Linear Precoding with Low-Resolution DACs for Massive MU-MIMO-OFDM Downlink

Sven Jacobsson *Student Member, IEEE*, Giuseppe Durisi *Senior Member, IEEE*,
Mikael Coldrey *Member, IEEE*, and Christoph Studer *Senior Member, IEEE*

Abstract—We consider the downlink of a massive multiuser (MU) multiple-input multiple-output (MIMO) system in which the base station (BS) is equipped with low-resolution digital-to-analog converters (DACs). In contrast to most existing results, we assume that the system operates over a frequency-selective wideband channel and uses orthogonal frequency division multiplexing (OFDM) to simplify equalization at the user equipments (UEs). Furthermore, we consider the practically relevant case of oversampling DACs. We theoretically analyze the uncoded bit error rate (BER) performance with linear precoders (e.g., zero forcing) and quadrature phase-shift keying using Bussgang's theorem. We also develop a lower bound on the information-theoretic sum-rate throughput achievable with Gaussian inputs, which can be evaluated in closed form for the case of 1-bit DACs. For the case of multi-bit DACs, we derive approximate, yet accurate, expressions for the distortion caused by low-precision DACs, which can be used to establish lower bounds on the corresponding sum-rate throughput. Our results demonstrate that, for a massive MU-MIMO-OFDM system with a 128-antenna BS serving 16 UEs, only 3–4 DAC bits are required to achieve an uncoded BER of 10^{-4} with a negligible performance loss compared to the infinite-resolution case at the cost of additional out-of-band emissions. Furthermore, our results highlight the importance of taking into account the inherent spatial and temporal correlations caused by low-precision DACs.

I. INTRODUCTION

Massive multiuser (MU) multiple-input multiple-output (MIMO) will be a key technology in future cellular communication systems [2]. Massive MU-MIMO equips the base station (BS) with hundreds of active antenna elements to serve simultaneously tens of user equipments (UEs), which results in

improved spectral efficiency and energy efficiency compared to traditional, small-scale MIMO systems [3]–[6]. Furthermore, all of these advantages can be achieved by means of simple signal processing schemes (e.g., linear precoding) at the BS [3], [4].

Increasing the number of active antenna elements at the BS by orders of magnitude increases the circuit power consumption. Data converters, i.e., analog-to-digital converters (ADCs) and digital-to-analog converters (DACs), are foreseen to be among the most dominant sources of power consumption in massive MU-MIMO systems. In today's state-of-the-art direct-conversion MIMO systems, a pair of high-resolution DACs (e.g., with 10 bits of resolution per real dimension or more) is used to generate the in-phase and quadrature components of the transmitted baseband signal at each antenna element at the BS. Scaling such high-resolution architectures to massive MU-MIMO systems, in which the number of antennas at the BS could be in the order of hundreds, would lead to excessively high power consumption. Hence, to maintain a reasonable power budget, the resolution of the DACs must be reduced. Furthermore, equipping the BS with hundreds of active antenna elements puts extreme data rate requirements on the interface connecting the baseband-processing unit to the radio unit (where the DACs are located), especially for wideband systems operating in the millimeter wave part of the wireless spectrum. By lowering the resolution of the DACs, one can—to some extent—mitigate this data rate bottleneck. In this work, we consider the massive MU-MIMO *downlink* (BS transmits data to multiple UEs) and focus on the case in which low-resolution DACs are used at the BS.

A. Relevant Prior Art

1) Transmitter Impairments and Out-of-Band Emissions:
The impact of *aggregate* radio frequency (RF) hardware impairments (i.e., from multiple sources of hardware impairments) in the massive MU-MIMO downlink has been investigated in, e.g., [7]–[9], using a model that treats hardware impairments as power-dependent additive white noise that is uncorrelated with the input signal. Such models have been validated experimentally [10] under the assumptions that (i) the system uses orthogonal frequency-division multiplexing (OFDM) and (ii) methods have been implemented at the BS to mitigate common hardware impairments. The latter indicates that the impairments in [7]–[9] capture only *residual* distortions that have not been entirely compensated. Furthermore, such aggregate models are not useful for analyzing out-of-band (OOB) emissions [11], which are of significant concern

in practice. In general, distortions (in-band and OOB) caused by transmitter impairments are beamformed, to some extent, in the direction of the useful signal (i.e., they are correlated over the BS antenna array). This behavior has been illustrated in, e.g., [12]–[14], which analyzed the specific distortion caused by nonlinear power amplifiers (PAs). In [15], focusing again on the distortion caused by nonlinear PAs, it was shown that, in some cases, ignoring the spatial correlation leads to marginal errors when characterizing spectral efficiency. As we shall see in Section V, similar conclusions hold for the distortion caused by DACs with medium-to-high resolution. However, for low-resolution (e.g., 1 bit) DACs, the correlation of the distortion should not be ignored when characterizing bit-error rate (BER) and spectral efficiency.

2) Low-Resolution ADCs: The impact on performance of using low-resolution ADCs to quantize the received signal at the BS in the massive MU-MIMO *uplink* (multiple UEs transmit to the BS) has been analyzed in, e.g., [16]–[20]. These analyses suggest that reliable communication is possible even in the extreme case of 1-bit ADCs. Furthermore, low-resolution ADCs can be deployed at the BS with negligible performance degradation compared to the infinite-resolution (i.e., no quantization) case, provided that the ratio between the number of BS antennas and UEs is sufficiently large. To facilitate an analytical performance analysis, simple approximate models that treat the distortion caused by the ADCs as a white additive noise that is uncorrelated with the input signal, have been used extensively in the literature (see, e.g., [21]–[23]). Although these models are accurate under certain assumptions on, e.g., the resolution and the sampling rate of the converters, and on the number of active users, they fail to take into account the inherent spatial and temporal correlation of the quantization distortion [24], [25]. To assess the accuracy of these simplified models, we will develop a distortion model that takes this correlation into account and compare it to simplistic distortion models.

3) Low-Resolution DACs: Linear precoders, such as maximal-ratio transmission (MRT), zero-forcing (ZF), and Wiener-filter (WF) precoding for narrowband systems with 1-bit DACs at the BS have been analyzed in, e.g., [26]–[29]. The results therein show that 1-bit-DAC massive MU-MIMO systems support low BERs and high sum-rate throughputs despite the severe nonlinearity caused by 1-bit DACs. Multi-bit DACs and linear precoding for the massive MU-MIMO downlink were considered in [29]. There, it was shown that DACs with few bits (e.g., 3–4 bits) yield a performance that is close to the infinite-resolution case if the BS has access to perfect channel state information (CSI). A linear WF-quantized (WFQ) precoder was proposed in [28] on the basis of an approximate model for the distortion caused by the DACs. This precoder is shown to outperform conventional precoders for small-to-moderate MIMO systems at high signal-to-noise ratio (SNR). However, as shown in [29], the performance gain of the WFQ precoder over conventional linear precoders (e.g., ZF) is marginal in the massive MU-MIMO case.

For the case of 1-bit DACs, more sophisticated nonlinear precoding strategies that significantly outperform linear precoders at the cost of an increased signal-processing complexity

were presented in, e.g., [29]–[39].

All existing results on low-resolution DACs reviewed so far dealt with the frequency-flat case. The first work to consider 1-bit DACs combined with (OFDM) for massive MU-MIMO over frequency-selective channels is [40]. There, by using an approximate model that treats the 1-bit-DAC distortion as white, it is shown that simple MRT precoding results in high signal-to-interference-noise-and-distortion ratio (SINDR) at the UEs provided that the number of BS antennas is significantly larger than the number of UEs. In the conference version of this paper [1], we used an exact model for the 1-bit-DAC distortion to characterize the BER achievable in an OFDM system with quadrature phase-shift keying (QPSK), as well as the sum-rate throughput achievable with a Gaussian codebook and nearest-neighbor decoding, for the case of linear precoding followed by oversampling 1-bit DACs. This paper complements the analysis previously reported in [1] by generalizing it to oversampling multi-bit DACs.

Finally, nonlinear precoding algorithms for the 1-bit-DAC case have been extended recently to OFDM and frequency-selective channels in [41]–[43].

B. Contributions

We analyze the performance achievable in the massive MU-MIMO-OFDM downlink with linear precoding and finite-resolution DACs. In contrast to existing results [26]–[37], which consider symbol-rate DACs and single-carrier modulation over frequency-flat channels, we focus on oversampling DACs and OFDM-based transmission over frequency-selective channels. We focus on conventional linear precoders, namely MRT and ZF precoding. These precoders are of practical interest because they entail relatively low complexity compared to nonlinear precoders [41]–[43]. Our main contributions can be summarized as follows:

- Using Bussgang's theorem [44], [45], we develop a lower bound on the information-theoretic sum-rate achievable with linear precoding and oversampling finite-resolution DACs. This lower bound can be achieved by Gaussian signaling and nearest-neighbor decoding at the UEs. We also show how to evaluate the lower bound for the case of 1-bit DACs using Van Vleck's arcsine law [46].
- For the case of multi-bit DACs, we develop an approximate, yet accurate, model for the distortion caused by the DACs, which takes into account the inherent spatial and temporal correlation in the DAC distortion. This approach yields an accurate approximation for the SINDR at the UE side, which we use to evaluate the sum-rate lower bound and to derive an approximation of the uncoded BER achievable with QPSK. We also briefly discuss how this approximation can be used to characterize OOB emissions caused by low-resolution DACs.
- We derive a simpler approximation for the DAC distortion, which treats this distortion as white (i.e., uncorrelated in both the spatial and the temporal domains). We show that such a crude model is accurate for medium-to-high resolution DACs and when the oversampling ratio (OSR) is not too high, but is not sufficient to accurately describe the distortion caused by low-resolution (e.g., 1 bit) DACs.

- We demonstrate through extensive numerical simulations that massive MU-MIMO-OFDM with low-resolution DACs enables excellent performance in terms of achievable rate and uncoded/coded BER. Specifically, we show that only a few DAC bits (e.g., 3–4 bits) are sufficient to approach the performance of systems that use infinite-resolution DACs.

C. Notation

Lowercase and uppercase boldface letters represent column vectors and matrices, respectively. For a matrix $\mathbf{A}$, we denote its complex conjugate, transpose, and Hermitian transpose by $\mathbf{A}^*$, $\mathbf{A}^T$, and $\mathbf{A}^H$, respectively. The entry on the kth row and on the ℓth column of the matrix $\mathbf{A}$ is denoted as $[\mathbf{A}]_{k,\ell}$. The kth entry of a vector $\mathbf{a}$ is denoted as $[\mathbf{a}]_k$. The trace and the main diagonal of $\mathbf{A}$ are $\mathrm{tr}(\mathbf{A})$ and $\mathrm{diag}(\mathbf{A})$, respectively. The square matrix $\mathrm{diag}(\mathbf{a})$ is diagonal with the elements of the vector $\mathbf{a}$ along its main diagonal. If $\mathbf{A}$ is an $M \times N$ matrix, then $\mathrm{vec}(\mathbf{A})$ is the MN-dimensional vector obtained by stacking the columns (taken from left to right) of $\mathbf{A}$. The $M \times M$ identity matrix, the $M \times N$ all-zeros matrix, and the $M \times N$ all-ones matrix are denoted by $\mathbf{I}_M$, $\mathbf{0}_{M \times N}$, and $\mathbf{1}_{M \times N}$, respectively. The real and the imaginary parts of a complex vector $\mathbf{a}$ are denoted by $\Re\{\mathbf{a}\}$ and $\Im\{\mathbf{a}\}$, respectively. We use $\|\mathbf{a}\|$ to denote the ℓ_2-norm of $\mathbf{a}$. The Kronecker product of two matrices $\mathbf{A}$ and $\mathbf{B}$ is $\mathbf{A} \otimes \mathbf{B}$. For equally-sized matrices $\mathbf{A}$ and $\mathbf{B}$, we denote by $\mathbf{A} \odot \mathbf{B}$ the Hadamard (element-wise) product. We use $\mathrm{sgn}(\cdot)$ to denote the signum function, which is applied entry-wise to vectors and defined as $\mathrm{sgn}(a) = 1$ if $a \geq 0$ and $\mathrm{sgn}(a) = -1$ if $a < 0$. We further use $\mathbb{1}_{\mathcal{A}}(a)$ to denote the indicator function, which is defined as $\mathbb{1}_{\mathcal{A}}(a) = 1$ for $a \in \mathcal{A}$ and $\mathbb{1}_{\mathcal{A}}(a) = 0$ for $a \notin \mathcal{A}$. The multivariate complex-valued circularly-symmetric Gaussian distribution with covariance matrix $\mathbf{K} \in \mathbb{C}^{N \times N}$ is denoted by $\mathcal{CN}(\mathbf{0}_{N \times 1}, \mathbf{K})$. We use $\mathbb{P}[\mathcal{E}]$ to denote the probability of the event $\mathcal{E}$ and $\mathbb{E}[x]$ to denote the expected value of x. The cumulative distribution function of the standard normal distribution is $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(-u^2/2) \, \mathrm{d}u$.

D. Paper Outline

The rest of the paper is organized as follows. In Section II, we introduce the system model as well as the considered linear precoders, and we describe the operation of the DACs. In Section III, we review Bussgang's theorem, derive the SINDR at the UEs, and develop a lower bound on the rate achievable with Gaussian inputs and nearest-neighbor decoding. In Section IV, we derive exact and approximate expressions for the distortion caused by finite-resolution DACs. We also use these expressions to evaluate the achievable-rate lower bound from Section III and the uncoded BER achievable with QPSK. In Section V, we provide numerical simulations that demonstrate the accuracy of our analytical results. We conclude in Section VI.

