



# LUND UNIVERSITY

## An Energy Efficiency Perspective on Massive MIMO Quantization

Sarajlic, Muris; Liu, Liang; Edfors, Ove

*Published in:*

Proceedings of the 50th Asilomar Conference on Signals, Systems and Computers

*DOI:*

[10.1109/ACSSC.2016.7869084](https://doi.org/10.1109/ACSSC.2016.7869084)

2017

*Document Version:*

Peer reviewed version (aka post-print)

[Link to publication](#)

*Citation for published version (APA):*

Sarajlic, M., Liu, L., & Edfors, O. (2017). An Energy Efficiency Perspective on Massive MIMO Quantization. In *Proceedings of the 50th Asilomar Conference on Signals, Systems and Computers* (pp. 473-478). IEEE - Institute of Electrical and Electronics Engineers Inc.. <https://doi.org/10.1109/ACSSC.2016.7869084>

*Total number of authors:*

3

### General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117  
221 00 Lund  
+46 46-222 00 00

# An Energy Efficiency Perspective on Massive MIMO Quantization

Muris Sarajlić, Liang Liu, and Ove Edfors

Department of Electrical and Information Technology, Lund University, Sweden

Email: (muris.sarajlic, liang.liu, ove.edfors)@eit.lth.se

**Abstract**—One of the basic aspects of Massive MIMO (MaMi) that is in the focus of current investigations is its potential of using low-cost and energy-efficient hardware. It is often claimed that MaMi will allow for using analog-to-digital converters (ADCs) with very low resolutions and that this will result in overall improvement of energy efficiency. In this contribution, we perform a parametric energy efficiency analysis of MaMi uplink for the entire base station receiver system with varying ADC resolutions. The analysis shows that, for a wide variety of system parameters, ADCs with intermediate bit resolutions (4 - 10 bits) are optimal in energy efficiency sense, and that using very low bit resolutions results in degradation of energy efficiency.

## I. INTRODUCTION

Massive MIMO [1] is an emerging wireless communication technique that promises substantial gains in both spectral and energy efficiency compared to traditional cellular systems. One feature of MaMi is its robustness to hardware imperfections [2], which implies that using MaMi will result in improved cost and energy efficiency. This fact has inspired a flurry of research activity focusing on diverse aspects of hardware impairments in MaMi and specific signaling and signal processing design tailored with the effects of the impairments in mind - all with the common goal of further improving the efficiency of the system.

A significant body of this research aims at analyzing and fine-tuning MaMi-based systems using ADCs with low resolution [3], [4]. The motivation for such investigations is well-established: the power consumption of ADCs grows at least linearly with sampling rate [15] and will prove to be a power consumption bottleneck in systems with very large bandwidths. Therefore, there is a huge interest in reducing the power consumption of ADCs by reducing the bit resolution as much as possible.

However, it is not perfectly clear whether choosing ADCs with extremely low resolutions will be beneficial from energy efficiency point of view, and analyses of this problem appear scarce. Only some very recent results [5], [6] seem to indicate that very low bit resolutions are not optimal in energy efficiency sense.

The analysis in this contribution aims at providing an insight in how the overall energy efficiency of a MaMi system - specifically, MaMi uplink - behaves when the ADC resolution is changed. The total power consumption of MaMi base station is parameterized, so that the analysis covers a wide variety of system architectures - from very economical to very power-hungry. Also, the ADC power consumption model that is

used attempts to reflect the functional dependencies that are found in actual ADC designs and as such aims to be close to hardware design reality.

## II. ENERGY EFFICIENCY METRIC

Energy efficiency of a base station (BS) receiver in the uplink of a MaMi system is defined as

$$\eta = \frac{C}{P_{tot}} \quad [\text{bits/Joule}], \quad (1)$$

with  $C$  [bits/s] being the uplink sumrate and  $P_{tot}$  [W] total power consumption of the MaMi BS (ADCs together with all other receiver blocks, analog and digital).

Dependencies of sumrate and power consumption on ADC bit resolution  $b$  need to be resolved separately. To this end, we first turn to finding an appropriate model for the impact of ADCs on system performance.

