Digital Polar Transmitters for Massive MIMO: Sum-Rate and Power Efficiency Analysis

Guixian Xu, *Member, IEEE*, Vesa Lampu, *Graduate Student Member, IEEE*, Marko Kosunen, *Member, IEEE*, Vishnu Unnikrishnan, *Member, IEEE*, Jussi Ryynänen, *Senior Member, IEEE*, Mikko Valkama, *Fellow, IEEE*, and Lauri Anttila, *Member, IEEE*

Abstract—In this article, we comprehensively investigate the potential of the digital polar radio transmitter architecture for multi-user massive multiple-input multiple-output orthogonal frequency-division multiplexing (MIMO-OFDM) downlink system. In terms of throughput performance, we derive a lower bound for the average sum-rate achievable with Gaussian signaling inputs and zero-forcing (ZF) precoding based on Bussgang decomposition. By diagonal approximation, we derive an approximate, yet accurate, model for the distortion caused by uniform polar quantization, which can be used to evaluate the corresponding sum-rate in closed form. To assess the power efficiency, we provide power consumption models with realistic parameters and values for the quantized polar and Cartesian transmitters, based on state-of-the-art integrated circuit (IC) designs and measurements. Extensive numerical results demonstrate that the proposed quantized polar transmitter can enable excellent performance in terms of average sum-rate, symbol error rate (SER), and out-of-band (OOB) emission level, compared to the Cartesian architecture. Furthermore, the power consumption comparisons show that the digital polar transmitter can save more than 36% in the energy consumption under 64-antenna setting in typical 5G enhanced mobile broadband use cases, thus making it highly appealing for future power-efficient massive MIMO transmitter implementations.

Index Terms—5G, massive multiple-input multiple-output (MIMO), orthogonal frequency-division multiplexing (OFDM), zero-forcing (ZF) precoding, digital polar transmitter, phase modulator, power efficiency, radio transmitters, sum-rate.

I. INTRODUCTION

MASSIVE multiple-input multiple-output (MIMO) is a key enabler in the fifth generation (5G) and beyond wireless communication systems. To this end, networks with large numbers of antennas at the base station (BS) could theoretically achieve multiple orders of spectral and energy efficiency gains over current 4G networks [1], [2]. Moreover, massive MIMO is regarded as a potential enabler of massive connectivity and ultra low latency in the context of 5G internet of things (IoT) [3]. However, the efficient implementation of a massive MIMO radio frequency frontend (RFFE) with hundreds or even thousands of antennas is a significant challenge [4]. One can directly implement as many conventional RF chains and independently controlled RF outputs as required, however, the power consumption, size, and cost of the mixed-signal components, i.e., analog-to-digital converters (ADCs) and digital-to-analog converters (DACs), and the RF parts, increase proportionally with the number of antennas. Furthermore, wideband orthogonal frequency-division multiplexing (OFDM) signals and high-order quadrature amplitude modulation (QAM) schemes utilized in 5G result in high peak-to-average power ratio (PAPR). This lowers the power efficiency of power amplifiers (PAs) in power-back-off operation [5]. Thus, it is highly desirable to develop custom BS radio transmitters and energy efficient transmission schemes [6] to increase hardware integration and output power, and to reduce the cost and power consumption of large-scale MIMO implementations at sub-6 GHz and mmWave bands [4].

A. Relevant Prior Art

1) Constant-Envelope and One-Bit Precoding: One key strategy for improving the cost and power efficiency of the transmitter is to cut the power consumption of the current power-hungry circuit components, such as high-resolution DACs and RF PAs. As an example, it has been demonstrated that massive MIMO systems with coarsely quantized DACs, even with one-bit quantization, can achieve satisfactory bit/symbol error rate (BER/SER) and achievable rate performance, especially if combined with tailored nonlinear precoding [7]–[10].

Another appealing option is to generate *constant envelope (CE)* waveforms with 0 dB PAPR, which would enable the highest efficiency for traditional PAs or the utilization of switch-mode PAs. In general, the complex baseband signals of the CE waveforms at each transmitting antenna are constant-amplitude and thus information can only be modulated in the phase of the transmitted signal [11]. This stringent constraint, however, degrades both the power and spectral efficiency, as well as introduces new signal processing challenges. To tackle this issue, through spatial processing, one-bit and CE precoding have been recently proposed and have since attracted wide interest in the literature [12], [13]. With CE precoding, the transmitted signal at each antenna is restricted to take an M-ary phase-shift keying (PSK) form or to have a continuous constant envelope [14]. It is also worth noting that one-bit precoding can be regarded as a special case of CE precoding from the precoding design point of view.

TABLE I
SUMMARY OF THE STATE-OF-THE-ART ONE-BIT/CE PRECODING APPROACHES FOR DOWNLINK MIMO-OFDM

Transmission scheme	Design principle	Performance	Complexity	Notes
One-bit DAC and linear precoding [17]	Block-level linear precoding with one-bit quantization	Poor BER/throughput and high OOB emission	Low	A special case in [18], closed-form analytical sum-rate expression
MU-CEP [5]	Nonlinear symbol-level, heuristic iterative algorithm	High power efficiency, similar data rate as ZF	High	An OFDM extension of [19]
SQUID-OFDM [20]	Nonlinear symbol-level, SQUID algorithm	Better BER than quantized MMSE	Moderate	CE with phase quantization
CESLP [21]	Nonlinear symbol-level, CCD algorithm	Improved BER for SQUID-OFDM algorithm	High	Not applicable to large-size problem
CD/GN CE [22]	Nonlinear symbol-level, CCD and GN algorithms	Better SER than ZF and CE precoding in [5]	High	Not applicable to large-size problem
MSM [23]	Nonlinear symbol-level, one-bit precoder for MSM	Better BER than quantized ZF	Moderate	An OFDM extension of [15], only for PSK

The abundant CE precoding designs can be classified into *mean squared error (MSE)-based* nonlinear precoding and *constellation-dependent* precoding designs. The former approach generates the constant envelope transmit vector directly by solving a nonlinear least-squares optimization problem, resulting in better BER performance compared to the quantized linear precoding [12]. The latter designs solve the well-reformulated optimization problems by taking into account the constellation structures for performance improvement [13]. However, the current algorithmic developments for interference exploitation with symbol-level precoding are intimately linked with factors such as the chosen design formulation, the scenario (e.g., single user or multi-user), and the utilized modulation constellation. In particular scenarios and with specific modulation constellations (e.g., low-order QAM and PSK), efficient algorithms or closed-form solutions have been developed for one-bit and CE precoding [15], [16].

For the above taxonomy, we note that one-bit/CE precoding for the narrow-band massive MIMO systems is well-studied [13], while in this work we focus on wideband OFDM transmission. Table I summarizes the state-of-the-art of one-bit/CE precoding for downlink MIMO-OFDM, while the one-bit quantized linear precoding such as maximal-ratio transmission (MRT) and zero-forcing (ZF) [17] is also listed for comparison. It was shown in [5] that the algorithm presented in [19] for CE single-carrier precoding over frequency-selective channels can be extended to OFDM transmission. The iterative time-domain processing includes the computation of the convolution at each iteration (see (64) in [5]), which leads to a high computational complexity. To reduce the complexity, [20] proposed a squared-infinity norm Douglas-Rachford splitting (SQUID) algorithm, called SQUID-OFDM. In [21], [22], the CE precoding designs are formulated into a constrained least-squares problem with a constant modulus constraint and solved using cyclic coordinate descent (CCD) based iterative algorithm, and more efficient Gauss-Newton (GN) algorithm. Based on the concept of constructive interference, a nonlinear precoder to transmit PSK signals in the one-bit quantized MIMO-OFDM systems is proposed in [23]. A maximizing the safety margin (MSM) problem is formulated to optimize the decision thresholds and reduce the quantization distortion. While the developed designs mentioned above have their distinct benefits, they also suffer from a loss in the achievable data rate, since a large share of the transmitted power is wasted into the channel null-space. Furthermore, these methods rely on a large number of iterations for optimization, at symbol-level, which is a considerable challenge for real-time implementation of wideband massive MIMO transmitters.

2) Digital Transmitter Architectures: The emerging paradigm of the digital-intensive RFFE design naturally leads to transmitter architectures relying on extensive *time-based signal processing*. In time-domain signal processing, the information is coded into the *phase or delay* of the signal while the dependency on amplitude resolution can be limited or even removed. Thus, digital logic can be extensively utilized to generate these signals [24], [25]. Then, the RFFE integrated on a single chip can also benefit from complementary metal-oxide-semiconductor (CMOS) process scaling, since smaller transistors allow faster switching time and thus increased time resolution while consuming less energy per transition [26], [27]. Moreover, time-based transmitter architectures have the potential to adapt well to phased-array and massive MIMO systems, where precise phase control of each antenna element is required. Well-known examples of time-based transmitter architectures are polar, outphasing and pulse-width modulation (PWM) transmitters. All these architectures have been invented decades ago, but they have only recently found new life thanks to advancements in *digital PAs (DPAs)* and wideband digital *phase modulators*. The emerging DPA designs that incorporate the functionality of DAC and PA in the same circuit can be used in combination with time-based transmitter architectures to leverage the advantages of switched-mode PAs [28]. Specifically, the switched-capacitor PA (SCPA) has become a popular DPA architecture in the RF integrated circuit (RFIC) domain, since it can achieve high accuracy and linearity by exploiting the precision of capacitance ratios that CMOS processes provide, while simultaneously achieving good power efficiency [29]. For further descriptions and implementations of the various DPA designs, the reader is referred to the overview article [30] and references therein.

In terms of power consumption, the digital polar architecture is the most promising transmitter architecture. It has been shown that polar architecture has an inherent power efficiency gain of up to 3 dB compared to the Cartesian architecture, due to the in-phase and quadrature (I/Q) signal summation in

the latter [31]. Several successful designs and chip prototypes [32]–[36] have demonstrated the feasibility and superiority of the digital polar transmitters. Specifically, [32] demonstrated an analysis and implementation of a polar quantizer in a CMOS process for wireless receivers. It showed, both analytically and experimentally, that the polar quantizer can achieve a higher signal-to-quantization-noise ratio (SQNR) compared to the rectangular (I/Q) quantizer. Then, a polar transmitter was implemented in [33] for mmWave communications, wherein the phase modulated signal was generated with upconversion mixers while the amplitude signal modulated a 4-bit RF-DAC. The first polar transmitter to meet either the linearity requirements of 256-QAM WLAN signals or the transmit signal quality requirement of aggregated 40 MHz LTE signal was reported in [34], in which the DPA delivers a peak output power of 21.9 dBm with 41% drain efficiency. In [35], a digital polar transmitter using a digital-to-time-converter (DTC) to enable high-efficiency RF-DAC for multi-band applications was implemented. The recent implementation of an all-digital polar transmitter for 2.5/5 GHz dual-band Wi-Fi 6 application [36] shows that the DPAs can reach a peak power/average power efficiency of 27 dBm/53% and 27 dBm/37% at low and high frequency bands, respectively. However, to the best of our knowledge, there are no implementations reported for large-scale MIMO transmitters. Furthermore, comprehensive theoretical analysis and performance characterization of applying the digital polar transmitter in massive MIMO-OFDM systems are still lacking. This paper is the first to address these issues, in terms of providing performance analysis tools and results related to the adoption of polar radio transmitters in massive MIMO systems. Such fundamental results and design insight related to, e.g., dimensioning the antenna array size in different deployment cases are of instrumental importance, paving the way towards actual hardware implementations.

B. Contributions

In contrast to the existing works [17], [18], [20], which consider the Cartesian transmitter with finite-resolution DACs and/or CE precoding in massive MIMO, we focus in this paper on the *polar radio transmitter architecture*. This architecture can benefit significantly from CMOS process downscaling and highly-efficient DPAs, contrary to the Cartesian architecture [37]. To better understand the potential advantages of the polar transmitter, we first investigate the average sum-rate, SER, and out-of-band (OOB) spectral emissions of the quantized polar transmitter in the wideband multi-user massive MIMO-OFDM downlink system. Then, we analyze the power efficiency of the quantized polar transmitter based on proposed power consumption models for the circuit components. Commonly used asymptotic analyses for massive MIMO systems are not applicable here, because the assumptions of (*i*) Gaussian multi-user interference and nonlinear distortion and (*ii*) Gaussian distribution-based channel models required for such analysis are not fulfilled in our system model. Thus, we follow a similar analysis framework as [18], which analyzed the achievable rate performance of the Cartesian architecture with low-resolution DACs in massive MIMO-OFDM using Bussgang decomposition, and then investigate and compare the performance of

both architectures. However, the analysis methods in [18] are only valid for the case of negligible overload distortion, whereas we take the overload distortion into account explicitly in this study. Compared to our earlier work in [38], here we develop more accurate and complete system analyses and performance evaluations, in the framework of the 5G enhanced mobile broadband (eMBB) applications. Specifically, in [38], a second-order Taylor approximation was utilized which yields good performance estimates with high phase resolution. Following up on this, in the current work, a more generic analysis is developed, being applicable also with lower phase resolutions.

The key contributions of the paper can be stated and summarized as follows:

- Firstly, a digital polar transmitter architecture for massive MIMO-OFDM systems is presented to improve the transmitter's power efficiency.
- We derive a lower bound for the average sum-rate achievable with Gaussian signaling and ZF precoding, utilizing Bussgang decomposition and a diagonal approximation which treats the polar quantization distortion as uncorrelated in both the spatial and temporal domains.
- We analyze and compare the total power consumption of the quantized polar transmitter and the traditional Cartesian architecture based on proposed power consumption models whose parameters/values in each required circuit block are extracted from state-of-the-art circuit designs and transmitter implementations.
- The extensive numerical results demonstrate that the quantized polar transmitter enables superior performance in terms of achievable average sum-rate, uncoded SER and OOB emission levels, compared to the Cartesian one [18] and CE precoding with the popular SQUID-OFDM algorithm [20]. Specifically, we show that only 3-4 amplitude bits and 5-6 phase bits are sufficient to approach the system performance without quantization.
- The power consumption comparison results show that the quantized polar transmitter architecture is more power-efficient than the Cartesian architecture in two typical 5G eMBB use cases. For a 64-antenna BS, the digital polar transmitter can save more than 36% in the total power consumption in both use cases. More broadly, the results suggest that the optimum array size for a given use case can be different for the different transmitter architectures.

