# Low-Resolution Massive MIMO Under Hardware Power Consumption Constraints

Italo Atzeni,[1] Antti Tölli,[1] and Giuseppe Durisi[2]

[1]Centre for Wireless Communications, University of Oulu, Finland
[2]Chalmers University of Technology, Sweden
Emails: {italo.atzeni, antti.tolli}@oulu.fi, durisi@chalmers.se

***Abstract*—We consider a fully digital massive multiple-input multiple-output architecture with low-resolution analog-to-digital/digital-to-analog converters (ADCs/DACs) at the base station (BS) and analyze the performance trade-off between the number of BS antennas, the resolution of the ADCs/DACs, and the bandwidth. Assuming a hardware power consumption constraint, we determine the relationship between these design parameters by using a realistic model for the power consumption of the ADCs/DACs and the radio frequency chains. Considering uplink pilot-aided channel estimation, we build on the Bussgang decomposition to derive tractable expressions for uplink and downlink ergodic achievable sum rates. Numerical results show that the ergodic performance is boosted when many BS antennas with very low resolution (i.e., 2 to 3 bits) are adopted in both the uplink and the downlink.**

## I. Introduction

To support the exponential growth in mobile data traffic, beyond-5G wireless systems are expected to exploit the large amount of bandwidth available in the sub-THz band (0.1–1 THz) [1]. In this context, fully digital architectures allow one to realize the full potential of massive multiple-input multiple-output (MIMO) by enabling highly flexible beamforming and large-scale spatial multiplexing. In fully digital massive MIMO, each base station (BS) antenna is connected to a radio-frequency (RF) chain that includes a pair of power-hungry analog-to-digital/digital-to-analog converters (ADCs/DACs). In this setting, the power consumption of each ADC/DAC scales linearly with the sampling rate and exponentially with the number of resolution bits [2], [3]. Hence, adopting low-resolution ADCs/DACs (even down to 1-bit [4], [5]) represents a necessary step towards the energy-efficient implementation of fully digital massive MIMO arrays comprising hundreds or thousands of antennas.

In this paper, we analyze the performance trade-off between the number of BS antennas, the resolution of the ADCs/DACs, and the bandwidth under a hardware power consumption constraint. The relationship between these design parameters is determined by means of a realistic model for the power consumption of the ADCs/DACs and the RF chains. In this setting, a lower resolution of the ADCs/DACs allows one to increase the number of BS antennas (and, consequently, the array gain and the spatial resolution) at the expense of a more significant quantization distortion during both the channel estimation and the uplink/downlink data transmission. To analyze this trade-off, we derive tractable expressions for uplink and downlink ergodic achievable sum rates with imperfect channel state information (CSI). Leveraging the Bussgang decomposition [6], we take into account the quantization distortion in the uplink/downlink data transmission as well as in the combining and precoding matrices, which are based on quantized channel estimates. To further simplify the resulting sum rate expressions, we assume that the BS adopts maximum ratio combining (MRC) in the uplink and maximum ratio transmission (MRT) in the downlink. In the evaluated scenarios, the most energy-efficient configurations consist of many antennas with 2 to 3 resolution bits in both the uplink and the downlink. A related study was presented in [7], [8], where the uplink and downlink outage rates were analyzed in the presence of a fronthaul constraint.

## II. System Model

We consider a massive MIMO system where a BS with $M$ antennas and low-resolution ADCs/DACs serves $K$ single-antenna user equipments (UEs) with infinite-resolution ADCs/DACs. We use $\mathbf{H} \triangleq [\mathbf{h}_1, \ldots, \mathbf{h}_K] \in \mathbb{C}^{M \times K}$ to denote the uplink channel matrix.[1] Each BS antenna is connected to two $b$-bit ADCs in the uplink (resp. to two $b$-bit DACs in the downlink), one for the in-phase and one for the quadrature component of the receive (resp. transmit) signal. Each ADC/DAC in the uplink/downlink consists of a $b$-bit quantizer characterized by a set of $2^b + 1$ quantization thresholds $\mathcal{T}_b \triangleq \{t_n\}_{n=0}^{2^b}$, with $-\infty = t_0 < \ldots < t_{2^b} = \infty$, and by a set of $2^b$ quantization labels $\mathcal{L}_b \triangleq \{\ell_n\}_{n=0}^{2^b-1}$, with $\ell_n \in (t_n, t_{n+1}]$. In this setting, we introduce the quantization function $Q_b(\cdot)$ such that, for $s \in \mathbb{C}$, $Q_b(s) = \ell_i + j\,\ell_k$ if $\mathrm{Re}[s] \in (t_i, t_{i+1}]$ and $\mathrm{Im}[s] \in (t_k, t_{k+1}]$; for a matrix or vector input, $Q_b(\cdot)$ is applied entrywise.

We assume that the number of BS antennas, the resolution of the ADCs/DACs, and the bandwidth are related through a hardware power consumption constraint. Let $P_{\text{HW}}$ be the maximum hardware power consumption allowed at the BS and let $P_{\text{ADC}}(b, B)$ and $P_{\text{DAC}}(b, B)$ denote the power consumed

The work of I. Atzeni was supported by the Marie Skłodowska-Curie Actions (MSCA-IF 897938 DELIGHT). The work of A. Tölli was supported by the Academy of Finland (318927 6Genesis Flagship and 319059 CCCWEE). The work of G. Durisi was supported by the Wallenberg AI, Autonomous Systems, and Software Program.

[1]We will consider the special case of i.i.d. Rayleigh fading channels in Section IV.

by each ADC and DAC, respectively, where $B$ represents the bandwidth. Both $P_{\text{ADC}}(b,B)$ and $P_{\text{DAC}}(b,B)$ grow linearly with the bandwidth and exponentially with the number of resolution bits, as in the power consumption model adopted in Section V-A. Hence, the number of BS antennas available in the uplink/downlink for a given resolution of the ADCs/DACs $b$ and a given bandwidth $B$ is

$$M = \begin{cases} \left\lfloor \frac{P_{\text{HW}}}{P_{\text{RF}}+2P_{\text{ADC}}(b,B)} \right\rfloor & \text{(uplink)}, \\ \left\lfloor \frac{P_{\text{HW}}}{P_{\text{RF}}+2P_{\text{DAC}}(b,B)} \right\rfloor & \text{(downlink)} \end{cases} \tag{1}$$

where $P_{\text{RF}}$ denotes the power consumed by each RF chain. Note that, in the following, $M$ assumes different values for the uplink and the downlink according to (1).

