Hardware Design and Optimal ADC Resolution for Uplink Massive MIMO Systems

Daniel Verenzuela, Emil Björnson and Michail Matthaiou

Hardware Design and Optimal ADC Resolution for Uplink Massive MIMO Systems

Daniel Verenzuela*, Emil Björnson*, and Michail Matthaiou†
*Department of Electrical Engineering (ISY), Linköping University, Linköping, Sweden.
†School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast, U.K.

***Abstract*—This work focuses on the hardware design for the efficient operation of Massive multiple-input multiple-output (MIMO) systems. A closed-form uplink achievable data rate expression is derived considering imperfect channel state information (CSI) and hardware impairments. We formulate an optimization problem to maximize the sum data rate subject to a constraint on the total power consumption. A general power consumption model accounting for the level of hardware impairments is utilized. The optimization variables are the number of base station (BS) antennas and the level of impairments per BS antenna. The resolution of the analog-to-digital converter (ADC) is a primary source of such impairments. The results show the trade-off between the number of BS antennas and the level of hardware impairments, which is important for practical hardware design. Moreover, the maximum power consumption can be tuned to achieve maximum energy efficiency (EE). Numerical results suggest that the optimal level of hardware impairments yields ADCs of 4 to 5 quantization bits.**

I. Introduction

The continuous increase in data traffic and power consumption of wireless networks have become important societal concerns; thus, major efforts are being made to improve the EE of communication systems [1], [2]. Massive MIMO has been proposed as a key enabler for the next generation wireless networks to improve spectral efficiency (SE) as well as EE [3], [4]. This technology is based on the use of a large number of BS antennas to serve tens of user equipments (UEs) simultaneously by exploiting spatial multiplexing techniques.

Several studies [5]–[7] indicate that the large number of BS antennas allow the use of low precision hardware at the BS since distortions average out in the receiver processing. However, as the number of BS antennas becomes higher, the total power consumption (including signal processing and antenna hardware) increases [8]. Since low precision hardware consumes less power, a trade-off must be found between hardware quality and number of BS antennas when designing an energy efficient Massive MIMO system. Despite the importance of this problem, the related literature is scarce.

An important source of hardware impairments and circuit power consumption is the ADC [9]. Each antenna commonly requires two separate ADCs for the in-phase and quadrature branches. Having low resolution ADCs reduces the power consumption of the hardware while degrading performance by introducing more quantization noise. In [10], [11] numerical results conclude that Massive MIMO can operate at low bit and packet error rates with 2-4 quantization bits, while having less than 1 dB of signal-to-noise ratio (SNR) performance loss. The works [12], [13] show that Massive MIMO can even operate with 1 bit ADCs. However, it is not clear yet whether operating with such low hardware precision is preferable in terms of data rate and power consumption performance.

The design of a wireless network entails weighing different performance metrics that are often mutually conflicting. In particular, we encounter this situation when our goal is to minimize the power consumption and maximize the data rates since to increase the data rates we need more power and vice versa. This trade-off has been studied in [14], [15] for the single antenna and millimeter wave MIMO systems respectively. In this work, we focus on the problem of selecting the number of BS antennas and level of hardware impairments to maximize the sum data rate while maintaining the power consumption below a threshold. To this end, we first derive an achievable rate expression per UE for the uplink communication of a single-cell Massive MIMO system. Second, we present a general power consumption model considering the quantization distortions of ADCs as the main source of hardware impairments. Third, we find the optimal number of BS antennas and level of hardware impairments that achieve maximum sum data rate subject to a total power consumption constraint. Finally, we present the EE in terms of the maximum power consumption allowed.

II. System Model

We consider the uplink communication of a single-cell Massive MIMO system where the BS is affected by hardware impairments and the UEs are assumed to have perfect hardware.[1] The system operates over a bandwidth B_{w} where K single-antenna UEs communicate with a BS that deploys M antennas. The channel is modeled as block-fading where each block has a coherence bandwidth B_c [Hz] and a coherence time T_c [s] that result in $S = B_c T_c$ symbols per block. In each block the channel is considered constant while having independent realizations in different blocks. We assume Rayleigh fading where the channel matrix between the UEs and the BS is denoted as $\mathbf{H} = [\mathbf{h}_1, ..., \mathbf{h}_K] \in \mathbb{C}^{M \times K}$. The channel between all BS antennas and the UE k is distributed as $\mathbf{h}_k \sim \mathcal{CN}(\mathbf{0}, \beta_k \mathbf{I}_M)$. The parameter β_k represents the large-scale fading component (path loss and shadow fading) of an arbitrary UE k. The BS needs to learn the channel in each coherence block to perform multi-user detection. For this purpose we assume that the UEs send orthogonal pilot sequences of length $B \geq K$ symbols to the BS to enable channel estimation, thus $S - B$ symbols are left for uplink data transmission.

