

Since all noise contributions in Equation (1) are assumed to be Gaussian we can combine them to form a combined noise $\boldsymbol{\eta}'_R[n]$ equal to

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

The receive signal is then reformulated to

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

We restrict the system to have M_C antennas exclusively connected to one RF front-end chain (see Fig. 1). Therefore, the matrix $\boldsymbol{W}_R$ modeling the analog combining at the receiver 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 after the analog combining $\boldsymbol{y}_C[n]$ can be calculated as:

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

For the case of digital beamforming the matrix $\boldsymbol{W}_R$ is simply 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

$$\boldsymbol{r}[n] = Q_b(\boldsymbol{y}_C[n]). \tag{6}$$

The quantization operator $Q_b(\boldsymbol{a})$ treats the I and 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:

$$r = Q_b(a) = q^j \ \forall \ a \ \in \left] q_l^{j-1} q_l^j \right]. \tag{7}$$

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. We thus limit our evaluation to this set of quantizers. For the theoretical evaluation we assume Gaussian signaling. Consequently, we use the stepsize to minimize the distortion for Gaussian signals shown in [30]. Since the actual receive power at each antenna can be different, an Automatic Gain Control (AGC) needs to adapt a Variable Gain Amplifier (VGA) to generate the minimal distortion. To simplify our model, we assume that the AGC always perfectly adapts to the current situation.

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 for 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 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 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.

A. Channel Model

The measurements in [31] show that for channels at 60 GHz, an exponential Power Delay Profile (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

$$\boldsymbol{H}[l] = \frac{1}{\sqrt{M_T}} \alpha(l) \boldsymbol{a}_r(\phi_r(l)) \boldsymbol{a}_t^T(\phi_t(l)), \tag{8}$$

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

For the numeric evaluation, we assume that the antennas of the transmitters and the receiver form a Uniform Linear Array (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 $\phi = d\ \mathrm{k} \sin(\theta) = \pi \sin(\theta)$. In this case we use the angular wave number k defined as $2\pi/\lambda$. The angle θ 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 $\phi_r(l)$ and $\phi_t(l)$ of path l depend on the angle of arrival $\theta_r(l)$ and departure $\theta_t(l)$ as follows:

$$\boldsymbol{a}_r^T(\phi_r(l)) = \left[1, e^{j\phi_r(l)}, e^{j2\phi_r(l)}, \cdots, e^{j(M_r-1)\phi_r(l)}\right]. \tag{9}$$

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

$$v_l = \mathbb{E}\left[|\alpha(l)|^2\right] = e^{-\beta l}. \tag{10}$$

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:

$$\boldsymbol{v}_n = \frac{\boldsymbol{v}}{||\boldsymbol{v}||^2}. \tag{11}$$

The SNR γ_u per user u is defined as:

$$\gamma_u = \frac{P_u\ \mathbb{E}\left[\left|\left|\sum_{l=0}^{L_u} \boldsymbol{H}_u[l]\boldsymbol{x}_u[n-l]\right|\right|_2^2\right]}{\mathbb{E}[||\boldsymbol{\eta}_R[n]||_2^2]}. \tag{12}$$

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 $\boldsymbol{x}_u[n]$ 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

$$Y_{(k,\ell,m)} = H_{(k,\ell,m)} X_{(k,\ell,m)} + \eta_{(k,\ell,m)}, \tag{13}$$

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. In [32] it was shown that a sufficient time and frequency synchronization can be achieved in a Multiple Input, Multiple Output (MIMO) system at low SNR in a multipath channel environment. To include channel estimation errors into the rate analysis, we evaluate the theoretical channel estimation performance. Since frequency domain channel estimation is equivalent to transform domain channel estimation in OFDM, we reformulate the theoretical Mean Square Error (MSE) expressions for our system. In [33] 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 [34] 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 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 Division Multiplexing (FDM) and a cyclic shift of the reference symbols. The amount of symbols dedicated for channel estimation is usually very limited, which limits the number of available orthogonal reference signals. Therefore, it is important to consider the reference overhead, estimation quality and number of orthogonal sequences during the system design. For current third Generation Partnership Project (3GPP) 5G New Radio (NR) type 1 and type 2 OFDM reference signals, only 8 and 12 orthogonal sequences are available, which already represent an overhead of 28.6 % [35]. 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

$$\hat{H}_{(p,q,m)} = Y_{(p,q,m)} X^*_{(p,q)} = H_{(p,q,m)} + \eta_{(p,q,m)}, \tag{14}$$

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

$$\hat{\boldsymbol{h}}_r = \left[\hat{H}_{(1,1,1)}, \hat{H}_{(2,1,1)}, \cdots, \hat{H}_{(K-1,L,M)}, \hat{H}_{(K,L,M)}\right]^T. \tag{15}$$

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

$$\hat{\boldsymbol{h}} = \boldsymbol{A}_{stf}\hat{\boldsymbol{h}}_r = \left(\boldsymbol{A}_s \otimes \boldsymbol{A}_t \otimes \boldsymbol{A}_f\right)\hat{\boldsymbol{h}}_r. \tag{16}$$

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 [34] 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 matrices 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 calculated as

$$\begin{aligned}&\frac{1}{KLM}\mathbb{E}\left[\left|\left|\hat{\boldsymbol{h}} - \boldsymbol{h}\right|\right|^2\right] = \\ &\frac{1}{KLM}\left(\mathbb{E}\left[\hat{\boldsymbol{h}}^H\hat{\boldsymbol{h}}\right] - 2\Re\left(\mathbb{E}\left[\hat{\boldsymbol{h}}^H\boldsymbol{h}\right]\right) + \mathbb{E}\left[\boldsymbol{h}^H\boldsymbol{h}\right]\right).\end{aligned} \tag{17}$$

We split the term in (17) into three components and calculate them separately.