II. SYSTEM MODEL

We consider the single-cell downlink massive MU-MIMO-OFDM system depicted in Fig. 1. The system consists of a BS with B antennas that simultaneously serves U single-antenna UEs in the same time-frequency resource using spatial multiplexing. Our model includes finite-resolution DACs at the BS. Specifically, the in-phase and quadrature components of the time-domain per-antenna transmitted signal at the BS are generated using a pair of finite resolution DACs. Let $\mathcal{L} = \{\ell_0, \ell_1, \ldots, \ell_{L-1}\}$, where $\ell_0 < \ell_1 < \cdots < \ell_{L-2} < \ell_{L-1}$, denote the set of L possible DAC outcomes, i.e., the set of possible amplitude (voltage) levels that are supported by the transcoder in the DACs. The DACs have finite resolution, and, hence, the set $\mathcal{L}$ has finite cardinality. We refer to $L = |\mathcal{L}|$ and to $\log_2(L)$ as the number of DAC *levels* and as the number of DAC *bits* per real dimension, respectively. We assume all DACs at the BS to be the same. Hence, the set of complex-valued DAC outputs at each antenna is $\mathcal{X} = \mathcal{L} \times \mathcal{L}$.

The system operates over a wideband channel and OFDM is used to deal with the frequency selectivity of the channel. Specifically, at the BS, the frequency-domain precoded vector is mapped to time domain by performing an inverse discrete Fourier transform (IDFT) at each antenna element. At the UEs, the time-domain received signal is transformed back to frequency domain through a discrete Fourier transform (DFT). As we will show in Section V-A, the nonlinearity introduced by the finite-resolution DACs will result in intercarrier interference (ICI).

We assume that each OFDM symbol consists of N time-domain samples. Let Δf be the subcarrier spacing and $f_s = N\Delta f$ be the sampling rate of the DACs. We use the disjoint sets $\mathcal{S}_d$ and $\mathcal{S}_g$ to denote the set of subcarriers designated for the $S \leq N$ data symbols (occupied subcarriers) and the set of $N - S$ guard subcarriers, respectively. Let $\mathbf{s}_k$ denote the U-dimensional data vector associated with the kth subcarrier ($k = 0, 1, \ldots, N - 1$). We assume that $\mathbf{s}_k = \mathbf{0}_{U \times 1}$ for $k \in \mathcal{S}_g$ and that $\mathbb{E}[\mathbf{s}_k \mathbf{s}_k^H] = \mathbf{I}_U$ for $k \in \mathcal{S}_d$. The case $S = N$ corresponds to symbol-rate-sampling DACs whereas $S < N$ corresponds to oversampling DACs. The OSR is $\xi = N/S$.

A. Channel Input-Output Relation

To isolate the performance impact of low-resolution DACs, we assume in our analysis that all RF circuitry other than the DACs (e.g., local oscillators, mixers, and PAs) are ideal and that the UEs are equipped with infinite resolution ADCs. We further assume that the sampling rate f_s of the DACs at the BS equals the sampling rate of the ADCs at the UEs and that the system is perfectly synchronized. Finally, we assume that the *reconstruction stage* (see, e.g., [47, Sec. 1.7]) of the DACs is an ideal low-pass filter with cutoff frequency $f_{\mathrm{cut}} = f_s/2$ so that the spectrum of the DAC output is contained within $[-f_s/2, f_s/2]$.[1]

1) Time Domain: With the above assumptions, the discrete-time baseband signal $\mathbf{y}_n \in \mathbb{C}^U$ received at the U UEs at time sample n can be written as

$$\mathbf{y}_n = \sum_{t=0}^{T-1} \mathbf{H}_t \mathbf{x}_{n-t} + \mathbf{w}_n \tag{1}$$

---

[1]A more realistic reconstruction stage, specifically a zero-order hold filter followed by nonideal low-pass filter, is considered in [48]. For this case, the spectrum of the DAC output will not be contained within $[-f_s/2, f_s/2]$.
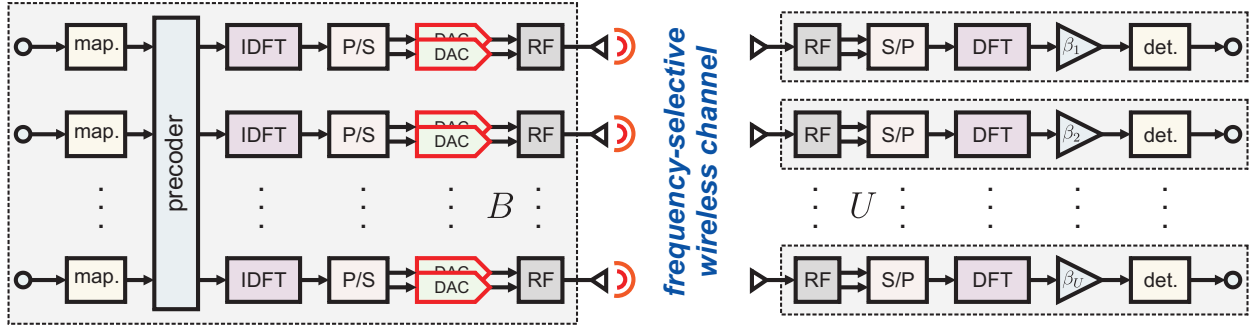
Fig. 1. Overview of the massive MU-MIMO-OFDM downlink with low-resolution DACs at the BS. Left: A B antenna BS performs linear precoding and generates the per-antenna OFDM time-domain signals that are passed through low-resolution DACs (highlighted with red color). Right: U single-antenna UEs perform independently OFDM demodulation and data detection. In the figure, "map." and "det." stand for mapper and detector, respectively.

for $n = 0, 1, \ldots, N - 1$. Here, $\mathbf{x}_n \in \mathcal{X}^B$ is the transmitted signal (i.e., the output of the DACs) at time sample n. The random vector $\mathbf{w}_n \sim \mathcal{CN}(\mathbf{0}_{U \times 1}, N_0 \mathbf{I}_U)$ models additive white Gaussian noise (AWGN) at the uth UE at time sample n. Here, N_0 is the power spectral density (PSD) of the AWGN. The matrix $\mathbf{H}_t \in \mathbb{C}^{U \times B}$ is the time-domain channel matrix associated with the tth tap of the frequency-selective channel ($t = 0, 1, \ldots, T - 1$, where T is the number of taps). We assume that this matrix has entries $[\mathbf{H}_t]_{u,b} \sim \mathcal{CN}(0, T^{-1})$ and that the entries of $\{\mathbf{H}_t\}_{t=0}^{T-1}$ are independent and remain constant over the duration of an OFDM symbol. Note that these assumptions yield a spatially white frequency-selective Rayleigh-fading channel with uniform power-delay profile.[2]

A cyclic prefix of length $T-1$ is prepended to the transmitted signal and is later discarded at the UEs. We shall not explicitly prepend the cyclic prefix to $\{\mathbf{x}_n\}$ to keep notation compact. The cyclic prefix makes the channel matrix circulant and, hence, diagonalizable in the infinite-resolution case through an IDFT and a DFT at the BS and at the UEs, respectively. The transmitted signal $\{\mathbf{x}_n\}$ satisfies the average power constraint

$$\mathbb{E}\left[\sum_{n=0}^{N-1} \|\mathbf{x}_n\|^2 \right] \leq PS. \qquad (2)$$

Here, $0 < P < \infty$ is the average transmit power at the BS. We define $\rho = P/N_0$ as the SNR.

2) Frequency Domain: Let $\mathbf{X} = [\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{N-1}] \in \mathcal{X}^{B \times N}$ and $\mathbf{Y} = [\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_{N-1}] \in \mathbb{C}^{U \times N}$ be the time-domain transmitted and received matrices over the N time samples, respectively. Furthermore, let $\widehat{\mathbf{X}} = \mathbf{X}\mathbf{F}_N^T$ and $\widehat{\mathbf{Y}} = \mathbf{Y}\mathbf{F}_N^T$ be the corresponding frequency-domain matrices. Here, $\mathbf{F}_N$ stands for the $N \times N$ DFT matrix, which satisfies $\mathbf{F}_N \mathbf{F}_N^H = \mathbf{I}_N$. Finally, let $\widehat{\mathbf{H}}_k = \sum_{t=0}^{T-1} \mathbf{H}_t \exp\left(-jk\frac{2\pi}{N}t\right)$ for $k = 0, 1, \ldots, N-1$, be the $U \times B$ frequency-domain channel matrix associated with the kth subcarrier. After discarding the cyclic prefix, we can write the frequency-domain input-output relation at the kth subcarrier as

$$\hat{\mathbf{y}}_k = \widehat{\mathbf{H}}_k \hat{\mathbf{x}}_k + \hat{\mathbf{w}}_k. \qquad (3)$$

Here, $\hat{\mathbf{x}}_k$ and $\hat{\mathbf{y}}_k$ are the kth column of $\widehat{\mathbf{X}}$ and $\widehat{\mathbf{Y}}$, respectively. Furthermore, $\hat{\mathbf{w}}_k \sim \mathcal{CN}(\mathbf{0}_{U \times 1}, N_0 \mathbf{I}_U)$ is the kth column of the matrix $\widehat{\mathbf{W}} = \mathbf{W}\mathbf{F}_N^T$, where $\mathbf{W} = [\mathbf{w}_0, \mathbf{w}_1, \ldots, \mathbf{w}_{N-1}]$. In the following two subsections, we shall relate the frequency-domain representation $\hat{\mathbf{x}}_k$ of the DAC output to the precoded transmitted symbols.

B. Uniform Quantization

For the discrete-time system model (1) considered in this paper (recall that this model assumes that the reconstruction stage in the DACs is an ideal low-pass filter), each DAC can be modeled simply as a *quantizer*,[3] i.e., a nonlinear device that maps a continuous-amplitude signal to a set of discrete numbers [50]. We characterize the quantizer by the set $\mathcal{L} = \{\ell_0, \ell_1, \ldots, \ell_{L-1}\}$ of L quantization labels and the set $\mathcal{T} = \{\tau_0, \tau_1, \ldots, \tau_L\}$, where $-\infty = \tau_0 < \tau_1 < \cdots < \tau_{L-1} < \tau_L = \infty$, of $L + 1$ quantization thresholds. We use the *quantization function* $\mathcal{Q}(\cdot) : \mathbb{C}^B \to \mathcal{X}^B$, which is applied entry-wise to a vector, to describe the joint operation of the $2B$ DACs at the BS. Let $\mathbf{z}_n \in \mathbb{C}^B$ denote the time-domain precoded vector at time sample n.[4] Also, let $z_{b,n} = [\mathbf{z}_n]_b$ and $x_{b,n} = [\mathbf{x}_n]_b$. Then,

$$x_{b,n} = \mathcal{Q}(z_{b,n}) \qquad (4)$$

$$= \sum_{i=0}^{L-1} \ell_i \mathbb{1}_{[\tau_i, \tau_{i+1})}(z_{b,n}^R) + j \sum_{i=0}^{L-1} \ell_i \mathbb{1}_{[\tau_i, \tau_{i+1})}(z_{b,n}^I) \qquad (5)$$

where $z_{b,n}^R = \Re\{z_{b,n}\}$ and $z_{b,n}^I = \Im\{z_{b,n}\}$. For simplicity, we shall model the DACs as symmetric uniform quantizers (the labels $\{\ell_i\}$ are equispaced and symmetric around zero) with step size Δ. For symmetric uniform quantizers, the quantization thresholds are $\tau_i = \Delta\left(i - \frac{L}{2}\right)$ for $i = 1, 2, \ldots, L - 1$ and $\tau_0 = -\infty$, $\tau_L = \infty$. Furthermore, the quantization labels are $\ell_i = \alpha\Delta\left(i - \frac{L}{2} + \frac{1}{2}\right)$ for $i = 0, 1, \ldots, L - 1$. Note that the quantization labels are scaled by a constant α to ensure that the transmit power constraint (2) is satisfied. If L is odd, then the quantizer has a label at zero; we shall refer to such quantizers

---

[2]Extensions to more general fading models that include, e.g., spatial correlation are immediate.

[3]We assume that the baseband processing unit at the BS uses floating-point arithmetic with infinite word length. The impact of having finite word lengths in massive MU-MIMO baseband processing has been investigated in [49].

[4]In Section II-C, we shall describe how the time-domain precoded vectors $\{\mathbf{z}_n\}$ for $n = 0, 1, \ldots, N - 1$ are obtained from the data symbols $\{\mathbf{s}_k\}$ for $k \in \mathcal{S}_d$.

as a *midtread* quantizers. If L is even, then the quantizer has a threshold at zero; we shall refer to such quantizers as a *midrise* quantizers.

The choice of the step size Δ determines the amount of distortion caused by the DACs. If Δ is too small, then there will be significant *overload* distortion (clipping or saturation); if Δ is too large, then there will be significant *granular* distortion. Specifically, let $A_{\text{clip}} = L\Delta/2$ be the *clipping level* of the uniform quantizer. The overload distortion is the error $\alpha^{-1}\mathcal{Q}(z) - z$ occurring if $|z| > A_{\text{clip}}$; the granular distortion is the error $\alpha^{-1}\mathcal{Q}(z) - z$ occurring if $|z| \leq A_{\text{clip}}$. It will turn out important for our analysis to keep the overload distortion negligible compared to the granular distortion. Therefore, we will chose the step size Δ such that the probability of the event $|z| > A_{\text{clip}}$ is "small" (we will discuss this important aspect in Sections IV-B and V). In the extreme case of 1-bit DACs, the quantization function (5) reduces to

$$x_{b,n} = \mathcal{Q}(z_{b,n}) = \frac{\alpha\Delta}{2}\big(\operatorname{sgn}(z_{b,n}^R) + j\operatorname{sgn}(z_{b,n}^I)\big). \quad (6)$$

Here, by setting $\alpha = \sqrt{2P/(\Delta^2\xi B)}$, we ensure that the power constraint (2) is satisfied with equality (recall that $\xi = N/S$ is the OSR).