## III. Ergodic Performance

In this section, we provide expressions for uplink and downlink ergodic achievable sum rates with $b$-bit ADCs and DACs, respectively, for general combining and precoding matrices. These are extended in Section IV to include the effect of imperfect CSI arising from the uplink pilot-aided channel estimation with low-resolution ADCs.

### A. Uplink Ergodic Achievable Sum Rate

Let $\mathbf{x}^{\text{ul}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_K)$ denote the uplink transmit symbol vector. At the BS, the normalized input signal at the ADCs is given by

$$\mathbf{y}^{\text{ul}} \triangleq \sqrt{\rho_{\text{BS}}}\mathbf{H}\mathbf{x}^{\text{ul}} + \mathbf{z}^{\text{ul}} \in \mathbb{C}^{M\times 1} \tag{2}$$

where $\mathbf{z}^{\text{ul}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$ is the additive white Gaussian noise (AWGN) vector; since the AWGN power is normalized to one, $\rho_{\text{BS}}$ represents the signal-to-noise ratio (SNR) at the BS. Then, at the output of the ADCs, we have

$$\mathbf{r}^{\text{ul}} \triangleq Q_b(\mathbf{y}^{\text{ul}}) \in \mathbb{C}^{M\times 1}. \tag{3}$$

We assume that the BS obtains a soft estimate of $\mathbf{x}^{\text{ul}}$ as

$$\hat{\mathbf{x}}^{\text{ul}} \triangleq [\hat{x}_1^{\text{ul}}, \ldots, \hat{x}_K^{\text{ul}}]^{\text{T}} \tag{4}$$

$$= \mathbf{V}^{\text{H}}\mathbf{r}^{\text{ul}} \in \mathbb{C}^{K\times 1} \tag{5}$$

where $\mathbf{V} \triangleq [\mathbf{v}_1, \ldots, \mathbf{v}_K] \in \mathbb{C}^{M\times K}$ is the combining matrix.

Using the Bussgang decomposition, we can write (3) as

$$\mathbf{r}^{\text{ul}} = \mathbf{G}_b^{\text{ul}}\mathbf{y}^{\text{ul}} + \mathbf{d}^{\text{ul}} \tag{6}$$

where $\mathbf{d}^{\text{ul}} \in \mathbb{C}^{M\times 1}$ is the zero-mean, non-Gaussian quantization distortion vector that is uncorrelated with $\mathbf{y}^{\text{ul}}$ and has covariance matrix $\mathbf{C}_{\mathbf{d}}^{\text{ul}} \triangleq \mathbb{E}\big[\mathbf{d}^{\text{ul}}(\mathbf{d}^{\text{ul}})^{\text{H}}\big] \in \mathbb{C}^{M\times M}$. In addition, $\mathbf{G}_b^{\text{ul}} \in \mathbb{C}^{M\times M}$ is a diagonal matrix given by (see, e.g., [2])

$$\mathbf{G}_b^{\text{ul}} \triangleq \big(\pi \text{Diag}(\mathbf{C}_{\mathbf{y}}^{\text{ul}})\big)^{-\frac{1}{2}} \sum_{n=0}^{2^b-1} \ell_n\Big(\exp\big(-t_n^2 \text{Diag}(\mathbf{C}_{\mathbf{y}}^{\text{ul}})^{-1}\big) - \exp\big(-t_{n+1}^2 \text{Diag}(\mathbf{C}_{\mathbf{y}}^{\text{ul}})^{-1}\big)\Big) \tag{7}$$

with $\mathbf{C}_{\mathbf{y}}^{\text{ul}} \triangleq \mathbb{E}\big[\mathbf{y}^{\text{ul}}(\mathbf{y}^{\text{ul}})^{\text{H}}\big] \in \mathbb{C}^{M\times M}$. Substituting (6) into (5), we obtain

$$\hat{\mathbf{x}}^{\text{ul}} = \mathbf{V}^{\text{H}}\big(\mathbf{G}_b^{\text{ul}}(\sqrt{\rho_{\text{BS}}}\mathbf{H}\mathbf{x}^{\text{ul}} + \mathbf{z}^{\text{ul}}) + \mathbf{d}^{\text{ul}}\big). \tag{8}$$

To derive an easy-to-evaluate expression for an uplink ergodic achievable sum rate, we use the use-and-then-forget bound [9, Thm. 4.4]. Specifically, we write the $k$th entry of (8) as

$$\hat{x}_k^{\text{ul}} = \sqrt{\rho_{\text{BS}}}\mathbb{E}[\mathbf{v}_k^{\text{H}}\mathbf{G}_b^{\text{ul}}\mathbf{h}_k]x_k^{\text{ul}} + \sqrt{\rho_{\text{BS}}}\big(\mathbf{v}_k^{\text{H}}\mathbf{G}_b^{\text{ul}}\mathbf{h}_k - \mathbb{E}[\mathbf{v}_k^{\text{H}}\mathbf{G}_b^{\text{ul}}\mathbf{h}_k]\big)x_k^{\text{ul}} + \sqrt{\rho_{\text{BS}}}\sum_{i\neq k}\mathbf{v}_k^{\text{H}}\mathbf{G}_b^{\text{ul}}\mathbf{h}_i x_i^{\text{ul}} + \mathbf{v}_k^{\text{H}}\mathbf{G}_b^{\text{ul}}\mathbf{z}^{\text{ul}} + \mathbf{v}_k^{\text{H}}\mathbf{d}^{\text{ul}}. \tag{9}$$

Treating $\mathbf{d}^{\text{ul}}$ as Gaussian noise with covariance matrix $\mathbf{C}_{\mathbf{d}}^{\text{ul}}$ and assuming that the BS is aware of the value of $\mathbb{E}[\mathbf{v}_k^{\text{H}}\mathbf{G}_b^{\text{ul}}\mathbf{h}_k]$, we obtain the effective uplink signal-to-interference-plus-noise-and-distortion ratio (SINDR) of UE $k$ as in (10) at the top of the next page. Finally, an uplink ergodic achievable sum rate is given by

$$R^{\text{ul}} \triangleq B\sum_{k=1}^{K}\log_2(1+\gamma_k^{\text{ul}}). \tag{11}$$