The third component can be calculated as

$$\mathbb{E}\left[\boldsymbol{h}^H\boldsymbol{h}\right] = \mathrm{tr}(\boldsymbol{R}_{\boldsymbol{hh}}) = \mathrm{tr}(\boldsymbol{R}^s_{\boldsymbol{hh}} \otimes \boldsymbol{R}^t_{\boldsymbol{hh}} \otimes \boldsymbol{R}^f_{\boldsymbol{hh}}). \tag{18}$$

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, depending on the channel statistics. The channel model chosen in this work allows this separation.

The first component can be calculated as

$$\begin{aligned}\mathbb{E}\left[\hat{\boldsymbol{h}}^H\hat{\boldsymbol{h}}\right] &= \mathrm{tr}\left(\boldsymbol{A}_{stf}\mathbb{E}\left[\hat{\boldsymbol{h}}_r\hat{\boldsymbol{h}}_r^H\right]\boldsymbol{A}_{stf}^H\right), \\ \mathbb{E}\left[\hat{\boldsymbol{h}}_r\hat{\boldsymbol{h}}_r^H\right] &= \sum_{p1\in\mathbb{P}}\sum_{p2\in\mathbb{P}}\left[\boldsymbol{R}_{\boldsymbol{hh}} + \boldsymbol{R}_{\boldsymbol{\eta\eta}}\right]_{p1,p2}\boldsymbol{e}_{p1}\boldsymbol{e}_{p2}^T,\end{aligned} \tag{19}$$

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 zeroes, and a one at the pth position. We assume it can also be separated into the sub matrices for space, time and frequency in the same way as the channel:

$$\boldsymbol{R}_{\boldsymbol{\eta\eta}} = \boldsymbol{R}^s_{\boldsymbol{\eta\eta}} \otimes \boldsymbol{R}^t_{\boldsymbol{\eta\eta}} \otimes \boldsymbol{R}^f_{\boldsymbol{\eta\eta}}. \tag{20}$$

The second component of (17) can be calculated in a similar fashion as the previous one

$$\begin{aligned}\mathbb{E}\left[\hat{\boldsymbol{h}}_r^H\boldsymbol{h}\right] &= \mathrm{tr}\left(\mathbb{E}\left[\boldsymbol{h}\hat{\boldsymbol{h}}_r^H\right]\boldsymbol{A}_{tf}^H\right), \\ \mathbb{E}\left[\boldsymbol{h}\hat{\boldsymbol{h}}_r^H\right] &= \sum_{p\in\mathbb{P}}\boldsymbol{R}_{\boldsymbol{hh}}\boldsymbol{e}_p\boldsymbol{e}_p^T,\end{aligned} \tag{21}$$

using that fact that the noise has zero mean.

Plugging (18), (19) and (21) into (17) we get the analytic MSE as

$$\begin{aligned}&\frac{1}{KLM}\mathbb{E}\left[\left|\left|\hat{\boldsymbol{h}} - \boldsymbol{h}\right|\right|^2\right] = \frac{1}{KLM}\Bigg[\\ &\mathrm{tr}\left(\boldsymbol{A}_{stf}\left(\sum_{p1\in\mathbb{P}}\sum_{p2\in\mathbb{P}}\left[\boldsymbol{R}_{\boldsymbol{hh}} + \boldsymbol{R}_{\boldsymbol{\eta\eta}}\right]_{p1,p2}\boldsymbol{e}_{p1}\boldsymbol{e}_{p2}^T\right)\boldsymbol{A}_{stf}^H\right) \\ &- 2\Re\left(\mathrm{tr}\left(\left(\sum_{p\in\mathbb{P}}\boldsymbol{R}_{\boldsymbol{hh}}\boldsymbol{e}_p\boldsymbol{e}_p^T\right)\boldsymbol{A}_{stf}^H\right)\right) \\ &+ \mathrm{tr}\left(\boldsymbol{R}_{\boldsymbol{hh}}\right)\Bigg].\end{aligned} \tag{22}$$

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:

$$\frac{1}{KLM}\mathbb{E}\left[\left|\left|\hat{\boldsymbol{h}} - \boldsymbol{h}\right|\right|^2\right] = \frac{1}{KLM}\left[C1 - 2\Re(C2) + C3\right], \tag{23}$$

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

$$\begin{aligned}C1 = \sum_{p1\in\mathbb{P}}\sum_{p2\in\mathbb{P}}\Big(&[\boldsymbol{R}^s_{\boldsymbol{hh}}]_{m1,m2}[\boldsymbol{R}^t_{\boldsymbol{hh}}]_{\ell1,\ell2}[\boldsymbol{R}^f_{\boldsymbol{hh}}]_{k1,k2} + \\ &[\boldsymbol{R}^s_{\boldsymbol{\eta\eta}}]_{m1,m2}[\boldsymbol{R}^t_{\boldsymbol{\eta\eta}}]_{\ell1,\ell2}[\boldsymbol{R}^f_{\boldsymbol{\eta\eta}}]_{k1,k2}\Big) \\ &\left([\boldsymbol{A}_s]^H_{m2}[\boldsymbol{A}_s]_{m1}\right)\left([\boldsymbol{A}_t]^H_{\ell2}[\boldsymbol{A}_t]_{\ell1}\right)\left([\boldsymbol{A}_f]^H_{k2}[\boldsymbol{A}_f]_{k1}\right) \\ C2 = \sum_{p\in\mathbb{P}}&[\boldsymbol{A}_s^H\boldsymbol{R}^s_{\boldsymbol{hh}}]_{m,m}[\boldsymbol{A}_t^H\boldsymbol{R}^t_{\boldsymbol{hh}}]_{\ell,\ell}[\boldsymbol{A}_f^H\boldsymbol{R}^f_{\boldsymbol{hh}}]_{k,k} \\ C3 = &\,\mathrm{tr}\left(\boldsymbol{R}^s_{\boldsymbol{hh}}\right)\mathrm{tr}\left(\boldsymbol{R}^t_{\boldsymbol{hh}}\right)\mathrm{tr}\left(\boldsymbol{R}^f_{\boldsymbol{hh}}\right),\end{aligned} \tag{24}$$