This paper has received funding from ELLIIT, CENIIT and the Swedish Foundation for Strategic Research.

[1] This assumption is made to pursue a tractable analysis of hardware impairments at the BS; however, it is important to mention that in real implementations the hardware impairments at the UEs are not negligible [5].

$$\mathrm{SINR}_k = \frac{\left(1-\epsilon^2\right)^2 M}{\epsilon^2(1-\epsilon^2)+\frac{2\epsilon^2\left(1-\epsilon^2\right)\left(\frac{\epsilon^2}{B}\sum_{i=1}^{K}\frac{\beta_i^2 p_i^2}{\beta_k^2 p_k^2}+2\left(1-\epsilon^2\right)\right)}{B\left((1-\epsilon^2)+\frac{\epsilon^2}{B}\sum_{i=1}^{K}\frac{\beta_i p_i}{\beta_k p_k}+\frac{\sigma^2}{B\beta_k p_k}\right)^2}+\left((1-\epsilon^2)+\frac{\epsilon^2}{B}\sum_{i=1}^{K}\frac{\beta_i p_i}{\beta_k p_k}+\frac{\sigma^2}{B\beta_k p_k}\right)\left(\sum_{i=1}^{K}\frac{\beta_i p_i}{\beta_k p_k}+\frac{\sigma^2}{p_k\beta_k}\right)} \tag{8}$$

The transmission power of the UEs over the complete bandwidth is p_k. We assume that the communication is affected by the noise $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \frac{\sigma^2}{B_\mathrm{w}}\mathbf{I}_M)$, where σ^2 is the noise power over the bandwidth B_w. Now let $\mathbf{x} = [x_1, ..., x_K]^T$ be the transmitted symbols from the K UEs to the BS which have zero mean and are mutually uncorrelated with covariance matrix $\mathbb{E}\left\{\mathbf{x}\mathbf{x}^H\right\} = \mathbf{I}_K$ where $\mathbb{E}\{\cdot\}$ and $(\cdot)^H$ denote expected value and conjugate transpose operators, respectively. Then, the $M \times 1$ received signal at the BS is

$$\mathbf{y} = \sqrt{1-\epsilon^2}\left(\sum_{k=1}^{K}\sqrt{\frac{p_k}{B_\mathrm{w}}}\mathbf{h}_k x_k + \mathbf{n}\right) + \mathbf{e}. \tag{1}$$

The influence of hardware impairments at the BS is modeled by a reduction of the received signal energy at each BS antenna by $1-\epsilon^2$ which is turned into distortion noise $\mathbf{e} \in \mathbb{C}^{M\times 1}$ and added back to the signal [5]. As a result $\mathbf{e}|\mathbf{H} \sim \mathcal{CN}\big(\mathbf{0}, \frac{\epsilon^2}{B_\mathrm{w}}\big(\sum_{k=1}^{K} p_k \mathbf{D}_{h_k} + \sigma^2\mathbf{I}_M\big)\big)$ where the notation $\mathbf{e}|\mathbf{H}\sim$ denotes the distribution of $\mathbf{e}$ conditioned on the channel realizations $\mathbf{H}$.[2] The term $\mathbf{D}_{h_k} = \mathrm{diag}\big\{|h_{1k}|^2, ..., |h_{Mk}|^2\big\}\ \forall k \in \{1, ..., K\}$, where h_{ij} is the $(i,\ j)$ element of $\mathbf{H}$. The parameter $0 < \epsilon < 1$ represents the level of hardware impairments, where $\epsilon = 0$ corresponds to ideal hardware, and can be related to the quantization distortion introduced by ADCs. Notice that we assume perfect automatic gain control (AGC) and therefore the same ϵ value for all BS antennas. We establish the following relationship between the number of quantization bits b and the hardware impairment parameter: $\epsilon = \zeta 2^{-b}$, where $1 < \zeta < 2$ is a constant related to the saturation level of the ADC [6], [16].[3]