### B. Downlink Ergodic Achievable Sum Rate

Let $\mathbf{x}^{\text{dl}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_K)$ denote the downlink transmit symbol vector. At the BS, the output signal at the DACs is given by

$$\mathbf{r}^{\text{dl}} \triangleq Q_b(\mathbf{W}\mathbf{x}^{\text{dl}}) \in \mathbb{C}^{M\times 1} \tag{12}$$

where $\mathbf{W} \triangleq [\mathbf{w}_1, \ldots, \mathbf{w}_K] \in \mathbb{C}^{M\times K}$ is the precoding matrix with $\mathbb{E}\big[\|\mathbf{W}\|_{\text{F}}^2\big] = 1$. Then, the normalized receive signal at the UEs is given by

$$\mathbf{y}^{\text{dl}} \triangleq [y_1^{\text{dl}}, \ldots, y_K^{\text{dl}}]^{\text{T}} \tag{13}$$

$$= \sqrt{\rho_{\text{UE}}}\mathbf{H}^{\text{H}}\mathbf{r}^{\text{dl}} + \mathbf{z}^{\text{dl}} \in \mathbb{C}^{K\times 1} \tag{14}$$

where $\mathbf{z}^{\text{dl}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_K)$ is the AWGN vector; since the AWGN power is normalized to one, $\rho_{\text{UE}}$ represents the SNR at the UEs.

Using the Bussgang decomposition, we can write (12) as

$$\mathbf{r}^{\text{dl}} = \mathbf{G}_b^{\text{dl}}\mathbf{W}\mathbf{x}^{\text{dl}} + \mathbf{d}^{\text{dl}} \tag{15}$$

where $\mathbf{d}^{\text{dl}} \in \mathbb{C}^{M\times 1}$ is the zero-mean, non-Gaussian quantization distortion vector that is uncorrelated with $\mathbf{W}\mathbf{x}^{\text{dl}}$ and has covariance matrix $\mathbf{C}_{\mathbf{d}}^{\text{dl}} \triangleq \mathbb{E}\big[\mathbf{d}^{\text{dl}}(\mathbf{d}^{\text{dl}})^{\text{H}}\big] \in \mathbb{C}^{M\times M}$. In addition, $\mathbf{G}_b^{\text{dl}} \in \mathbb{C}^{M\times M}$ is a diagonal matrix obtained by replacing $\mathbf{C}_{\mathbf{y}}^{\text{ul}}$ in (7) with $\mathbb{E}[\mathbf{W}\mathbf{W}^{\text{H}}]$. Substituting (15) into (14), we obtain

$$\mathbf{y}^{\text{dl}} = \sqrt{\rho_{\text{UE}}}\mathbf{H}^{\text{H}}(\mathbf{G}_b^{\text{dl}}\mathbf{W}\mathbf{x}^{\text{dl}} + \mathbf{d}^{\text{dl}}) + \mathbf{z}^{\text{dl}}. \tag{16}$$

To derive an easy-to-evaluate expression for a downlink ergodic achievable sum rate, we use the channel hardening bound [9, Thm. 4.6]. Specifically, we write the $k$th entry of (16) as

$$y_k^{\text{dl}} = \sqrt{\rho_{\text{UE}}}\mathbb{E}[\mathbf{h}_k^{\text{H}}\mathbf{G}_b^{\text{dl}}\mathbf{w}_k]x_k^{\text{dl}} + \sqrt{\rho_{\text{UE}}}\big(\mathbf{h}_k^{\text{H}}\mathbf{G}_b^{\text{dl}}\mathbf{w}_k - \mathbb{E}[\mathbf{h}_k^{\text{H}}\mathbf{G}_b^{\text{dl}}\mathbf{w}_k]\big)x_k^{\text{dl}} + \sqrt{\rho_{\text{UE}}}\sum_{i\neq k}\mathbf{h}_k^{\text{H}}\mathbf{G}_b^{\text{dl}}\mathbf{w}_i x_i^{\text{dl}} + \sqrt{\rho_{\text{UE}}}\mathbf{h}_k^{\text{H}}\mathbf{d}^{\text{dl}} + z_k^{\text{dl}}. \tag{17}$$

Treating $\mathbf{d}^{\text{dl}}$ as Gaussian noise with covariance matrix $\mathbf{C}_{\mathbf{d}}^{\text{dl}}$ and assuming that UE $k$ is aware of the value of $\mathbb{E}[\mathbf{h}_k^{\text{H}}\mathbf{G}_b^{\text{dl}}\mathbf{w}_k]$, we obtain the effective downlink SINDR of UE $k$ as in (18) at the top of the page. Finally, a downlink ergodic achievable sum rate is given by

$$R^{\text{dl}} \triangleq B\sum_{k=1}^{K}\log_2(1+\gamma_k^{\text{dl}}). \tag{19}$$

## IV. Ergodic Performance with Imperfect CSI

The ergodic achievable sum rate expressions derived in Section III take into account the quantization distortion in the uplink/downlink data transmission. In practice, however, the quantization also affects the combining and precoding matrices, which are based on quantized channel estimates. In this section, we include the effect of imperfect CSI arising from the uplink

$$\gamma_k^{\text{ul}} \triangleq \frac{\rho_{\text{BS}}\left|\mathbb{E}[\mathbf{v}_k^{\text{H}}\mathbf{G}_b^{\text{ul}}\mathbf{h}_k]\right|^2}{\rho_{\text{BS}}\sum_{i=1}^{K}\mathbb{E}\left[|\mathbf{v}_k^{\text{H}}\mathbf{G}_b^{\text{ul}}\mathbf{h}_i|^2\right]-\rho_{\text{BS}}\left|\mathbb{E}[\mathbf{v}_k^{\text{H}}\mathbf{G}_b^{\text{ul}}\mathbf{h}_k]\right|^2+\mathbb{E}\left[\|\mathbf{G}_b^{\text{ul}}\mathbf{v}_k\|^2\right]+\mathbb{E}[\mathbf{v}_k^{\text{H}}\mathbf{C}_{\mathbf{d}}^{\text{ul}}\mathbf{v}_k]}, \tag{10}$$