C. Linear Precoding

At the BS, the data symbols for the U UEs are mapped to the antenna array by a precoder. We focus, in this paper, only on linear precoders because of their low computational complexity and because a performance analysis is analytically tractable.

We assume that the BS has access to perfect CSI,[5] i.e., it has perfect knowledge of the realizations of the frequency-domain channel matrices $\{\widehat{\mathbf{H}}_k\}$ for $k \in \mathcal{S}_d$. With linear precoding, the transmitted vector $\mathbf{x}_n$ can be written as

$$\mathbf{x}_n = \mathcal{Q}(\mathbf{z}_n) \quad (7)$$

where $\mathbf{z}_n \in \mathbb{C}^B$ denotes the time-domain precoded vector at time sample n, which is obtained from the data symbols $\{\mathbf{s}_k\}$ for $k \in \mathcal{S}_d$ as

$$\mathbf{z}_n = \frac{1}{\sqrt{N}}\sum_{k \in \mathcal{S}_d} \widehat{\mathbf{P}}_k \mathbf{s}_k \exp\left(jk\frac{2\pi}{N}n\right) \quad (8)$$

for $n = 0, 1, \ldots, N-1$. In words, the data symbols on the kth subcarrier ($k \in \mathcal{S}_d$) are multiplied in the frequency domain with per-subcarrier *precoding matrices* $\widehat{\mathbf{P}}_k \in \mathbb{C}^{B \times U}$. The resulting frequency-domain precoded vector is then mapped to time domain by an IDFT. We use the convention that $\widehat{\mathbf{P}}_k = \mathbf{0}_{B \times U}$ for $k \in \mathcal{S}_g$. In what follows, we will focus on two linear precoders that are commonly studied in the infinite-resolution case, namely MRT and ZF precoding.

1) MRT Precoding: The MRT precoder maximizes the power directed towards each UE, ignoring MU interference. The MRT precoding matrices are given by

$$\widehat{\mathbf{P}}_k^{\text{MRT}} = \frac{1}{\beta_{\text{MRT}}B}\widehat{\mathbf{H}}_k^H \quad (9)$$

---

for $k \in \mathcal{S}_d$, where

$$\beta^{\text{MRT}} = \sqrt{\frac{1}{PSB^2}\sum_{k \in \mathcal{S}_d} \operatorname{tr}(\widehat{\mathbf{H}}_k\widehat{\mathbf{H}}_k^H)} \quad (10)$$

ensures that the power constraint (2) is satisfied (in the infinite-resolution case)

2) ZF Precoding: With ZF precoding, the BS nulls (in the infinite-resolution case) the MU interference by choosing the pseudo-inverse of the channel matrix as precoding matrix. For $k \in \mathcal{S}_d$, the ZF precoding matrices are given by

$$\widehat{\mathbf{P}}_k^{\text{ZF}} = \frac{1}{\beta^{\text{ZF}}}\widehat{\mathbf{H}}_k^H\left(\widehat{\mathbf{H}}_k\widehat{\mathbf{H}}_k^H\right)^{-1} \quad (11)$$

for $k \in \mathcal{S}_d$, where

$$\beta^{\text{ZF}} = \sqrt{\frac{1}{PS}\sum_{k \in \mathcal{S}_d} \operatorname{tr}\left(\widehat{\mathbf{H}}_k\widehat{\mathbf{H}}_k^H\right)^{-1}} \quad (12)$$

ensures that the power constraint (2) is satisfied (in the infinite-resolution case).

III. PERFORMANCE ANALYSIS

In the infinite-resolution case, the frequency-domain received signal $\hat{\mathbf{y}}_k$ in (3) can be written as

$$\hat{\mathbf{y}}_k = \widehat{\mathbf{H}}_k\widehat{\mathbf{P}}_k\mathbf{s}_k + \hat{\mathbf{w}}_k. \quad (13)$$

In words, the received signal on subcarrier $k \in \mathcal{S}_d$ depends only on $\mathbf{s}_k$ and not on the data symbols transmitted on other subcarriers. Hence, each subcarrier can be analyzed separately. In the finite-resolution-DAC case, however, the nonlinearity introduced by the finite-resolution DACs through (5) makes the received signal on one subcarrier depend, in general, on the data symbols transmitted on all other subcarriers. To enable a performance analysis, it is convenient to write the frequency-domain received signal $\widehat{\mathbf{Y}}$ in vectorized form $\hat{\mathbf{y}} = \operatorname{vec}(\widehat{\mathbf{Y}}) \in \mathbb{C}^{UN}$ as

$$\hat{\mathbf{y}} = \widehat{\mathbf{H}}\hat{\mathbf{x}} + \hat{\mathbf{w}}. \quad (14)$$

Here, $\hat{\mathbf{x}} = \operatorname{vec}(\widehat{\mathbf{X}}) \in \mathbb{C}^{BN}$, $\hat{\mathbf{w}} = \operatorname{vec}(\widehat{\mathbf{W}}) \in \mathbb{C}^{UN}$, and $\widehat{\mathbf{H}}$ is the $UN \times BN$ block-diagonal matrix that has the matrices $\widehat{\mathbf{H}}_0, \widehat{\mathbf{H}}_1, \ldots, \widehat{\mathbf{H}}_{N-1}$ on its main diagonal. Next, we rewrite $\hat{\mathbf{x}}$ as

$$\hat{\mathbf{x}} = \operatorname{vec}(\widehat{\mathbf{X}}) = \operatorname{vec}(\mathbf{X}\mathbf{F}_N^T) = (\mathbf{F}_N \otimes \mathbf{I}_B)\mathbf{x} \quad (15)$$

where $\mathbf{x} = \operatorname{vec}(\mathbf{X}) = \mathcal{Q}(\mathbf{z}) \in \mathcal{X}^{BN}$, with $\mathbf{z} = \operatorname{vec}(\mathbf{Z}) \in \mathbb{C}^{BN}$ and $\mathbf{Z} = [\mathbf{z}_0, \mathbf{z}_1, \ldots, \mathbf{z}_{N-1}]$. Now let $\widehat{\mathbf{P}} \in \mathbb{C}^{BN \times UN}$ denote the block-diagonal matrix that contains the matrices $\widehat{\mathbf{P}}_0, \widehat{\mathbf{P}}_1, \ldots, \widehat{\mathbf{P}}_{N-1}$ on its main diagonal. With these definitions, we can compactly write the discrete-time precoded vector $\mathbf{z}$ as

$$\mathbf{z} = \left(\mathbf{F}_N^H \otimes \mathbf{I}_B\right)\widehat{\mathbf{P}}\mathbf{s} \quad (16)$$

where $\mathbf{s} = \operatorname{vec}(\mathbf{S})$ and $\mathbf{S} = [\mathbf{s}_0, \mathbf{s}_1, \ldots, \mathbf{s}_{N-1}]$. Now, using (15) and (16) in (14), we obtain

$$\hat{\mathbf{y}} = \widehat{\mathbf{H}}(\mathbf{F}_N \otimes \mathbf{I}_B)\mathbf{x} + \hat{\mathbf{w}} \quad (17)$$

$$= \widehat{\mathbf{H}}(\mathbf{F}_N \otimes \mathbf{I}_B)\mathcal{Q}\left(\left(\mathbf{F}_N^H \otimes \mathbf{I}_B\right)\widehat{\mathbf{P}}\mathbf{s}\right) + \hat{\mathbf{w}}. \quad (18)$$

Comparing (13) and (18), we see how the nonlinearity introduced by $\mathcal{Q}(\cdot)$ complicates the input-output relation, thus,

preventing a straightforward evaluation of the performance of the MU-MIMO-OFDM downlink system. We next use Bussgang's theorem [44] to decompose (18) into a form that enables an analytical performance analysis and that captures the ICI caused by the DACs.

A. Decomposition Using Bussgang's Theorem

The finite-resolution DACs introduce a quantization error $\mathbf{e} \in \mathbb{C}^{BN}$,

$$\mathbf{e} = \alpha^{-1}\mathbf{x} - \mathbf{z} = \alpha^{-1}\mathcal{Q}(\mathbf{z}) - \mathbf{z}. \tag{19}$$

The multiplicaction by α^{-1} is necessary because we scaled the quantization labels by α to satisfy the power constraint (2). Note that $\mathbf{e}$ is *correlated* with the quantizer input $\mathbf{z}$. For Gaussian inputs, Bussgang's theorem [44] allows us to decompose $\mathcal{Q}(\mathbf{z})$ into two components: a linear function in $\mathbf{z}$ and a distortion that is *uncorrelated* with $\mathbf{z}$. Specifically, assume that $\mathbf{s}_k \sim \mathcal{CN}(\mathbf{0}_{U\times 1}, \mathbf{I}_U)$ for all $k \in \mathcal{S}_d$, and that the $\{\mathbf{s}_k\}$ for $k \in \mathcal{S}_d$ are independent. Let x_m and z_m be the mth element of $\mathbf{x}$ and $\mathbf{z}$, respectively. According to Bussgang's theorem [44], it holds that

$$\mathbb{E}[x_m z_n^*] = g_m \mathbb{E}[z_m z_n^*] \tag{20}$$

where $g_m = \sigma_m^{-2}\mathbb{E}[x_m z_m^*]$ and $\sigma_m^2 = \mathbb{E}[|z_m|^2]$. Define now $\mathbf{C_{xz}} = \mathbb{E}[\mathbf{xz}^H] \in \mathbb{C}^{BN\times BN}$ and $\mathbf{C_z} = \mathbb{E}[\mathbf{zz}^H] \in \mathbb{C}^{BN\times BN}$. It follows from (20) that $\mathbf{C_{xz}} = \mathbf{GC_z}$ with $\mathbf{G} = \mathrm{diag}([g_1, g_2, \ldots, g_{BN}]^T)$. Consequently, the transmitted time-domain vector $\mathbf{x}$ can be written as

$$\mathbf{x} = \mathcal{Q}(\mathbf{z}) = \mathbf{Gz} + \mathbf{d} \tag{21}$$

where the distortion $\mathbf{d} \in \mathbb{C}^{BN}$ is uncorrelated with $\mathbf{z}$, i.e., $\mathbb{E}[\mathbf{zd}^H] = \mathbf{0}_{BN\times BN}$. Indeed,

$$\mathbb{E}[\mathbf{dz}^H] = \mathbb{E}[(\mathbf{x} - \mathbf{Gz})\mathbf{z}^H] \tag{22}$$

$$= \mathbf{C_{xz}} - \mathbf{GC_z} = \mathbf{0}_{BN\times BN}. \tag{23}$$

Here, we used that $\mathbf{C_{xz}} = \mathbf{GC_z}$. It turns out that the real-valued diagonal matrix $\mathbf{G}$ can be evaluated in closed form. Indeed, by following the same steps as in [29, App. A], we find that

$$\mathbf{G} = \mathbf{I}_N \otimes \mathrm{diag}(\mathbf{g}) \tag{24}$$

where $\mathrm{diag}(\mathbf{g}) \in \mathbb{R}^{B\times B}$ is given by

$$\mathrm{diag}(\mathbf{g}) = \frac{\alpha\Delta}{\sqrt{\pi}} \mathrm{diag}\left(\frac{1}{N}\sum_{k\in\mathcal{S}_d}\widehat{\mathbf{P}}_k\widehat{\mathbf{P}}_k^H\right)^{-1/2}$$
$$\times \sum_{i=1}^{L-1}\exp\left(-\Delta^2\left(i - \frac{L}{2}\right)^2\right)$$
$$\times\mathrm{diag}\left(\frac{1}{N}\sum_{k\in\mathcal{S}_d}\widehat{\mathbf{P}}_k\widehat{\mathbf{P}}_k^H\right)^{-1}\right). \tag{25}$$

For the special case of 1-bit DACs, (25) reduces to

$$\mathrm{diag}(\mathbf{g}) = \sqrt{\frac{2P}{\pi\xi B}}\,\mathrm{diag}\left(\frac{1}{N}\sum_{k\in\mathcal{S}_d}\widehat{\mathbf{P}}_k\widehat{\mathbf{P}}_k^H\right)^{-1/2}. \tag{26}$$

Inserting (21) into (18), we obtain

$$\hat{\mathbf{y}} = \widehat{\mathbf{H}}(\mathbf{F}_N \otimes \mathbf{I}_B)\left(\mathbf{G}(\mathbf{F}_N^H \otimes \mathbf{I}_B)\widehat{\mathbf{P}}\mathbf{s} + \mathbf{d}\right) + \hat{\mathbf{w}} \tag{27}$$

$$= \widehat{\mathbf{H}}\mathbf{G}\widehat{\mathbf{P}}\mathbf{s} + \widehat{\mathbf{H}}(\mathbf{F}_N \otimes \mathbf{I}_B)\mathbf{d} + \hat{\mathbf{w}} \tag{28}$$

where the last step holds because $(\mathbf{F}_N \otimes \mathbf{I}_B)\mathbf{G}(\mathbf{F}_N^H \otimes \mathbf{I}_B) = \mathbf{G}$, as a consequence of (24).