C. Organization and Notation

Organization: The remainder of this paper is organized as follows. In Section II, we introduce the system model and the concept of the quantized polar transmitter. In Section III, we derive the signal-to-interference-noise-and-distortion ratio (SINDR) at the user equipments (UEs) and develop a lower bound for the average sum-rate achievable with Gaussian signaling and ZF precoding. Then, we derive a diagonal approximated covariance of the distortion caused by polar quantization, which we use to evaluate the sum-rate. In Section IV, we analyze the power consumption of polar and Cartesian transmitters based on proposed power consumption models. In
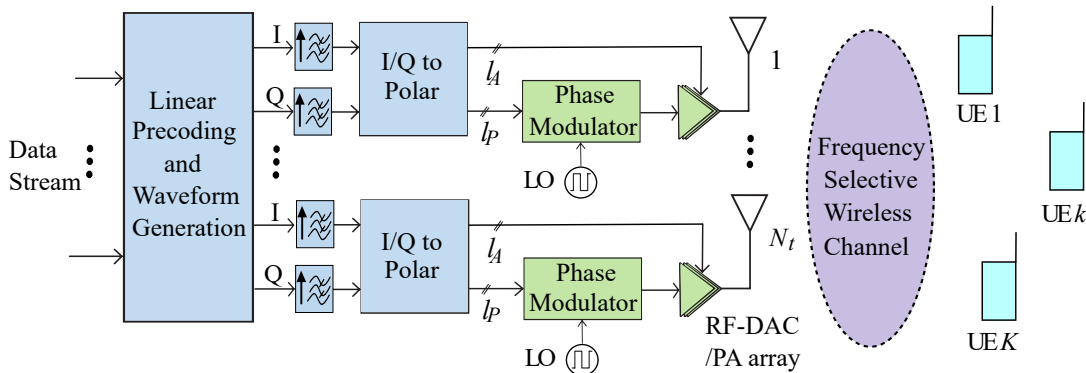
Fig. 1. The block diagram of the considered massive MIMO-OFDM multi-user downlink transmission with the linear precoding and the quantized polar transmitter at the BS.

Section V, we provide extensive numerical results to demonstrate the system performance and the power consumption of the quantized polar transmitter. Section VI concludes the paper.

Notation: Boldfaced lowercase and uppercase letters are used to denote vectors and matrices. $\mathbb{C}^n$ and $\mathbb{R}^n$ represents n-dimensional complex and real vector space, respectively. Superscripts $(\cdot)^T$ and $(\cdot)^H$ stand for the transpose and conjugate transpose operators, respectively, while $\text{vec}(\cdot)$ and $\text{blkdiag}(\cdot)$ denote the vectorization and block diagonal operator, respectively. The operators $\mathbb{E}\{\cdot\}$, $\text{Tr}(\cdot)$, $\lfloor \cdot \rfloor$ and $\| \cdot \|$ stand for the statistical expectation, trace, floor, and vector norm of a vector. Kronecker product is denoted by $\otimes$, while the Hadamard (element-wise) product of two equally-sized matrices is denoted as $\odot$. $\Re\{\cdot\}$ and $\Im\{\cdot\}$ represent the real and imaginary parts of a complex input, respectively. $\mathbf{I}_n$ denotes the $n \times n$ identity matrix and $[\mathbf{X}]_{i,j}$ denotes the (i-th, j-th) element of matrix $\mathbf{X}$. $\mathcal{CN}(\cdot)$ denotes the complex Gaussian distribution. The function $\exp(\cdot)$ denotes the natural exponential function.

II. SYSTEM MODEL

We consider a single-cell massive MIMO-OFDM downlink system as depicted in Fig. 1. The system consists of a BS with N_t antennas, which simultaneously serve K single-antenna UEs over a frequency selective channel, where N_t is significantly larger than K. At the BS, the discrete-time OFDM signal at each antenna is generated by applying the inverse fast Fourier transform (IFFT) to the frequency-domain precoded vector of data symbols, before passing the signal through the polar transmitter. At the UEs, the time-domain received signal is transformed back to frequency domain through FFT followed by OFDM demodulation to obtain the received data symbols. We assume that the subcarrier spacing is Δf and the sampling rate $f_s = \kappa N_{FFT} \Delta f$, where κ is the oversampling ratio and N_{FFT} is the corresponding FFT size. Then, each OFDM symbol consists of $N = \kappa N_{FFT}$ time-domain samples, with a sample interval of $T_c = \frac{1}{N\Delta f}$. We also assume that a cyclic prefix (CP) of length larger than D is added to remove inter-symbol interference, with D denoting the channel impulse response length in samples (cf. (6)).

A. Baseband Precoding

The symbol vector $\mathbf{u}_l = [u_{1,l}, \cdots, u_{K,l}]^T \in \mathcal{O}^K$ associated with the l-th subcarrier ($l = 0, \cdots, N-1$) contains the data symbols for K UEs, where $\mathcal{O}$ represents the set of QAM constellations. Let the disjoint sets $\mathcal{S}_u$ and $\mathcal{S}_g$ be the set of S active subcarriers associated with intended data symbols and the set of $N - S$ guard subcarriers, respectively. Hence, we set $\mathbf{u}_{k,l} = \mathbf{0}_{K \times 1}$ for $l \in \mathcal{S}_g$ and $\mathbb{E}\{\|\mathbf{u}_l\|^2\} = \mathbf{I}_K$ for $l \in \mathcal{S}_u$. In order to cancel multi-user interference, frequency-domain precoded vectors are generated by multiplying $\mathbf{u}_l$ with the per-subcarrier precoding matrices $\mathbf{P}_l \in \mathbb{C}^{N_t \times K}$. The transmitted time-domain signal $\mathbf{z}_n$ is then generated by applying normal OFDM processing at each transmit branch, expressed as

$$\mathbf{z}_n = \frac{1}{\sqrt{N}} \sum_{l \in \mathcal{S}_u} \mathbf{P}_l \mathbf{u}_l \exp(j\frac{2\pi}{N}ln) \tag{1}$$

for $n = 0, 1, \cdots, N-1$. Moreover, the transmitted signal $\mathbf{z}_n$ is assumed to satisfy the total transmit power constraint

$$\mathbb{E}\Big\{ \sum_{n=0}^{N-1} \|\mathbf{z}_n\|^2 \Big\} = P_{\text{total}}, \tag{2}$$

where P_{total} denotes the total transmit power at the BS. In the following, we will focus on linear ZF precoding, which is commonly studied in the Cartesian case in massive MIMO systems.

Assuming that perfect channel state information (CSI) is available at the BS[1], the corresponding precoding matrices for the active subcarrier $l \in \mathcal{S}_u$ are given by

$$\mathbf{P}_l = \frac{1}{\beta} \widehat{\mathbf{H}}_l^H (\widehat{\mathbf{H}}_l \widehat{\mathbf{H}}_l^H)^{-1}, \tag{3}$$

where the associated power scaling constant is given by [18]

$$\beta = \sqrt{\frac{1}{P_{\text{total}}} \sum_{l \in \mathcal{S}_u} \text{Tr}(\widehat{\mathbf{H}}_l \widehat{\mathbf{H}}_l^H)^{-1}}, \tag{4}$$

[1]Note that imperfect CSI will not affect the performance of the polar transmitter architecture any differently than that of other RF transmitter architectures (such as the Cartesian architecture). This is due to the imperfect CSI affecting only the baseband precoder, which is the same in both the polar and Cartesian architectures. This is true also when quantized transmitter architectures are considered, since the quantization takes place after the precoding. This will be concretely demonstrated in Section V-B through numerical results.

and the channel matrix $\widehat{\mathbf{H}}_l = \boldsymbol{\Gamma}^{\frac{1}{2}} \widetilde{\mathbf{H}}_l \in \mathbb{C}^{K \times N_t}$ includes the small-scale fading at the l-th subcarrier $\widetilde{\mathbf{H}}_l$ and the large-scale fading coefficient between the BS and UEs. The diagonal large-scale fading matrix $[\boldsymbol{\Gamma}]_{k,k} = \gamma_k$ contains the path loss, penetration loss and shadow fading for the k-th UE. For the small-scale fading, we adopt a cluster-based channel model to address the frequency-selectivity and spatial correlation characteristics of the massive MIMO channels. We note that 3GPP also utilizes a similar channel modeling approach in 5G mobile radio standardization [39].

Thus, the delay-d channel vector for the UE k can be expressed as

$$\mathbf{h}_k[d] = \sum_{c=1}^{N_{cl}} \sum_{r=1}^{N_{ray}} \alpha_{rc} f_{\text{flt}}(dT_c - \kappa_c - \kappa_r)$$
$$\times a_{\text{UE}}(\phi_c - \varphi_r)\mathbf{a}_{\text{BS}}(\theta_c - \vartheta_r), \qquad (5)$$

where there are N_{cl} clusters and each cluster consists of N_{ray} rays. Each cluster c has a time-delay κ_c, angle of arrival (AoA) ϕ_c, and angle of departure (AoD) θ_c, while each ray r with the ray delay κ_r and the corresponding AoA and AoD are denoted by φ_r and ϑ_r, respectively. The complex gain $\alpha_{rc} \sim \mathcal{CN}(0, \sigma_c^2)$, where σ_c^2 represents the average power of the cluster c. The function $f_{\text{flt}}(\cdot)$ is a pulse-shaping function for T_c-spaced signaling. $\mathbf{a}_{\text{BS}}(\cdot)$ denotes the antenna array response vector of the BS, and $a_{\text{UE}}(\cdot)$ accounts for the phase between the clusters and UEs. For the uniform linear array (ULA), $\mathbf{a}_{\text{BS}}(\phi) = \frac{1}{\sqrt{N_t}}[1, e^{j\frac{2\pi}{\lambda} d_s \sin(\phi)}, \cdots, e^{j(N_t-1)\frac{2\pi}{\lambda} d_s \sin(\phi)}]^T$, where λ and d_s denote the carrier wavelength and inter-antenna spacing, respectively. Thus, the compound multi-user channel matrix is $\mathbf{H}[d] = [\mathbf{h}_1[d], \cdots, \mathbf{h}_K[d]]^T \in \mathbb{C}^{K \times N_t}$. Lastly, the corresponding multi-user frequency-domain response at l-th subcarrier is given by

$$\widetilde{\mathbf{H}}_l = \sum_{d=0}^{D-1} \mathbf{H}[d] \exp(-j\frac{2\pi}{N} ld), \qquad (6)$$

where D denotes the channel impulse response (CIR) length in samples. A line-of-sight (LOS) component can also be added in (5) to account for a propagation environment with a LOS path.

B. Quantized Polar Transmitter Architecture

Instead of the conventional and popular Cartesian architecture, a quantized polar transmitter architecture is adopted at the BS. As shown in Fig. 1, the time-domain per-antenna transmitted I/Q signal is first upsampled and pulse-shaped. Then, these I/Q signal components are converted to a polar representation in the digital domain with, for example, the coordinate rotation digital computer (CORDIC) algorithm. Thus, the amplitude and phase signals are given by

$$\mathbf{A}_n = \sqrt{\Re\{\mathbf{z}_n\}^2 + \Im\{\mathbf{z}_n\}^2}, \qquad (7)$$

$$\boldsymbol{\Psi}_n = \arctan\left(\frac{\Im\{\mathbf{z}_n\}}{\Re\{\mathbf{z}_n\}}\right), \qquad (8)$$

for $n = 0, \cdots, N - 1$, where the I/Q signal branches of $\mathbf{z}_n$ follow independent identically distributed (i.i.d.) zero-mean asymptotic Gaussian distributions for a sufficiently large number of subcarriers. Thus, the amplitude signal $\mathbf{A}_n$ follows the Rayleigh distribution and the phase signal $\boldsymbol{\Psi}_n$ follows a uniform distribution in the interval $[0, 2\pi)$.

After that, uniform polar quantization is adopted in which the amplitude and phase signals are quantized independently. Then, the quantized amplitude signal directly modulates an l_A-bit resolution (i.e., $L_A = 2^{l_A}$ quantization levels) RF-DAC, acting also as the PA. The phase signal, quantized to l_P bits, modulates the phase of the RF carrier in the phase modulator. The phase modulator can be realized with digitally controlled circuits, for example, a digital phase-locked loop (PLL) [40] or a delay-line based approach [24], [25], [35], [36], [41]. In the last stage, the amplitude signal and the phase modulated signal are combined in the RF-DAC to generate the per-antenna transmitted signal at the desired carrier frequency.

In this work, we assume that RF components are ideal without impairments, aside from the quantization in the DACs and phase modulators. Moreover, perfect timing between the amplitude and phase paths as well as perfect synchronization between the BS and UEs are assumed. However, there are some issues that need to be considered for the practical implementation of the wideband digital polar transmitter. Firstly, the bandwidths of the amplitude and phase signals in polar transmitter are significantly higher than that of the original I/Q signal due to the nonlinear signal conversion from I/Q to polar. Hence, the polar transmitter needs to be able to accommodate wider internal bandwidth or the composite RF signal will be distorted. Secondly, any time delay mismatches between the amplitude and phase signals will result in erroneous restoration of the transmit signal, causing linearity degradation. These are crucial challenges in real-life wideband polar transmitter implementations, especially for massive MIMO-OFDM systems. Nonetheless, it is encouraging that recent studies and implementations mentioned above have demonstrated the feasibility of the digital polar transmitter and phase modulator with more than 100 MHz RF signal bandwidth.