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 [34] 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 $\boldsymbol{R}^t_{\boldsymbol{hh}}$ 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 [33]. Based on the correlation matrix $\boldsymbol{R}_{\mathrm{CIR}}$ of the CIR we can calculate the correlation matrix in the frequency domain $\boldsymbol{R}^f_{\boldsymbol{hh}}$ as

$$\boldsymbol{R}^f_{\boldsymbol{hh}} = \boldsymbol{W}\boldsymbol{R}_{\mathrm{CIR}}\boldsymbol{W}^H, \tag{25}$$

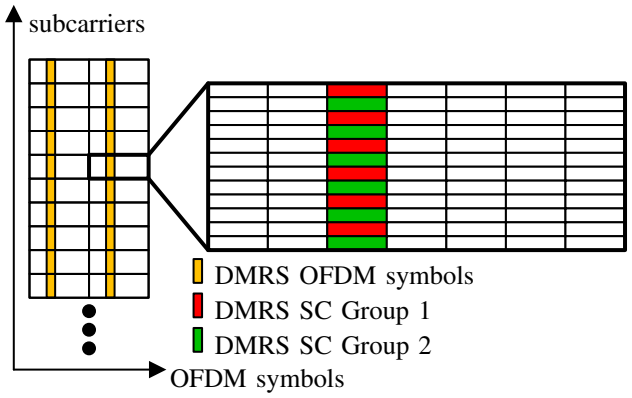

Fig. 2. 3GPP NR OFDM type one reference signal pattern for up to 4 UEs.

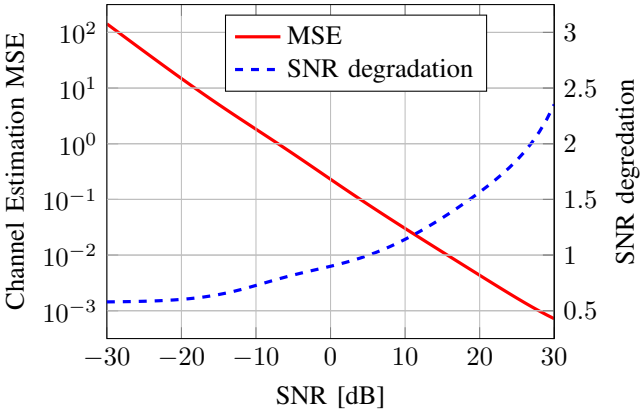

Fig. 3. Channel estimation MSE and resulting SNR degradation dependent on input SNR.

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

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 ULA with element spacing of $\lambda/2$ is employed, so that the elements of the spatial correlation matrix can be calculated as:

$$[\boldsymbol{R}^s_{\boldsymbol{hh}}]_{m1,m2} = E\left[e^{j\pi\sin(\theta)(m1-m2)}\right] = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{j\pi\sin(\theta)(m1-m2)}d\theta. \tag{26}$$

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

$$[\boldsymbol{R}^s_{\boldsymbol{hh}}]_{m1,m2} = \mathrm{J}_0(\pi(m1-m2)). \tag{27}$$

The simple model in (26) is just one possible choice for $\boldsymbol{R}^s_{\boldsymbol{hh}}$; our analysis methodology is general enough to accommodate any model for $\boldsymbol{R}^s_{\boldsymbol{hh}}$; here. Since our focus is on comparing the spectral and energy efficiency of two different system architectures, we choose a relatively simple channel model in order to not overly complicate the analysis. The specific choice of $\boldsymbol{R}^s_{\boldsymbol{hh}}$ will not significantly impact our conclusions. 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 $\boldsymbol{R}^s_{\boldsymbol{hh}}$ 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

$$\boldsymbol{A}_s = \boldsymbol{R}^s_{\boldsymbol{hh}}(\boldsymbol{R}^s_{\boldsymbol{hh}} + \sigma^2_\eta \boldsymbol{I})^{-1}. \tag{28}$$

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 an orthogonal cover code of a Gold sequence as in the future 5G NR standard [35].

This reference signal scheme can only provide sufficient orthogonality for 4 users. If more users are necessary it is possible to use an additional OFDM symbol for reference signal transmission and employ Code Division Multiplexing (CDM). This would generate orthogonal sequences for 8 users and is called type 1 OFDM DeModulation Reference Signals (DMRS) for up to 8 users in the 3GPP NR context. Another option is type 2 OFDM DMRS which can provide a maximum of 12 orthogonal signals for different users. The only difference from type 1 is that there are now 3 groups of reference signals orthogonalized via FDM instead of 2.

As shown in Fig. 2, in contrast to LTE, the 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 use the same channel 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. We use the described channel model with an exponential PDP and a Doppler-spread of 5 Hz.

C. Power Model

For modeling the power of the different RF frontends we use the model described in [36]. 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 Metal-Oxide-Semiconductor (CMOS) implementations. Table I shows the power consumption of the different components.

The power consumption of the ADC is dependent on the Effective Number of Bits (ENOB) and the sampling frequency f_s. Since we do not model all noise and non-linear effects of an ADC, the ADC resolution b is directly used as the ENOB. To support the 2 GHz bandwidth available for 802.11ad and leaving some room for the transition of the analog filters, we use a sampling frequency f_s of 2.5 GHz for the ADC.