B. Achievable Sum-Rate with Gaussian Inputs

Let $\hat{y}_{u,k} = [\hat{\mathbf{y}}_k]_u$ denote the received signal on the kth subcarrier and at the uth UE. It follows from (18) that

$$\hat{y}_{u,k} = \left[\widehat{\mathbf{H}}_k\,\mathrm{diag}(\mathbf{g})\,\widehat{\mathbf{P}}_k\right]_{u,u}s_{u,k}$$
$$+\sum_{v\neq u}\left[\widehat{\mathbf{H}}_k\,\mathrm{diag}(\mathbf{g})\,\widehat{\mathbf{P}}_k\right]_{u,v}s_{v,k}$$
$$+[\widehat{\mathbf{H}}(\mathbf{F}_N \otimes \mathbf{I}_B)\mathbf{d}]_{u+kU} + \hat{w}_{u,k} \tag{29}$$

where $s_{u,k} = [\mathbf{s}_k]_u$ and $\hat{w}_{u,k} = [\hat{\mathbf{w}}_k]_u$. The first term on the right-hand side (RHS) of (29) corresponds to the desired signal; the second term captures the MU interference; the third term describes the distortion introduced by the DACs; the fourth term represents the AWGN.

Let now $\gamma_{u,k}(\widehat{\mathbf{H}})$ be the SINDR on the kth subcarrier for the uth UE. Using (29) and assuming that $\mathbf{s}_k \sim \mathcal{CN}(\mathbf{0}_{U\times 1}, \mathbf{I}_U)$ for $k \in \mathcal{S}_d$, we can express $\gamma_{u,k}(\widehat{\mathbf{H}})$ as

$$\gamma_{u,k}(\widehat{\mathbf{H}}) = \frac{\left[\left|\widehat{\mathbf{H}}_k\,\mathrm{diag}(\mathbf{g})\,\widehat{\mathbf{P}}_k\right|^2\right]_{u,u}}{\sum_{v\neq u}\left[\left|\widehat{\mathbf{H}}_k\mathrm{diag}(\mathbf{g})\widehat{\mathbf{P}}_k\right|^2\right]_{u,v} + D_{u,k}(\widehat{\mathbf{H}}) + N_0} \tag{30}$$

where

$$D_{u,k}(\widehat{\mathbf{H}})$$
$$= \left[\widehat{\mathbf{H}}(\mathbf{F}_N \otimes \mathbf{I}_B)\,\mathbf{C_d}(\mathbf{F}_N^H \otimes \mathbf{I}_B)\,\widehat{\mathbf{H}}^H\right]_{u+kU,u+kU}. \tag{31}$$

Here, $\mathbf{C_d} = \mathbb{E}[\mathbf{dd}^H] \in \mathbb{C}^{BN\times BN}$ is the covariance of the distortion $\mathbf{d}$. For the case of symbol-rate-sampling DACs and single-carrier transmission ($S = N = 1$) over a frequency-flat channel ($T = 1$), the SINDR in (30) simplifies to the SINDR reported in [29, Eq. (26)].

Due to the nonlinearity introduced by the DACs, the distortion $\mathbf{d}$ is not Gaussian distributed, which makes it challenging to compute exactly the achievable rate. It is, however, possible to derive a lower bound the achievable rate using the so-called "auxiliary-channel lower bound" [51, p. 3503]. Through standard manipulations of the mutual information (see, e.g., [29, Sec. III-D]), we obtain the following lower bound on the ergodic sum rate[6] that is explicit in the SINDR (30)

$$R_{\mathrm{sum}} = \frac{1}{S}\mathbb{E}\left[\sum_{u=1}^{U}\sum_{k\in\mathcal{S}_d}\log_2\left(1 + \gamma_{u,k}(\widehat{\mathbf{H}})\right)\right] \tag{32}$$

where the expectation is over the channel matrix $\widehat{\mathbf{H}}$. It follows from a generalized mutual information analysis, similar to the one reported in [52], [53], that the lower bound in (32)

---

[6]We assume that coding can be performed over sufficiently many independent realization of $\{\mathbf{H}_t\}$ for $t = 0, 1, \ldots, T - 1$.

corresponds to the ergodic sum rate achievable using a Gaussian codebook and a mismatched scaled nearest-neighbor decoder at the UEs under the assumption that the channel gains[7] $\left[\widehat{\mathbf{H}}_k \operatorname{diag}(\mathbf{g}) \widehat{\mathbf{P}}_k\right]_{u,u}$ for $k \in \mathcal{S}_d$ are perfectly known to the uth UE, $u = 1, 2, \ldots, U$.

IV. EXACT AND APPROXIMATE DISTORTION MODELS

We next tackle the problem of evaluating the covariance matrix $\mathbf{C}_{\mathbf{d}}$ of the distortion $\mathbf{d}$, which is required to compute (30) and, hence, the achievable rate (32). As in the previous section, we assume Gaussian signaling, i.e., that $\mathbf{s}_k \sim \mathcal{CN}(\mathbf{0}_{U \times 1}, \mathbf{I}_U)$ for $k \in \mathcal{S}_d$. This implies that $\mathbf{z} \sim \mathcal{CN}(\mathbf{0}_{BN \times 1}, \mathbf{C_z})$ where

$$\mathbf{C_z} = \left(\mathbf{F}_N^H \otimes \mathbf{I}_B\right) \widehat{\mathbf{P}} \widehat{\mathbf{P}}^H (\mathbf{F}_N \otimes \mathbf{I}_B). \tag{33}$$

It follows from (21) that

$$\mathbf{C_d} = \mathbb{E}\left[\mathbf{dd}^H\right] = \mathbf{C_x} - \mathbf{G C_z G}. \tag{34}$$

Here, we used that $\mathbf{G}^H = \mathbf{G}$ since $\mathbf{G}$ is a real-valued diagonal matrix, and that $\mathbf{C_{xz}} = \mathbf{G C_z}$. Hence, to evaluate $\mathbf{C_d}$, one has to compute the covariance $\mathbf{C_x} = \mathbb{E}\left[\mathbf{xx}^H\right] \in \mathbb{C}^{BN \times BN}$ of the DAC output. We next discuss how to evaluate $\mathbf{C_x}$.

A. Computation of $\mathbf{C_x}$

Let $x_m = [\mathbf{x}]_m$ and $x_n = [\mathbf{x}]_n$. We can write the entry on the mth row and nth column of $\mathbf{C_x}$ as $\mathbb{E}[x_m x_n^*]$. Let now $x_m^R = \Re\{x_m\}$ and $x_m^I = \Im\{x_m\}$ denote the real and imaginary components of x_m, respectively. Since the input to the DACs is a circularly-symmetric Gaussian random variable, $\mathbb{E}[x_m x_n^*]$ can be as expanded as follows:

$$\mathbb{E}[x_m x_n^*] = 2\Big(\mathbb{E}\left[x_m^R x_n^R\right] + j \mathbb{E}\left[x_m^I x_n^R\right]\Big). \tag{35}$$

Let now z_m be the mth element of $\mathbf{z}$ and let $z_m^R = \Re\{z_m\}$. For the case $n = m$, for which it holds that $\mathbb{E}\left[x_m^I x_m^R\right] = 0$, (35) reduces to

$$\mathbb{E}\left[|x_m|^2\right] = 2 \sum_{i=0}^{L-1} \ell_i^2 \, \mathbb{P}[x_m^R = \ell_i] \tag{36}$$

$$= 2 \sum_{i=0}^{L-1} \ell_i^2 \, \mathbb{P}[\tau_i \le z_m^R < \tau_{i+1}] \tag{37}$$

$$= 2 \sum_{i=0}^{L-1} \ell_i^2 \left(\Phi\left(\frac{\sqrt{2}\,\tau_{i+1}}{\sigma_m}\right) - \Phi\left(\frac{\sqrt{2}\,\tau_i}{\sigma_m}\right) \right) \tag{38}$$

$$= \frac{\alpha^2 \Delta^2}{2}(L-1)^2$$
$$\quad -4\alpha^2 \Delta^2 \sum_{i=1}^{L-1} \left(i - \frac{L}{2}\right) \Phi\left(\frac{\sqrt{2}}{\sigma_m}\left(i - \frac{L}{2}\right)\right) \tag{39}$$

where (38) follows because $z_m \sim \mathcal{CN}(0, \sigma_m^2)$ with $\sigma_m^2 = \mathbb{E}\left[|z_m|^2\right] = [\mathbf{C_z}]_{m,m}$. To derive (39), we used that, for the uniform quantizer considered in this work, we have $\ell_{i+1} =$

$\ell_i + \alpha\Delta$ for $i \in \{0, 1, \ldots, L-2\}$. For the case $m \ne n$, the expectation $\mathbb{E}\left[x_m^C x_n^R\right]$, where $C \in \{R, I\}$, can be written as

$$\mathbb{E}\left[x_m^C x_n^R\right]$$
$$= \sum_{a=0}^{L-1} \sum_{b=0}^{L-1} \ell_a \ell_b \, \mathbb{P}[x_m^C = \ell_a, x_n^R = \ell_b] \tag{40}$$
$$= \sum_{a=0}^{L-1} \sum_{b=0}^{L-1} \ell_a \ell_b \, \mathbb{P}[\tau_a \le z_m^C < \tau_{a+1}, \tau_b \le z_n^R < \tau_{b+1}]. \tag{41}$$

Unfortunately, (41) does not have a known closed-form expression and hence, has to be evaluated using numerical methods (cf. [25]). One exception is the special case of 1-bit DACs ($L = 2$), for which it holds that

$$\mathbb{E}[x_m x_n^*]$$
$$= \frac{2P}{\pi \xi B} \left(\arcsin\left(\frac{\sigma_{m,n}^R}{\sigma_m \sigma_n}\right) + j \arcsin\left(\frac{\sigma_{m,n}^I}{\sigma_m \sigma_n}\right) \right). \tag{42}$$

Here, we have defined $\sigma_{m,n}^R = \Re\{[\mathbf{C_z}]_{m,n}\}$ and $\sigma_{m,n}^I = \Im\{[\mathbf{C_z}]_{m,n}\}$. This well-known result, reported first by Van Vleck and Middleton [46], is commonly referred to as the *arcsine law*. Writing (42) in matrix form, we obtain, for the 1-bit-DAC case,

$$\mathbf{C_x} = \frac{2P}{\pi \xi B} \Big(\arcsin\Big(\operatorname{diag}(\mathbf{C_z})^{-\frac{1}{2}} \, \Re\{\mathbf{C_z}\} \operatorname{diag}(\mathbf{C_z})^{-\frac{1}{2}}\Big)$$
$$+ j \arcsin\Big(\operatorname{diag}(\mathbf{C_z})^{-\frac{1}{2}} \, \Im\{\mathbf{C_z}\} \operatorname{diag}(\mathbf{C_z})^{-\frac{1}{2}}\Big) \Big). \tag{43}$$

By inserting (43) into (34) we find the desired covariance matrix $\mathbf{C_d}$, which allow us to compute the SINDR (30) and the ergodic sum rate (32) for the 1-bit-DAC case.

To evaluate $\mathbf{C_d}$ in (34) for the case $L > 2$, one has to compute (41) using numerical integration, which is time consuming and offers limited insights.[8] In what follows, we shall present two closed-form approximations for $\mathbf{C_x}$, which trade accuracy for complexity in distinct ways.

B. Rounding Approximation

First, we present a *rounding approximation* $\mathbf{C_d^{\mathrm{round}}}$ of $\mathbf{C_d}$. This approximation takes into account the correlation between the entries of $\mathbf{d}$ and turns out to be accurate as long as the step size Δ of the DACs is set so that the overload distortion is negligible compared to the granular distortion. To derive this approximation, we start by noting from (19) that $\mathbf{C_x}$ can be written as

$$\mathbf{C_x} = \alpha^2\big(\mathbf{C_z} + \mathbf{C_{ze}} + \mathbf{C_{ze}}^H + \mathbf{C_e}\big) \tag{44}$$

where $\mathbf{C_e} = \mathbb{E}\left[\mathbf{ee}^H\right] \in \mathbb{C}^{BN \times BN}$ is the covariance of $\mathbf{e}$ and $\mathbf{C_{ze}} = \mathbb{E}\left[\mathbf{ze}^H\right] \in \mathbb{C}^{BN \times BN}$; this last matrix can be

---

[7]These are the per-subcarrier scaling factors used in the scaled nearest-neighbor decoding rule.