C. Uniform Polar Quantization

Uniform polar quantization is adopted in the polar transmitter, where two uniform real-valued scalar mid-rise quantizers act on both the amplitude and phase paths. Due to the Rayleigh distribution of the amplitude signal, the amplitude quantization is more complicated. In general, one should choose the best step size for minimizing the quantization distortion caused by the finite amplitude resolution. However, too small a step size will result in significant overload distortion. Here, we set the clipping level $\mathcal{A}_{\text{clip}}$ of the amplitude quantizer such that the amplitudes of samples that exceed this level will be clipped. The clipping operation creates signal distortion, resulting in unwanted in-band and out-of-band emissions that degrade error vector magnitude (EVM) and interfere neighboring channel users.

With L_A quantization levels, the step size $\Delta = \frac{\mathcal{A}_{\text{clip}}}{L_A}$ and the quantization thresholds and labels of the amplitude quantizer

are given by

$$\tau_i = (i-1)\Delta, \ i=1,\cdots,L_A, \tau_{L_A+1} = \infty, \quad (9)$$

$$\ell_i = (i-\frac{1}{2})\Delta, \ i=1,\cdots,L_A. \quad (10)$$

Then, the quantized amplitude signal is given by

$$\mathbf{A}_n^Q = \left(\left\lfloor \frac{\mathbf{A}_n}{\mathcal{A}_{\text{clip}}/L_A} \right\rfloor + \frac{1}{2}\right) \frac{\mathcal{A}_{\text{clip}}}{L_A} \quad (11a)$$

$$\triangleq \mathbf{A}_n + \mathbf{e}_{\mathbf{A}_n}, n=0,\cdots,N-1, \quad (11b)$$

where $\mathbf{e}_{\mathbf{A}_n}$ is the quantization error of the amplitude signal. Similarly, the quantization thresholds of the L_P-level (i.e., $L_P = 2^{l_P}$) phase quantizer $\upsilon_p = (p-1)\frac{2\pi}{L_P}$, $p = 1,\cdots,L_P, \upsilon_{L_P+1} = 2\pi$, and the output of the phase quantizer can be expressed as

$$\mathbf{\Psi}_n^Q = \left(\left\lfloor \frac{\mathbf{\Psi}_n}{2\pi/L_P} \right\rfloor + \frac{1}{2}\right) \frac{2\pi}{L_P} \quad (12a)$$

$$\triangleq \mathbf{\Psi}_n + \mathbf{e}_{\mathbf{\Psi}_n}, \ n=0,\cdots,N-1, \quad (12b)$$

where $\mathbf{e}_{\mathbf{\Psi}_n}$ denotes the quantization error of the phase signal.

Then, the quantized transmitted signal at each BS antenna over the n-th time sample can be written as

$$\mathbf{x}_n = \mathbf{A}_n^Q \odot \exp(j\mathbf{\Psi}_n^Q). \quad (13)$$

From input-output perspective, the polar quantizer can be viewed as a single instantaneous nonlinear transformation. Thus, as is demonstrated in the next section, we can apply the Bussgang decomposition in analyzing the average sum-rate.

III. AVERAGE SUM-RATE ANALYSIS

A. Channel Input-Output Relationship

The received discrete time baseband signal $\mathbf{y}_n \in \mathbb{C}^K$ for K UEs is given by

$$\mathbf{y}_n = \sum_{d=0}^{D-1} \mathbf{\Gamma}^{\frac{1}{2}} \mathbf{H}[d]\mathbf{x}_{n-d} + \mathbf{n}_n, \quad (14)$$

where $\mathbf{n}_n \sim \mathcal{CN}(\mathbf{0}, N_0\mathbf{I}_K)$ denotes additive white Gaussian noise (AWGN) at the UEs at time instance n and N_0 is the noise variance. Then, the transmit SNR is defined as P_{total}/N_0.

For analytical convenience, we transfer the time-domain signal model in (14) into frequency domain. Let $\mathbf{F}_N$ be the normalized $N \times N$ discrete Fourier transform (DFT) matrix with the property $\mathbf{F}_N\mathbf{F}_N^H = \mathbf{I}_N$. Then the transmitted and received signal matrices in the frequency-domain are $\widehat{\mathbf{X}} = \mathbf{X}\mathbf{F}_N$ and $\widehat{\mathbf{Y}} = \mathbf{Y}\mathbf{F}_N$, respectively, where $\mathbf{X} = [\mathbf{x}_0, \mathbf{x}_1, \cdots, \mathbf{x}_{N-1}]$ and $\mathbf{Y} = [\mathbf{y}_0, \mathbf{y}_1, \cdots, \mathbf{y}_{N-1}]$ are the corresponding time-domain transmitted and received signal matrices over N time samples. After discarding the CP, the received frequency-domain signal $\widehat{\mathbf{y}}_l \in \mathbb{C}^K$ over K UEs at the l-th subcarrier can be expressed as

$$\widehat{\mathbf{y}}_l = \widehat{\mathbf{H}}_l\widehat{\mathbf{x}}_l + \widehat{\mathbf{n}}_l, \quad (15)$$

where $\widehat{\mathbf{y}}_l$ and $\widehat{\mathbf{x}}_l$ are the l-th column of $\widehat{\mathbf{Y}}$ and $\widehat{\mathbf{X}}$, respectively, and $\widehat{\mathbf{n}}_l \sim \mathcal{CN}(\mathbf{0}, N_0\mathbf{I}_K)$ is the l-th column of the matrix $\widehat{\mathbf{N}} = \mathbf{N}\mathbf{F}_N$, where $\mathbf{N} = [\mathbf{n}_0, \mathbf{n}_1, \cdots, \mathbf{n}_{N-1}]$.

Next, let us define the compact forms for the following vectors and matrices:

$$\widehat{\mathbf{y}} \triangleq \text{vec}([\widehat{\mathbf{y}}_0, \widehat{\mathbf{y}}_1, \cdots, \widehat{\mathbf{y}}_{N-1}]) \in \mathbb{C}^{KN}, \quad (16a)$$

$$\widehat{\mathbf{H}} \triangleq \text{blkdiag}([\widehat{\mathbf{H}}_0, \widehat{\mathbf{H}}_1, \cdots, \widehat{\mathbf{H}}_{N-1}]) \in \mathbb{C}^{KN \times N_tN}, \quad (16b)$$

$$\widehat{\mathbf{P}} \triangleq \text{blkdiag}([\mathbf{P}_0, \mathbf{P}_1, \cdots, \mathbf{P}_{N-1}]) \in \mathbb{C}^{N_tN \times KN}, \quad (16c)$$

$$\widehat{\mathbf{u}} \triangleq \text{vec}([\mathbf{u}_0, \mathbf{u}_1, \cdots, \mathbf{u}_{N-1}]) \in \mathbb{C}^{KN}, \quad (16d)$$

$$\widehat{\mathbf{n}} \triangleq \text{vec}([\widehat{\mathbf{n}}_0, \widehat{\mathbf{n}}_1, \cdots, \widehat{\mathbf{n}}_{N-1}]) \in \mathbb{C}^{KN}, \quad (16e)$$

where $\mathbf{P}_l = \mathbf{0}_{N_t \times K}$ for $l \in \mathcal{S}_g$ in (16c). By using the simplified Bussgang decomposition [18], [42], the transmitted frequency-domain vector $\widehat{\mathbf{x}}$ can be written as

$$\widehat{\mathbf{x}} = \text{vec}(\widehat{\mathbf{X}}) = (\mathbf{F}_N \otimes \mathbf{I}_{N_t})(\mathbf{B}\mathbf{z} + \mathbf{d}), \quad (17)$$

where $\mathbf{B} = \text{diag}(b_1,\ldots,b_{N_t}) \otimes \mathbf{I}_N \in \mathbb{R}^{N_tN}$ is the Bussgang gain matrix, and the quantization error $\mathbf{d} = \text{vec}([\mathbf{d}_0, \mathbf{d}_1, \cdots, \mathbf{d}_{N-1}]) \in \mathbb{C}^{N_tN}$ is uncorrelated with the discrete-time precoded vector

$$\mathbf{z} = (\mathbf{F}_N^H \otimes \mathbf{I}_{N_t})\widehat{\mathbf{P}}\widehat{\mathbf{u}}, \quad (18)$$

which is known to converge to Gaussian as the number of subcarriers goes to infinity. Finally, the frequency-domain received signal over N subcarriers can be rewritten as

$$\widehat{\mathbf{y}} = \widehat{\mathbf{H}}\widehat{\mathbf{x}} + \widehat{\mathbf{n}}. \quad (19)$$

By substituting (17) and (18) into (19), we have

$$\begin{aligned}
\widehat{\mathbf{y}} &= \widehat{\mathbf{H}}(\mathbf{F}_N \otimes \mathbf{I}_{N_t})(\mathbf{B}\mathbf{z} + \mathbf{d}) + \widehat{\mathbf{n}} \\
&= \widehat{\mathbf{H}}(\mathbf{F}_N \otimes \mathbf{I}_{N_t})[\mathbf{B}(\mathbf{F}_N^H \otimes \mathbf{I}_{N_t})\widehat{\mathbf{P}}\widehat{\mathbf{u}} + \mathbf{d}] + \widehat{\mathbf{n}} \\
&= \widehat{\mathbf{H}}\mathbf{B}\widehat{\mathbf{P}}\widehat{\mathbf{u}} + \widehat{\mathbf{H}}(\mathbf{F}_N \otimes \mathbf{I}_{N_t})\mathbf{d} + \widehat{\mathbf{n}}. \quad (20)
\end{aligned}$$

Let $\widehat{y}_{k,l} = [\widehat{\mathbf{y}}_l]_k$ be the received signal for the k-th UE at the l-th subcarrier, then we have

$$\widehat{y}_{k,l} = [\widehat{\mathbf{H}}_l\text{diag}(\mathbf{b})\mathbf{P}_l]_{k,k}u_{k,l} + \sum_{m \neq k}[\widehat{\mathbf{H}}_l\text{diag}(\mathbf{b})\mathbf{P}_l]_{k,m}u_{m,l}$$
$$+ [\widehat{\mathbf{H}}(\mathbf{F}_N \otimes \mathbf{I}_{N_t})\mathbf{d}]_{k+lK,k+lK} + \widehat{n}_{k,l}, \quad (21)$$

where $\mathbf{b} = [b_1,\ldots,b_{N_t}]$, data symbol $u_{k,l} = [\mathbf{u}_l]_k$, and noise $\widehat{n}_{k,l} = [\widehat{\mathbf{n}}_l]_k$.

B. Achievable Sum-rate with Gaussian Signaling

With Gaussian signaling inputs[2], i.e., $\mathbf{u}_l \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_K)$ for $l \in \mathcal{S}_u$, the SINDR for UE k at the l-th subcarrier can be expressed as [18]

$$SINDR_{k,l}(\widehat{\mathbf{H}}) = \frac{|\widehat{\mathbf{H}}_l\text{diag}(\mathbf{b})\mathbf{P}_l|^2_{k,k}}{\sum_{m \neq k}|\widehat{\mathbf{H}}_l\text{diag}(\mathbf{b})\mathbf{P}_l|^2_{k,m} + D_{k,l}(\widehat{\mathbf{H}}) + N_0}, \quad (22)$$

where

$$D_{k,l}(\widehat{\mathbf{H}}) = [\widehat{\mathbf{H}}(\mathbf{F}_N \otimes \mathbf{I}_{N_t})\mathbf{C}_\mathbf{d}(\mathbf{F}_N^H \otimes \mathbf{I}_{N_t})\widehat{\mathbf{H}}^H]_{k+lK,k+lK}$$

---

[2]According to information theory, one of the necessary conditions to achieve Shannon capacity/throughput is that the channel inputs must be continuous Gaussian random variables. However, this is not true in modern wireless communication systems, in which the channel inputs generally take their values from a finite alphabet (constellation) with equal probability.

and $\mathbf{C_d} = \mathbb{E}\{\mathbf{dd}^H\}$ denotes the covariance of the distortion $\mathbf{d}$.

It is challenging to derive the exact achievable rate due to the non-Gaussian distributed residual multi-user interference and quantization distortion. Nevertheless, a lower bound on the achievable rate can be derived by using the so-called "auxiliary-channel lower bound" [7], [18] approach. Specifically, the lower bound of the average sum-rate can be written as

$$R_{\text{sum}} = \frac{1}{S}\mathbb{E}_{\widehat{\mathbf{H}}}\left\{\sum_{k=1}^{K}\sum_{l\in\mathcal{S}_u}\log_2(1 + SINDR_{k,l}(\widehat{\mathbf{H}}))\right\}. \quad (23)$$

From (17), the time-domain transmit signal $\mathbf{x} \triangleq \text{vec}(\mathbf{X}) = \mathbf{Bz} + \mathbf{d}$. Then, we have

$$\mathbf{C_d} = \mathbf{C_x} - \mathbf{BC_zB}, \quad (24)$$

where $\mathbf{C_x} = \mathbb{E}\{\mathbf{xx}^H\} \in \mathbb{C}^{N_t N \times N_t N}$ and $\mathbf{C_z}$ denotes the covariance matrix of the input Gaussian signal $\mathbf{z}$, which is given by (cf. (18))

$$\mathbf{C_z} = (\mathbf{F}_N^H \otimes \mathbf{I}_{N_t})\widehat{\mathbf{P}}\widehat{\mathbf{P}}^H(\mathbf{F}_N \otimes \mathbf{I}_{N_t}). \quad (25)$$

Due to the overload distortion, the rounding approximation method, used for example in [18], is not valid. In addition, the computation of $\mathbf{C_x}$ is more challenging compared to I/Q signal quantization, due to the asymmetry of polar quantization. For tractable analysis, we employ a diagonal approximation for deriving $\mathbf{C_d}$, which is widely used when analyzing the performance of low-resolution converters in massive MIMO systems (see, e.g., [42]). Doing this, we obtain

$$\mathbf{C_d}^{\text{diag}} = \mathbf{C_x}^{\text{diag}} - \mathbf{BC_z}^{\text{diag}}\mathbf{B}. \quad (26)$$

C. Computation of $\mathbf{C_d}^{\text{diag}}$

Let now $z_m = [\mathbf{z}_n]_m$ be the m-th element of the input signal $\mathbf{z}_n$, with magnitude $a_m = [\mathbf{A}_n]_m$ and phase angle $\psi_m = [\mathbf{\Psi}_n]_m$. Then, its probability density function in the polar coordinates can be expressed as

$$p(a_m, \psi_m) = g(a_m)h(\psi_m), \quad (27)$$

where

$$g(a_m) = \frac{a_m}{\sigma_m^2}\exp\left(-\frac{a_m^2}{2\sigma_m^2}\right), \; 0 \le a_m < \infty, \quad (28)$$

$$h(\psi_m) = \frac{1}{2\pi}, \; 0 \le \psi_m < 2\pi, \quad (29)$$

wherein $\sigma_m^2 = \frac{1}{2}[\mathbb{E}\{\mathbf{z}_n\mathbf{z}_n^H\}]_{m,m} = \frac{1}{2}[\mathbf{P}_n\mathbf{P}_n^H]_{m,m}$.