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	one-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

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

$$\begin{aligned} P_R = P_{LO} + M_R \left(P_{LNA} + P_H + 2P_M\right) + \\ \text{flag}_C \left(M_R P_{PS}\right) + \\ M_h \left(\neg\text{flag}_{\text{1bit}} \left(2P_{VGA} + 2P_{ADC1}\right) + \text{flag}_{\text{1bit}} \left(2P_{LA}\right)\right) + \\ M_t \left(\neg\text{flag}_{\text{1bit}} \left(2P_{VGA} + 2P_{ADC2}\right) + \text{flag}_{\text{1bit}} \left(2P_{LA}\right)\right), \end{aligned} \tag{29}$$

where flag_C indicates if analog combining is used:

$$\text{flag}_C = \begin{cases} 0, & M_{RFE} = M_h + M_t = M_R, M_C = 1 \\ 1, & \text{otherwise} \end{cases}. \tag{30}$$

The variable $\text{flag}_{\text{1bit}}$ indicates if one-bit or higher resolution quantization is used. The operator $\neg$ represents a logic negation. In the case of one-bit quantization, the power consumption of the VGA is replaced by that of the Limiting Amplifier (LA) and the power consumption of the one-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) [37], [38].

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 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 [36], if the receive antennas form a 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 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:

$$\mathbb{B} = \{\phi_1, \phi_2, \cdots, \phi_{4M_C}\}, \phi_j = -\pi + \frac{j\pi}{2M_C}. \tag{31}$$

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

$$\begin{aligned} p(j) &= \sum_{l=0}^{L-1} \left|\left|\boldsymbol{w}_j^H \boldsymbol{H}_u^i[l]\right|\right|_2^2, \\ [\boldsymbol{P}]_{u,i} &= \max_j p(j), \\ [\boldsymbol{J}]_{u,i} &= \arg\max_j p(j), \end{aligned} \tag{32}$$

with the vector w_j defined as:

$$\boldsymbol{w}_j = \left[1, e^{\phi_j}, \cdots, e^{(M_C-1)\phi_j}\right]^H. \tag{33}$$

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

$$\mathbb{U} = \{1, \cdots, U\}, \quad \mathbb{I} = \{1, \cdots, M_{\text{RFE}}\}. \tag{34}$$

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

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 $\boldsymbol{e}$:

$$\boldsymbol{r}[n] = Q(\boldsymbol{y}_C[n]) \approx \boldsymbol{F}\boldsymbol{y}_C[n] + \boldsymbol{e}[n], \tag{35}$$

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 $\boldsymbol{w}_R^i \ \forall i = \{1, \ldots, M_{\text{RFE}}\}$

with $\boldsymbol{y}_C[n]$ representing 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 [36]. To simplify the notation we use the operands defined in Appendix A:

$$\begin{aligned} \boldsymbol{F} &= \text{TF}(Q^1(\cdot), \cdots, Q^{M_{\text{RFE}}}), \\ \boldsymbol{R}_{\boldsymbol{rr}} &= \text{T}\left(\boldsymbol{R}_{\boldsymbol{y}_C\boldsymbol{y}_C}, Q^1(\cdot), \cdots, Q^{M_{\text{RFE}}}\right). \end{aligned} \tag{36}$$

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

$$\boldsymbol{R}_{\boldsymbol{ee}} = \boldsymbol{R}_{\boldsymbol{rr}} - \boldsymbol{F}\boldsymbol{R}_{\boldsymbol{y}_C\boldsymbol{y}_C}\boldsymbol{F}. \tag{37}$$

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:

$$\boldsymbol{H}'[l] = \boldsymbol{F}\boldsymbol{W}_R^H\boldsymbol{H}[l], \tag{38}$$

and

$$\boldsymbol{R}_{\boldsymbol{\eta}'\boldsymbol{\eta}'} = \boldsymbol{F}\boldsymbol{W}_R^H\boldsymbol{R}_{\boldsymbol{\eta}'_R\boldsymbol{\eta}'_R}\boldsymbol{W}_R\boldsymbol{F}^H + \boldsymbol{R}_{\boldsymbol{ee}}. \tag{39}$$

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 [36], including the off-diagonal elements in the evaluation can have a dramatic impact on the overall performance. Therefore, in this evaluation we generalized our previously derived formulas for the case with different quantization functions to also include the off-diagonal elements.

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 the 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 received signal power as well as 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 SNRs as low as -10 dB for a SISO system is possible [39]. 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.

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

$$\boldsymbol{R}_{\boldsymbol{w}[f]\boldsymbol{w}[f]} = \sum_{u=1}^{U} \boldsymbol{R}_{\boldsymbol{w}_u[f]\boldsymbol{w}_u[f]}, \tag{40}$$

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:

$$\left[\boldsymbol{R}_{\boldsymbol{w}_u[f]\boldsymbol{w}_u[f]}\right]_{m,m} = |[\boldsymbol{h}_u[f]]_m|^2\sigma_u^2. \tag{41}$$

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

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

$$\boldsymbol{R}_{\boldsymbol{w}[f]\boldsymbol{w}[f]} = \text{TE}(\boldsymbol{H}[f], \boldsymbol{R}_{\boldsymbol{yy}}, \boldsymbol{R}_{\boldsymbol{\eta\eta}}). \tag{42}$$

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:

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

- Assume $\boldsymbol{x}[f]$ is Gaussian
- Beamforming vectors $\boldsymbol{w}_R^i$ are selected from the derived finite set separately for each antenna group based on an SNR criteria
- 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 [40]. Since otherwise the overhead for beamtraining is very large for most theoretical work as well as practical implementations, the vectors are drawn from a predefined set [41], [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 [42] 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.