[8]For the uplink case, focusing the quantization distortion caused by symbol-rate-sampling ADCs, the required covariance matrix was found by means of numerical integration in [25]. This approach, however, is not practical in our setup as the number of entries in $\mathbf{C_x}$ scales quadratically in BN, where B could be in the order of hundreds and N could be in the order of thousands.

expressed also as

$$\mathbf{C_{ze}} = \mathbb{E}\left[\mathbf{z}\big(\alpha^{-1}\mathbf{x} - \mathbf{z}\big)^H\right] \tag{45}$$

$$= \mathbb{E}\left[\mathbf{z}\big(\alpha^{-1}\mathbf{Gz} + \alpha^{-1}\mathbf{d} - \mathbf{z}\big)^H\right] \tag{46}$$

$$= \mathbf{C_z}(\alpha^{-1}\mathbf{G} - \mathbf{I}_{BN}). \tag{47}$$

Here, we have used that $\mathbb{E}\left[\mathbf{zd}^H\right] = \mathbf{0}_{BN \times BN}$. Inserting (47) into (44), we obtain

$$\mathbf{C_x} = \alpha(\mathbf{GC_z} + \mathbf{C_zG}) + \alpha^2(\mathbf{C_e} - \mathbf{C_z}). \tag{48}$$

Note that the only unknown quantity in (48) is $\mathbf{C_e}$. Not surprisingly, evaluating $\mathbf{C_e}$ is just as difficult as evaluating $\mathbf{C_x}$ and, in general, no closed-form expression is known. However, if the step size Δ of the DACs is set such that the overload distortion is negligible compared to the granular distortion, the error $\mathbf{e}$ can be accurately approximated by

$$\mathbf{e} = \alpha^{-1}\mathcal{Q}(\mathbf{z}) - \mathbf{z} \approx \mathcal{R}(\mathbf{z}) - \mathbf{z} \tag{49}$$

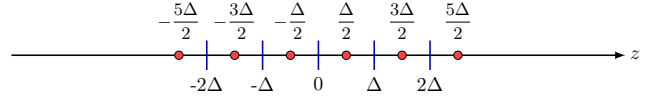where the *rounding function* $\mathcal{R}(\cdot)$ is defined as follows:

$$\mathcal{R}(z) = \begin{cases} \Delta\left\lfloor \dfrac{z}{\Delta} + \dfrac{1}{2} \right\rfloor, & \text{if } L \text{ is odd,} \\[3mm] \Delta\left\lfloor \dfrac{z}{\Delta} \right\rfloor + \dfrac{\Delta}{2}, & \text{if } L \text{ is even.} \end{cases} \tag{50}$$

Some comments on (49) and (50) are in order. The rounding function $\mathcal{R}(\cdot)$ describes a uniform symmetric quantizer with the same step size Δ as the uniform symmetric quantizer described by $\alpha^{-1}\mathcal{Q}(\cdot)$ but with an infinite number of quantization levels. If the input z lies within the granular region of the uniform quantizer described by $\alpha^{-1}\mathcal{Q}(\cdot)$, i.e., if $|z| \leq A_{\text{clip}}$, then $\mathcal{R}(z) = \alpha^{-1}\mathcal{Q}(z)$. If, however, the amplitude of the input z exceeds the clipping level A_{clip}, i.e., if $|z| > A_{\text{clip}}$, then $\mathcal{R}(z) \neq \alpha^{-1}\mathcal{Q}(z)$. The difference between $\alpha^{-1}\mathcal{Q}(\cdot)$ and $\mathcal{R}(\cdot)$ is illustrated in Fig. 2. Replacing $\alpha^{-1}\mathcal{Q}(\cdot)$ with $\mathcal{R}(\cdot)$ is convenient because the statistical theory of the quantizer (50) is well-investigated (see, e.g., [54]–[57]) and the corresponding covariance matrix $\mathbf{C_e}$ is known. Specifically, let e_m and e_n denote the mth and the nth entry of $\mathbf{e}$, respectively. Furthermore, assume that (49) holds with equality. For a midtread quantizer (L odd), Sripad and Snyder showed that [55, Eq. (23)]
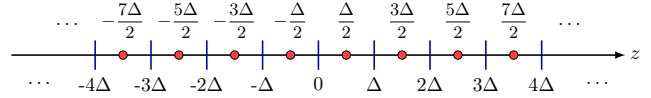
$$\mathbb{E}[e_m e_n^*]$$
$$= \frac{2\Delta^2}{\pi^2} \sum_{a=1}^{\infty} \sum_{b=1}^{\infty} \frac{(-1)^{a+b}}{ab} \exp\left(-\frac{\pi^2\big(a^2\sigma_m^2 + b^2\sigma_n^2\big)}{\Delta^2}\right)$$
$$\times \left(\sinh\left(\frac{2\pi^2 ab\,\sigma_{m,n}^R}{\Delta^2}\right) + j\sinh\left(\frac{2\pi^2 ab\,\sigma_{m,n}^I}{\Delta^2}\right)\right). \tag{51}$$

For midrise quantizers (L even) it further holds that (see Appendix A)

$$\mathbb{E}[e_m e_n^*]$$
$$= \frac{2\Delta^2}{\pi^2} \sum_{a=1}^{\infty} \sum_{b=1}^{\infty} \frac{1}{ab} \exp\left(-\frac{\pi^2\big(a^2\sigma_m^2 + b^2\sigma_n^2\big)}{\Delta^2}\right)$$
$$\times \left(\sinh\left(\frac{2\pi^2 ab\,\sigma_{m,n}^R}{\Delta^2}\right) + j\sinh\left(\frac{2\pi^2 ab\,\sigma_{m,n}^I}{\Delta^2}\right)\right). \tag{52}$$

(a) $\alpha^{-1}\mathcal{Q}(z)$: uniform midrise quantizer with $L = 6$ levels.

(b) $\mathcal{R}(z)$: uniform midrise quantizer with an infinite number of levels.

Fig. 2. Comparison between the quantization levels and thresholds associated with the rules $\alpha^{-1}\mathcal{Q}(z)$ and $\mathcal{R}(z)$; the red circles correspond to quantization labels, and the blue vertical lines correspond to the quantization thresholds.

Using (51) and (52), we can write the covariance of $\mathbf{e}$ for every $L \geq 2$ as

$$\mathbf{C_e} = \frac{2\Delta^2}{\pi^2} \sum_{a=1}^{\infty} \sum_{b=1}^{\infty} \frac{\cos(\pi L)^{a+b}}{ab}$$
$$\times \exp\left(-\frac{\pi^2 a^2}{\Delta^2}\text{diag}(\mathbf{C_z})\,\mathbf{1}_{BN \times BN}\right.$$
$$\left.-\frac{\pi^2 b^2}{\Delta^2}\mathbf{1}_{BN \times BN}\,\text{diag}(\mathbf{C_z})\right)$$
$$\odot\left(\sinh\left(\frac{2\pi^2 ab}{\Delta^2}\,\Re\{\mathbf{C_z}\}\right)\right.$$
$$\left.+j\sinh\left(\frac{2\pi^2 ab}{\Delta^2}\,\Im\{\mathbf{C_z}\}\right)\right). \tag{53}$$

We obtain the desired rounding approximation for $\mathbf{C_x}$ by inserting (53) into (48). Interestingly, for the case $L = 2$, it is possible to retrieve Van Vleck's arcsine law (43) from (48) and (53). We formalize this result in the following theorem; a proof is given in Appendix B.

Theorem 1: Assume $L = 2$, insert (53) into (48), and let $\Delta \to \infty$. Then, we obtain Van Vlecks's arcsine law (43).

Finally, by inserting (53) and (48) into (34), we obtain rounding approximation $\mathbf{C_d^{\text{round}}}$ of $\mathbf{C_d}$:

$$\mathbf{C_d^{\text{round}}} = \frac{2\alpha^2\Delta^2}{\pi^2} \sum_{a=1}^{\infty} \sum_{b=1}^{\infty} \frac{\cos(\pi L)^{a+b}}{ab}$$
$$\times \exp\left(-\frac{\pi^2 a^2}{\Delta^2}\text{diag}(\mathbf{C_z})\,\mathbf{1}_{BN \times BN}\right.$$
$$\left.-\frac{\pi^2 b^2}{\Delta^2}\mathbf{1}_{BN \times BN}\,\text{diag}(\mathbf{C_z})\right)$$
$$\odot\left(\sinh\left(\frac{2\pi^2 ab}{\Delta^2}\,\Re\{\mathbf{C_z}\}\right)\right.$$
$$\left.+j\sinh\left(\frac{2\pi^2 ab}{\Delta^2}\,\Im\{\mathbf{C_z}\}\right)\right)$$
$$+\alpha(\mathbf{GC_z} + \mathbf{C_zG}) - \alpha^2\mathbf{C_z} - \mathbf{GC_zG}. \tag{54}$$

As we will see in Section V, the rounding approximation (54) is accurate independently of the resolution of the DACs, provided that Δ is chosen so that clipping occurs with low probability.

Note that evaluating (54) involves computing two infinite sums. The series, however, converges rapidly because the terms decay exponentially in the variables a and b. For the numerical

results reported in Section V, we evaluate (54) by summing only over the first 30 terms of both sums; this turns out to be sufficient to obtain accurate results.

C. Diagonal Approximation

Recall from Section IV-A that the diagonal elements of $\mathbf{C_x}$ can be computed exactly using (39). Building on this observation, we present next a *diagonal approximation* $\mathbf{C_d^{diag}}$ of $\mathbf{C_d}$ in which the distortion caused by the DACs is modeled as a white process, both in space and time, by assuming that the off-diagonal elements of $\mathbf{C_d}$ are zero. Specifically,

$$\mathbf{C_d^{diag}} = \mathrm{diag}(\mathbf{C_x}) - \mathbf{G}\,\mathrm{diag}(\mathbf{C_z})\,\mathbf{G} \tag{55}$$

$$= \frac{\alpha^2 \Delta^2}{2}(L-1)^2\,\mathbf{I}_{BN} - \mathbf{G}\,\mathrm{diag}(\mathbf{C_z})\,\mathbf{G}$$

$$- 4\alpha^2 \Delta^2 \sum_{i=1}^{L-1} \left(i - \frac{L}{2}\right)$$

$$\times \Phi\left(\sqrt{2}\,\mathrm{diag}(\mathbf{C_z})^{-\frac{1}{2}}\left(i - \frac{L}{2}\right)\right). \tag{56}$$

Here, we used (34) to obtain (55), and (39) to obtain (56). Note that $\mathbf{C_d^{diag}} = \mathbf{C_d} \odot \mathbf{I}_{BN}$ (cf. [15, Eq. (26)]), which implies that the elements on the diagonal of $\mathbf{C_d}$ in (56), for Gaussian inputs, are exact.

For the uplink case, an approximate model called the *additive quantization noise model* (AQNM) is commonly used to characterize the quantization distortion caused by low-resolution ADCs (see, e.g., [21], [22]). Similarly to the diagonal approximation in (56), in the AQNM the covariance matrix of the quantization distortion is approximated by a diagonal matrix. The AQNM model, however, assumes that the the power of the input to the quantizer is equal on all antennas (which can be achieved in the uplink by using an automatic gain control circuit) and that quantization labels are rescaled to minimize the MSE between the quantizer input and output. These assumptions are not valid for the massive MU-MIMO-OFDM downlink scenario considered is this paper. Indeed, the power of the transmitted signal on different antennas is not equal and the quantization labels are uniformly separated.[9]

Despite its simplicity, the diagonal approximation (56) is accurate for DACs with medium-to-high resolution and if the OSR is relatively small (e.g., when $L \geq 4$ and $\xi \leq 4$). However, as we will demonstrate in Section V, the diagonal approximation is not sufficiently accurate for DACs with low resolution and if the OSR is high (e.g., when $L < 4$ and $\xi > 4$), which implies that the correlation within the distortion caused by the DACs should not be ignored.

V. NUMERICAL RESULTS

We focus on a massive MU-MIMO-OFDM system in which the number of BS antennas is $B = 128$ and the number of UEs is $U = 16$.[10] We consider a frequency-selective Rayleigh fading channel with $T = 4$ nonzero taps and a uniform power delay profile. The OFDM parameters are inspired by a 5 MHz LTE system [58]. Specifically, the number of occupied subcarriers is $S = 300$ and the subcarrier spacing is $\Delta f = 15$ kHz. Furthermore, we assume that the occupied subcarriers are the first 150 to the left and to the right of the DC subcarrier (the DC subcarrier is not occupied). The total number of subcarriers (the size of the DFT) is $N = 1024$. Hence, the sampling rate of the DACs is $f_s = N\Delta f = 1024 \cdot 15 \cdot 10^3 = 15.36$ MHz and the OSR is $\xi = N/S = 1024/300 \approx 3.4$.[11]

Under the assumption that the input to the DACs on the bth antenna $(b = 1, 2, \ldots, B)$ is $z_b \sim \mathcal{CN}(0, P/(\xi B))$ and that the clipping level is set to $A_{\mathrm{clip}} = \sqrt{P/(2\xi B)}\left(1 - \Phi^{-1}(P_{\mathrm{clip}}/2)\right)$, the DACs will clip the signal with probability P_{clip}. The corresponding step size is $\Delta = 2A_{\mathrm{clip}}/L$. In what follows, we have set the clipping level of the DACs so that $P_{\mathrm{clip}} = 0.1\%$. This choice is not necessarily optimal, but ensures that the clipping distortion caused by the DACs is small compared to the granular distortion for the values of L considered here, which will enable us to use the rounding approximation (53). Furthermore, we shall see that this particular choice yields near-optimal (infinite resolution) performance also for DACs with low resolution.

A. Power Spectral Density

To demonstrate the accuracy of the approximations in Section IV-B and Section IV-C, we plot in Fig. 3 the (normalized) PSD of the transmitted signal (averaged over the BS antennas and over 100 channel realizations) and the (normalized) PSD of the received signal (averaged over the UEs and over 100 channel realizations). Here, the BS uses ZF precoding and the data vectors $\mathbf{s}_k$ for $k \in \mathcal{S}_d$ contain QPSK symbols.[12] Numerical simulations are compared with analytic results obtained by computing the PSD of the transmitted vector $\mathbf{x}$ and of the corresponding received vector $\mathbf{y}$ using the rounding approximation and the diagonal approximation presented in Section IV-B and Section IV-C, respectively. We note that the rounding approximation yields accurate results independently of the resolution of the DACs. We also note that the diagonal approximation results in a poor approximation of the PSD for the case $L = 2$ but yields a more accurate approximation as the number of bits increase. Indeed, when the number of bits increases, the distortion caused by the DACs becomes more

---

[9]For the uplink case, Mezghani and Nossek derived an approximation for the covariance matrix of the quantization distortion caused by symbol-rate-sampling ADCs that, similarly to the proposed rounding approximation, takes into account the correlation in the signal transmitted on different antennas [24]. The approximation in [24], however, assumes that the power of the input to the quantizer is equal on all antennas and that the quantization labels are rescaled to minimize the MSE between the quantizer input and output (these are the same assumptions that are used to derive the AQNM approximation [21], [22]). Neither of these assumptions are valid for the massive MU-MIMO-OFDM downlink scenario considered here.

[10]Our simulation framework is available for download from GitHub (https://github.com/quantizedmassivemimo/1bit_linear_precoding_ofdm). The purpose is to enable interested readers to perform simulations with different system parameters than the ones reported here.

[11]In Section V-E, we shall investigate the impact of the OSR on performance.