Hence, the Bussgang gain corresponding to the m-th component for the polar quantization of z_m is given by

$$b_m = \frac{\mathbb{E}\{z_m^Q z_m^*\}}{\mathbb{E}\{z_m z_m^*\}}$$

$$= \frac{\mathbb{E}\{a_m^Q \exp(j\psi_m^Q)a_m \exp(-j\psi_m)\}}{\mathbb{E}\{a_m^2\}}$$

$$= \frac{\mathbb{E}\{(a_m^2 + a_m e_{a_m})\exp(je_{\psi_m})\}}{\mathbb{E}\{a_m^2\}}$$

$$= \frac{\left(2\sigma_m^2 + \mathbb{E}\{a_m e_{a_m}\}\right)\left(\exp(j\frac{\pi}{L_P}) - \exp(-j\frac{\pi}{L_P})\right)L_P}{j4\pi\sigma_m^2},$$

$$m = 1, \dots, N_t, \quad (30)$$

where the superscript $(\cdot)^*$ denotes the complex conjugation operation, $a_m^Q = [\mathbf{A}_n^Q]_m$ and $\psi_m^Q = [\mathbf{\Psi}_n^Q]_m$ are the quantized amplitude and phase, respectively; $e_{a_m} = [\mathbf{e}_{\mathbf{A}_n}]_m$ denotes the amplitude quantization error while the phase error $e_{\psi_m} = [\mathbf{e}_{\psi_n}]_m$ is assumed to be uniformly distributed over the interval $[-\frac{\pi}{L_P}, \frac{\pi}{L_P}]$.

Using (11b), we have

$$\mathbb{E}\{(a_m^Q)^2\} = \mathbb{E}\{(a_m + e_{a_m})^2\}$$
$$= \mathbb{E}\{a_m^2 + 2a_m e_{a_m} + e_{a_m}^2\}. \quad (31)$$

Thus, we obtain

$$\mathbb{E}\{a_m e_{a_m}\} = \frac{1}{2}\left(\mathbb{E}\{(a_m^Q)^2\} - 2\sigma_m^2 - \mathbb{E}\{e_{a_m}^2\}\right). \quad (32)$$

With Rayleigh distributed amplitude signal, $\mathbb{E}\{(a_m^Q)^2\}$ can be expressed as

$$\mathbb{E}\{(a_m^Q)^2\} = \sum_{i=0}^{L_A-1}\ell_i^2\mathbb{P}[\tau_i \le a_m \le \tau_{i+1}]$$

$$= \sum_{i=0}^{L_A-1}\ell_i^2\int_{\tau_i}^{\tau_{i+1}}\frac{a_m}{\sigma_m^2}\exp(-\frac{a_m^2}{2\sigma_m^2})da_m$$

$$= \sum_{i=0}^{L_A-1}\ell_i^2\left(\exp(-\frac{\tau_i^2}{2\sigma_m^2}) - \exp(-\frac{\tau_{i+1}^2}{2\sigma_m^2})\right). \quad (33)$$

The overall amplitude distortion $\mathbb{E}\{e_{a_m}^2\}$ includes both granular and overload distortion. It can be written as

$$\mathbb{E}\{e_{a_m}^2\} = \frac{\mathcal{A}_{\text{clip}}^2}{12L_A^2} + \int_{\mathcal{A}_{\text{clip}}}^{\infty}\frac{a_m}{\sigma_m^2}(a_m - a_{max})^2\exp(-\frac{a_m^2}{2\sigma_m^2})da_m$$

$$= \frac{\mathcal{A}_{\text{clip}}^2}{12L_A^2} + \int_{\mathcal{A}_{\text{clip}}}^{\infty}\frac{a_m}{\sigma_m^2}\left(a_m - \mathcal{A}_{\text{clip}}(1 - \frac{1}{2L_A})\right)^2$$

$$\times \exp(-\frac{a_m^2}{2\sigma_m^2})da_m$$

$$= \frac{\mathcal{A}_{\text{clip}}^2}{12L_A^2} + \left(2\sigma_m^2 + \frac{\mathcal{A}_{\text{clip}}^2}{4L_A^2}\right)\exp(-\frac{\mathcal{A}_{\text{clip}}^2}{2\sigma_m^2})$$

$$- \frac{(2L_A - 1)\mathcal{A}_{\text{clip}}\sqrt{2\pi}\sigma_m}{L_A}Q(\frac{\mathcal{A}_{\text{clip}}}{\sigma_m}), \quad (34)$$

where the granular distortion is given by $\frac{\mathcal{A}_{\text{clip}}^2}{12L_A^2}$ under the assumption of uniformly distributed granular distortion, $a_{max} = A_{\text{clip}}(1 - \frac{1}{2L_A})$ denotes the maximum quantization output of the amplitude quantizer, and the function $Q(x) \overset{def}{=} \frac{1}{\sqrt{2\pi}}\int_x^{\infty}\exp(-\frac{v^2}{2})dv$ denotes the Q-function.

Finally, by inserting (34), (33) and (32) into (30), we can find the b_m in the Bussgang gain matrix $\mathbf{B}$. Then, the diagonal approximation $\mathbf{C_d}^{\text{diag}} = \mathbf{C_d} \odot \mathbf{I}_{NN_t}$ can be derived based on (26), wherein the m-th diagonal entry of $\mathbf{C_x}^{\text{diag}}$ is $\mathbb{E}\{(a_m^Q)^2\}$.

IV. POWER EFFICIENCY ANALYSIS AND COMPARISON

A. Methodology

In general, while there are several IC and prototyping implementations for Cartesian-based 5G transmitters [43]–[45] and digital polar transmitter architectures [33]–[36], directly comparing the power efficiency of two transmitter

architectures is difficult [46]. This is because the reported implementations may have different design targets and omit some implementation details due to only some circuit components being integrated. Moreover, the implementations may use different silicon technologies and circuit components. As depicted in Fig. 1, the quantized polar transmitter at each antenna consists of block components including I/Q to polar converter (e.g., CORDIC), local oscillator (LO), digital phase modulator, and RF-DAC. On the other hand, the RF chain at each antenna with conventional Cartesian architecture mainly consists of the DAC, low-pass filter (LPF), I/Q mixer, LO, a $90°$ hybrid with buffers, and a linear-class PA (class-A/B/AB), which is typically an integrated CMOS PA in the large-scale MIMO context. To make these two architectures comparable, we need to establish realistic power consumption models of the required circuit blocks in each transmitter, based on an extensive survey of the state-of-the-art circuit designs and measurements.

For a fair comparison, each transmit architecture has to deliver the same target spectral efficiency (SE). Thus, the total power consumption P_{All} for each transmitter architecture can be expressed as

$$P_{\text{All}} = P_{\text{total}}/\eta + P_{\text{cir}}, \qquad (35)$$

where P_{total} is the required total transmit power to meet the target SE, η is the average power efficiency of the power amplification stage, and P_{cir} denotes the power consumption of the circuit blocks, which depends on the chosen transmitter architecture. We are limiting our focus to the power consumption of the RF transmitter and RFFE. The baseband power consumption will be omitted due to the same baseband processing being applied in both architectures. The power consumption of any active cooling is also omitted. Moreover, three test environments (indoor hotspot, dense urban, rural) for eMBB are defined in 5G standardisation [47]. In this work, the rural and dense urban environments are considered, since the power consumption of wide-area BSs is most critical from the overall network power efficiency point of view.

B. P_{total} Estimation

In general, the output power of each PA is split into K parts due to spatial multiplexing and even power allocation, while N_t transmit antennas with beamforming provide N_t^2 times increased power gain. Hence, the maximum output signal power after coherent summation is $\frac{P_{\text{total}}N_t}{K}$. Note that increasing the transmit power P_{total} and array size N_t both improve signal power gain in (22), which effectively provides higher effective isotropic radiated power (EIRP) and helps achieve the target SINR. However, we can only rely on Monte Carlo simulations to estimate the actual power terms in (22) instead of quantitatively analyzing them with respect to the system parameters. Then, the corresponding estimated SINDR with the given channel realization $\hat{\mathbf{H}}$ can be denoted as

$$SINDR_{k,l}^{\text{est}} = \frac{|G_{k,l}|^2}{I_{k,l}^2 + D_{k,l}^2 + \sigma_{\text{rx}}^2}, \qquad (36)$$

where the signal power gain $G_{k,l}$ is given by $G_{k,l} = \arg\min_{g_{k,l}} \mathbb{E}\{\|g_{k,l}\hat{y}_{k,l} - u_{k,l}\|^2\}$, and the interference plus distortion power $I_{k,l}^2 + D_{k,l}^2 = \mathbb{E}\{\|G_{k,l}\hat{y}_{k,l} - u_{k,l}\|^2\}$. All the receiver thermal noise power σ_{rx}^2 is treated as constant in each use case. Moreover, all signal, interference plus distortion, and noise powers are relative powers, referred to the maximum BS transmit power (see more in Table 5 of [48]). As a result, we obtain the expected signal power $\bar{G}^2$, and multi-user interference plus distortion power $\bar{I}^2 + \bar{D}^2$ over all active subcarriers via Monte Carlo simulation. Then, to reach the targeted SINRs in each use case, the average SINDR per UE yields

$$\frac{P_{\text{adjust}}\bar{G}^2}{P_{\text{adjust}}(\bar{I}^2 + \bar{D}^2) + \bar{\sigma}_{\text{rx}}^2} \geq SINR_{\text{req}}, \qquad (37)$$

where P_{adjust} is the power adjustment factor, $SINR_{\text{req}}$ is the required SINR in each use case, and the thermal noise power $\bar{\sigma}_{\text{rx}}^2$ at the UE side in each use case can be reproduced by link budget estimation from Table II in [46] with corresponding parameter values [39], [48]. When the equality holds, we have

$$P_{\text{adjust}}^{\star} = \frac{\bar{\sigma}_{\text{rx}}^2 \times SINR_{\text{req}}}{\bar{G}^2 - (\bar{I}^2 + \bar{D}^2)SINR_{\text{req}}}. \qquad (38)$$

Lastly, the required transmit power for different system parameters is obtained, leading to e.g., P_{total} (dBm) $= 49\,\text{dBm} + 10\log_{10}(P_{\text{adjust}}^{\star})$ in the rural eMBB use case.

C. Hardware Power Consumption Modeling

In this subsection, we analyze the power consumption of the circuit components in both the polar and Cartesian transmitter architectures. Additionally, we select some realistic power consumption values or parameters for each circuit block based on the state-of-the-art reported circuit designs and measurements.

Firstly, we consider a direct-conversion Cartesian transmitter and denote the power consumption of the DAC, filters, I/Q mixer and LO as P_{DAC}, P_{F}, P_{M} and P_{LO}, respectively.

1) *Power Consumption of DAC*: The DAC power consumption is mainly determined by the sampling rate and quantization resolution. An empirical model for computing the total power consumption in each DAC is given by [46]

$$P_{\text{DAC}} = \text{FOM}_{\text{DAC}} \times (2^{l_{DAC}} \times \text{BW}_{\text{OSR}}) + P_{\text{buffer}}, \qquad (39)$$

where FOM_{DAC} and BW_{OSR} denote the figure-of-merit (FOM) and oversampled data rate, respectively, and P_{buffer} is a constant hardware overhead for signal amplification, e.g., we assume $P_{\text{buffer}} = 10$ mW for –14 dBm output signal power. We specify the FOM of the DAC to be $\text{FOM}_{\text{DAC}} = 0.08$ PJ/conversion as in [46] and $\text{BW}_{\text{OSR}} = 614$ MS/s for 100 MHz signal bandwidth.

2) *LO, Mixer and Filters*: LO signal generation and distribution design pose a challenge to massive MIMO transceiver design, due to large number of RF chains and a high frequency range. Currently, there are three approaches for LO generation and distribution that have been proposed in the literature for massive MIMO [49]. Firstly, LO signal can be generated by an independent

TABLE II
SUMMARY OF CIRCUIT BLOCKS IN THE CARTESIAN AND THE QUANTIZED POLAR TRANSMITTER ARCHITECTURES

Cartesian transmitter			Quantized polar transmitter		
Component	Notation	Value	Component	Notation	Value
PA efficiency [46], [53]	η	15%	PA efficiency [36]	η	25%
DAC FOM [46]	FOM_{DAC}	0.08 PJ/Conv	CORDIC [55]	P_{CORDIC}	20 mW
DAC buffer [46]	P_{buffer}	10 mW	Phase modulator [35]	P_{PM}	27 mW
LO [50]	P_{LO}	25 mW	LO	P_{LO}	25 mW
Mixers [51]	P_{M}	5 mW	Note: 25 mW for LO includes LO buffer, overhead.		
Filters [52]	P_{F}	20 mW			

oscillator locally for each transceiver. Secondly, a low-frequency reference is first distributed to each transceiver and a PLL is then used to generate the desired LO frequency. Thirdly, in the centralized LO distribution architecture, high-frequency LO signal is directly distributed from a common source to all RF chains. In this work, we assume that the independent LO generation approach is utilized in both transmitter architectures. The state-of-the-art oscillator designs facilitate phase noises lower than -110 dBc/Hz at 1 MHz offset such that the system performance is not affected, and they consume less than 12.5 mW of direct current (DC) power [50]. Thus, we consider that the LO components consume 25 mW of power, when the $90°$ phase shifters and LO buffer for driving the passive up-conversion mixers are considered. We assume the power consumption of the mixers to be $P_M = 5$ mW [51]. Regarding the required power to run filter components, we selected the value of $P_F = 20$ mW from [52].