III. Achievable rate with imperfect CSI

To analyze the system performance we derive an achievable rate for the UEs based on estimated CSI. The pilot sequences are grouped into a matrix $\mathbf{V} = [\mathbf{v}_1, ..., \mathbf{v}_K] \in \mathbb{C}^{B\times K}$ with elements having $|\mathrm{v}_{ik}| = 1$ and the columns corresponding to each distinct orthogonal pilot sequence used by the K UEs; that is, $\mathbf{V}^H\mathbf{V} = B\mathbf{I}_B$. By using an equivalent model to (1), the received signal from all pilot sequences sent by all UEs is

$$\mathbf{Y}_p = \sqrt{1-\epsilon^2}\left(\mathbf{H}\mathbf{D}_p^{\frac{1}{2}}\mathbf{V}^T + \mathbf{N}\right) + \boldsymbol{\Upsilon} \in \mathbb{C}^{M\times B} \tag{2}$$

where $\mathbf{D}_p = \mathrm{diag}\left\{\frac{p_1}{B_\mathrm{w}}, ..., \frac{p_K}{B_\mathrm{w}}\right\}$ and $p_k\frac{B}{S}$ is the total energy of the pilot sequence transmitted by UE k.[4] The noise is represented by $\mathbf{N} \in \mathbb{C}^{M\times B}$ with i.i.d elements distributed as $n_{ij} \sim \mathcal{CN}(0, \frac{\sigma^2}{B_\mathrm{w}})$. The term $\boldsymbol{\Upsilon} \in \mathbb{C}^{M\times B}$ represents the distortion noise from hardware impairments which has independent elements $\upsilon_{ij}|\mathbf{H} \sim \mathcal{CN}\left(0, \frac{\epsilon^2}{B_\mathrm{w}}\left(\sum_{k=1}^{K}|h_{ik}|^2 p_k + \sigma^2\right)\right)$. The following lemma gives the linear minimum mean-squared error (LMMSE) estimate of the channel $\mathbf{H}$.

Lemma 1. *The LMMSE estimate of the channel $\mathbf{H}$, from the received signal $\mathbf{Y}_p$, between all UEs and all BS antennas is*

$$\hat{\mathbf{H}} = \frac{1}{B}\mathbf{Y}_p\mathbf{V}^*\mathbf{D}_p^{-\frac{1}{2}}\mathrm{diag}\left\{c_1, ..., c_K\right\} \tag{3}$$

where

$$c_k = \frac{\sqrt{1-\epsilon^2}B\beta_k p_k}{\left(1-\epsilon^2\right)B\beta_k p_k + \epsilon^2\sum_{i=1}^{K}\beta_i p_i + \sigma^2},\ \forall k \in \{1, ..., K\}. \tag{4}$$

The estimation error is defined as $\boldsymbol{\mathcal{E}} = \mathbf{H} - \hat{\mathbf{H}}$. The columns of $\boldsymbol{\mathcal{E}} = [\boldsymbol{\varepsilon}_1, ..., \boldsymbol{\varepsilon}_K]$ and $\hat{\mathbf{H}} = \left[\hat{\mathbf{h}}_1, ..., \hat{\mathbf{h}}_K\right]$ are uncorrelated random vectors with zero mean and covariance matrices

$$\mathbf{D}_{\boldsymbol{\varepsilon}_k} = \frac{\beta_k\left(\epsilon^2\sum_{i=1}^{K}\beta_i p_i + \sigma^2\right)}{\left(1-\epsilon^2\right)B\beta_k p_k + \epsilon^2\sum_{i=1}^{K}\beta_i p_i + \sigma^2}\mathbf{I}_M, \tag{5}$$

$$\mathbf{D}_{\hat{\mathbf{h}}_k} = \frac{\left(1-\epsilon^2\right)B\beta_k^2 p_k}{\left(1-\epsilon^2\right)B\beta_k p_k + \epsilon^2\sum_{i=1}^{K}\beta_i p_i + \sigma^2}\mathbf{I}_M, \tag{6}$$

respectively.