[12]Recall that in Section III and Section IV, we assumed that the per-antenna DAC input is a zero-mean Gaussian random variable. This assumption holds approximately true also for the case of finite-cardinality constellations and linear precoding, because the per-antenna DAC input can be written as sum of US independent and identically distributed random variables with zero mean and finite variance. Hence, as US grow large, the per-antenna DAC input converges to a Gaussian random variable by the central limit theorem.

(a) Transmitted, 1-bit DACs ($L = 2$).　(b) Transmitted, 2-bit DACs ($L = 4$).　(c) Transmitted, 3-bit DACs ($L = 8$).

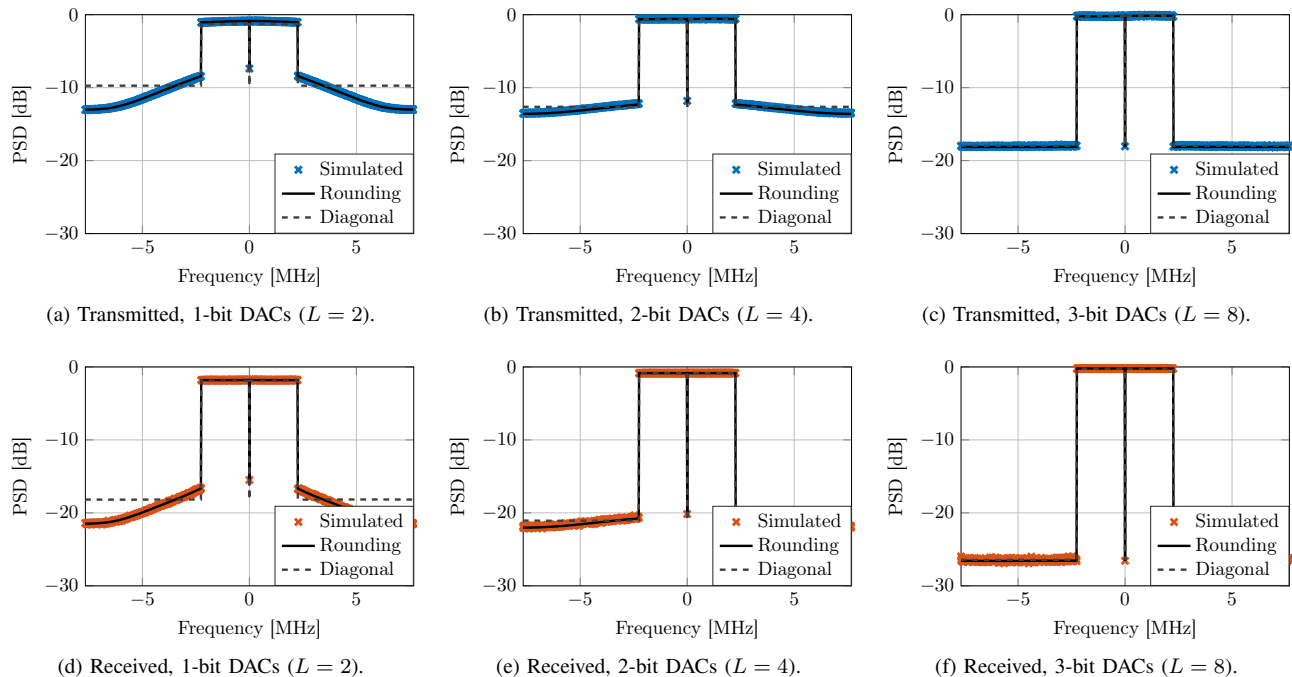(d) Received, 1-bit DACs ($L = 2$).　(e) Received, 2-bit DACs ($L = 4$).　(f) Received, 3-bit DACs ($L = 8$).

Fig. 3. PSD of the transmitted and received signal; $B = 128$ and $U = 16$. The markers correspond to simulated values, the solid lines correspond to the rounding approximation presented in Section IV-B and the dashed lines correspond to the diagonal approximation presented in Section IV-C.

spectrally white (compare Fig. 3a and Fig. 3c). We also see from the figure that the low-resolution DACs cause severe OOB distortion at the BS, which is captured accurately by the rounding approximation for all values of L. Interestingly, the relative amount of OOB distortion is smaller at the UEs than at the BS, which is in line with recent findings reported for PAs in [13]. Nevertheless, the OOB distortion caused by the low-resolution DACs is a significant issue in practical systems as it may cause interference to UEs operating in adjacent frequency bands.

B. Error-Rate Performance

1) Uncoded BER: If the elements of the symbol vector $\mathbf{s}_k$ for $k \in \mathcal{S}_d$ are drawn independently from a QPSK constellation, we can approximate the uncoded BER by

$$BER = 1 - \frac{1}{US} \sum_{k \in \mathcal{S}_d} \sum_{u=1}^{U} \mathbb{E}\left[\Phi\left(\gamma_{u,k}^{1/2}(\widehat{\mathbf{H}})\right)\right] \qquad (57)$$

where $\gamma_{u,k}(\widehat{\mathbf{H}})$ is given in (30). We evaluate this quantity by using the rounding approximation (54) and the diagonal approximation (56). Note that for the approximations to be accurate, the number of BS antennas B does not need to be large. To illustrate this aspect, we show in Fig. 4a the uncoded BER with QPSK and ZF for the single-input single-output (SISO) case (i.e, when $U = 1$ and $B = 1$) as a function of the SNR and the number of DAC bits.[13] We compare simulated BER values with the analytical BER in (57) and note that the rounding approximation is accurate over the entire range of SNR values. We note that the diagonal approximation become

[13]In the SISO-OFDM case, ZF precoding reduces to channel inversion.

more accurate as the number of DAC bits increase. We further note that, in the SISO-OFDM case, low uncoded BERs are not supported with QPSK and low-resolution DACs. Indeed, 7–8 DAC bits are required to achieve a target BER of 10^{-4} without a significant performance degradation compared to the infinite-resolution case.

In the massive MU-MIMO-OFDM case, the large number of antennas at the BS enables a considerable reduction of the resolution of the DACs compared to the SISO-OFDM case. To illustrate this, we show in Fig. 4b the uncoded BER with QPSK and ZF as a function of the SNR and the number of DAC bits for the case $B = 128$ and $U = 16$. In contrast to the SISO-OFDM case, low uncoded BERs are now supported by low-resolution DACs. Indeed, an uncoded BER below 10^{-4} is supported in the 1-bit-DAC case provided that the SNR exceeds 9 dB. Furthermore, only 3–4 DAC bits are necessary to approach infinite-resolution performance for a target BER of 10^{-4}. We again note that the rounding approximation (54) is accurate over the entire range of SNR values and independently of the number of DAC bits, and that the diagonal approximation (56) becomes more accurate as the number of DAC bits increase. However, the diagonal approximation significantly overestimates the BER performance with 1-bit DACs for high values of SNR, as this approximation does not take into account the inherent (spatial and temporal) correlation in the DAC distortion. Hence, one must be careful when using approximations that ignore the correlation in the quantization distortion as they are accurate only in some scenarios (e.g., for low SNR or when the resolution of the quantizer is sufficiently high).

2) Coded BER: In Fig. 5, we show the coded BER with ZF as a function of the SNR and the number of DAC bits for the case of QPSK and 16 quadrature amplitude

(a) Uncoded BER with QPSK; $B = 1$ and $U = 1$.

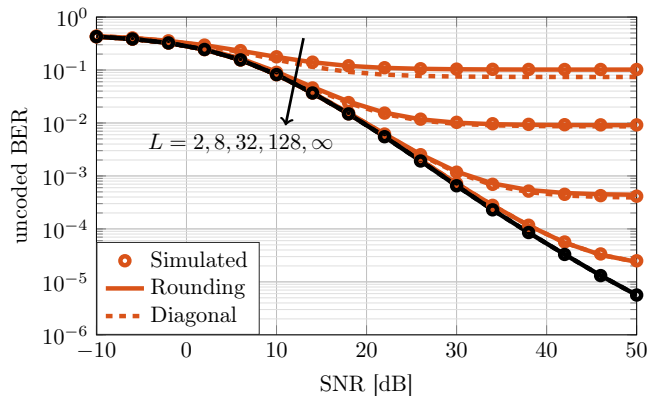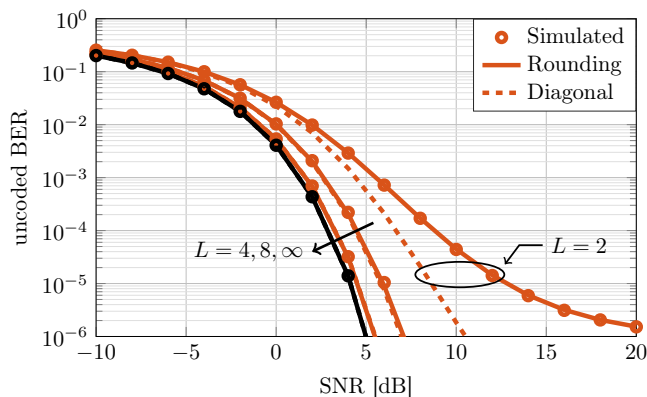(a) Coded BER with QPSK; $B = 128$ and $U = 16$.

(b) Uncoded BER with QPSK; $B = 128$ and $U = 16$.

(b) Coded BER with 16-QAM; $B = 128$ and $U = 16$.

Fig. 4. Uncoded BER with ZF and QPSK signaling. The markers correspond to simulated values, the solid lines correspond to the BER (57) computed using the rounding approximation in Section IV-B, and the dashed lines correspond to the BER (57) computed using the diagonal approximation in Section IV-C. The black lines correspond to the infinite-resolution case.

Fig. 5. Coded BER (simulated) with ZF for the case of QPSK and 16-QAM signaling. The information streams are encoded using a rate-5/6 convolutional code spanning 10 OFDM symbols. The UEs use nearest-neighbor detection and soft-input max-log BCJR decoding. The black lines correspond to the infinite-resolution case.

modulation (QAM) signaling. Here, we consider the case $B = 128$ and $U = 16$, and show only simulated BER values. The BS uses a (weak) rate-5/6 convolutional code with random interleaving to encode the information bits (separately for each UE) over 10 OFDM symbols. Hence, a codeword spans 3000 symbols. Each UE performs soft-input max-log Bahl, Cocke, Jelinek and Raviv (BCJR) decoding to estimate the transmitted information streams. We note that high-order constellations, such as 16-QAM, are supported with linear precoding in the massive MU-MIMO-OFDM case despite the low-resolution DACs, and that only few DAC bits are needed to close the gap to the infinite-resolution performance.

In this work, for analytical tractability, we have focused exclusively on linear precoding. However, it is well-known (see, e.g., [29]–[39]) that is possible to achieve superior performance, at the cost of an increased computational complexity, by using nonlinear precoders. For the case of OFDM transmission with 1-bit DACs, a comparison between linear and nonlinear precoding, in terms of uncoded/coded BER and computational complexity, is provided in [41].
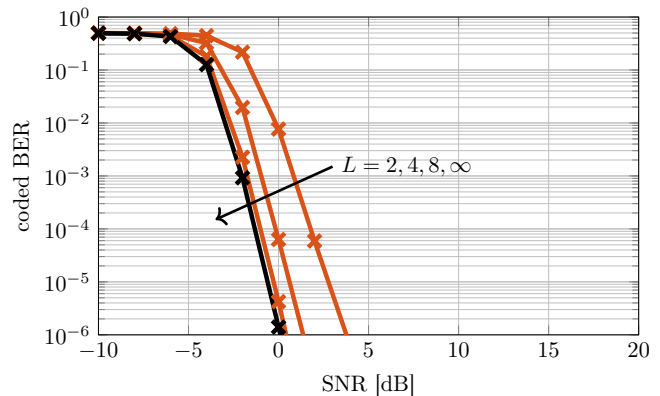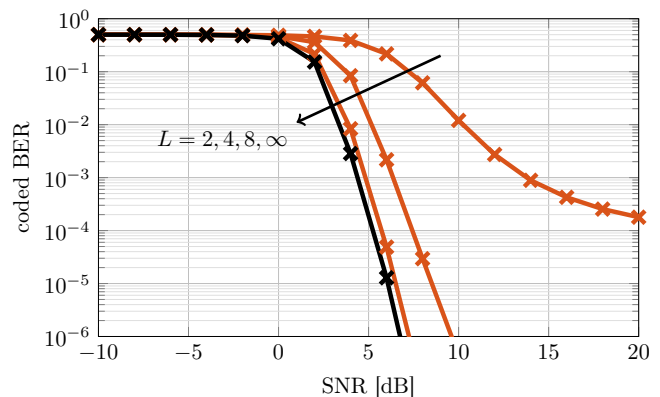
C. Achievable Rate

In Fig. 6, we show the achievable sum rate with Gaussian signaling and ZF precoding as a function of the SNR and the number of DAC levels. Recall that this sum rate is achieved by using a Gaussian codebook and nearest-neighbor decoding at the UEs (see Section III-B). The sum rate in (32) is evaluated by computing (30) using the rounding approximation and the diagonal approximation in (54) and (56), respectively. We note that the diagonal approximation is accurate for the case $L \geq 4$ but significantly overestimates the achievable rate for $L = 2$. We further note that high sum-rate throughputs are supported by the massive MU-MIMO-OFDM system, despite the low-resolution DACs at the BS.