3) *CMOS PA*: PA consumes large amount of power in current BSs operating in sub-6 GHz. It is tricky to balance the output power, linearity, and efficiency in practical PA designs. In [53], a comprehensive PA performance survey is reported. We note that PA efficiency is related to the operating frequency, output power and semiconductor technologies. In this work, we assume that CMOS-based PAs are utilized. Assuming a PAPR of 8 dB for the clipped OFDM signals, the 49 dBm total transmit power leads to 18.9 dBm output power per PA with an extremely large-scale array of 1024 antennas at the BS. Thus, the required saturated power of the PA is 26.9 dBm, which poses a huge challenge for CMOS PA design. Additionally, contrary to the optimistic values of $\eta \approx 40\%$ found in the massive MIMO energy efficiency analyses [54], we opt for a more conservative PA efficiency of $\eta = 15\%$, based on the required saturated output power and the state-of-the-art CMOS PA implementations.

The power consumption details of the quantized polar transmitter are, in turn, listed below.

1) *CORDIC*: The main principle of CORDIC is carrying out calculations based on shift-registers and adders instead of multiplications, which saves hardware resources. In [55], parallel rotation and vectoring CORDIC are designed to perform I/Q signal to polar conversion with 5-bit amplitude and 7-bit phase signals, with an energy metric (Power/Frequency/Bits) of 0.41 PJ/bit. Thus, we estimate the maximum power consumption of CORDIC (with integrated pulse shaping filter) to be $P_{\text{CORDIC}} = 20$ mW.

2) *Phase Modulator*: In the quantized polar transmitter, a digital-intensive RF phase modulator can be implemented using digital PLL or delay-based architectures such as DTC or DIPM. In this work, we assume that DTC-based phase modulators are adopted and the power model is $P_{\text{PM}} = 27$ mW with 100 MHz bandwidth [35].

3) *RF-DAC*: Combining the RF-DAC and PA functionality on a single die constitutes fully integrated RF-DPA implementations, capable of delivering high output powers. In addition, by utilizing switched-capacitor circuits, power efficiency, linearity and scalability can be further improved. The state-of-the-art average PA efficiency of DPA in polar transmitter implementations is around 25% [36].

Finally, we assume that the power consumption of the LO is the same for both the Cartesian and the quantized polar transmitter architectures.

Table II shows a summary of the specifications of the circuits blocks for Cartesian and polar transmitter architectures. Note that the estimated power consumption from the proposed power model matches well with the reported power consumptions in the surveyed designs, or is at least comparable to key circuit blocks in [36], [43]–[45].

V. NUMERICAL RESULTS

In this section, we first present the quantized polar transmitter -based massive MIMO system performance in terms of OOB emission levels, average sum-rate, and uncoded average SER. Then, we present the total power consumption results and comparison between the polar transmitter and the conventional Cartesian transmitter architecture in rural eMBB and dense urban eMBB use cases.

We focus on a massive MIMO-OFDM multi-user downlink transmission in which the BS is equipped with $N_t = 128$ half-wavelength spaced antennas. There are $K = 16$ single antenna UEs which are mutually far apart (≥ 10 m mutual distance) and uniformly distributed within the respective cell with predefined cell radius in [48]. The RMa-NLOS and UMa-LOS propagation channels are assumed, where the large scale fading coefficients are calculated based on [39] and the small scale fading coefficients are generated through cluster-based channel models (cf. (5)). We assume that the channel between the BS and each UE consists of N_{cl} clusters with uniformly distributed AoAs/AoDs in $[0, 2\pi)$ and a LOS path with the Ricean K-factor of 9 dB, if it exists. We also assume that all clusters have equal power. Moreover, each cluster consists of N_{ray} rays with Laplacian distributed AoAs/AoDs, and the

TABLE III
SIMULATION PARAMETERS FOR CLUSTER-BASED CHANNEL MODELS IN (5)

Chanel model	RMa-NLOS	UMa-LOS
Number of clusters N_{cl}	10	12
Number of rays per cluster N_r	20	20
AoAs/AoDs per cluster ϕ_c, θ_c	Uniformly distributed over $[0, 2\pi)$	
AoAs/AoDs per ray φ_r, ϑ_r	Laplacian distribution	
Angle spread (degree)	33.1	47.5
Maximum delay spread (ns)	60	60
Ricean K-factor (dB)	N/A	9

maximum delay spread of 60 ns on each path. The detailed channel parameters are listed in Table III. Similar to [56], channel normalization (*Normalization 1* in [56]) is utilized such that the imbalances of channel attenuations (e.g., path loss variations) between UEs are removed, while variations over antenna elements and frequencies remain. With this normalization, the channel vectors of each UE are normalized such that the average energy over all antenna ports and effective subcarriers is equal to one. As a result, equality between average per-UE received SNRs is obtained in all scenarios.

The assumed 100 MHz M-QAM OFDM baseband signal for 3.5 GHz carrier frequency is in-line with 5G NR specifications [57]. There are $S = 1584$ effective subcarriers with subcarrier spacing of 60 kHz and the normal CP length, and total time-domain samples of $N = 5 \times 2048$ per main OFDM symbol where the oversampling factor $\kappa = 5$ and $N_{FFT} = 2048$. The whole signal consists of 10 OFDM symbols, and time-domain windowing is adopted to improve the spectral containment of the OFDM signals.

A. Power Spectral Density

To demonstrate the OOB emission levels of the quantized polar transmitter, we plot in Fig. 2 the normalized power spectral density (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) for the case of the RMa-NLOS scenario (similar transmit spectrum in the UMa-LOS scenario). We also plot the PSD of a traditional Cartesian transmitter with the low-resolution DACs using ZF precoding and SQUID-OFDM algorithm. In the simulations, the total transmit power is set mutually equal in all cases. We set $\mathcal{A}_{\text{clip}} = 1$ and set the average clipping probability of the amplitude signals to 10^{-2}. Then, under the assumption that the I/Q signal in each antenna is $z_m \sim \mathcal{CN}(0, \frac{P_{\text{total}}}{\kappa N_t})$, $m = 1, \cdots, N_t$, the total transmit power is set to $P_{\text{total}} = \kappa N_t/(-\ln(10^{-2}))$.

We can see that the OOB emission levels of the quantized polar transmitter with 3-bit amplitude and 3-bit phase quantization are slightly higher as compared to the Cartesian one with 3-bit DACs.

Higher phase resolution (e.g., 6-bit) can be used to reduce the emission levels, since smaller phase errors will improve the reconstructed signal integrity. However, the emission levels will be eventually saturated when increasing the number of amplitude and phase bits, due to the overload distortion (see Fig. 2a). We also note that the relative level of OOB emissions
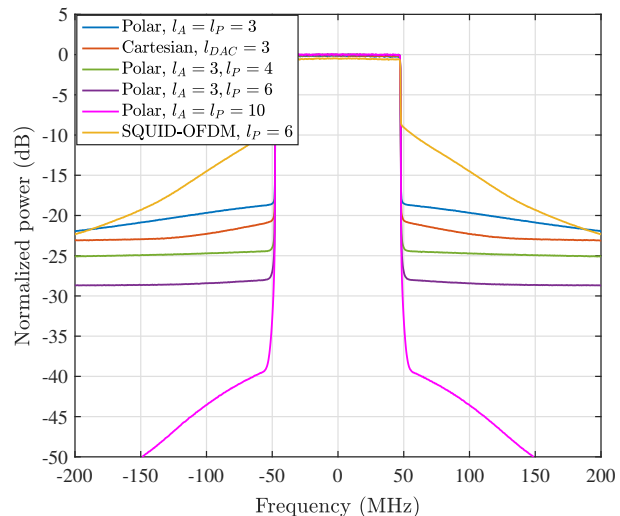
(a) Transmitted spectrum
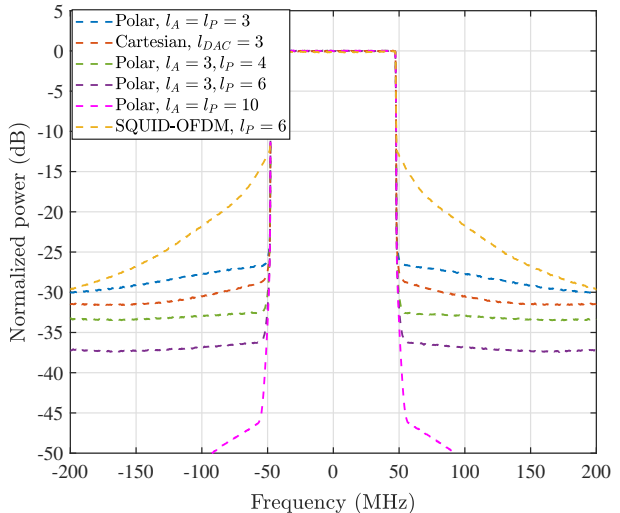
(b) Received spectrum

Fig. 2. Normalized PSD of the transmitted and received signals in the RMa-NLOS scenario; 64-QAM OFDM, $N_t = 128$, and $K = 16$.

is lower at the UEs than at the BS, which is in line with recent findings reported in [18]. Lastly, we can observe that the OOB distortions with CE precoding (SQUID-OFDM) are a significant issue in practical systems as they may cause interference to users in adjacent frequency bands. This is due to CE precoding reducing the in-band power while OOB power increases under the fixed average total transmit power at the BS. Further analyses of the effects of the number of served UEs, BS antennas and different channel models on the

(a) RMa-NLOS ($l_P = 6$)

(b) RMa-NLOS (SNR = 20 dB)

(c) UMa-LOS ($l_P = 6$)
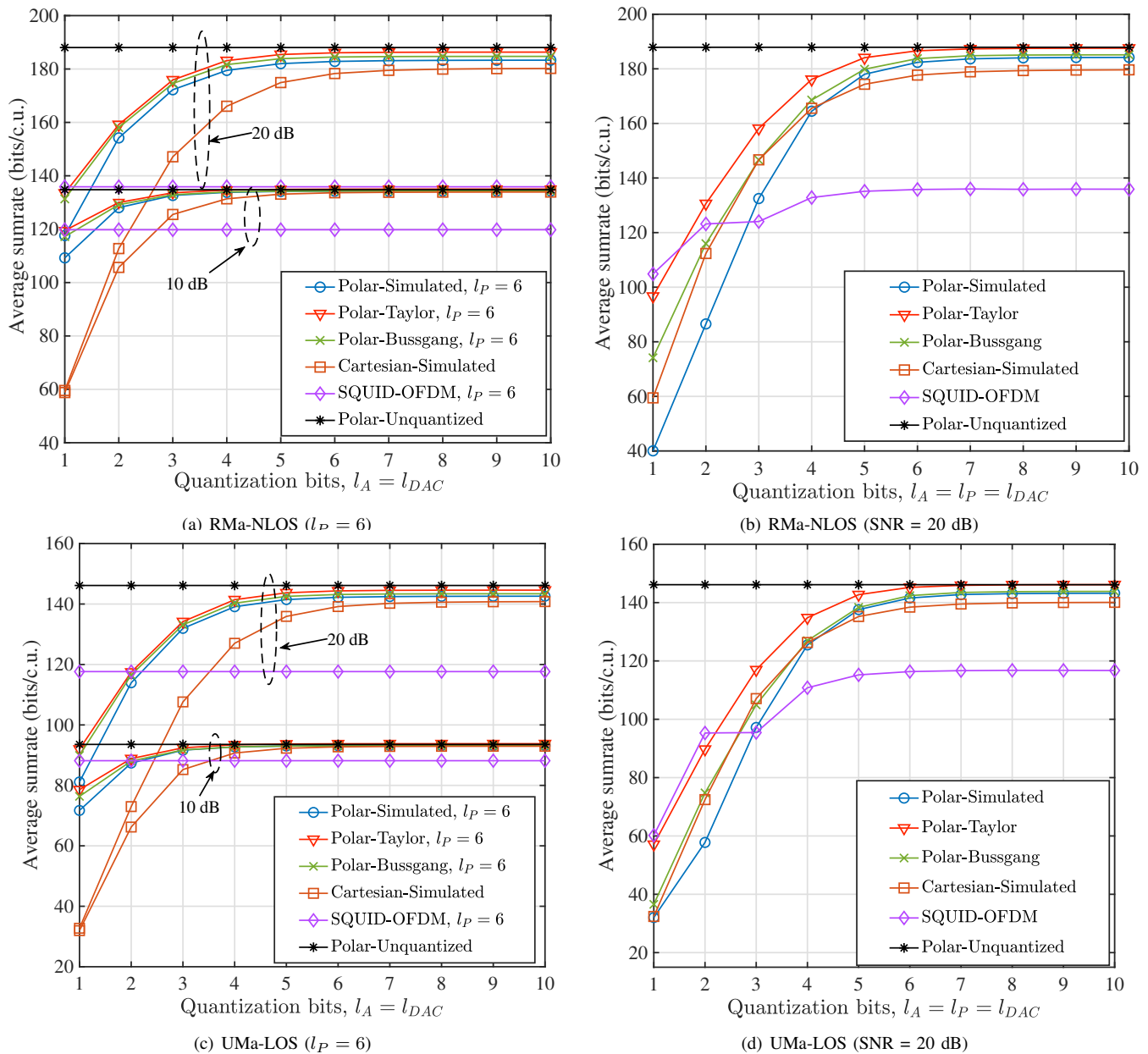
(d) UMa-LOS (SNR = 20 dB)

Fig. 3. Average sum-rate with ZF and Gaussian signaling versus different amplitude and phase resolutions; 1% clipping probability, $N_t = 128$ and $K = 16$; perfect CSI case.