$$\gamma_k^{\text{dl}} \triangleq \frac{\rho_{\text{UE}}\left|\mathbb{E}[\mathbf{h}_k^{\text{H}}\mathbf{G}_b^{\text{dl}}\mathbf{w}_k]\right|^2}{\rho_{\text{UE}}\sum_{i=1}^{K}\mathbb{E}\left[|\mathbf{h}_k^{\text{H}}\mathbf{G}_b^{\text{dl}}\mathbf{w}_i|^2\right]-\rho_{\text{UE}}\left|\mathbb{E}[\mathbf{h}_k^{\text{H}}\mathbf{G}_b^{\text{dl}}\mathbf{w}_k]\right|^2+\rho_{\text{UE}}\mathbb{E}[\mathbf{h}_k^{\text{H}}\mathbf{C}_{\mathbf{d}}^{\text{dl}}\mathbf{h}_k]+1} \tag{18}$$

pilot-aided channel estimation with low-resolution ADCs. To further simplify the resulting sum rate expressions, we consider the special case of i.i.d. Rayleigh fading channels, where the entries of $\mathbf{H}$ are independent $\mathcal{CN}(0,1)$ random variables, and we assume that the BS adopts MRC in the uplink and MRT in the downlink. The full derivations are omitted due to the space limitations and will be included in an extended version of this paper, which will also consider correlated channels.

### A. Channel Estimation with $b$-Bit ADCs

During the uplink pilot-aided channel estimation phase, the UEs simultaneously transmit their uplink pilots of length $\tau$. Let $\mathbf{P} \triangleq [\mathbf{p}_1,\ldots,\mathbf{p}_K] \in \mathbb{C}^{\tau\times K}$ denote the pilot matrix, where $\mathbf{p}_k \in \mathbb{C}^{\tau\times 1}$ represents the pilot sequence assigned to UE $k$. In this context, we assume $\tau \geq K$ and that the pilot sequences are orthogonal, so that $\mathbf{P}^{\text{H}}\mathbf{P} = \tau\mathbf{I}_K$. Furthermore, we define $\bar{\mathbf{P}}_k \triangleq \mathbf{p}_k \otimes \mathbf{I}_M \in \mathbb{C}^{M\tau\times M}$ and $\bar{\mathbf{P}} \triangleq [\bar{\mathbf{P}}_1,\ldots,\bar{\mathbf{P}}_K] \in \mathbb{C}^{M\tau\times MK}$. At the BS, the normalized input signal at the ADCs is given by

$$\mathbf{y}^{\text{ce}} \triangleq \sqrt{\rho_{\text{BS}}}\bar{\mathbf{P}}^*\mathbf{h} + \mathbf{z}^{\text{ce}} \in \mathbb{C}^{M\tau\times 1} \tag{20}$$

with $\mathbf{h} \triangleq \text{vec}(\mathbf{H}) \in \mathbb{C}^{MK\times 1}$ and where $\mathbf{z}^{\text{ce}} \sim \mathcal{CN}(\mathbf{0},\mathbf{I}_{M\tau})$ is the AWGN vector. Then, at the output of the ADCs, we have

$$\mathbf{r}^{\text{ce}} \triangleq Q_b(\mathbf{y}^{\text{ce}}) \in \mathbb{C}^{M\tau\times 1}. \tag{21}$$

We assume that the BS obtains an estimate of $\mathbf{h}$ (up to a scaling factor) as

$$\hat{\mathbf{h}} \triangleq \frac{1}{\sqrt{\rho_{\text{BS}}}\tau}\bar{\mathbf{P}}^{\text{T}}\mathbf{r}^{\text{ce}} \in \mathbb{C}^{MK\times 1}. \tag{22}$$

Using the Bussgang decomposition, we can write (21) as

$$\mathbf{r}^{\text{ce}} = \mathbf{G}_b^{\text{ce}}\mathbf{y}^{\text{ce}} + \mathbf{d}^{\text{ce}} \tag{23}$$

where $\mathbf{d}^{\text{ce}} \in \mathbb{C}^{M\tau\times 1}$ is the zero-mean, non-Gaussian quantization distortion vector that is uncorrelated with $\mathbf{y}^{\text{ce}}$ and has covariance matrix $\mathbf{C}_{\mathbf{d}}^{\text{ce}} \triangleq \mathbb{E}\left[\mathbf{d}^{\text{ce}}(\mathbf{d}^{\text{ce}})^{\text{H}}\right] \in \mathbb{C}^{M\tau\times M\tau}$. In addition, $\mathbf{G}_b^{\text{ce}} \in \mathbb{C}^{M\tau\times M\tau}$ is a diagonal matrix obtained by replacing $\mathbf{C}_{\mathbf{y}}^{\text{ul}}$ in (7) with $\mathbf{C}_{\mathbf{y}}^{\text{ce}} \triangleq \mathbb{E}\left[\mathbf{y}^{\text{ce}}(\mathbf{y}^{\text{ce}})^{\text{H}}\right] \in \mathbb{C}^{M\tau\times M\tau}$. Substituting (23) into (22), we obtain

$$\hat{\mathbf{h}} = \frac{1}{\tau}\left(\bar{\mathbf{P}}^{\text{T}}\mathbf{G}_b^{\text{ce}}\bar{\mathbf{P}}^*\mathbf{h} + \frac{1}{\sqrt{\rho_{\text{BS}}}}\bar{\mathbf{P}}^{\text{T}}\mathbf{G}_b^{\text{ce}}\mathbf{z}^{\text{ce}} + \frac{1}{\sqrt{\rho_{\text{BS}}}}\bar{\mathbf{P}}^{\text{T}}\mathbf{d}^{\text{ce}}\right). \tag{24}$$