Proof: It follows from applying standard LMMSE estimation [17, Ch. 12] to the problem at hand. ■

Notice that this result is closely related to the LMMSE estimator derived in [6] since both models consider the presence of distortion noise that is dependent on the channel realizations; however, in our model we consider the reduction factor $1-\epsilon^2$ of the received signal. Once we have the channel estimates, we can combine the received signals to obtain separate data streams from each UE. In Massive MIMO, the use of large number of BS antennas creates asymptotically favorable propagation conditions where the channel directions $\mathbf{h}_k/\|\mathbf{h}_k\|$ of different UEs become mutually orthogonal as M increases. This phenomenon makes the use of linear detection techniques asymptotically optimal as M grows large [3]. Thus, we consider simple maximum ratio combining (MRC) to minimize the computational complexity and reduce the circuit power consumption. As a result the processed received signal is given by $\mathbf{r} = \hat{\mathbf{H}}^H\mathbf{y}$. Then, an achievable rate for the performance of each UE is given in the following theorem.

Theorem 1. *For MRC, a lower bound on the ergodic capacity of an arbitrary UE k is*

$$\mathrm{R}(\mathrm{SINR}_k) = B_\mathrm{w}\left(1-\frac{B}{S}\right)\log_2\left(1+\mathrm{SINR}_k\right)\ \textit{[bit/s]} \tag{7}$$

where SINR_k is the SINR of UE k given in (8) *at the top of the page.* [5]

Proof: This theorem follows from finding a lower bound on the mutual information with imperfect CSI at the receiver

[2] It is worth mentioning that this is a simplified model for the hardware impairments that allows a tractable analysis. However, in [7] it is shown that it is a good approximation to perform analysis in Massive MIMO systems.

[3] The saturation level of the ADC is the maximum level of signal amplitude that can be quantized by the ADC.

[4] Notice that in our analysis we assume the same transmission energy for pilot and data symbols as $\frac{p_k}{B_\mathrm{w}}$.

[5] The proof of this result does not require the distortions to be circularly symmetric complex Gaussian when conditioned on the channel. The distortions can have any distribution that has circular symmetry and shares the same first, second and fourth order moments when conditioned on the channel.

as in [18] and only using the knowledge of the average effective channel $\mathbb{E}\{\|\hat{\mathbf{h}}_k\|^2\}$ while treating the deviation from the average received channel as worst-case Gaussian noise with variance $\mathbb{E}\{\|\hat{\mathbf{h}}_k\|^4\} - \mathbb{E}\{\|\hat{\mathbf{h}}_k\|^2\}^2$ as done in [6]. ■

Notice that since the added distortion of hardware impairments is not Gaussian, to the best of our knowledge, the lower bound in (7) is the tightest capacity bound that exists. The expression (8) can be further simplified by considering statistical channel inversion power allocation and using K symbols for pilot transmission which is summarized as follows.

Corollary 1. *If $p_k = \frac{\rho}{\beta_k}$ and $B = K$, then the achievable rate per UE is expressed as $\bar{\mathrm{R}} = \mathrm{R}(\mathrm{SINR})$ [bit/s/UE] where*

$$\mathrm{SINR} = \frac{\left(1-\epsilon^2\right)^2 M}{\epsilon^2(1-\epsilon^2) + \frac{2\epsilon^2\left(1-\epsilon^2\right)\left(2-\epsilon^2\right)}{K\left(1+\frac{\sigma^2}{K\rho}\right)^2} + K\left(1+\frac{\sigma^2}{K\rho}\right)^2} \tag{9}$$

and the term $\rho > 0$ is a design parameter to control the average transmit power per UE.

In the SINR expression (8) and (9), the numerator is related to the coherent gain and the denominator accounts for the effects of inter-user interference, channel estimation errors, hardware impairment distortions, channel gain uncertainty (which results from the bounding technique used in Theorem 1) and noise. Notice that in (9) the SINR is user-independent providing a general average data rate per UE which will be used to perform the optimization of the hardware impairments.