OOB emissions, as well as the spatial radiation pattern of the distortions, are out of the scope of this paper, but form an important future work item.

B. Average Sum-Rate

Next, we will present results of the average sum-rate for the massive MIMO-OFDM systems. We will compare the performance of the proposed quantized polar transmitter paired with ZF precoding to Cartesian transmitter with low-resolution DACs, as proposed in [18], and to CE precoding system employing the SQUID-OFDM algorithm [20]. For SQUID-OFDM, we use l_P to denote the phase quantization bits while the constant amplitude signal satisfies the total transmit power. Due to the overload distortion being omitted in [18], we use the simulated results for the Cartesian transmitter rather than

the analytical ones. In addition, analytical results with the second order Taylor approximation method from [38] are also presented for comparison. The results in Fig. 3 are the averages over 100 channel realizations.

1) Perfect CSI: We can see from Fig. 3 that the proposed diagonal approximation with Bussgang decomposition attains very good accuracy for the case of $l_A = l_P \geq 4$ bits (see Fig. 3b and Fig. 3d) and $l_A \geq 2$, $l_P = 6$ bits (see Fig. 3a and Fig. 3c) for the average sum-rate performance of the polar transmitter. We note that the Taylor approximation is accurate for the case of higher phase resolution $l_P = 6$, but significantly overestimates the achievable rate for the small phase resolution case. This is due, in large part, to the quantization noise being temporally non-white, similarly to the case of Cartesian transmitter with low-resolution DACs [18]. We can see that the

average sum-rates of the unquantized polar transmitter can be approached by using very few amplitude bits and a moderate number of phase bits. For instance, the performance loss can be neglected in both scenarios when $l_A = 3\text{-}4$ amplitude bits and $l_P = 6$ phase bits are utilized at the SNR of 10 dB. On the contrary, there is a substantial average sum-rate loss for the SQUID-OFDM scheme. Thus, like other CE precoding schemes [14], [19], extra power is required to achieve the sum-capacity of the Gaussian broadcast channel under average total power constraint since a notable share of the transmitted power is lost into the channel null-space. Furthermore, the SQUID-OFDM algorithm requires further parameter optimization to improve its robustness, see Fig. 3b and Fig. 3d.

Overall, we note that better average sum-rate performance can be attained with the digital polar transmitter with $l_P = 6$ and low amplitude resolution, i.e., 2-4 bits, compared to the Cartesian one with higher-resolution DACs. Moreover, the polar transmitter has a better average sum-rate performance compared to the Cartesian transmitter also with equal amplitude and phase resolutions. This is due to the properties of polar quantization with complex Gaussian input, which have been analyzed and demonstrated in [32].

2) Imperfect CSI: Next, we investigate the impact of imperfect CSI on the average sum-rate performance. Specifically, we consider a case where the BS acquires a noisy time domain compound multi-user CSI of $\mathbf{H}^{\mathrm{est}}[d]$, which can be modelled or expressed as [18]

$$\mathbf{H}^{\mathrm{est}}[d] = \sqrt{1-\epsilon}\,\mathbf{H}[d] + \sqrt{\epsilon}\,\mathbf{E}, \tag{40}$$

where $\epsilon \in [0,1]$ is the uncertainty factor, and $\mathbf{E}$ is the channel estimation error, wherein the elements are random complex values that are i.i.d. Gaussian distributed with zero mean, having a variance equal to the variance of the multi-path components of $\mathbf{H}[d]$. The case $\epsilon = 0$ corresponds to perfect CSI, while values lower than 1 indicate that only partial CSI is available. In a TDD system, the value of the channel uncertainty ϵ is dependent on the length of pilot sequences that have been transmitted during the uplink training phase. Then, the corresponding multi-user frequency-domain response at l-th subcarrier is given by $\widetilde{\mathbf{H}}_l^{\mathrm{est}} = \sum_{d=0}^{D-1} \mathbf{H}^{\mathrm{est}}[d] \exp(-j\frac{2\pi}{N}ld)$.

In Fig. 4, we show the average sum-rate for massive MIMO-OFDM system with different transmitter architectures in the two considered channel cases, and with uncertainty factor ϵ having values of 0.1, 0.2, and 0.5. We note that the imperfect CSI does indeed deteriorate the average sum-rate similarly for each considered transmitter architecture, no matter the channel type. This is due to the imperfect CSI only affecting the ZF precoding, which in the quantized polar and Cartesian transmitters is the same, and the quantization occurs only after the precoding.

C. Symbol Error-Rate Performance

In Fig. 5 and Fig. 6, we present the uncoded average SER performances of the different transmitter architectures with ZF precoding and different modulation schemes for the case of $l_A = 2, 3, 4$ and $l_P = 6$ in the RMa-NLOS and UMa-LOS scenarios, respectively. We can see that only a few amplitude
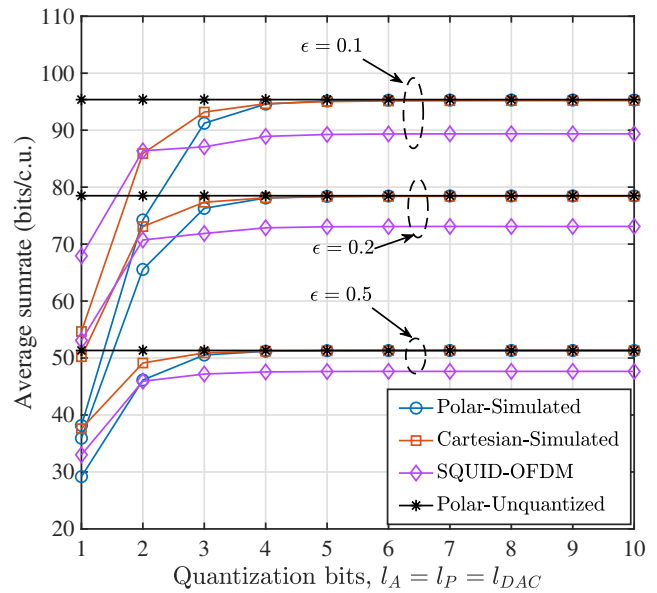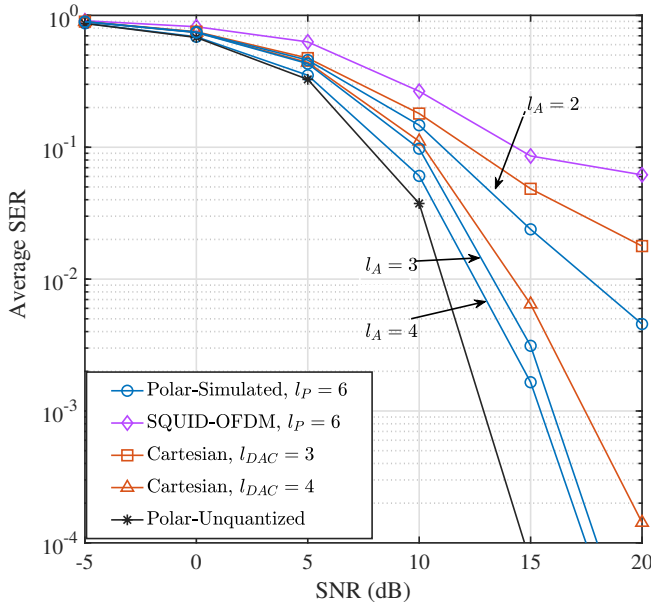

(a) RMa-NLOS (SNR = 20 dB)
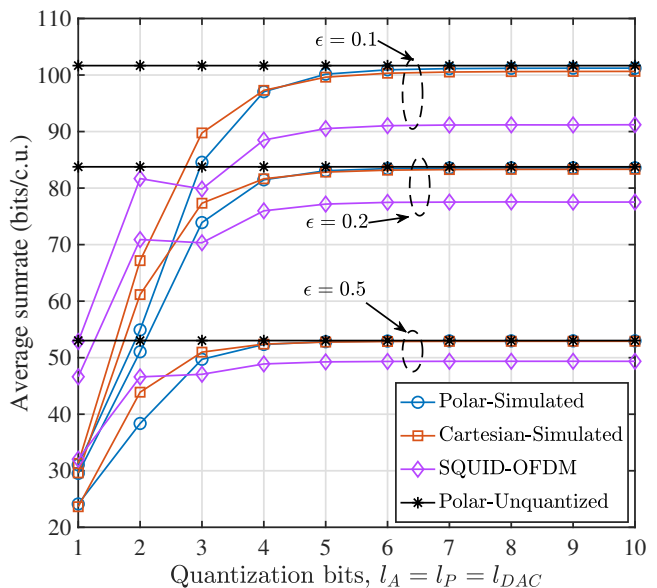
(b) UMa-LOS (SNR = 20 dB)

Fig. 4. Average sum-rate with Gaussian signaling and various values for CSI uncertainty factor ϵ versus different quantization resolutions; 1% clipping probability, $N_t = 128$ and $K = 16$.

bits are sufficient to approach the optimal SER performance of the unquantized one in both scenarios. For instance, only 3-4 amplitude bits are required in the polar transmitter to achieve an uncoded SER of 10^{-4} with a negligible performance loss for 64-QAM in RMa-NLOS and 16-QAM in UMa-LOS scenarios as compared to the infinite-resolution case. As expected, less amplitude bits are needed for polar transmitter with 6-bit phase resolution to outperform the Cartesian with more DAC bits, e.g., $l_A = 2$ while $l_{DAC} = 3$. Moreover, the SER performance of the quantized polar transmitter is significantly improved by increasing the amplitude resolution from $l_A = 2$ to $l_A = 3$. This is because the polar quantization error of the low amplitude input signals is dominated by
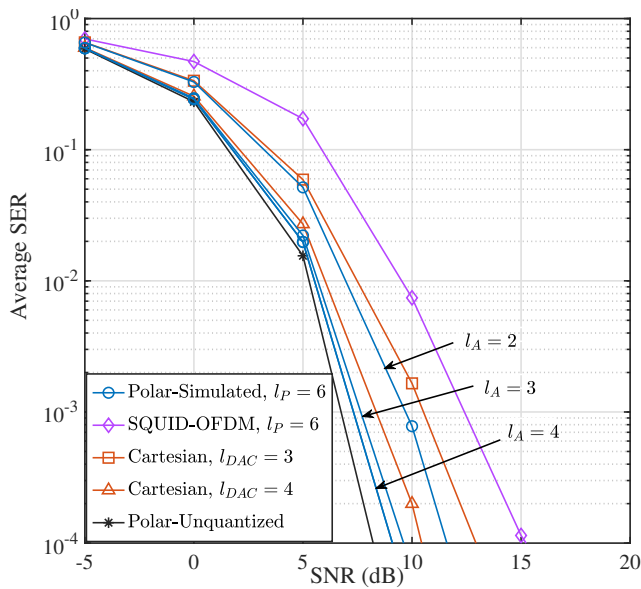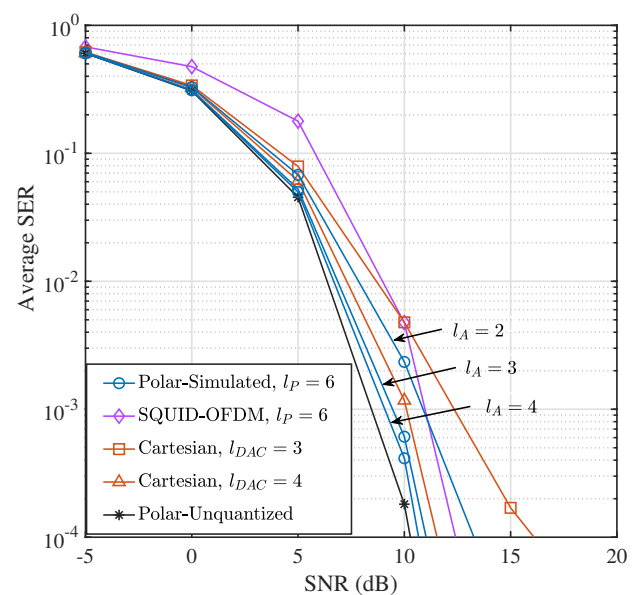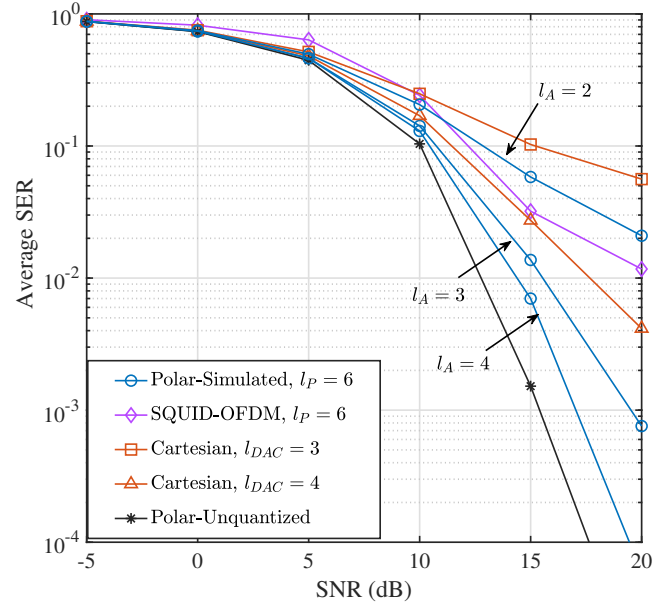
(a) Uncoded SER with 64-QAM

(a) Uncoded SER with 16-QAM

(b) Uncoded SER with 256-QAM

(b) Uncoded SER with 64-QAM

Fig. 5. Average uncoded SER performance for different transmit architectures operating in the RMa-NLOS scenario with ZF and different modulation schemes versus different amplitude and phase resolutions; $l_P = 6$, 1% clipping probability, $N_t = 128$ and $K = 16$; perfect CSI.