## III. ADC PERFORMANCE MODELING

This analysis assumes ADCs with bit resolution  $b$  that perform scalar quantization and are uniform with  $N_q = 2^b$  quantization levels. Uniform quantization was chosen because it is both close to hardware implementation reality [7] and allows for simple and tractable modeling. Additionally, sampling is assumed to be performed at Nyquist rate.

*Quantizer mapping rule:* given the ADC resolution  $b$  and a real positive scalar  $X_{ol}$ , the *quantization step* of the quantizer is defined as  $\Delta = 2X_{ol}/2^b$ , and the values of  $N_q$  quantization levels are assigned as  $q_i = i\Delta - (N_q + 1)\Delta/2$ ,  $i = 1, \dots, N_q$ . Additionally, the real line segment  $[-X_{ol}, X_{ol}]$  can be divided into  $N_q$  equal sub-segments  $[-X_{ol}, T_1], [T_1, T_2], \dots, [T_{N_q-1}, X_{ol}]$  with sub-segment boundaries (thresholds)  $T_i = q_i + \frac{\Delta}{2}$ . Given a discrete-time input  $x[n]$ , the input-output characteristic of the quantizer is defined as

$$Q(x[n]) = \begin{cases} q_1, & x \leq T_1 \\ q_i, & T_{i-1} < x \leq T_i, \quad i = 2, \dots, N_q - 1 \\ q_{N_q}, & x > T_{N_q-1} \end{cases} \quad (2)$$

Quantization  $Q(x[n])$  is a nonlinear mapping of  $x[n] \in \mathbb{R}$  to a discrete set that results in additive distortion

$$Q(x[n]) = x[n] + q[n]. \quad (3)$$

The nature of distortion  $q[n]$  is twofold (in the follow-up, time index  $n$  is dropped for clarity). If  $|x| > X_{ol}$ ,  $x$  is represented

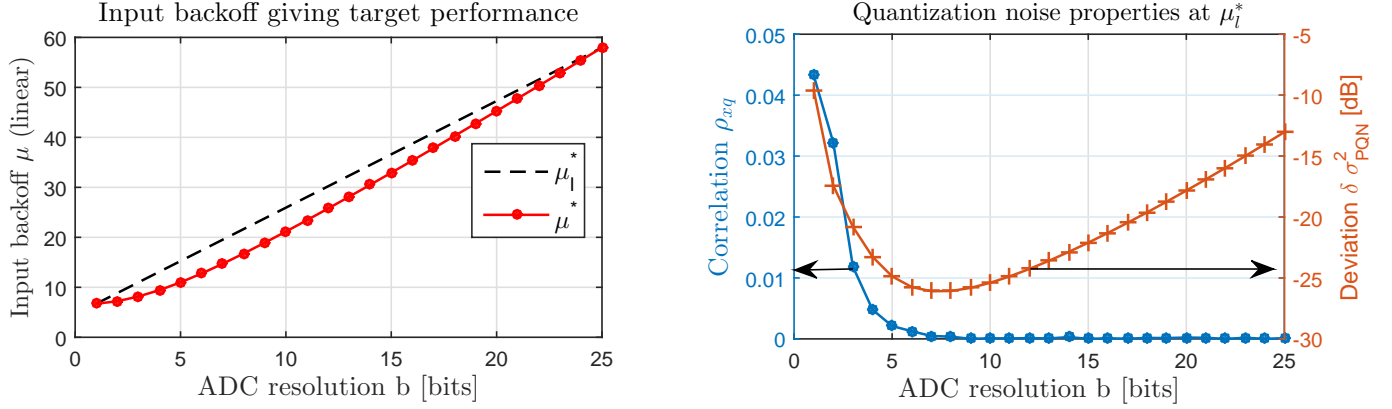


Fig. 1: Left: input backoff  $\mu$  achieving -13 dB deviation from PQN model. Right: actual deviation and input-distortion correlation for the linear fit

by one of the outer quantization levels  $q_1$  or  $q_{N_q}$  and we say that the signal is “clipped”. Consequently,  $q$  is referred to as *clipping* or *overload distortion* with variance  $\sigma_{ol}^2$ . On the other hand, if  $|x| \leq X_{ol}$ , the amplitude of distortion  $q$  is bounded by  $\frac{\Delta}{2}$ ; distortion  $q$  is then referred to as *granular noise*.