A. Power consumption model

The ADCs are responsible for a large portion of the power consumption and distortion noise; thus, we consider them as the main source of power consumption that is related to the hardware impairment parameter ϵ. Another important consideration is that the computational complexity of baseband operations is proportional to the word length which equals the number of quantization bits. This means that the amount of power used for signal processing is also dependent on the number of quantization bits used by the ADCs. We define the power consumption affected by the hardware impairments as

$$\tilde{P}(\epsilon) = 2MP_{\mathrm{ADC}} + P_{\mathrm{CE}} + P_{\mathrm{LP}}. \tag{10}$$

In (10), P_{ADC} accounts for the ADC power consumption at each BS antenna, which we model as

$$P_{\mathrm{ADC}} = \mathcal{D}_1 2^{\varphi b} = \mathcal{D}_1 \zeta^{\varphi} \epsilon^{-\varphi} \tag{11}$$

where $\mathcal{D}_1 \geq 0$ and $\varphi \geq 0$ are set constants. Typical values are $\mathcal{D}_1 \in [10^{-8}, 10^{-7}]$ [W/conversion step], $\varphi \in [1, 2]$.[6] This general model is chosen to encompass different architectures and implementations of ADCs found in the literature [9], [14], [19], [20]. The terms P_{CE} and P_{LP} represent the power consumed by channel estimation and linear processing, respectively, which can be modeled as in [4]. We calculate the number of bit operations required for channel estimation and linear processing tasks, from numerical linear algebra computations, and divide it by the computational efficiency (in [bit operations/Joule]) of current microprocessors. Channel estimation involves a matrix multiplication operation between the received signal $\mathbf{Y}_p$ and the pilots of all UEs $\mathbf{V}$, consuming

$$P_{\mathrm{CE}} = \frac{B_{\mathrm{W}}}{L}\frac{B}{S} 2MKb = \frac{B_{\mathrm{W}}}{L}\frac{B}{S} 2MK \log_2\left(\frac{\zeta}{\epsilon}\right) \tag{12}$$

where L is the computational efficiency of the hardware at the BS expressed in [bit operations/Joule]. For linear processing the BS combines the received signal as $\mathbf{r} = \hat{\mathbf{H}}^H \mathbf{y}$, consuming

$$P_{\mathrm{LP}} = \frac{B_{\mathrm{W}}}{L}\left(1 - \frac{B}{S}\right) 2MK \log_2\left(\frac{\zeta}{\epsilon}\right) \tag{13}$$

In general the elements of $\hat{\mathbf{H}}$ are represented with a higher bit length that $\mathbf{y}$. However, notice that since the elements of $\mathbf{V}$ have unit modulus we can represent them with the same bit length as the elements of $\mathbf{Y}_p$, for instance we could use a Hadamard matrix with entries $+1$ or -1 [21]. Thus, the matrix multiplication $\mathbf{V}^T\mathbf{Y}_p^H\mathbf{y}$ (done in $\hat{\mathbf{H}}^H\mathbf{y}$) involves elements with the same bit length. The rest of the terms in $\hat{\mathbf{H}}$ are scaling factors (see (3)) which do not to affect the SINR calculations. The total power consumption is obtained by combining (10) with the model found in [4], resulting in

$$P(\epsilon, M) = \frac{1}{\eta}\sum_{i=1}^{K} p_i + \mathcal{C}_0 + \mathcal{C}_1 K + \mathcal{D}_0 M + \tilde{P}(\epsilon) \tag{14}$$

$$= \tilde{\mathcal{C}}_0 + \mathcal{D}_0 M + \mathcal{D}_1 2M\zeta^{\varphi}\epsilon^{-\varphi} + \frac{B_{\mathrm{W}}}{L} 2MK \log_2\left(\frac{\zeta}{\epsilon}\right), \tag{15}$$

where

$$\tilde{\mathcal{C}_0} = \frac{1}{\eta}\sum_{i=1}^{K} p_i + \mathcal{C}_0 + \mathcal{C}_1 K\,. \tag{16}$$

The first term in (16) represents the transmission power consumption where η is the efficiency of the power amplifier. $\mathcal{C}_0 > 0$ accounts for the static power consumption (e.g., site cooling) and $\mathcal{C}_1 > 0$, $\mathcal{D}_0 > 0$ are the power consumption per UE and BS antenna hardware respectively (independent of ϵ).