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

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 f separately:

$$R \leq \int_{f1}^{f2} \max_{\boldsymbol{R}_{\boldsymbol{x}(f)\boldsymbol{x}(f)}} I(\boldsymbol{x}(f), \boldsymbol{r}(f)|\boldsymbol{H}'(f))df \quad \text{s.t.}\ \ \mathbb{E}[||\boldsymbol{x}(f)||_2^2] \leq P_{Tx}\ \forall f \in [f_1, f_2], \tag{43}$$

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:

$$I(\boldsymbol{x}(f), \boldsymbol{r}(f)|\boldsymbol{H}'(f)) = \log_2\left(\det\left(\boldsymbol{I} + \boldsymbol{R}^{-1}_{\boldsymbol{\eta'\eta'}}\boldsymbol{H}'(f)\boldsymbol{R}_{\boldsymbol{x}(f)\boldsymbol{x}(f)}\boldsymbol{H}'^H(f)\right)\right). \tag{44}$$

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

The procedure for calculating the sum rate is summarized in Algorithm 2. To simplify the calculation we us a discrete Fourier transformat with a sufficiently large number of bins.

IV. Simulation Results

Here we describe the chosen evaluation setup and the corresponding results. A basestation with 64 antennas ($M_R = 64$) 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 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.

To access this result from the theoretical side we need to look at the maximum SNR of the combined signal from

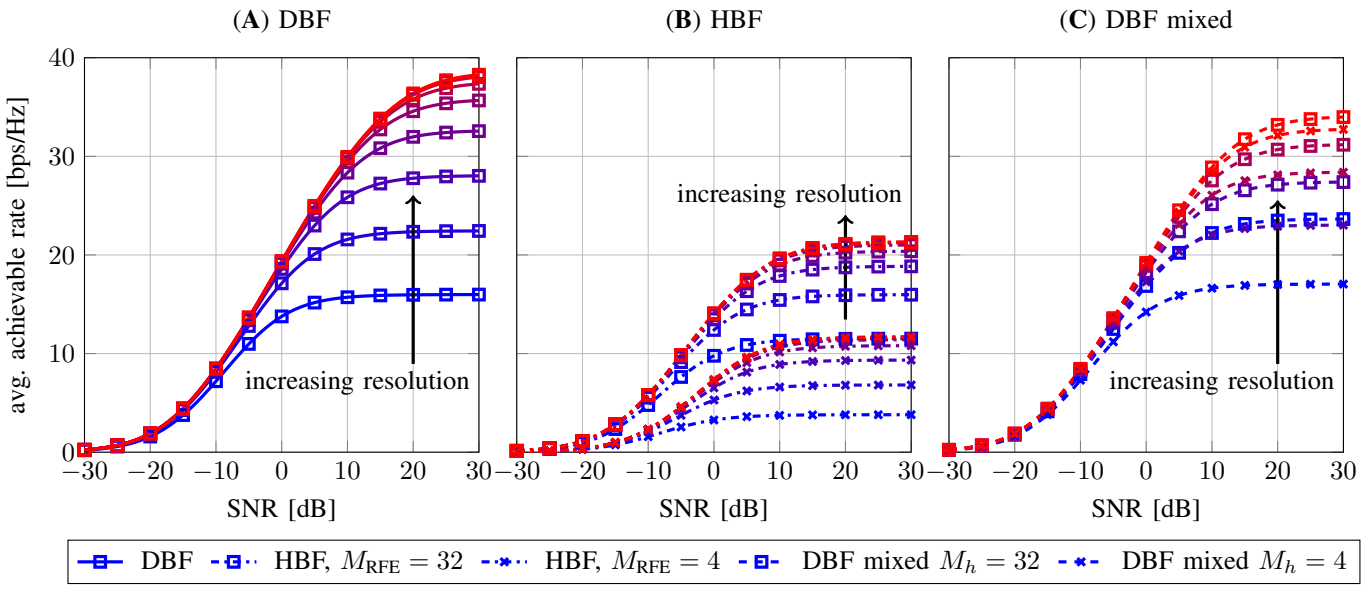

Fig. 4. DBF, HBF and DBF mixed average achievable rate for $M_R = 64$, $U = 4$, $M_{\text{RFE}} \in \{4, 32\}$, $M_h \in \{4, 32\}$ and ADC resolution $b \in \{1, \cdots, 8\}$.

each user and compare this to the quantization noise dependent on the resolution. The Signal-to-Quantization-Noise-Ratio (SQNR) in dB of an ADC can be well approximated by

$$\text{SQNR} \approx 6.02 \cdot b, \tag{45}$$

with b being the resolution of the ADC. This approximation is similar to the relationship shown in [30] between distortion introduced by an ADC assuming a Gaussian input. 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

$$P_n = U \cdot 10^{\text{EVM}/10} P_u = 4 \cdot 10^{-25/10} P_u = 1.26 \cdot 10^{-2} P_u, \tag{46}$$

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

$$P_s = U^2 P_u = 16.0 P_u. \tag{47}$$

Therefore the maximum SNR in dB is equal to

$$\text{SNR}_{\text{max}} = 10 \log_{10} \left(\frac{U^2 P_u}{U \cdot 10^{\text{EVM}/10} P_u} \right) = 31.0 \text{ dB}. \tag{48}$$

Since the quantization noise should be sufficiently smaller than the received noise we can state that SQNR $>$ SNR$_{\text{max}}$. For our simulation parameters this leads to the condition that $b > 5.15$ bits, which excatly matches our simulation results in Fig. 4 (**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. 4 (**B**).