In practical systems, the ADC is usually preceded by an automatic gain control (AGC) variable gain amplifier that is used to conveniently adjust the dynamic range of the signal  $x$ . The primary purpose of the AGC is to minimize overload distortion. A welcome consequence of a properly controlled dynamic range of the input signal is a particularly convenient model for the ADC distortion.

The model in question is described as follows: it was shown in [8] that, for a uniform quantizer with normally distributed  $x$ , the distortion  $q$  can very well be approximated as being uniformly distributed, uncorrelated with the input and white, with the variance of  $q$  being

$$\mathbb{E}\{q^2\} \approx \frac{1}{3} X_{ol}^2 2^{-2b} \triangleq \sigma_{PQN}^2. \quad (4)$$

This commonly used model is usually referred to as the pseudoquantization noise (PQN) model. The approximations in the PQN model can be made extremely tight if the dynamic range of  $x$  is set properly.

A commonly used design parameter for the AGC is the *input backoff*  $\mu = X_{ol}^2 / \mathbb{E}\{x^2\}$ . In this work,  $\mu$  is set so that the normalized deviation of the distortion variance from  $\sigma_{PQN}^2$ ,  $\delta\sigma_{PQN}^2 = |\mathbb{E}\{q^2\} - \sigma_{PQN}^2| / \sigma_{PQN}^2$  is less or equal to -13 dB (indicating that the main source of distortion is granular noise and that it can be safely modeled by the PQN model). The resulting  $\mu^*(b)$  was then approximated linearly by a chord  $\mu_i^*(b)$ . Deviation  $\delta\sigma_{PQN}$  and input-distortion crosscorrelation  $\rho_{xq} = \mathbb{E}\{xq\} / (\sqrt{\mathbb{E}\{x^2\}}\sqrt{\mathbb{E}\{q^2\}})$  were obtained by simulations for  $b \in [1, 25]$  and  $\mu_i^*(b)$ . The results are shown in Fig. 1 and show that the PQN model applies well even for very low bit resolutions (1 bit) if the AGC backoff is set properly.

#### IV. SYSTEM MODEL AND SUMRATE CALCULATION

By having determined an appropriate model for the effects of ADC, we now employ it in the overall system model for the uplink, which is constructed assuming the following system setup:

- Uplink of a single-cell MaMi system with  $M$  antennas and  $K$  users;
- i.i.d. Rayleigh block fading over  $T$  symbols;
- Least-squares channel estimation is performed using orthogonal pilot sequences of length  $\tau$  in the uplink;
- Channel estimates are used for linear receiver processing. Maximum ratio combining (MRC) and zero-forcing (ZF) receivers are considered.

A system model of the uplink, where ADCs are substituted by quantization noise sources following the PQN model and AGCs precede ADCs, is illustrated in Fig. 2.

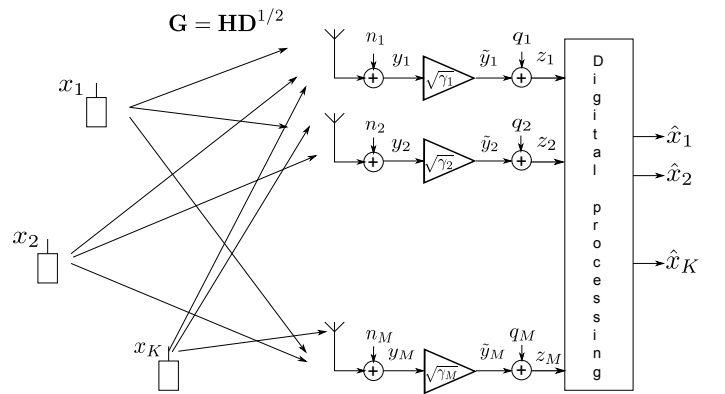


Fig. 2: Uplink system model with quantization noise

User  $k$  sends a data symbol  $x_k \in \mathbb{C}$ . User symbols are collected in vector  $\mathbf{x} = (x_1 \ x_2 \ \dots \ x_k)^T$ , with  $\mathbb{E}\{\mathbf{x}\mathbf{x}^H\} = \mathbf{I}_K$ . Single-carrier, narrowband transmission is assumed, and thus the propagation channel is represented by the  $M \times K$  matrix  $\mathbf{G} = \mathbf{H}\mathbf{D}^{1/2}$ , with the elements of  $M \times K$  matrix  $\mathbf{H}$  modeling small-scale (SS) fading coefficients. The elements

of  $\mathbf{H}$  are zero mean, circularly symmetric, complex Gaussian random variables with variance 1.