D. Impact of Imperfect CSI

Until now, we have assumed that perfect CSI is available at the BS. We now relax this assumption and investigate the impact on performance of imperfect CSI. Specifically, we consider the case in which the BS has access only to noisy versions $\{\mathbf{H}_t^{\text{est}}\}$ of $\{\mathbf{H}_t\}$ for $t = 0, 1, \ldots, T - 1$. Specifically, $\mathbf{H}_t^{\text{est}} = \sqrt{1 - \varepsilon}\mathbf{H}_t + \sqrt{\varepsilon}\mathbf{H}_t^{\text{err}}$ for $t = 0, 1, \ldots, T - 1$, where $\varepsilon \in [0, 1]$ and where the entries of $\mathbf{H}_t^{\text{err}}$ are uncorrelated with
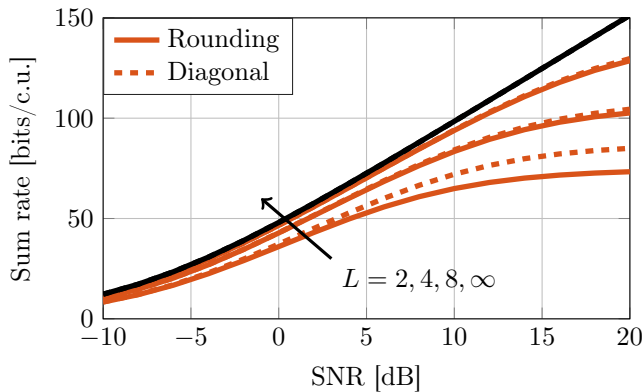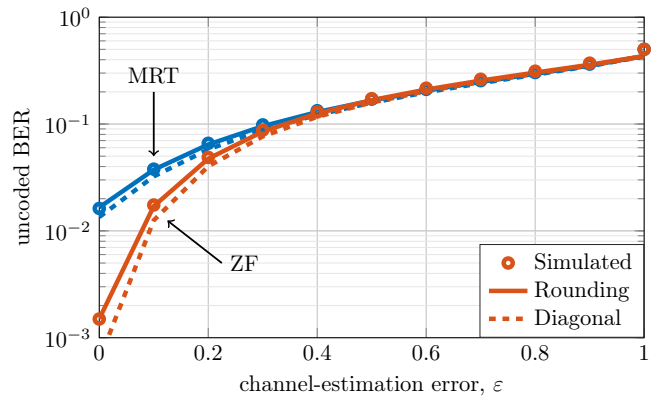
Fig. 6. Sum-rate with ZF and Gaussian signaling; $B = 128$ and $U = 16$. The solid lines correspond to the rounding approximation in Section IV-B, the dashed lines correspond to the diagonal approximation in Section IV-C. The black lines, which overlap, correspond to the sum rate achievable in the infinite-resolution case.
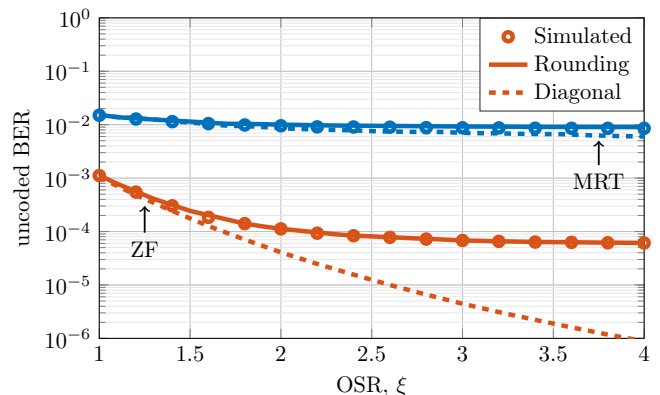
zero mean and unit variance. The corresponding frequency-domain estimate for the kth subcarrier ($k \in \mathcal{S}_d$) is $\widehat{\mathbf{H}}_k^{\text{est}} = \sum_{t=0}^{T-1} \mathbf{H}_t^{\text{est}} \exp\left(-jk\frac{2\pi}{N}t\right)$; the corresponding MRT and ZF precoding vectors are obtained from this estimate using (9) and (11), respectively. The case $\varepsilon = 0$ corresponds to perfect CSI (i.e., the case considered in Fig. 4b) and the case $\varepsilon = 1$ corresponds to no CSI. In a time-division duplex system, the value of the channel-estimation error ε will depend on the pilot sequence that have been transmitted during the uplink phase and on the resolution of the ADCs at the BS (see, e.g., [17]–[20]). In what follows, for simplicity, we assume that the entries of $\mathbf{H}_t^{\text{err}}$ are $\mathcal{CN}(0,1)$ distributed for $t = 0, 1, \ldots, T-1$. In Fig. 7a, we show, for the 1-bit-DAC case, the uncoded BER with QPSK as a function of the channel-estimation error ε for MRT and ZF precoding. We note that the rounding approximation is accurate also for the case of imperfect CSI. We further note that ZF outperforms MRT for $\varepsilon < 0.4$.

E. Impact of Oversampling

In Fig. 7b, we investigate the impact of the OSR on the uncoded BER for the case of 1-bit DACs and perfect CSI. Specifically, we plot the uncoded BER for the case of uncoded QPSK with MRT and ZF precoding as a function of the OSR. The SNR is set to $\rho = 10$ dB. We note that, for ZF, the uncoded BER can be considerably improved by operating the 1-bit DACs at a sampling rate higher than the symbol rate. Indeed, the uncoded BER with ZF can be decreased by an order of magnitude compared to the symbol-rate sampling case ($\xi = 1$) by operating the 1-bit DACs at twice the symbol rate. However, further increasing the sampling rate yields only marginal performance gains. For MRT, the OSR has little impact on BER performance as, in this case, the system is limited by the MU interference rather than by the DAC distortion. We also note that, in the 1-bit-DAC case, the diagonal approximation is accurate for small OSRs (e.g., for $\xi < 1.4$) but underestimates significantly the BER for larger values of ξ. For higher resolution DACs, the diagonal approximation is accurate also for high OSRs (see, e.g., Fig. 4b).

(a) Impact of Imperfect CSI; $\rho = 5$ dB, $\xi = 3.4$.

(b) Impact of OSR; $\rho = 10$ dB, $\varepsilon = 0$.

Fig. 7. Impact of the imperfect CSI and OSR on the uncoded BER for the 1-bit-DAC case; $B = 128$ and $U = 16$. The markers correspond to simulated values, the solid lines correspond to the BER (57) according to the rounding approximation in Section IV-B, the dashed lines correspond to the BER (57) according to the diagonal approximation in Section IV-C.

VI. CONCLUSIONS

We have characterized the performance in terms of un-coded/coded BER and achievable sum rate of a massive MU-MIMO-OFDM downlink system, in which the BS is equipped with finite-resolution DACs and uses linear precoding. Using Bussgang's theorem, we have derived a lower bound on the achievable rate and an accurate approximation for the uncoded BER with QPSK.

We have developed two approximations for the distortion caused by the low-resolution DACs. The *rounding approximation* ignores the overload distortion caused by the DACs and is accurate for DACs of arbitrary resolution and for any OSR. We also prove that the rounding approximation can be made exact in the 1-bit-DAC case by taking Δ in (53) to infinity. The *diagonal approximation* is an easy-to-evaluate approximation that assumes that the DAC distortion is spatially and temporally white. This approximation is accurate for moderate-to-high DAC resolutions and if the OSR is not too high, but significantly overestimates the performance (in terms of BER and achievable rate) for low-resolution (e.g., 1-bit) DACs. Our results highlight the importance of taking into account the correlation of the distortion caused by the quantizer in the DACs. Similar findings have been reported recently in [25] for the distortion caused

by ADCs in the uplink.

In practice, OOB emissions caused by the low-resolution DACs could prevent their use in practical systems. In this work, for simplicity, we have assumed that the reconstruction stage in the DACs is an ideal low-pass filter. In the recent paper [48], we have analyzed, using the models developed in this paper, spectral and spatial emissions under more realistic assumptions on the reconstruction stage in the DACs.

APPENDIX A
DERIVATION OF (52)

It can be shown (see, e.g., [57, Eq. (9.7)]) that for midtread rounding quantizers it holds that

$$
\mathbb{E}\big[e_m^C e_n^R\big] = \frac{\Delta^2}{2\pi^2} \sum_{a=1}^{\infty} \sum_{b=1}^{\infty} \frac{(-1)^{a+b}}{ab}
$$
$$
\times \bigg(\Re\bigg\{ \varphi_{z_m^C, z_n^R}\bigg(\frac{2\pi a}{\Delta}, -\frac{2\pi b}{\Delta} \bigg) \bigg\}
$$
$$
- \Re\bigg\{ \varphi_{z_m^C, z_n^R}\bigg(\frac{2\pi a}{\Delta}, \frac{2\pi b}{\Delta} \bigg) \bigg\} \bigg) \tag{58}
$$

where $\varphi_{z_m^C, z_n^R}(u,v) = \mathbb{E}\big[e^{j(uz_m^C + vz_n^R)}\big]$ is the characteristic function of $[z_m^C, z_n^R]^T$, for $C \in \{R, I\}$.

Note that by adding a constant $\Delta/2$ to the input of a midtread rounding quantizer, the output of said midtread rounding quantizer equals exactly the output of a midrise rounding quantizer for the case when no constant has been added. Further note that $\varphi_{z_m^C + \Delta/2, z_n^R + \Delta/2}(u,v) = e^{j\Delta(u+v)/2}\varphi_{z_m^C, z_n^R}(u,v)$. Hence, for midrise rounding quantizers, it holds that

$$
\mathbb{E}\big[e_m^C e_n^R\big] = \frac{\Delta^2}{2\pi^2} \sum_{a=1}^{\infty} \sum_{b=1}^{\infty} \frac{(-1)^{a+b}}{ab}
$$
$$
\times \bigg(\Re\bigg\{ e^{j\pi(a-b)}\varphi_{z_m^C, z_n^R}\bigg(\frac{2\pi a}{\Delta}, -\frac{2\pi b}{\Delta} \bigg) \bigg\}
$$
$$
- \Re\bigg\{ e^{j\pi(a+b)}\varphi_{z_m^C, z_n^R}\bigg(\frac{2\pi a}{\Delta}, \frac{2\pi b}{\Delta} \bigg) \bigg\} \bigg). \tag{59}
$$

In our case, $[z_m^C, z_n^R]^T$ is a zero-mean Gaussian random vector for which it holds that $\varphi_{z_m^C, z_n^R}(u,v) = \exp\big(-\frac{1}{4}\big(u\sigma_m^2 + 4uv\sigma_{m,n}^C + v\sigma_n^2\big)\big)$, which we use to simplify (59) as follows:

$$
\mathbb{E}\big[e_m^C e_n^R\big]
$$
$$
= \frac{\Delta^2}{2\pi^2} \sum_{a=1}^{\infty} \sum_{b=1}^{\infty} \frac{1}{ab} \exp\bigg(-\frac{\pi^2\big(a^2\sigma_m^2 + b^2\sigma_n^2\big)}{\Delta^2} \bigg)
$$
$$
\times \bigg(\exp\bigg(\frac{2\pi^2 ab\,\sigma_{m,n}^C}{\Delta^2} \bigg) - \exp\bigg(-\frac{2\pi^2 ab\,\sigma_{m,n}^C}{\Delta^2} \bigg) \bigg) \tag{60}
$$
$$
= \frac{\Delta^2}{\pi^2} \sum_{a=1}^{\infty} \sum_{b=1}^{\infty} \frac{1}{ab} \exp\bigg(-\frac{\pi^2\big(a^2\sigma_m^2 + b^2\sigma_n^2\big)}{\Delta^2} \bigg)
$$
$$
\times \sinh\bigg(\frac{2\pi^2 ab\,\sigma_{mn}^C}{\Delta^2} \bigg). \tag{61}
$$

Here, (60) holds as $\Re\big\{e^{j\pi(a-b)}\big\} = \cos(\pi(a-b)) = (-1)^{a+b}$ and $\Re\big\{e^{j\pi(a+b)}\big\} = \cos(\pi(a+b)) = (-1)^{a+b}$ for $a \in \mathbb{Z}$ and $b \in \mathbb{Z}$. We obtain (61) by noting that $2\sinh(x) = e^x - e^{-x}$.

From (61) we finally obtain the desired result (52) by noting that $\mathbb{E}[e_m e_n^*] = 2\big(\mathbb{E}\big[e_m^R e_n^R\big] + j\,\mathbb{E}\big[e_m^I e_n^R\big]\big)$, which holds as the input to the DACs is a circularly-symmetric Gaussian random variable.

APPENDIX B
PROOF OF THEOREM 1

Recall that, for the case $L = 2$, $\alpha = \sqrt{2P/(\Delta^2\xi B)}$ and $\mathrm{diag}(\mathbf{g})$ is given by (26). Hence, since $P < \infty$, it follows from (48) that

$$
\lim_{\Delta \to \infty} \mathbf{C_x} = \lim_{\Delta \to \infty} \frac{2P}{\Delta^2 \xi B} \mathbf{C_e}. \tag{62}
$$

Hence, to prove Theorem 1, we need to show that the entry on the mth row and on the nth column of the RHS of (62) equals the RHS of (42) if we use (53) to evaluate $\mathbf{C_e}$. Define $u_a = a/\Delta$ and $v_b = b/\Delta$. Then, we can write the entry on the mth row and nth column of the RHS of (62) as follows:

$$
\lim_{\Delta \to \infty} [\mathbf{C_x}]_{m,n}
$$
$$
= \lim_{\Delta \to \infty} \frac{4P}{\pi^2 \Delta^2 \xi B}
$$
$$
\times \sum_{a=1}^{\infty} \sum_{b=1}^{\infty} \frac{1}{u_a v_b} \exp\big(-\pi^2\big(u_a^2\sigma_m^2 + v_b^2\sigma_n^2\big)\big)
$$
$$
\times \big(\sinh\big(2\pi^2 u_a v_b\,\sigma_{m,n}^R\big) + j\sinh\big(2\pi^2 u_a v_b\,\sigma_{m,n}^I\big) \big) \tag{63}
$$
$$
= \frac{4P}{\pi^2 \xi B} \int_0^\infty \int_0^\infty \frac{1}{uv} \exp\big(-\pi^2\big(u^2\sigma_m^2 + v^2\sigma_n^2\big)\big)
$$
$$
\times \big(\sinh\big(2\pi^2 uv\sigma_{m,n}^R\big) + j\sinh\big(2\pi^2 uv\sigma_{m,n}^I\big) \big) \,\mathrm{d}u\,\mathrm{d}v \tag{64}
$$
$$
= \frac{2P}{\pi \xi B} \bigg(\arcsin\bigg(\frac{\sigma_{m,n}^R}{\sigma_m \sigma_n} \bigg) + j\arcsin\bigg(\frac{\sigma_{m,n}^I}{\sigma_m \sigma_n} \bigg) \bigg). \tag{65}
$$

Here, to obtain (63), we replaced $[\mathbf{C_e}]_{m,n}$ with (52). To obtain (64), we used that (63) is a two-dimensional Riemann sum, which, by definition, can be written as the two-dimensional integral in (64). Finally, to obtain (65), we used that

$$
\frac{2}{\pi} \int_0^\infty \int_0^\infty \frac{1}{uv} \exp\big(-\pi^2\big(u^2\sigma_m^2 + v^2\sigma_n^2\big)\big)
$$
$$
\times \sinh\big(2\pi^2 uv\,\sigma_{m,n}^C\big) \,\mathrm{d}u\,\mathrm{d}v = \arcsin\bigg(\frac{\sigma_{m,n}^C}{\sigma_m \sigma_n} \bigg). \tag{66}
$$

We note that (65) is equal to the RHS of (42), which concludes the proof.