IV. Optimal Hardware Impairments Trade-off

Our goal is to find the optimal level of hardware impairments and the number of BS antennas to maximize the sum data rate, denoted as $\mathrm{R}_{\mathrm{SUM}} = K\bar{\mathrm{R}}$ where $\bar{\mathrm{R}}$ is given in Corollary 1, subject to a total power consumption constraint. We define the following optimization problem:

$$\begin{array}{ll} \underset{\epsilon\in[0,1],\, M\in\mathbb{Z}_+}{\text{maximize}} & \mathrm{R}_{\mathrm{SUM}}(\epsilon,\, M), \\ \text{subject to} & P(\epsilon,\, M) \leq \gamma \end{array} \tag{17}$$

where γ is the maximum value of the power consumption and $\mathbb{Z}_+$ is the set of non-negative integers.

By using simple calculus we can verify that both the sum data rate $\mathrm{R}_{\mathrm{SUM}}(\epsilon,\, M)$ and the total power consumption $P(\epsilon,\, M)$ are monotonically decreasing functions of $\epsilon \in [0, 1]$. Moreover, $\mathrm{R}_{\mathrm{SUM}}(\epsilon,\, M)$ increases logarithmically with M whereas $P(\epsilon,\, M)$ increases linearly.

Lemma 2. *Assume that the optimization problem* (17) *is feasible, then the maximum $\mathrm{R}_{\mathrm{SUM}}$ is found when $P(\epsilon, M) = \gamma$, thus the optimal values of ϵ and M satisfy*

$$M(\epsilon) = \frac{\gamma - \tilde{\mathcal{C}_0}}{\mathcal{D}_0 + 2\mathcal{D}_1\zeta^{\varphi}\epsilon^{-\varphi} + \frac{B_{\mathrm{W}}}{L} 2K \log_2\left(\frac{\zeta}{\epsilon}\right)}. \tag{18}$$

[6]The choice of $\mathcal{D}_1$ and φ can be related to the figure-of-merit (FoM) of ADCs where smaller $\mathcal{D}_1$ are paired with larger φ and vice versa [19].

Table I. SIMULATION PARAMETERS

Parameter	Symbol	Value
System Bandwidth	B_W	20 MHz
Const. related to ADC saturation level	ζ	1.6
Consts. related to the ADC FoM	$\mathcal{D}_1$	10^{-7} W/conv-step
	φ	1
Power amplifier efficiency	η	0.39
Static power consumption	$\mathcal{C}_0$	10 W
Circuit power per active UE	$\mathcal{C}_1$	0.1 W
Circuit power per BS antenna (ind. of ϵ)	$\mathcal{D}_0$	0.1 W
Computational efficiency	L/B_W	20480 [bit op./W]
Noise power	σ^2	$2 \cdot 10^{-13}$ W

Proof: Suppose that $\{\epsilon^*, M^*\}$ is the optimal solution to (17) and that $P(\epsilon^*, M^*) < \gamma$. Since $P(\epsilon, M)$ is a decreasing function of ϵ, let us select a value $\epsilon_1 < \epsilon^*$ such that $P(\epsilon_1, M^*) = \gamma$. However since $\mathrm{R_{SUM}}(\epsilon, M)$ is a decreasing function of ϵ we have $\mathrm{R_{SUM}}(\epsilon_1, M^*) > \mathrm{R_{SUM}}(\epsilon^*, M^*)$. This is a contradiction to the presumption $P(\epsilon^*, M^*) < \gamma$, thus there can only be optimal solutions with $P(\epsilon^*, M^*) = \gamma$. ■

To solve the optimization problem in (17) we can conduct an exhaustive search algorithm over a finite set of M values and for each M we can solve (18) to obtain the corresponding ϵ. For this purpose we introduce the following lemma.

Lemma 3. *For a fixed value of M the corresponding ϵ that satisfies $P(\epsilon, M) = \gamma$ is $\epsilon = \zeta 2^{-b}$ where*

$$b = -\frac{W\left(\frac{\varphi \ln(2)\mathcal{D}_1 \exp\left(-\frac{\varphi \ln(2)(\tilde{\mathcal{C}}_0 + \mathcal{D}_0 M - \gamma)}{\frac{B_W}{L} 2MK}\right)}{\frac{B_W}{L} K}\right)}{\varphi \ln(2)} - \frac{\tilde{\mathcal{C}}_0 + \mathcal{D}_0 M - \gamma}{\frac{B_W}{L} 2MK} \tag{19}$$

and $W(\cdot)$ denotes the Lambert function, found in [22].