Fig. 6. Average uncoded SER performance for different transmit architectures operating in the UMa-LOS scenario with ZF and different modulation schemes versus different amplitude and phase resolutions; $l_P = 6$, 1% clipping probability, $N_t = 128$ and $K = 16$; perfect CSI.

the amplitude resolution and that of large amplitude input signals is dominated by the phase resolution [32]. On the other hand, the performance of the SQUID-based scheme shows clear dependence on the channel characteristics and the modulation order, and can only be seen to have comparable performance in the UMa-LOS case with 16-QAM at high SNR. We also note that the superiority of the quantized polar transmitter is more significant for the high-order modulation schemes (e.g., 256-QAM and 64-QAM in RMa-NLOS and UMa-LOS, respectively), compared to the traditional Cartesian architecture and the SQUID-based CE precoding design with the same phase resolution. This is an important finding of high practical relevance, as the modulation orders supported in 5G

NR downlink at sub-6GHz have been recently extended to cover even 1024-QAM.

### D. Power Consumption Comparison

In this subsection, the required total transmit power P_{total} to reach the targeted SE and the total power consumption P_{All} for the two transmitter architectures operating in the rural and dense urban eMBB use cases are presented. We assume the RMa-NLOS and UMa-LOS propagation environments for the rural eMBB and the dense urban eMBB use cases, respectively [48]. From Fig. 7, we can see that the increase of the array size is an effective way to reduce the transmit power requirement
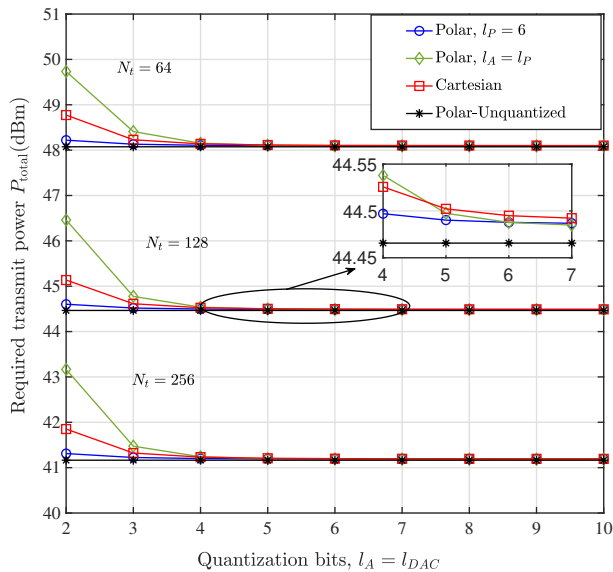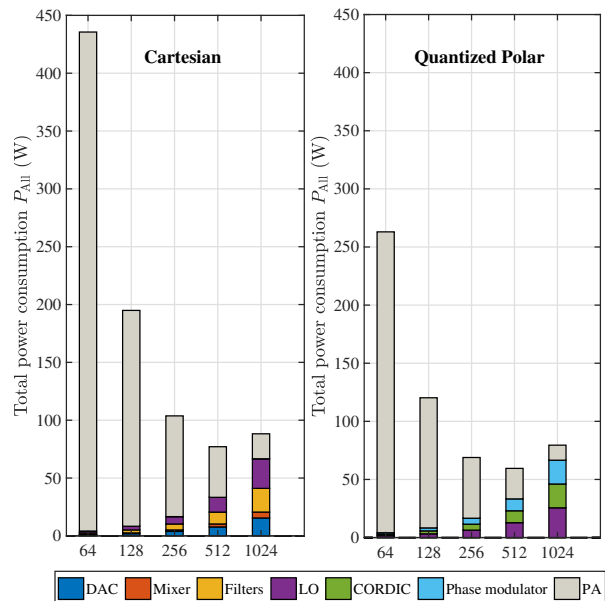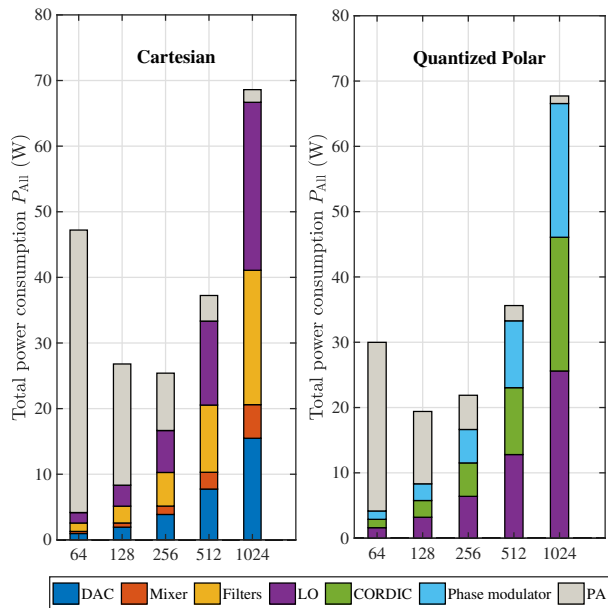
(a) Rural eMBB (Target SE 3.3 bps/Hz)
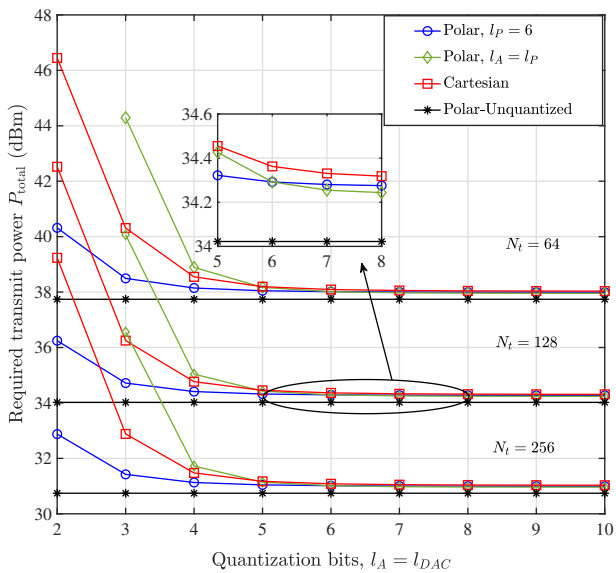
(b) Dense urban eMBB (Target SE 7.8 bps/Hz)

Fig. 7. The required total transmit power P_{total} with different number of transmit antennas versus the amplitude and phase resolutions to reach the target SE in the rural eMBB and dense urban eMBB use cases; ZF with Gaussian signaling and $K = 16$; perfect CSI.

(a) Rural eMBB

(b) Dense urban eMBB

Fig. 8. Total power consumption P_{All} for the traditional Cartesian and the quantized polar transmitter architectures in the Rural and dense urban eMBB use cases; ZF with Gaussian signaling, $N_t \in \{64, 128, 256, 512, 1024\}$ and $K = 16$.

in both use cases since it helps to improve both signal gain and interference control by forming narrower beams. We note that the required total transmit power of the quantized polar transmitter with $l_A = 3$ and $l_P = 6$ is almost same as the Cartesian transmitter with $l_{DAC} \geq 4$ in rural eMBB use case. However, the polar transmitter with smaller amplitude and phase bits requires more power to meet the target SE for the case of $l_A = l_P = l_{DAC}$. In addition, it is not feasible to estimate the P_{adjust} for the polar transmitter with 2-bit amplitude and phase resolutions in dense urban use case since the term $\bar{G} - (\bar{I}^2 + \bar{D}^2)SINR_{\text{req}} < 0$ in (38). This is because the SINDR is limited by the quantization distortion due to coarse quantization. For a fair comparison, we consider the total required transmit power with $l_A = l_P = l_{DAC} = 6$

for the total power consumption comparison between the two transmitter architectures.

The total power consumption P_{All} in rural eMBB use case is presented in Fig. 8a, while also pushing the antenna count up to 1024. We can see that the quantized polar transmitter architecture is highly power efficient and it saves up to 39.6% in power consumption compared to the conventional Cartesian architecture in the 64-antenna setting. The difference is reduced when increasing the array size, because the power consumption per PA, which is the main factor in the power efficiency gain of the polar architecture, is reduced. Thus, since the individual circuit components are assumed to have a

constant power consumption in our analysis, the corresponding total circuit power consumption, P_{cir}, becomes more dominant and eventually the bottleneck in terms of the total power consumption. Specifically, with 1024 antennas, the power consumption of the PAs is already becoming minor compared to the total circuit power consumption. Therefore, to make large arrays feasible and practical, the circuit power consumption must be scaled down as the array size is increased.

The power-saving superiority of the quantized polar transmitter is also demonstrated in Fig. 8b for dense urban use case. In this use case, the polar transmitter architecture is still highly power efficient with up to 36.4% power saving compared to Cartesian one. Similarly to the rural eMBB case, with the increase of the number of BS antennas, the output power per PA is becoming smaller and the power consumption of the other circuit blocks gradually starts to dominate the total power consumption. Interestingly, the optimum array sizes to minimize the total power consumption are different in this case – 256 for Cartesian and 128 for quantized polar transmitter.

Overall, we can conclude that a power efficient system requires large array sizes in both use cases, however, there is a clear limit beyond which increasing the array size is not anymore directly useful. Similar to the analysis in [46], the exact turning points of total power consumption for different transmitter architectures depend on the use case, the assumed circuit power consumption model, and the targeted SE.

VI. CONCLUSION

The quantized polar transmitter for a multi-user downlink massive MIMO-OFDM system was presented and analyzed. We derived a lower bound on the achievable sum-rate based on Bussgang decomposition. To express and evaluate this lower bound, we presented a diagonal approximation for the covariance of the polar quantization distortion. In addition, we presented two realistic circuit power consumption models for the quantized polar and Cartesian transmitters, and analyzed the total transmit power by considering two 5G eMBB use cases. The quantized polar transmitter in a massive MIMO system was shown to outperform the traditional Cartesian-based system in terms of sum-rate, SER, and OOB emissions. Moreover, the power consumption comparison demonstrated the superior power efficiency of the quantized polar transmitter architecture compared to the Cartesian transmitter in two representative 5G eMBB use cases. Altogether, these results show that the requirements for the array size, transmit power, and hardware specifications of the RF chains can be relaxed compared to traditional Cartesian transmitter architecture. Additionally, the results show that the optimum array size varies depending on the use case, and that the optimum array size for a given use case can be different for the different transmitter architectures. The tools and analysis methods provided in the article allow for assessing such optimum array size for any given deployment scenario.

REFERENCES

[1] 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.

[2] C. Kong, C. Zhong, A. K. Papazafeiropoulos, M. Matthaiou, and Z. Zhang, "Sum-rate and power scaling of massive MIMO systems with channel aging," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4879–4893, Dec. 2015.

[3] A.-S. Bana, E. De Carvalho, B. Soret, T. Abrao, J. C. Marinello, E. G. Larsson, and P. Popovski, "Massive MIMO for internet of things (IoT) connectivity," *Phys. Commun.*, vol. 37, pp. 1–17, 2019.

[4] C. Fager, T. Eriksson, F. Barradas, K. Hausmair, T. Cunha, and J. C. Pedro, "Linearity and efficiency in 5G transmitters: New techniques for analyzing efficiency, linearity, and linearization in a 5G active antenna transmitter context," *IEEE Microw. Mag.*, vol. 20, pp. 35–49, Apr. 2019.

[5] C. Mollén, E. G. Larsson, and T. Eriksson, "Waveforms for the massive MIMO downlink: Amplifier efficiency, distortion, and performance," *IEEE Trans. Commun.*, vol. 64, no. 12, pp. 5050–5063, Dec. 2016.

[6] C. Fang, B. Makki, J. Li, and T. Svensson, "Hybrid precoding in cooperative millimeter wave networks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 8, pp. 5373–5388, Aug. 2021.

[7] 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.

[8] 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.

[9] J. Xu, W. Xu, and F. Gong, "On performance of quantized transceiver in multiuser massive MIMO downlinks," *IEEE Wireless Commun. Lett.*, vol. 6, no. 5, pp. 562–565, Oct. 2017.

[10] C. Kong, A. Mezghani, C. Zhong, A. L. Swindlehurst, and Z. Zhang, "Multipair massive MIMO relaying systems with one-bit ADC and DAC," *IEEE Trans. Signal Process.*, vol. 66, no. 11, pp. 2984–2997, Jun. 2018.

[11] S. C. Thompson, A. U. Ahmed, J. G. Proakis, J. R. Zeidler, and M. J. Geile, "Constant envelope OFDM," *IEEE Trans. Commun.*, vol. 56, no. 8, pp. 1300–1312, Aug. 2008.

[12] 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," *IEEE Trans. Signal Process.*, vol. 67, no. 20, pp. 5309–5324, Oct. 2019.

[13] A. Li, C. Masouros, A. L. Swindlehurst, and W. Yu, "1-bit massive MIMO transmission: Embracing interference with symbol-level precoding," *IEEE Commun. Mag.*, vol. 59, no. 5, pp. 121–127, May 2021.

[14] S. K. Mohammed and E. G. Larsson, "Per-antenna constant envelope precoding for large multi-user MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 3, pp. 1059–1071, Mar. 2013.

[15] H. Jedda, A. Mezghani, A. L. Swindlehurst, and J. A. Nossek, "Quantized constant envelope precoding with PSK and QAM signaling," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8022–8034, Dec. 2018.

[16] A. Li and C. Masouros, "Interference exploitation precoding made practical: Optimal closed-form solutions for PSK modulations," *IEEE Trans. Wireless Commun.*, vol. 17, no. 11, pp. 7661–7676, Nov. 2018.

[17] S. Jacobsson, G. Durisi, M. Coldrey, and C. Studer, "Massive MU-MIMO-OFDM downlink with one-bit DACs and linear precoding," in *Proc. IEEE GLOBECOM*, Singapore, Dec. 4-8, 2017, pp. 1–6.

[18] ——, "Linear precoding with low-resolution DACs for massive MU-MIMO-OFDM downlink," *IEEE Trans. Wireless Commun.*, vol. 18, pp. 1595–1609, Mar. 2019.