We can also calculate the minimum resolution at which the ADC is not limiting the performance for HBF, but we need a little more 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

$$P_s = (M_C + U - 1)^2 P_u = 361 P_u. \tag{49}$$

Since the major noise component is the transmit noise and the analog combiner is adapted to this users 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

$$P_n = \left(M_C^2 + U - 1 \right) 10^{\text{EVM}/10} * P_u = 0.819 P_u. \tag{50}$$

The maximum SNR before the ADC can now be calculate as

$$\begin{aligned} \text{SNR}_{\text{max}} &= 10 \log_{10} \left(\frac{(M_C + U - 1)^2 P_u}{\left(M_C^2 + U - 1 \right) 10^{\text{EVM}/10} * P_u} \right) \\ &= 26.4 dB. \end{aligned} \tag{51}$$

Since again the statement SQNR $>$ SNR$_{\text{max}}$ has to hold for the system not being limited by the quantization. We then arrive at the condition that the resolution b should be greater than 4.38 bit. This is the same result as in Fig. 4 (**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

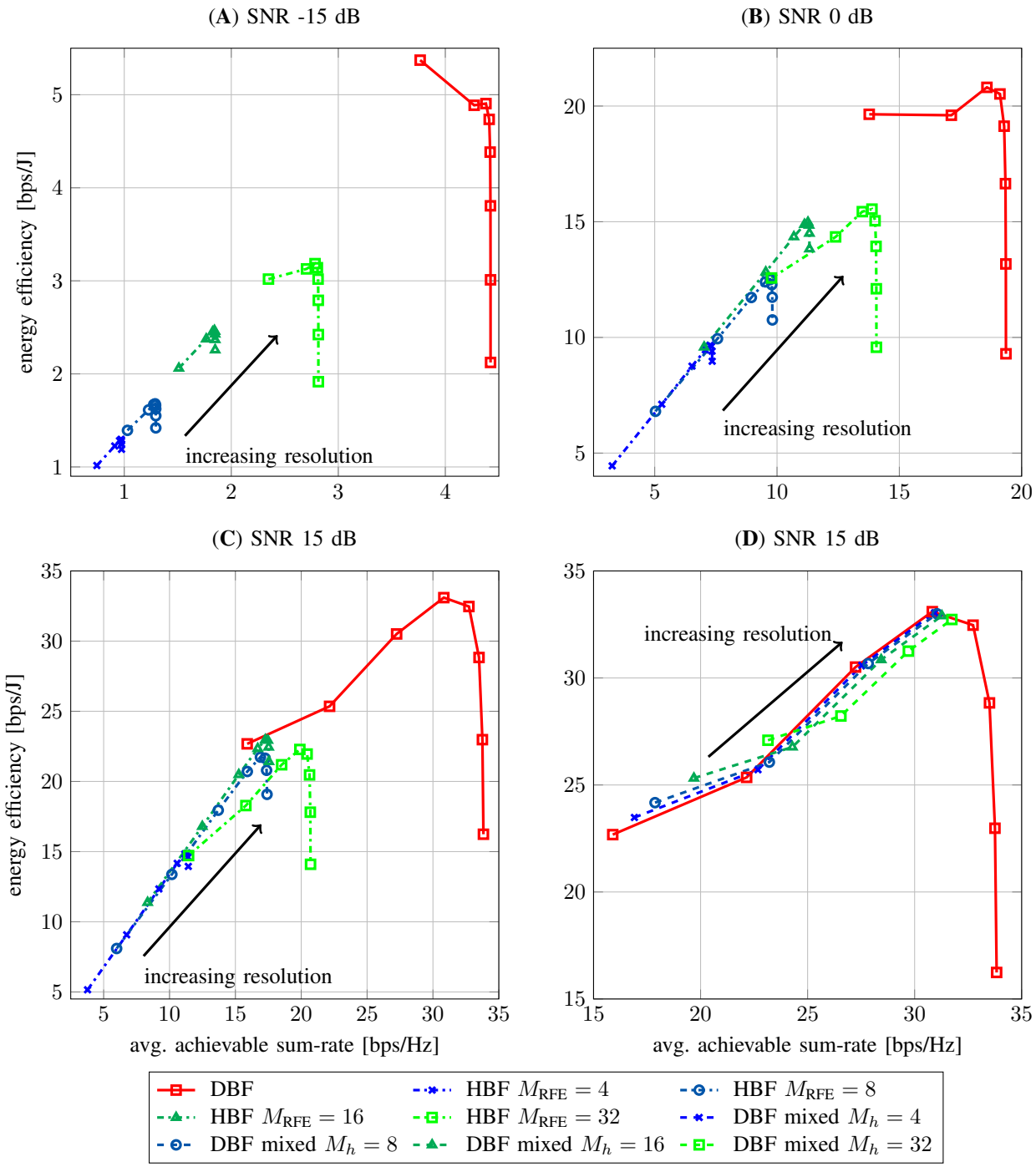

Fig. 5. Spectral and energy efficiency of digital beamforming with/without mixed ADC configuration and hybrid beamforming with $M_R = 64$, $U = 4$, $M_{\text{RFE}} \in \{4, 8, 16, 32\}$, $M_{\text{h}} \in \{4, 8, 16, 32\}$ and ADC resolution $b \in \{1, \cdots, 8\}$, $b_l \in \{1, \cdots, 4\}$ and $b_h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

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. 4 (**A**) are on top of each other.

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.

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

$$\text{energy efficiency} = \frac{R}{P_R}. \tag{52}$$

The scenarios in Fig. 5 (**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. 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 achieve a higher rate than hybrid beamforming systems for multiuser scenarios. 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 rates 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 an interesting scenario to evaluate.