The  $K \times K$  matrix  $\mathbf{D}^{1/2}$  is a diagonal matrix of amplitude path gains and large-scale (LS) fading coefficients taken jointly. The  $(m, k)$  element of  $\mathbf{G}$  can be written as  $g_{mk} = h_{mk}\sqrt{\beta_k}$ , with  $h_{mk}$  being the narrowband small-scale fading coefficient between the  $k$ th user and  $m$ th antenna and  $\beta_k$  the joint path power gain and large-scale fading coefficient. It should be pointed out that, if some uplink power control is employed, its effects will also be modeled by  $\beta$ . In the case of ideal uplink power control, all  $\beta_k = 1$ .

Assuming that every user transmits with equal transmit power  $p_u$ , the signal at the receive antennas is

$$\mathbf{y} = \sqrt{p_u}\mathbf{G}\mathbf{x} + \mathbf{n} = \sqrt{p_u}\mathbf{H}\mathbf{D}^{1/2}\mathbf{x} + \mathbf{n} \quad (5)$$

where  $\mathbf{n} = (n_1 \ n_2 \ \dots \ n_m)^T$  is the vector of input-referred thermal noise at each antenna. Thermal noise powers at all antennas are assumed equal to  $p_n$ .

Received signal  $y_i$  will experience variations of average power due to SS and LS fading. This power is averaged over both SSF and LSF and combined with  $\mu_i^*$  to find the AGC gains that result in target performance as described in Section III. The AGC gain per I/Q branch of  $i$ th receiver chain is found to be

$$\gamma_i = \frac{2}{\mu_i^* \left( p_u \sum_{k=1}^K \beta_k + p_n \right)}. \quad (6)$$

Amplitude AGC gains  $\sqrt{\gamma_i}$  can be conveniently collected in a diagonal matrix  $\mathbf{\Gamma}^{1/2}$ .

The signal after the AGC is

$$\begin{aligned} \tilde{\mathbf{y}} &= \mathbf{\Gamma}^{1/2}\mathbf{y} = \sqrt{p_u} \mathbf{\Gamma}^{1/2}\mathbf{H}\mathbf{D}^{1/2}\mathbf{x} + \mathbf{\Gamma}^{1/2}\mathbf{n} \\ &= \sqrt{p_u} \tilde{\mathbf{H}}\mathbf{x} + \tilde{\mathbf{n}}. \end{aligned} \quad (7)$$

Finally, quantization noise is added. Assuming  $X_{ol} = 1$ , variance of complex quantization noise in the  $i$ th chain is

$$p_{q,i} = \mathbb{E} [|q_i|^2] = \frac{2}{3} 2^{-2b_i}. \quad (8)$$

and the signal model after the ADC becomes

$$\mathbf{z} = \tilde{\mathbf{y}} + \mathbf{q} = \sqrt{p_u} \tilde{\mathbf{H}}\mathbf{x} + \tilde{\mathbf{n}} + \mathbf{q}, \quad (9)$$

where  $\mathbf{q}$  holds the complex quantization noise samples from all antennas.

*Channel estimation* in the uplink is performed using pilot sequences that are spatially orthogonal and  $\tau$  symbols long. More precisely, pilot sequences for all  $K$  users are represented by a  $K \times \tau$  matrix  $\mathbf{\Phi} = \sqrt{p_u}\bar{\tau}\mathbf{\Psi}$ , where in turn  $\mathbf{\Psi}$  is a  $K \times \tau$  matrix with orthonormal rows:  $\mathbf{\Psi}\mathbf{\Psi}^H = \mathbf{I}_{K \times K}$ . Sequences  $\mathbf{\Phi}$  are optimal for least-squares pilot-based channel estimation [9].