Notice that ϵ in Lemma 3 is a decreasing function of b and the optimal solution is found for M being an increasing function of ϵ (see (18)). Thus, the largest feasible M is found for $b = 1$ which gives us the interval $M \in \{1, 2, ..., \lfloor M(\zeta/2) \rfloor\}$ to conduct the exhaustive search over. Finally, once we have obtained the optimal ϵ^* and M^* for a given γ we can also define the EE $= \frac{\mathrm{R_{SUM}}(\epsilon^*, M^*)}{\gamma}$ and find the power consumption threshold γ that yields maximum EE. This allows us to obtain further insights on how to design Massive MIMO networks.

V. NUMERICAL RESULTS

This section provides numerical results drawn from Monte-Carlo (MC) simulations in order to validate the analytical results found in the previous sections. Table I gives the selected values of the main parameters found in [4], [19].[7]

Figure 1 shows the sum data rate $\mathrm{R_{SUM}}$ as a function of ϵ. We see that the theoretical result in Corollary 1 is indistinguishable of the MC simulations, validating our analysis. Figure 2 depicts a 3D graph of $\mathrm{R_{SUM}}$ in terms of ϵ and M. We see that $\mathrm{R_{SUM}}$ is monotonically decreasing with ϵ and increasing with M. For a maximum power consumption of $\gamma = \{22, 26, 30\}$ [W] the maximum rate is obtained at $M = \{87, 126, 164\}$ and $\epsilon = \{0.055, 0.056, 0.056\}$ respectively, resulting in $b \approx 4$ or 5 quantization bits. Figure 3 shows

[7]To model the large-scale fading we use the model $\beta = w^{-1} d^{-\alpha}$ where $w = 35$ dB accounts for propagation loss (e.g. wall penetration), $\alpha = 3.76$ is the path loss exponent and d is the distance between UEs and the BS based on uniformly distributed UEs in a circular cell of radius 250 m.

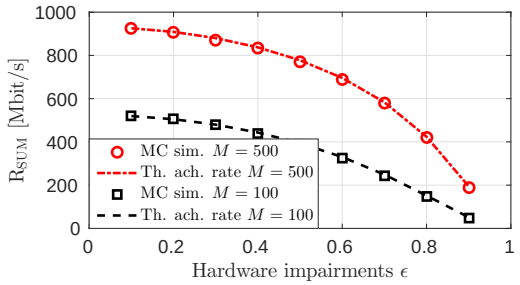

Figure 1. Sum data rate [Mbit/s] as a function of ϵ for $M = \{100, 500\}$, $K = 10$ and $\rho = 0.3162\sigma^2$ (SNR= -5 dB).

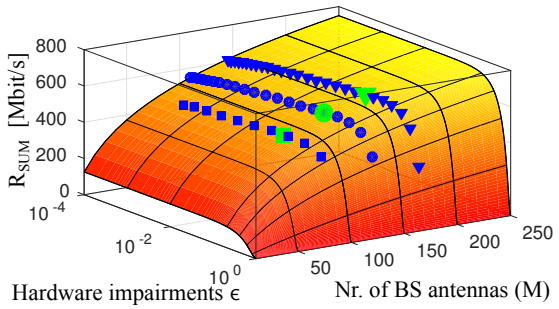

Figure 2. Sum data rate [Mbit/s] as a function of M and ϵ for $K = 10$ and $\rho = 0.3162\sigma^2$ (SNR= -5 dB). The blue squared, circle and triangle points represent the curves where $P(\epsilon, M) = \gamma$ for $\gamma = \{22, 26, 30\}$ [W] respectively, and the green points are the corresponding maximum data rates.

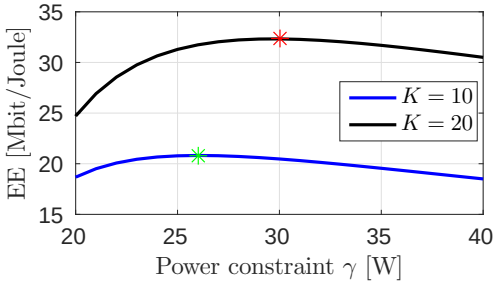

Figure 3. EE [Mbit/Joule] as a function of γ for $K = \{10, 20\}$ and $\rho = 0.3162\sigma^2$ (SNR $= -5$ dB). The star represents the maximum EE located at $\gamma^* = \{26, 30\}$ [W], which corresponds to $M^* = \{126, 164\}$ and $\epsilon^* = \{0.056, 0.0795\}$ respectively.

the EE as a function of the maximum power consumption γ. Notice that the EE has a unique maximum meaning that we should select γ appropriately to maximize EE.