### B. Uplink Ergodic Achievable Sum Rate with MRC

Assuming that MRC is used at the BS, we have $\mathbf{V} = \mathbf{G}_b^{\text{ul}}\hat{\mathbf{H}}$ and the combining vector corresponding to UE $k$ is given by

$$\mathbf{v}_k = \frac{1}{\tau}\mathbf{G}_b^{\text{ul}}\left(\bar{\mathbf{P}}_k^{\text{T}}\mathbf{G}_b^{\text{ce}}\bar{\mathbf{P}}^*\mathbf{h} + \frac{1}{\sqrt{\rho_{\text{BS}}}}\bar{\mathbf{P}}_k^{\text{T}}\mathbf{G}_b^{\text{ce}}\mathbf{z}^{\text{ce}} + \frac{1}{\sqrt{\rho_{\text{BS}}}}\bar{\mathbf{P}}_k^{\text{T}}\mathbf{d}^{\text{ce}}\right). \tag{25}$$

Under the i.i.d. Rayleigh fading assumption, we have $\mathbf{C}_{\mathbf{y}}^{\text{ul}} = (\rho_{\text{BS}}K+1)\mathbf{I}_M$ and $\mathbf{C}_{\mathbf{y}}^{\text{ce}} = \rho_{\text{BS}}\bar{\mathbf{P}}^*\bar{\mathbf{P}}^{\text{T}} + \mathbf{I}_{M\tau}$ (with $\text{Diag}(\mathbf{C}_{\mathbf{y}}^{\text{ul}}) = (\rho_{\text{BS}}K+1)\mathbf{I}_{M\tau}$), and we can thus write $\mathbf{G}_b^{\text{ul}}$ and $\mathbf{G}_b^{\text{ce}}$ as scaled identity matrices, i.e., $\mathbf{G}_b^{\text{ul}} = G_b^{\text{ul}}\mathbf{I}_M$ and $\mathbf{G}_b^{\text{ce}} = G_b^{\text{ce}}\mathbf{I}_{M\tau}$ (see (7)). Moreover, we introduce the following preliminary definitions:

$$A_k \triangleq \text{tr}(\mathbf{C}_{\mathbf{d}}^{\text{ce}}\bar{\mathbf{P}}_k^*\bar{\mathbf{P}}_k^{\text{T}}), \tag{26}$$

$$B_k \triangleq \text{tr}(\mathbf{C}_{\mathbf{d}}^{\text{ce}}\bar{\mathbf{P}}_k^*\mathbf{C}_{\mathbf{d}}^{\text{ul}}\bar{\mathbf{P}}_k^{\text{T}}). \tag{27}$$

Then, we can express each expectation term in (10) as

$$\mathbb{E}[\mathbf{v}_k^{\text{H}}\mathbf{G}_b^{\text{ul}}\mathbf{h}_k] = G_b^{\text{ce}}(G_b^{\text{ul}})^2M, \tag{28}$$

$$\mathbb{E}\left[|\mathbf{v}_k^{\text{H}}\mathbf{G}_b^{\text{ul}}\mathbf{h}_i|^2\right] = \mathbb{E}\left[\|\mathbf{G}_b^{\text{ul}}\mathbf{v}_k\|^2\right] \tag{29}$$

$$= \left(1+\frac{1}{\rho_{\text{BS}}\tau}\right)(G_b^{\text{ce}})^2(G_b^{\text{ul}})^4M + \frac{1}{\rho_{\text{BS}}\tau^2}(G_b^{\text{ul}})^4A_k \qquad (i\neq k), \tag{30}$$

$$\mathbb{E}\left[|\mathbf{v}_k^{\text{H}}\mathbf{G}_b^{\text{ul}}\mathbf{h}_k|^2\right] = \left(M+1+\frac{1}{\rho_{\text{BS}}\tau}\right)(G_b^{\text{ce}})^2(G_b^{\text{ul}})^4M + \frac{1}{\rho_{\text{BS}}\tau^2}(G_b^{\text{ul}})^4A_k, \tag{31}$$

$$\mathbb{E}[\mathbf{v}_k^{\text{H}}\mathbf{C}_{\mathbf{d}}^{\text{ul}}\mathbf{v}_k] = \left(1+\frac{1}{\rho_{\text{BS}}\tau}\right)(G_b^{\text{ce}})^2(G_b^{\text{ul}})^2\text{tr}(\mathbf{C}_{\mathbf{d}}^{\text{ul}}) + \frac{1}{\rho_{\text{BS}}\tau^2}(G_b^{\text{ul}})^2B_k^{\text{ul}}. \tag{32}$$

This allows us to write the effective uplink SINDR with imperfect CSI and MRC as in (33) at the top of the next page. Finally, we substitute the resulting expression into (11) to obtain an uplink ergodic achievable sum rate.

### C. Downlink Ergodic Achievable Sum Rate with MRT

Assuming that MRT is used at the BS, we have[2] $\mathbf{W} = \frac{1}{\sqrt{\delta}}\hat{\mathbf{H}}$, with $\delta \triangleq \mathbb{E}\left[\|\hat{\mathbf{h}}\|^2\right]$ and $\hat{\mathbf{H}} \triangleq \text{unvec}(\hat{\mathbf{h}}) \in \mathbb{C}^{M\times K}$, and the precoding vector corresponding to UE $k$ is given by

$$\mathbf{w}_k = \frac{1}{\sqrt{\delta}\tau}\left(\bar{\mathbf{P}}_k^{\text{T}}\mathbf{G}_b^{\text{ce}}\bar{\mathbf{P}}^*\mathbf{h} + \frac{1}{\sqrt{\rho_{\text{BS}}}}\bar{\mathbf{P}}_k^{\text{T}}\mathbf{G}_b^{\text{ce}}\mathbf{z}^{\text{ce}} + \frac{1}{\sqrt{\rho_{\text{BS}}}}\bar{\mathbf{P}}_k^{\text{T}}\mathbf{d}^{\text{ce}}\right). \tag{34}$$

Under the i.i.d. Rayleigh fading assumption, we have

$$\mathbb{E}[\mathbf{W}\mathbf{W}^{\text{H}}] = \frac{1}{\delta}\left(K\left(1+\frac{1}{\rho_{\text{BS}}\tau}\right)(G_b^{\text{ce}})^2\mathbf{I}_M + \frac{1}{\rho_{\text{BS}}\tau^2}\sum_{k=1}^{K}\bar{\mathbf{P}}_k^{\text{T}}\mathbf{C}_{\mathbf{d}}^{\text{ce}}\bar{\mathbf{P}}_k^*\right) \tag{35}$$