When a block of pilot symbols  $\mathbf{\Phi}$  is transmitted, the received signal is

$$\mathbf{Z} = \tilde{\mathbf{H}}\mathbf{\Phi} + \tilde{\mathbf{N}} + \mathbf{\Xi}, \quad (10)$$

where the columns of matrices  $\tilde{\mathbf{N}} = [\tilde{\mathbf{n}}_1 \ \tilde{\mathbf{n}}_2 \ \dots \ \tilde{\mathbf{n}}_\tau]$  and  $\mathbf{\Xi} = [\mathbf{q}_1 \ \mathbf{q}_2 \ \dots \ \mathbf{q}_\tau]$  are thermal and quantization noise vectors for

each channel use (symbol). The least-squares channel estimate is then

$$\widehat{\tilde{\mathbf{H}}} = \mathbf{Z}\mathbf{\Phi}^\dagger = \tilde{\mathbf{H}} + (\tilde{\mathbf{N}} + \mathbf{\Xi})\mathbf{\Phi}^\dagger = \tilde{\mathbf{H}} + \widehat{\mathbf{H}}_\epsilon. \quad (11)$$

*Linear processing* matrices for the uplink are formed using the channel estimates:

- MRC:  $\hat{\mathbf{A}}_{\text{MRC}} = \widehat{\tilde{\mathbf{H}}}$ ,
- ZF:  $\hat{\mathbf{A}}_{\text{ZF}} = \widehat{\tilde{\mathbf{H}}} \left( \widehat{\tilde{\mathbf{H}}}^H \widehat{\tilde{\mathbf{H}}} \right)^{-1}$ .

The MIMO receiver applies the processing matrix to estimate the vector of symbols sent by the users as

$$\hat{\mathbf{x}} = \hat{\mathbf{A}}^H \mathbf{z} = \sqrt{p_u} \hat{\mathbf{A}}^H \tilde{\mathbf{H}}\mathbf{x} + \hat{\mathbf{A}}^H \tilde{\mathbf{n}} + \hat{\mathbf{A}}^H \mathbf{q}. \quad (12)$$

It can be shown that  $\hat{\mathbf{A}}$  can be split into a sum of two terms, one being the “true” processing matrix (based solely on the actual channel  $\tilde{\mathbf{H}}$ ) and the other an error term that is a consequence of channel estimation errors, namely

- MRC:  $\hat{\mathbf{A}}_{\text{MRC}} = \mathbf{A}_{\text{MRC}} + \mathbf{A}_{\text{MRC},\epsilon} = \tilde{\mathbf{H}} + \mathbf{A}_{\text{MRC},\epsilon}$ , and
- ZF:  $\hat{\mathbf{A}}_{\text{ZF}} = \mathbf{A}_{\text{ZF}} + \mathbf{A}_{\text{ZF},\epsilon} = \tilde{\mathbf{H}} \left( \tilde{\mathbf{H}}^H \tilde{\mathbf{H}} \right)^{-1} + \mathbf{A}_{\text{ZF},\epsilon}$ .

This fact holds for both MRC (follows directly from (11)) and ZF [10].

This simple decomposition allows for splitting the estimate of user data symbol  $x_k$ , pertaining to  $k$ th user, into a wanted signal term and a noise term

$$\hat{x}_k = x_k^{(w)} + w_k = \sqrt{p_u} \mathbf{a}_k^H \tilde{\mathbf{h}}_k x_k + w_k, \quad (13)$$

where  $\mathbf{a}_k$  is the  $k$ th column of  $\mathbf{A}$  and  $\tilde{\mathbf{h}}_k$  the  $k$ th column of  $\tilde{\mathbf{H}}$ . Additive noise term  $w_k$  contains interuser interference and effects of thermal and quantization noise during channel estimation and data transmission phases.

One important observation (the proof of which is omitted here) is that the constituent terms of  $w_k$  are all *uncorrelated and Gaussian*. This is a consequence of several factors, namely: quantization noise being uncorrelated with the input to the ADC, noise in channel estimation phase being independent from the one in data transmission phase, and a large number of antennas (so that the central limit theorem applies).

The signal-to-interference-thermal-and-quantization-noise ratio for  $k$ th user is then calculated as

$$\text{SINQR}_k = \frac{\mathbb{E}_{x,n,q} \{|x_k^{(w)}|^2\}}{\mathbb{E}_{x,n,q} \{|w_k|^2\}}. \quad (14)$$

The ergodic sumrate of the system is the sum of achievable rates for each user, averaged over channel realizations:

$$C = B \frac{T - \tau}{T} \sum_{k=1}^K \mathbb{E}_{\mathbf{H}} \{ \log_2(1 + \text{SINQR}_k) \}, \quad (15)$$

with  $B$  being the bandwidth of the system.