VI. CONCLUSION

Massive MIMO systems operate with hundreds of BS antennas, thereby making the hardware power consumption a potential bottleneck. We have studied the tradeoff between having many BS antennas and high-quality hardware. This was formulated as a sum rate maximization problem under a total power constraint. This problem was solved optimally when using a new closed-form uplink achievable rate expression. The solution manifests the fundamental tradeoff between the number of BS antennas and level of hardware impairments. The numerical results suggest that hardware quality corresponding to 4 or 5 bit ADCs is the preferred choice, while lower resolutions will greatly reduce the data rates.

REFERENCES

[1] GreenTouch Green Meter Research Study, "Reducing the net energy consumption in communications networks by up to 90% by 2020," Tech. Rep., Jun. 2013.

[2] D. Feng, C. Jiang, L. Gubong, L. J. Cimini Jr., G. Feng, and G. Y. Li, "A survey of energy-efficient wireless communications," *IEEE J. Commun. Surveys & Tutorials*, vol. 15, no. 1, pp. 167–178, Feb. 2013.

[3] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[4] 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.

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

[6] E. Björnson, M. Matthaiou, and M. Debbah, "Massive MIMO with non-ideal arbitrary arrays: Hardware scaling laws and circuit-aware design," *IEEE Trans. Wireless Commun.*, vol. 14, no. 8, pp. 4353–4368, Aug. 2015.

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

[8] H. Yang and T. L. Marzetta, "Total energy efficiency of cellular large scale antenna system multiple access mobile networks," in *Proc. IEEE GreenCom*, Oct. 2013, pp. 27–32.

[9] A. Mezghani and J. A. Nossek, "Modeling and minimization of transceiver power consumption in wireless networks," in *Proc. IEEE/ITG WSA*, Feb. 2011, pp. 1–8.

[10] C. Desset and L. V. der Perre, "Validation of low-accuracy quantization in massive MIMO and constellation EVM analysis," in *Proc. EuCNC*, Jun. 2015, pp. 21–25.

[11] C. Studer and G. Durisi, "Quantized massive MU-MIMO-OFDM uplink," *in press IEEE Trans. Commun.*, 2016.

[12] C. Risi, D. Persson, and E. G. Larsson, "Massive MIMO with 1-bit ADC." [Online]. Available: http://arxiv.org/abs/1404.7736

[13] S. Jacobsson, G. Durisi, M. Coldrey, U. Gustavsson, and C. Studer, "Massive MIMO with low-resolution ADCs." [Online]. Available: http://arxiv.org/abs/1602.01139

[14] A. Mezghani, N. Damak, and J. A. Nossek, "Circuit aware design of power-efficient short range communication systems," in *Proc. IEEE ISWCS*, Sept. 2010, pp. 869–873.

[15] O. Orhan, E. Erkip, and S. Rangan, "Low power analog-to-digital conversion in millimeter wave systems: Impact of resolution and bandwidth on performance," in *Proc. IEEE ITA*, Feb 2015, pp. 191–198.

[16] A. Mezghani, M. Rouatbi, and J. A. Nossek, "An iterative receiver for quantized MIMO systems," in *Proc. IEEE MELECON*, Mar. 2012, pp. 1049–1052.

[17] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993.

[18] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 933–946, 2000.

[19] B. Murmann, "ADC performance survey 1997-2015," [Online]. Available: http://web.stanford.edu/~murmann/adcsurvey.html.

[20] B. Murmann, "Limits on ADC power dissipation," in *Analog Circuit Design*, A. H. M. Roermund and J. v. Huijsing, Eds. Springer, 2006.

[21] A. Hedayat and W. D. Wallis, "Hadamard matrices and their applications," *Ann. Statist.*, vol. 6, no. 6, pp. 1184–1238, Nov. 1978.

[22] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, "On the lambertW function," *Adv. in Comput. Math.*, vol. 5, no. 1, pp. 329–359, 1996.