Appendix

A. Output Correlation after Quantization

In [43] we showed how to calculate the output correlation of a quantized system from the input correlation of a Gaussian signal. However in [36] 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

$$\rho_o = \sum_{l=1}^{N_a-1} \sum_{j=1}^{N_c-1} a_l^r c_j^r \int_0^{\rho_i'} f_{ac}(a_l^s, c_j^s, \rho_i)\, d\rho_i, \tag{53}$$

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

$$f_{ac}(a, c, \rho_i) = \frac{1}{2\pi\sigma_a\sigma_c\sqrt{1-\rho_i^2}} \exp\left(-\frac{1}{2(1-\rho_i^2)}\left[\frac{a^2}{\sigma_a^2} + \frac{c^2}{\sigma_c^2} - \frac{2\rho_i ac}{\sigma_a\sigma_c}\right]\right). \tag{54}$$

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 represent the 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 an AGC loop. Since we need to perform this transformation for every antenna pair, calculating the integral for every point is a large overhead. Therefore, we generate 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 such a way that the change of ρ_o between adjacent points in the grid does 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

$$[\boldsymbol{R_{rr}}]_{i,i} = (1 - \sigma_{qi}^2)\,[\boldsymbol{R_{yy}}]_{i,i}\,, \tag{55}$$

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 (53) 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 $\mathrm{T}(\cdot)$

$$\boldsymbol{R_{rr}} = \mathrm{T}\left(\boldsymbol{R_{yy}}, Q^1(\cdot), \cdots, Q^{M_{\mathrm{RFE}}}\right). \tag{56}$$

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

$$[\boldsymbol{F}]_{i,i} = (1 - \sigma_{qi}). \tag{57}$$

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

$$\boldsymbol{F} = \mathrm{TF}(Q^1(\cdot), \cdots, Q^{M_{\mathrm{RFE}}}). \tag{58}$$

B. List of Symbols

The list of symbols can be found in Table II.

TABLE II
LIST OF SYMBOLS

symbol	meaning
U	number of users
u	user index
M_T	number of transmit antennas
M_C	number of antennas per subarray
M_{RFE}	number of RF chains
M_R	number of receive antennas
n	discrete time index
$\boldsymbol{x}_u[n]$	Tx signal
$\boldsymbol{\eta}_u[n]$	Tx impairments
$\boldsymbol{H}_u[l]$	channel to the basestation
l	delay index
L_u	maximum delay of the channel
$\boldsymbol{\eta}_R[n]$	noise at the receiver
$\boldsymbol{y}[n]$	receive signal
P_u	transmit power
$\boldsymbol{\eta}'_R[n]$	combination of all noise components at the receiver
$\boldsymbol{W}_R$	analog combining matrix
i	subarray index
$\boldsymbol{w}_R^i$	analog combining vector
$\boldsymbol{y}_C[n]$	receive signal after analog combining
$\boldsymbol{r}[n]$	signal after quantization
b	resolution in bits
$Q_b(\cdot)$	quantization operation
M_h	number of high resolution ADCs
b_h	resolution of high resolution ADCs
M_l	number of low resolution ADCs
b_l	resolution of low resolution ADCs
γ_u	average per-antenna SNR
k	subcarrier index
ℓ	OFDM symbol index
m	receive antenna index
$Y_{(k,\ell,m)}$	receive signal
$X_{(k,\ell,m)}$	transmit signal
$H_{(k,\ell,m)}$	channel
$\eta_{(k,\ell,m)}$	noise
$\hat{\boldsymbol{h}}$	vectorized channel estimate
$\boldsymbol{h}$	vectorized channel
$\boldsymbol{A}_s$	spatial interpolation matrix
$\boldsymbol{A}_t$	temporal interpolation matrix
$\boldsymbol{A}_f$	frequency interpolation matrix
$\boldsymbol{A}_{stf}$	combined interpolation matrix
$\boldsymbol{R}_{\boldsymbol{hh}}^s$	spatial correlation matrix of the channel
$\boldsymbol{R}_{\boldsymbol{hh}}^t$	temporal correlation matrix of the channel
$\boldsymbol{R}_{\boldsymbol{hh}}^f$	frequency correlation matrix of the channel
$\boldsymbol{R}_{\boldsymbol{hh}}$	combined correlation matrix of the channel
P_R	power consumption of the receiver front-end
R	sum data rate

ACKNOWLEDGMENT

This work has been performed in the framework of the Horizon 2020 project ONE5G (ICT-760809) receiving funds from the European Union. H. Pirzade and A. Swindlehurst were supported by the U.S. National Science Foundation under Grants ECCS-1547155 and CCF-1703635. A. Swindlehurst was also supported by a Hans Fischer Senior Fellowship from the Technische Universität München Institute for Advanced Study. 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 et all, "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, 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 et all., "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 et all., "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 et all., "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] D. C. Daly et al., "Through the looking glass - the 2017 edition: Trends in solid-state circuits from isscc," *IEEE Solid-State Circuits Magazine*, vol. 9, no. 1, pp. 12–22, 2017.

[25] A. M. Niknejad et al., "A circuit designer's guide to 5g mm-wave," in *2015 IEEE Custom Integrated Circuits Conference (CICC)*, Sept 2015, pp. 1–8.

[26] E. Bjrnson et al., "Massive mimo systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," *IEEE Transactions on Information Theory*, vol. 60, no. 11, pp. 7112–7139, Nov. 2014.

[27] C. S. et al., "Mimo transmission with residual transmit-rf impairments," in *2010 International ITG Workshop on Smart Antennas (WSA)*, Feb. 2010, pp. 189–196.

[28] P. Zetterberg, "Experimental investigation of tdd reciprocity-based zero-forcing transmit precoding," *EURASIP Journal on Advances in Signal Processing*, vol. 2011, no. 1, p. 137541, Dec 2010. [Online]. Available: https://doi.org/10.1155/2011/137541

[29] M. Wenk, *MIMO-OFDM Testbed: Challenges, Implementations, and Measurement Results*, ser. Series in microelectronics. ETH, 2010. [Online]. Available: https://books.google.pl/books?id=pTPQZwEACAAJ

[30] J. Max, "Quantizing for minimum distortion," *IRE Transactions on Information Theory*, vol. 6, no. 1, pp. 7–12, Mar. 1960.