REFERENCES

[1] S. Jacobsson, G. Durisi, M. Coldrey, and C. Studer, "Massive MU-MIMO-OFDM downlink with one-bit DACs and linear precoding," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Singapore, Singapore, Dec. 2017.

[2] F. Boccardi, R. W. Heath Jr., A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[3] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[4] F. Rusek, D. Persson, B. Kiong, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[5] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[6] L. Lu, G. Ye Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.

[7] U. Gustavsson, C. Sanchéz-Perez, T. Eriksson, F. Athley, G. Durisi, P. Landin, K. Hausmair, C. Fager, and L. Svensson, "On the impact of hardware impairments on massive MIMO," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Austin, TX, USA, Dec. 2014, pp. 294–300.

[8] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah, "Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," *IEEE Trans. Inf. Theory*, vol. 11, no. 60, pp. 7112–7139, Nov. 2014.

[9] X. Zhang, M. Matthaiou, M. Coldrey, and E. Björnson, "Impact of residual transmit RF impairments on training-based MIMO systems," *IEEE Trans. Commun.*, vol. 63, no. 8, pp. 2899–2911, Aug. 2015.

[10] C. Studer, M. Wenk, and A. Burg, "MIMO transmission with residual transmit-RF impairments," in *Proc. Int. ITG Workshop on Smart Antennas (WSA)*, Bremen, Germany, Feb. 2010, pp. 189–196.

[11] E. G. Larsson and L. Van der Perre, "Out-of-band radiation from antenna arrays clarified," *IEEE Wireless Commun. Lett.*, vol. 7, no. 4, pp. 610–613, Feb. 2018.

[12] N. N. Moghadam, P. Zetterberg, P. Händel, and H. Hjalmarsson, "Correlation of distortion noise between the branches of MIMO transmit antennas," in *Proc. IEEE Int. Symp. Personal, Indoor, Mobile Radio Commun. (PIMRC)*, Sydney, NSW, Australia, Sep. 2012, pp. 2079–2084.

[13] C. Mollén, J. Choi, E. G. Larsson, and R. W. Heath Jr., "Achievable uplink rates for massive mimo with coarse quantization," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, New Orleans, LA, USA, Nov. 2017, pp. 6488–6492.

[14] S. Blandino, C. Desset, A. Bourdoux, L. Van der Perre, and S. Pollin, "Analysis of out-of-band interference from saturated power amplifiers in massive MIMO," in *Proc. Eur. Conf. Netw. Commun. (EuCNC)*, Oulu, Finland, Jul. 2017.

[15] E. Björnson, L. Sanguinetti, and J. Hoydis, "Can hardware distortion correlation be neglected when analyzing uplink SE in massive MIMO?" in *Proc. IEEE Int. Workshop Signal Process. Advances Wireless Commun. (SPAWC)*, Kalamata, Greece, Jun. 2018.

[16] C.-K. Wen, C.-J. Wang, S. Jin, K.-K. Wong, and P. Ting, "Bayes-optimal joint channel-and-data estimation for massive MIMO with low-precision ADCs," *IEEE Trans. Signal Process.*, vol. 64, no. 10, pp. 2541–2556, Jul. 2015.

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

[18] 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 Trans. Signal Process.*, vol. 65, no. 15, pp. 4075–4089, Aug. 2017.

[19] S. Jacobsson, G. Durisi, M. Coldrey, U. Gustavsson, and C. Studer, "Throughput analysis of massive MIMO uplink with low-resolution ADCs," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 4038–4051, Jun. 2017.

[20] C. Mollén, J. Choi, E. G. Larsson, and R. W. Heath Jr., "Uplink performance of wideband massive MIMO with one-bit ADCs," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 87–100, Oct. 2016.

[21] L. Fan, S. Jin, C.-K. Wen, and H. Zhang, "Uplink achievable rate for massive MIMO systems with low-resolution ADC," *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2186 – 2189, Dec. 2015.

[22] 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 *Proc. Inf. Theory Applicat. Workshop (ITA)*, San Diego, CA, USA, Feb. 2015, pp. 191–198.

[23] M. Sarajalic, L. Liu, and O. Edfors, "When are low resolution ADCs energy efficient in massive MIMO?" *IEEE Access*, vol. 5, pp. 14 837–14 853, Nov. 2017.

[24] A. Mezghani and J. A. Nossek, "Capacity lower bound of MIMO channels with output quantization and correlated noise," in *IEEE Int. Symp. Inf. Theory (ISIT)*, Cambridge, MA, USA, Jul. 2012.

[25] B. Li, L. N., and W. Zhang, "On transmission model for massive MIMO under low-resolution output quantization," in *Proc. IEEE Veh. Technol. Conf. Spring (VTC-Spring)*, Sydney, NSW, Australia, Jun. 2017.

[26] A. K. Saxena, I. Fijalkow, and A. L. Swindlehurst, "Analysis of one-bit quantized precoding for the multiuser massive MIMO downlink," *IEEE Trans. Signal Process.*, vol. 65, no. 17, pp. 4624–4634, Sep. 2017.

[27] Y. Li, C. Tao, A. L. Swindlehurst, A. Mezghani, and L. Liu, "Downlink achievable rate analysis in massive MIMO systems with one-bit DACs," *IEEE Commun. Lett.*, vol. 21, no. 7, pp. 1669–1672, Jul. 2017.

[28] A. Mezghani, R. Ghiat, and J. A. Nossek, "Transmit processing with low resolution D/A-converters," in *Proc. IEEE Int. Conf. Electron., Circuits, Syst. (ICECS)*, Yasmine Hammamet, Tunisia, Dec. 2009, pp. 683–686.

[29] S. Jacobsson, G. Durisi, M. Coldrey, T. Goldstein, and C. Studer, "Quantized precoding for massive MU-MIMO," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 4670–4684, Nov. 2017.

[30] ——, "Nonlinear 1-bit precoding for massive MU-MIMO with higher-order modulation," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Nov. 2016, pp. 763–767.

[31] H. Jedda, J. A. Nossek, and A. Mezghani, "Minimum BER precoding in 1-bit massive MIMO systems," in *IEEE Sensor Array and Multichannel Signal Process. Workshop (SAM)*, Rio de Janeiro, Brazil, Jul. 2016.

[32] A. L. Swindlehurst, A. K. Saxena, A. Mezghani, and I. Fijalkow, "Minimum probability-of-error perturbation precoding for the one-bit massive MIMO downlink," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, New Orleans, LA, USA, Mar. 2017, pp. 6483–6487.

[33] O. Castañeda, S. Jacobsson, G. Durisi, M. Coldrey, T. Goldstein, and C. Studer, "1-bit massive MU-MIMO precoding in VLSI," *IEEE J. Emerging Sel. Topics Circuits Syst.*, vol. 7, no. 4, pp. 508–522, Dec. 2017.

[34] L. Landau and R. C. de Lamare, "Branch-and-bound precoding for multiuser MIMO systems with 1-bit quantization," *IEEE Wireless Commun. Lett.*, vol. 6, no. 6, pp. 770–773, Dec. 2017.

[35] S. Jacobsson, W. Xu, G. Durisi, and C. Studer, "MSE-optimal 1-bit precoding for multiuser MIMO via branch and bound," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Calgary, AB, Canada, Apr. 2018, pp. 3589–3593.

[36] A. Li, C. Masouros, F. Liu, and A. L. Swindlehurst, "Massive MIMO 1-bit DAC transmission: A low-complexity symbol scaling approach," *IEEE Trans. Wireless Commun.*, vol. 17, no. 11, pp. 7559–7575, Sep. 2018.

[37] M. Shao, Q. Li, and W.-K. Ma, "One-bit massive MIMO precoding via minimum symbol-error probability design," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Calgary, AB, Canada, Mar. 2018, pp. 3579–3583.

[38] M. Shao, Q. Li, W.-K. Ma, and A. M.-C. So, "A framework for one-bit and constant-envelope precoding over multiuser massive MISO channels," Oct. 2018. [Online]. Available: https://arxiv.org/abs/1810.03159

[39] F. Sohrabi, Y.-F. Liu, and W. Yu, "One-bit precoding and constellation range design for massive MIMO with QAM signaling," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 3, pp. 557–570, Jun. 2018.

[40] R. D. J. Guerreiro and P. Montezuma, "Use of 1-bit digital-to-analogue converters in massive MIMO systems," *IEEE Electron. Lett.*, vol. 52, no. 9, pp. 778–779, Apr. 2016.

[41] S. Jacobsson, O. Castañeda, C. Jeon, G. Durisi, and C. Studer, "Nonlinear precoding for phase-quantized constant-envelope massive MU-MIMO-OFDM," in *Proc. IEEE Int. Conf. Telecommunications (ICT)*, St. Malo, France, Jun. 2018, pp. 367–372.

[42] A. Nedelcu, F. Steiner, M. Staudacher, G. Kramer, W. Zirwas, R. Sisava Ganesan, P. Baracca, and S. Wesemann, "Quantized precoding for multi-antenna downlink channels with MAGIQ," in *Int. ITG Workshop on Smart Antennas (WSA)*, Bochum, Germany, Mar. 2017.

[43] H. Jedda, A. Mezghani, J. A. Nossek, and A. L. Swindlehurst, "Massive MIMO downlink 1-bit precoding for frequency selective channels," in *Int. Workshop Comput. Advances in Multi-Sensor Adaptive Process. (CAMSAP)*, Curacao, Curacao, Dec. 2017.

[44] J. J. Bussgang, "Crosscorrelation functions of amplitude-distorted Gaussian signals," Res. Lab. Elec., Cambridge, MA, USA, Tech. Rep. 216, Mar. 1952.

[45] H. E. Rowe, "Memoryless nonlinearities with Gaussian inputs: Elementary results," *Bell Labs Tech. J.*, vol. 61, no. 7, pp. 1519–1525, Sep. 1982.

[46] J. H. Van Vleck and D. Middleton, "The spectrum of clipped noise," *Proc. IEEE*, vol. 54, no. 1, pp. 2–19, Jan. 1966.

[47] F. Maloberti, *Data converters.* Springer, 2007.

[48] S. Jacobsson, M. Coldrey, G. Durisi, and C. Studer, "On out-of-band emissions of quantized precoding in massive MU-MIMO-OFDM," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Oct.–Nov. 2017, pp. 21–26.

[49] S. Gunnarsson, M. Bortas, Y. Huang, C.-M. Chen, L. Van der Perre, and O. Edfors, "Lousy processing increases energy efficiency in massive MIMO systems," in *Proc. Eur. Conf. Netw. Commun. (EuCNC)*, Oulu, Finland, Jun. 2017.

[50] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2325–2383, Oct. 1998.

[51] D. M. Arnold, H.-A. Loeliger, P. O. Vontobel, A. Kavcic, and W. Zeng, "Simulation-based computation of information rates for channels with memory," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3498–3508, Aug. 2006.

[52] A. Lapidoth, "Nearest neighbor decoding for additive non-Gaussian noise channels," *IEEE Trans. Inf. Theory*, vol. 42, no. 5, pp. 1520–1529, Sep. 1996.

[53] W. Zhang, "A general framework for transmission with transceiver distortion and some applications," *IEEE Trans. Commun.*, vol. 60, no. 2, pp. 384–399, Feb. 2012.

[54] B. Widrow, I. Kollár, and M.-C. Liu, "Statistical theory of quantization," *IEEE Trans. Instrum. Meas.*, vol. 45, no. 2, pp. 353–361, Apr. 1996.

[55] A. B. Sripad and D. L. Snyder, "A necessary and sufficient condition for quantization errors to be uniform and white," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 25, no. 5, pp. 442–448, Oct. 1977.

[56] S. P. Lipshitz, R. A. Wannamaker, and J. Vanderkooy, "Quantization and dither: A theoretical survey," *J. Audio Eng. Soc.*, vol. 40, no. 5, pp. 355–375, May 1992.

[57] B. Widrow and I. Kollár, *Quantization Noise: Roundoff Error in Digital Computation, Signal Processing, Control, and Communications*. Cambridge Univ. Press, 2008.

[58] 3GPP, "LTE; evolved universal terrestrial radio access (E-UTRA); user equipment (UE) radio transmission and reception," Apr. 2017, TS 36.101 version 14.3.0 Rel. 14.