(with $\text{Diag}\big(\mathbb{E}[\mathbf{W}\mathbf{W}^{\text{H}}]\big) = \frac{1}{M}\mathbf{I}_M$), and we can thus write $\mathbf{G}_b^{\text{dl}}$ as a scaled identity matrix, i.e., $\mathbf{G}_b^{\text{dl}} = G_b^{\text{dl}}\mathbf{I}_M$ (see (7)). Then, we can express each expectation term in (18) as

$$\mathbb{E}[\mathbf{h}_k^{\text{H}}\mathbf{G}_b^{\text{dl}}\mathbf{w}_k] = \frac{1}{\sqrt{\delta}}G_b^{\text{ce}}G_b^{\text{dl}}M, \tag{36}$$

$$\mathbb{E}\left[|\mathbf{h}_k^{\text{H}}\mathbf{G}_b^{\text{dl}}\mathbf{w}_i|^2\right] = \frac{1}{\delta}\left(\left(1+\frac{1}{\rho_{\text{BS}}\tau}\right)(G_b^{\text{ce}})^2(G_b^{\text{dl}})^2M + \frac{1}{\rho_{\text{BS}}\tau^2}(G_b^{\text{dl}})^2A_i\right) \qquad (i\neq k) \tag{37}$$

[2]We use $\text{unvec}(\cdot)$ as the inverse of the $\text{vec}(\cdot)$ operator.

$$\gamma_k^{\text{ul}} = \frac{\rho_{\text{BS}}(G_b^{\text{ce}})^2(G_b^{\text{ul}})^2M^2}{(\rho_{\text{BS}}K+1)\big(1+\frac{1}{\rho_{\text{BS}}\tau}\big)(G_b^{\text{ce}})^2(G_b^{\text{ul}})^2M + (\rho_{\text{BS}}K+1)\frac{1}{\rho_{\text{BS}}\tau^2}(G_b^{\text{ul}})^2A_k + \big(1+\frac{1}{\rho_{\text{BS}}\tau}\big)(G_b^{\text{ce}})^2\text{tr}(\mathbf{C}_{\mathbf{d}}^{\text{ul}}) + \frac{1}{\rho_{\text{BS}}\tau^2}B_k}, \tag{33}$$

$$\gamma_k^{\text{dl}} = \frac{\rho_{\text{UE}}(G_b^{\text{ce}})^2(G_b^{\text{dl}})^2M^2}{\rho_{\text{UE}}K\big(1+\frac{1}{\rho_{\text{BS}}\tau}\big)(G_b^{\text{ce}})^2(G_b^{\text{dl}})^2M + \rho_{\text{UE}}\frac{1}{\rho_{\text{BS}}\tau^2}(G_b^{\text{dl}})^2\sum_{i=1}^K A_i + \delta\rho_{\text{UE}}\text{tr}(\mathbf{C}_{\mathbf{d}}^{\text{dl}}) + \delta} \tag{40}$$

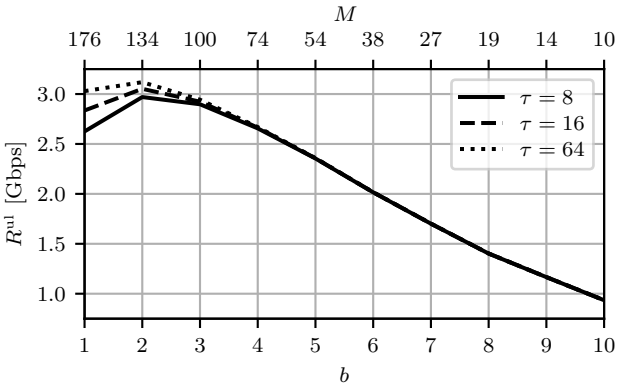



Fig. 1: Uplink ergodic achievable sum rate versus the resolution of the ADCs with $B = 0.1$ GHz and $P_{\text{HW}} = 10\big(P_{\text{RF}} + 2P_{\text{ADC}}(10 \text{ bits}, 0.1 \text{ GHz})\big)$.

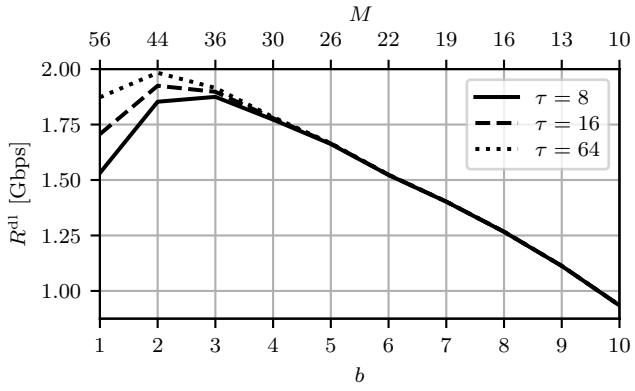



Fig. 2: Downlink ergodic achievable sum rate versus the resolution of the DACs with $B = 0.1$ GHz and $P_{\text{HW}} = 10\big(P_{\text{RF}} + 2P_{\text{DAC}}(10 \text{ bits}, 0.1 \text{ GHz})\big)$.

and

$$\mathbb{E}\big[|\mathbf{h}_k^{\text{H}}\mathbf{G}_b^{\text{dl}}\mathbf{w}_k|^2\big] = \frac{1}{\delta}\Bigg(\bigg(M+1+\frac{1}{\rho_{\text{BS}}\tau}\bigg)(G_b^{\text{ce}})^2(G_b^{\text{dl}})^2M + \frac{1}{\rho_{\text{BS}}\tau^2}(G_b^{\text{dl}})^2A_i\Bigg), \tag{38}$$

$$\mathbb{E}[\mathbf{h}_k^{\text{H}}\mathbf{C}_{\mathbf{d}}^{\text{dl}}\mathbf{h}_k] = \text{tr}(\mathbf{C}_{\mathbf{d}}^{\text{dl}}). \tag{39}$$

This allows us to write the effective downlink SINDR with imperfect CSI and MRT as in (40) at the top of the page. Finally, we substitute the resulting expression into (19) to obtain a downlink ergodic achievable sum rate.