[19] S. K. Mohammed and E. G. Larsson, "Constant-envelope multi-user precoding for frequency-selective massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 2, no. 5, pp. 547–550, Oct. 2013.

[20] 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 ICT*, Saint-Malo, France, Jun. 26-28, 2018, pp. 367–372.

[21] C. G. Tsinos, S. Domouchtsidis, S. Chatzinotas, and B. Ottersten, "Symbol level precoding with low resolution DACs for constant envelope OFDM MU-MIMO systems," *IEEE Access*, vol. 8, pp. 12 856–12 866, Jan. 2020.

[22] S. Domouchtsidis, C. G. Tsinos, S. Chatzinotas, and B. Ottersten, "Constant envelope MIMO-OFDM precoding for low complexity large-scale antenna array systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 7973–7985, Dec. 2020.

[23] F. Askerbeyli, H. Jedda, and J. A. Nossek, "1-bit precoding in Massive MU-MISO-OFDM downlink with linear programming," in *Proc. 23rd Int. ITG Workshop Smart Antennas (WSA)*, Vienna, Austria, Apr. 24-26, 2019, pp. 1–5.

[24] J. Lemberg, M. Kosunen, E. Roverato, M. Martelius, K. Stadius, L. Anttila, M. Valkama, and J. Ryynänen, "Digital interpolating phase

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2023.3282158

16

modulator for wideband outphasing transmitters," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 63, no. 5, pp. 705–715, May 2016.

[25] M. Kosunen, J. Lemberg, M. Martelius, E. Roverato, T. Nieminen, M. Englund, K. Stadius, L. Anttila, J. Pallonen, M. Valkama, and J. Ryynänen, "13.5 A 0.35-to-2.6 GHz multilevel outphasing transmitter with a digital interpolating phase modulator enabling up to 400 MHz instantaneous bandwidth," in *Proc. IEEE ISSCC*, San Francisco, CA, Feb. 5-9, 2017, pp. 224–225.

[26] S. Shinjo, K. Nakatani, K. Tsutsumi, and H. Nakamizo, "Integrating the front end: A highly integrated RF front end for high-SHF wide-band massive MIMO in 5G," *IEEE Microw. Mag.*, vol. 18, no. 5, pp. 31–40, July-Aug. 2017.

[27] R. B. Staszewski, "State-of-the-art and future directions of high-performance all-digital frequency synthesis in nanometer CMOS," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 58, pp. 1497–1510, Jul. 2011.

[28] J. S. Walling, S.-M. Yoo, and D. J. Allstot, "Digital power amplifier: A new way to exploit the switched-capacitor circuit," *IEEE Commun. Mag.*, vol. 50, no. 4, pp. 145–151, Apr. 2012.

[29] A. R. Saleem, K. Stadius, M. Kosunen, L. Anttila, M. Valkama, and J. Ryynänen, "A 30-GHz switched-capacitor power amplifier for 5G SoCs," in *Proc. ICECS*, Glasgow, UK, Nov. 23-25, 2020, pp. 1–4.

[30] X. Luo, H. J. Qian, Y. Yin, and H. Xu, "Empowering multifunction: Digital power amplifiers, the last RF frontier of the analog and digital kingdoms," *IEEE Microw. Mag.*, vol. 21, no. 12, pp. 47–67, Dec. 2020.

[31] C. Preissl, P. Preyler, T. Mayer, A. Springer, and M. Huemer, "Analysis of spectral degradation and error compensation in 5G NR digital polar transmitters," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 67, no. 8, pp. 2719–2729, Aug. 2020.

[32] P. Nazari, B.-K. Chun, F. Tzeng, and P. Heydari, "Polar quantizer for wireless receivers: Theory, analysis, and CMOS implementation," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 61, no. 3, pp. 877–887, Nov. 2014.

[33] K. Khalaf *et al.*, "Digitally modulated CMOS polar transmitters for highly-efficient mm-wave wireless communication," *IEEE J. Solid-State Circuits*, vol. 51, no. 7, pp. 1579–1592, Jul. 2016.

[34] Q. Zhu, S. Yu, S. Wang, L. Huang, Z. Wang, X. Zhang, and Y. Xu, "A digital polar transmitter with DC–DC converter supporting 256-QAM WLAN and 40-MHz LTE-A carrier aggregation," *IEEE J. Solid-State Circuits*, vol. 52, no. 5, pp. 1196–1209, May 2017.

[35] Palaskas *et al.*, "A cellular multiband DTC-based digital polar transmitter with -153 dBc/Hz noise in 14-nm FinFET," in *Proc. IEEE ESSCIRC*, Cracow, Poland, Sep. 23-26, 2019, pp. 179–182.

[36] Ben-Bassat *et al.*, "A fully integrated 27-dBm dual-band all-digital polar transmitter supporting 160 MHz for Wi-Fi 6 applications," *IEEE J. Solid-State Circuits*, vol. 55, no. 12, pp. 3414–3425, Dec. 2020.

[37] R. Winoto, "Digital radio-frequency transmitters: An introduction and tutorial," *IEEE Solid-State Circuits Mag.*, vol. 10, no. 4, pp. 70–80, Fall 2018.

[38] G. Xu, V. Lampu, M. Kosunen, K. Stadius, J. Ryynänen, M. Valkama, and L. Anttila, "Quantized polar transmitters for power efficient massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 10, no. 4, pp. 859–862, April 2020.

[39] 3GPP Tech. Rep. 38.901, "Study on channel model for frequencies from 0.5 to 100 GHz," *v15.0.0*, Jun. 2018.

[40] D. Cherniak *et al.*, "A 250-Mb/s direct phase modulator with- 42.4-dB EVM based on a 14-GHz digital PLL," *IEEE Solid-State Circuits Lett.*, vol. 3, pp. 126–129, Aug. 2020.

[41] J. Lemberg, M. Martelius, E. Roverato, Y. Antonov, T. Nieminen, K. Stadius, L. Anttila, M. Valkama, M. Kosunen, and J. Ryynänen, "A 1.5–1.9-GHz all-digital tri-phasing transmitter with an integrated multilevel class-D power amplifier achieving 100-MHz RF bandwidth," *IEEE J. Solid-State Circuits*, vol. 54, no. 6, pp. 1517–1527, Jun. 2019.

[42] O. T. Demir and E. Björnson, "The Bussgang decomposition of nonlinear systems: Basic theory and MIMO extensions," *IEEE Signal Process. Mag.*, vol. 38, no. 1, pp. 131–136, Jan. 2021.

[43] B. Jann *et al.*, "21.5 a 5G sub-6GHz zero-IF and mm-Wave IF transceiver with MIMO and carrier aggregation," in *Proc. ISSCC*, San Francisco, CA, Feb. 17-21, 2019, pp. 352–354.

[44] J. Lee *et al.*, "A sub-6-GHz 5G new radio RF transceiver supporting EN-DC with 3.15-Gb/s DL and 1.27-Gb/s UL in 14-nm FinFET CMOS," *IEEE J. Solid-State Circuits*, vol. 54, no. 12, pp. 3541–3552, Sep. 2019.

[45] M.-D. Tsai *et al.*, "10.6 a 4G/5G cellular transmitter in 12nm FinFET with harmonic rejection," in *Proc. ISSCC*, San Francisco, CA, Feb. 16-20, 2020, pp. 182–184.

[46] H. Yan, S. Ramesh, T. Gallagher, C. Ling, and D. Cabric, "Performance, power, and area design trade-offs in millimeter-wave transmitter beamforming architectures," *IEEE Circuits Syst. Mag.*, vol. 19, no. 2, pp. 33–58, May 2019.

[47] J. Navarro-Ortiz, P. Romero-Diaz, S. Sendra, P. Ameigeiras, J. J. Ramos-Munoz, and J. M. Lopez-Soler, "A survey on 5G usage scenarios and traffic models," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 2, pp. 905–929, 2nd Quart. 2020.

[48] ITU-R M.2412, "Guidelines for evaluation of radio interface technologies for IMT-2020," *ITU-R WP 5D*, Oct. 2017.

[49] R. M. Vaghefi *et al.*, "Achieving phase coherency and gain stability in active antenna arrays for sub-6 GHz FDD and TDD FD-MIMO: Challenges and solutions," *IEEE Access*, vol. 8, pp. 152 680–152 696, Aug. 2020.

[50] M. Shahmohammadi, M. Babaie, and R. B. Staszewski, "Tuning range extension of a transformer-based oscillator through common-mode colpitts resonance," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 64, no. 4, pp. 836–846, Mar. 2017.

[51] J.-D. Chen and J.-b. Qian, "Low-power up-conversion folded CMOS mixer for 2–12 GHz ultra-wideband applications," *IET Microw. Antennas Propag.*, vol. 14, no. 15, pp. 1975–1982, 2020.

[52] H. Kim, C.-B. Chae, G. De Veciana, and R. W. Heath, "A cross-layer approach to energy efficiency for adaptive MIMO systems exploiting spare capacity," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4264–4275, Aug. 2009.

[53] H. Wang, F. Wang, H. Nguyen, S. Li, T. Huang, A. Ahmed, M. Smith, N. Mannem, and J. Lee, "Power amplifiers performance survey 2000-present," *Georgia Tech Electronics and Micro-System Lab (GEMS)*, 2018.

[54] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?" *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3059–3075, Jun. 2015.

[55] C. Li, Y. Huang, K. Khalaf, A. Bourdoux, M. Verhelst, L. Van Der Perre, and S. Pollin, "Energy-efficient digital front-end processor for 60 GHz polar transmitter," *Journal of Signal Processing Systems*, vol. 90, no. 5, pp. 777–789, 2018.

[56] X. Gao, O. Edfors, F. Rusek, and F. Tufvesson, "Massive MIMO performance evaluation based on measured propagation data," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3899–3911, Jul. 2015.

[57] 3GPP Tech. Spec. 38.104, "NR; base station (BS) radio transmission and reception," *v15.4.0, (Release 15)*, Dec. 2018.

Guixian Xu received the Ph.D. degree in communications and information systems from Beijing University of Posts and Telecommunications (BUPT), China, in 2017. From 2015 to 2016, he was a visiting Ph.D. student at the National Tsing Hua University in Hsinchu, Taiwan. From 2017 to 2018, he was a postdoctoral researcher with the connectivity section in the Department of Electronic Systems at Aalborg University, Denmark. Since 2020, he has been working as a postdoc at Tampere University in Finland. His research interests include digital transmitter architectures, digital predistortion technology, and machine learning for wireless communication networks.

Vesa Lampu received the B.Sc. and M.Sc. degrees in electrical engineering from Tampere University, Tampere, Finland, in 2017 and 2019, respectively, where he is currently pursuing the D.Sc. degree. He has been with the Department of Electrical Engineering, Tampere University, since 2018. His current research has two main focus areas: energy-efficient transmitter systems and nonlinear system identification, with applications in full-duplex and intermodulation cancellation.

Marko Kosunen (S'97−M'07) received his M.Sc, L.Sc and D.Sc (with honors) degrees from Helsinki University of Technology, Espoo, Finland, in 1998, 2001 and 2006, respectively. He is currently a Senior Researcher at Aalto University, Department of Electronics and Nanoengineering. Academic years 2017-2019 he visited Berkeley Wireless Reserarch Center, UC Berkeley, on Marie Sklodowska-Curie grant from European Union. He has authored and co-authored more than hundred journal and conference papers and holds several patents. His current research interests include programmatic circuit design methodologies, digital intensive and time-based data converters and transceiver circuits.

Lauri Anttila received the D.Sc. (Tech.) degree (Hons.) from the Tampere University of Technology, Tampere, Finland, in 2011. Since 2016, he has been a Senior Research Fellow with the Department of Electrical Engineering, Tampere University. From 2016 to 2017, he was a Visiting Research Fellow with the Department of Electronics and Nanoengineering at Aalto University, Finland. His research interests are in radio communications and signal processing, with a focus on the radio implementation challenges in systems such as 5G/6G, large-scale antenna systems, full-duplex radio, and integrated sensing and communication systems.

Vishnu Unnikrishnan (Member IEEE) received the M.Sc. degree in electrical engineering and the Ph.D. degree in integrated circuits and systems, both from Linköping University, Sweden, in 2012 and 2016 respectively. He is currently an Assistant Professor at the Department of Electrical Engineering, Tampere University, Finland. From 2017 to 2021, he was a Postdoctoral Researcher at the Department of Electronics and Nanoengineering, Aalto University, Finland. From 2004 to 2009, he was with Bosch, India. His research interests include energy-efficient high-performance integrated circuits and systems, radio/wireline transceivers, and digital implementation/enhancement of analog/mixed-signal functions in integrated circuits.

Jussi Ryynänen (S'99–M'04–SM'16) was born in Ilmajoki, Finland, in 1973. He received the M.Sc. and D.Sc. degrees in electrical engineering from the Helsinki University of Technology, Espoo, Finland,in 1998 and 2004, respectively. He is a full professor and the Head of the Department of Electronics and Nanoengineering, Aalto University, Espoo. He has authored or co-authored more than 200 refereed journal and conference papers in analog and RF circuit design. He holds seven patents on RF circuits. His research interests are integrated transceiver circuits for wireless applications. Prof. Ryynänen has served as a TPC Member for the European Solid-State Circuits Conference (ESSCIRC) and the IEEE International Solid-State Circuits Conference (ISSCC), and as a Guest Editor for the IEEE Journal of Solid-State Circuits

Mikko Valkama [S'00, M'01, SM'15, F'22] received his M.Sc. (Tech.) and D.Sc. (Tech.) degrees (both with honors) from Tampere University of Technology, Finland, in 2000 and 2001, respectively. In 2003, he was with the Communications Systems and Signal Processing Institute at SDSU, San Diego, CA, as a visiting research fellow. Currently, he is a Full Professor and the Head of the Unit of Electrical Engineering at the newly formed Tampere University, Finland. His general research interests include radio communications, radio localization, and radio-based sensing, with particular emphasis on 5G and 6G mobile radio networks.