[31] "Deliverable D2.2 Measurement Results and Final mmMAGIC Channel Models," 2017. [Online]. Available: https://bscw.5g-mmmagic.eu/pub/bscw.cgi/d202656/mmMAGIC_D2-2.pdf

[32] A. N. Mody and G. L. Stuber, "Synchronization for mimo ofdm systems," in *Global Telecommunications Conference, 2001. GLOBECOM '01. IEEE*, vol. 1, 2001, pp. 509–513 vol.1.

[33] 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.

[34] P. Hoeher et al., "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.

[35] 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

[36] 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)*, Mar. 2017.

[37] *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.

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

[39] R. Ratasuk et al., "Nb-iot system for m2m communication," in *Wireless Communications and Networking Conference (WCNC) 2016*, April 2016, pp. 1–5.

[40] G. D. Forney and G. Ungerboeck, "Modulation and coding for linear Gaussian channels," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2384–2415, Oct 1998.

[41] D. De Donno et al., "Millimeter-wave beam training acceleration through low-complexity hybrid transceivers," *IEEE Transactions on Wireless Communications*, vol. 16, no. 6, pp. 3646–3660, June 2017.

[42] W. Feller, "The Fundamental Limit Theorems in Probability," *Bull. Amer. Math. Soc.*, vol. 51, pp. 800–832, 1945.

[43] 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.

Kilian Roth is since 2014 pursuing his Ph.D. degree at the Department of Electrical and Computer Engineering at the Technical University Munich (TUM). In 2014, he joined the Next Generation and Standards (NGS) at Intel as a Ph.D. student. From 2013 to 2014 he was an undergrad intern at Intel working on synchronization, channel estimation and demodulation aspects of LTE ICIC and LTE new carrier type. He graduated from Munich University of Applied Sciences (MUAS) in 2011 with a B.Eng. and obtained a M.Sc. from TUM in 2014. He was an active participant in the projects Flex5GWare and mmMagic and is currently working in the project ONE5G, all funded by the EC under the H2020 Framework Program. His main research interests are signal processing algorithms and systems concepts to enable energy efficient systems at millimeter and centimeter wave frequencies.

Hessam Pirzadeh received the B.Sc. degree in engineering science from the University of Tehran, Tehran, Iran, and the M.Sc. degree in electrical engineering from Iran University of Science and Technology (IUST), Tehran, Iran, in 2013 and 2016, respectively. Since 2016, he has been working toward the Ph.D. degree in the Center for Pervasive Communications and Computing, University of California Irvine, Irvine, CA, USA. His research interests span topics in signal processing and wireless communications with an emphasis on massive MIMO systems.

A. LEE SWINDLEHURST received the B.S. (1985) and M.S. (1986) degrees in Electrical Engineering from Brigham Young University (BYU), and the PhD (1991) degree in Electrical Engineering from Stanford University. He was with the Department of Electrical and Computer Engineering at BYU from 1990-2007, where he served as Department Chair from 2003-06. During 1996-97, he held a joint appointment as a visiting scholar at Uppsala University and the Royal Institute of Technology in Sweden. From 2006-07, he was on leave working as Vice President of Research for ArrayComm LLC in San Jose, California. Since 2007 he has been a Professor in the Electrical Engineering and Computer Science Department at the University of California Irvine, where he served as Associate Dean for Research and Graduate Studies in the Samueli School of Engineering from 2013-16. During 2014-17 he was also a Hans Fischer Senior Fellow in the Institute for Advanced Studies at the Technical University of Munich. His research focuses on array signal processing for radar, wireless communications, and biomedical applications, and he has over 300 publications in these areas. Dr. Swindlehurst is a Fellow of the IEEE and was the inaugural Editor-in-Chief of the IEEE Journal of Selected Topics in Signal Processing. He received the 2000 IEEE W. R. G. Baker Prize Paper Award, the 2006 IEEE Communications Society Stephen O. Rice Prize in the Field of Communication Theory, the 2006 and 2010 IEEE Signal Processing Societys Best Paper Awards, and the 2017 IEEE Signal Processing Society Donald G. Fink Overview Paper Award.

Josef A. Nossek (S'72M'74SM'81F'93LF'13) received the Dipl.-Ing. and Dr. techn. degrees in electrical engineering from Vienna University of Technology, Austria, in 1974 and 1980, respectively. He joined Siemens AG, Germany, in 1974, where he was engaged in filter design for communication systems. From 1987 to 1989, he was the Head of the Radio Systems Design Department, where he was instrumental in introducing high-speed VLSI signal processing into digital microwave radio. From 1989 to 2016, he has been the Head of the Institute of Circuit Theory and Signal Processing (NWS) at the Technical University Munich, Germany. In 2016, he joined the Universidade Federal do Ceara, Fortaleza, Brasil as a Full Professor. He was the President Elect, President, and Past President of the IEEE Circuits and Systems Society in 2001, 2002, and 2003, respectively. He was the President of Verband der Elektrotechnik, Elektronik, und Informationstechnik (VDE) in 2007 - 2008 and the President of the Convention of National Associations of Electrical Engineers of Europe (EUREL) in 2013. He received the ITG Best Paper Award in 1988, the Mannesmann Mobilfunk (currently Vodafone) Innovations Award in 1998, and the Award for Excellence in Teaching from the Bavarian Ministry for Science, Research and Art in 1998. From the IEEE Circuits and Systems Society, he received the Golden Jubilee Medal for Outstanding Contributions to the Society in 1999 and the Education Award in 2008. He received the Order of Merit of the Federal Republic of Germany (Bundesverdienstkreuz am Bande) in 2008. In 2009 he became member of the National Academy of Engineering in Germany (acatech), in 2011 he received the IEEE Guillemin-Cauer Best Paper Award and in 2013 the honorary doctorate (Dr. h. c.) from the Peter Pazmany Catholic University, Hungary. He received the VDE Ring of Honor in 2014 and the TUM Emeritus of Excellence in 2016.