## V. Numerical Results and Discussion

We now evaluate the performance trade-off between the number of BS antennas, the resolution of the ADCs/DACs, and the bandwidth under hardware power consumption constraints. To do so, we use the uplink and downlink ergodic achievable sum rate expressions derived in Sections IV-B and IV-C, respectively, which consider imperfect CSI and MRC/MRT. Throughout we disregard the channel estimation overhead, which can be readily included by multiplying (11) and (19) by $\big(1-\frac{\tau}{T}\big)$ for a given coherence time $T$ (see, e.g., [9, Sec. 4]).

In the numerical results that follow, we consider $K = 8$ UEs and assume that $\mathbf{P}$ comprises the first $K$ columns of the $\tau$-dimensional discrete Fourier transform matrix. The SNR at the BS in the uplink (resp. at the UEs in the downlink) includes the transmit power at the UEs (resp. at the BS), the pathloss, and the thermal noise power. Hence, we have $\rho_{\text{BS}} = \frac{P_{\text{UE}}\beta_k}{\sigma_{\text{BS}}^2}$, where $P_{\text{UE}} = 20$ dBm is the transmit power at the UEs, $\beta_k$ represents the pathloss between the BS and UE $k$, and $\sigma_{\text{BS}}^2$ is the thermal noise power at the BS. Likewise, we have $\rho_{\text{UE}} = \frac{P_{\text{BS}}\beta_k}{\sigma_{\text{UE}}^2}$, where $P_{\text{BS}} = 30$ dBm is the transmit power at the BS and $\sigma_{\text{UE}}^2$ is the thermal noise power at the UEs. The pathloss for each UE $k$ is expressed as $\beta_k = d_k^{-\alpha}$, where $d_k$ is the distance between the BS and UE $k$ and $\alpha = 4$ is the pathloss exponent. For simplicity, all the UEs are placed at the same distance from the BS, with $\{d_k = 100 \text{ m}\}_{k=1}^K$. The thermal noise power at both the BS and the UEs is given by $\sigma_{\text{BS}}^2 = \sigma_{\text{UE}}^2 = \nu - 174 + 10\log_{10}(B)$ [dBm], where $\nu = 13$ dB is the noise figure; note that a larger bandwidth corresponds to an increased thermal noise power. Lastly, the covariance matrices of the quantization distortion $\mathbf{C}_{\mathbf{d}}^{\text{ul}}$, $\mathbf{C}_{\mathbf{d}}^{\text{dl}}$, and $\mathbf{C}_{\mathbf{d}}^{\text{ce}}$ are computed via Monte Carlo simulations with $10^6$ independent channel realizations.

### A. ADCs/DACs Setting

We assume that the quantization thresholds and labels for each $b$-bit ADC/DAC are computed following the same procedure as in [2]. For the ADCs, the Lloyd-Max algorithm is applied to the Gaussian distribution with variance $\rho_{\text{UE}}K + 1$, obtaining the thresholds $\mathcal{T}_b$ and a set of labels $\tilde{\mathcal{L}}_b \triangleq \{\tilde{\ell}_n\}_{n=0}^{2^b-1}$.[3] Then, we obtain the labels $\mathcal{L}_b = \zeta\tilde{\mathcal{L}}_b$, with

$$\zeta \triangleq \eta^{\text{ul}}\Bigg(\sum_{n=0}^{2^b-1}\tilde{\ell}_n^2\Big(\Phi\big(\eta^{-1}|t_{n+1}|\big) - \Phi\big(\eta^{-1}|t_n|\big)\Big)\Bigg)^{-\frac{1}{2}} \tag{41}$$

with $\eta^{\text{ul}} \triangleq \sqrt{\frac{\rho_{\text{UE}}K+1}{2}}$ and where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard Gaussian distribution. Note that such a rescaling of the labels ensures that the variance of each entry of $\mathbf{r}^{\text{ul}}$ in (3) and $\mathbf{r}^{\text{ce}}$ in (21) is $\rho_{\text{UE}}K + 1$. We follow a similar procedure for the DACs, where the Lloyd-Max algorithm is applied to the Gaussian distribution with variance $\frac{1}{M}$ and $\eta^{\text{ul}}$ in (41) is replaced by $\eta^{\text{dl}} \triangleq \frac{1}{\sqrt{2M}}$.

To determine the number of BS antennas as a function of the resolution of the ADCs/DACs under a hardware power consumption constraint (see (1)), we use the following power consumption model for the ADCs and DACs [10]:

$$P_{\text{ADC}}(b, B) = 3V_{\text{dd}}^2L_{\text{min}}(2B + f_{\text{cor}})10^{0.1525b-4.838}, \tag{42}$$

$$P_{\text{DAC}}(b, B) = \frac{1}{2}V_{\text{dd}}I_0(2^b - 1) + bC_{\text{p}}(2B + f_{\text{cor}})V_{\text{dd}}^2 \tag{43}$$

[3] This approach minimizes the mean squared error between the non-quantized and quantized signal.

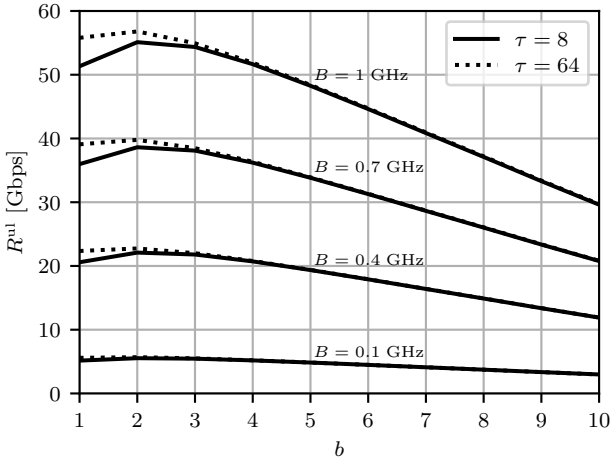



Fig. 3: Uplink ergodic achievable sum rate versus the resolution of the ADCs with $B \in \{0.1, 0.4, 0.7, 1\}$ GHz and $P_{\text{HW}} = 10\big(P_{\text{RF}} + 2P_{\text{ADC}}(10 \text{ bits}, 1 \text{ GHz})\big)$.

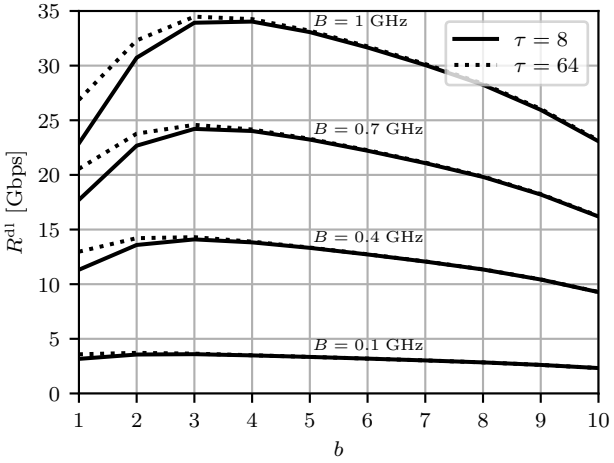



Fig. 4: Downlink ergodic achievable sum rate versus the resolution of the DACs with $B \in \{0.1, 0.4, 0.7, 1\}$ GHz and $P_{\text{HW}} = 10\big(P_{\text{RF}} + 2P_{\text{DAC}}(10 \text{ bits}, 1 \text{ GHz})\big)$.

where $V_{\text{dd}} = 3$ V, $L_{\text{min}} = 0.5$ $\mu$m, $f_{\text{cor}} = 1$ MHz, $I_0 = 10$ $\mu$A, and $C_{\text{p}} = 1$ pF are chosen as in [10, App. E].

### B. Performance Evaluation

Let us assume $P_{\text{RF}} = 40$ mW in the uplink (as in [11]) and $P_{\text{RF}} = 10$ mW in the downlink. In Fig. 1 and Fig. 2, we plot the uplink and downlink ergodic achievable sum rates versus the resolution of the ADCs and DACs, respectively, with $B = 0.1$ GHz and where $P_{\text{HW}}$ is set such that it can supply 10 BS antennas with 10-bit ADCs/DACs at such bandwidth; note that the value of $M$ corresponding to each value of $b$ is detailed at the top of the figures. In general, lowering the resolution of the ADCs/DACs under a hardware power consumption constraint has a two-fold effect: on the one hand, the reduced power consumption allows one to increase the number of BS antennas, which enhances the array gain; on the other hand, the quantization distortion becomes more significant during both the channel estimation and the uplink/downlink data transmission. In the uplink, the ergodic performance is maximized with 2-bit ADCs for all the considered pilot lengths. Moreover, in the downlink, the best ergodic performance is achieved with 3-bit DACs for $\tau = 8$ and with 2-bit DACs for $\tau \geq 16$. In both the uplink and the downlink, the impact of imperfect CSI is more evident for lower resolutions and the extreme case of 1-bit ADCs is nearly optimal in the uplink when accurate channel estimates are available. In Fig. 3 and Fig. 4, we plot the uplink and downlink ergodic achievable sum rates versus the resolution of the ADCs and DACs, respectively, with different values of $B$ and where $P_{\text{HW}}$ is set such that it can supply 10 BS antennas with 10-bit ADCs/DACs at $B = 1$ GHz. In the uplink, the ergodic performance is maximized with 2-bit ADCs for all the considered bandwidths. In the downlink, the optimal number of resolution bits increases with the bandwidth.

## VI. Conclusions

Considering a fully digital massive MIMO architecture with low-resolution ADCs/DACs at the BS, we derive tractable expressions for uplink and downlink ergodic achievable sum rates with imperfect CSI. These are used to analyzes the performance trade-off between the number of BS antennas, the resolution of the ADCs/DACs, and the bandwidth under a hardware power consumption constraint, where the relationship between these design parameters follows from a realistic model for the power consumption of the ADCs/DACs and the RF chains. In the evaluated scenarios, the most energy-efficient configurations feature 2 resolution bits in the uplink and 2 to 3 resolution bits in the downlink. Future work will consider correlated channels as well as power consumption models reflecting the latest advancements in ADC/DAC technologies.

## References


[1] N. Rajatheva, I. Atzeni, E. Björnson *et al.*, "White paper on broadband connectivity in 6G," Jun. 2020. [Online]. Available: http://jultika.oulu.fi/files/isbn9789526226798.pdf

[2] S. Jacobsson, G. Durisi, M. Coldrey, U. Gustavsson, and C. Studer, "Throughput analysis of massive MIMO uplink with low-resolution ADCs," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 1304–1309, Jun. 2017.

[3] Y. Li, C. Tao, G. Seco-Granados, A. Mezghani, A. L. Swindlehurst, and L. Liu, "Channel estimation and performance analysis of one-bit massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 65, no. 15, pp. 4075–4089, Aug. 2017.

[4] I. Atzeni and A. Tölli, "Channel estimation and data detection analysis of massive MIMO with 1-bit ADCs," *IEEE Trans. Wireless Commun.*, 2022.

[5] ——, "Uplink data detection analysis of 1-bit quantized massive MIMO," in *Proc. IEEE Int. Workshop Signal Process. Adv. in Wireless Commun. (SPAWC)*, Lucca, Italy, Sep. 2021.

[6] J. J. Bussgang, "Crosscorrelation functions of amplitude-distorted Gaussian signals," Res. Lab. Electron., Massachusetts Inst. Technol., Tech. Rep. 216, Mar. 1952.

[7] S. Jacobsson, Y. Ettefagh, G. Durisi, and C. Studer, "All-digital massive MIMO with a fronthaul constraint," in *IEEE Statistical Signal Process. Workshop (SSP)*, Freiburg, Germany, Jun. 2018.

[8] Y. Ettefagh, S. Jacobsson, A. Hu, G. Durisi, and C. Studer, "All-digital massive MIMO uplink and downlink rates under a fronthaul constraint," in *Proc. Asilomar Conf. Signals, Syst., and Comput. (ASILOMAR)*, Pacific Grove, CA, USA, Nov. 2019.

[9] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Foundations and Trends® Signal Process.*, vol. 11, no. 3–4, pp. 154–655, Nov. 2017.

[10] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2349–2360, Sep. 2005.

[11] S. Joshi and M. Juntti, "RF-chain ADC resolution trade-off in MIMO hybrid architecture," in *Proc. Asilomar Conf. Signals, Syst., and Comput. (ASILOMAR)*, Pacific Grove, CA, USA, Nov. 2018.