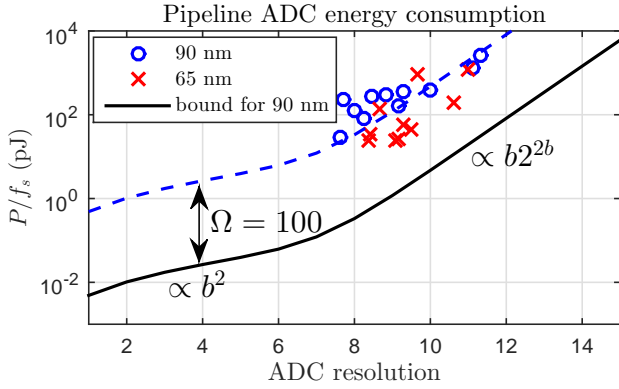


Fig. 3: ADC power consumption model, compared with actual ADC designs

## V. POWER CONSUMPTION MODEL

In this work, system setup choices and models aim to be as close to hardware reality as possible. To this end, we focus on a particular type of ADC - the pipeline ADC. This type of ADC is typically designed for intermediate bit resolutions, medium to high sampling rates  $f_s$  and has power consumption that is comparatively superior to other types of ADCs when observed over a wide range of operating resolutions (very low to very high) [11], [12], [13].

For the power consumption model of the ADC, this work adopts the model described in [14]. It represents a theoretical bound on power dissipation of pipeline ADCs that was nevertheless shown to correctly predict the trends observed in actual designs. As such, it can be of use in a parametric energy efficiency analysis, where the *character of functional dependency* between  $b$  and power consumption is of primary interest.

As shown in Fig. 3, where the model from [14] is compared with selected pipeline ADC designs collected in [15], the functional dependency in the model matches the trend exemplified by state-of-the-art pipeline architectures. Notwithstanding, there is a gap (about two orders of magnitude wide) between the bound and the designs. This implies that a correction factor  $\Omega$  can be used if the model is to be matched to state-of-the-art; this is illustrated in Fig. 3, where the dashed blue line represents the bound from [14] offset by  $\Omega = 100$ . The power consumption of pipeline ADCs can therefore be calculated as

$$P_{ADC} = \Omega (c_1 b + c_2 b^2 + c_3 2^{2b} + c_4 b 2^{2b}) f_s. \quad (16)$$

Another characteristic of this model that is worth pointing out is that the power consumption is linear in sampling rate  $f_s$ . This trend is also shown to be correct by analyzing actual ADCs in [15]; it only breaks down for high sampling rates (on the order of 400 - 500 MHz). This results in energy efficiency being *independent of system bandwidth*, since passband system bandwidth in the numerator and Nyquist sampling rate of the ADC in the denominator of (1) cancel out.

In order to paint the whole picture of energy efficiency of a MaMi base station in the uplink, power consumption

of remaining blocks (analog and digital) needs to be taken into account. This proves to be an extremely challenging task due to wide variability of available system architectures and apparent lack of unifying theoretical information. Therefore, this work chooses a *parametric approach* to modeling of the total power consumption.

Power consumption of the blocks excluding ADCs, denoted by  $P_{rest}$ , is normalized by  $P_{ADC, ref}$  - power consumption of ADCs across all RF chains at an arbitrary bit resolution  $b_{ref}$ . Total power consumption of the BS in the uplink can therefore be expressed as

$$P_{tot} = 2MP_{ADC} + P_{rest} = 2M(P_{ADC} + \alpha P_{ADC, ref}), \quad (17)$$

where the quantity

$$\alpha = \frac{P_{rest}}{2MP_{ADC, ref}} \quad (18)$$

is referred to as the *architecture factor*. Introduction of the architecture factor helps tie together the power consumption of the ADCs and remaining receiver blocks: moreover, it allows for a parameterized analysis that covers a wide range of system architectures.

## VI. RESULTS

The aim of this contribution was to provide an initial overview of the energy efficiency trends as various system parameters change. To provide this initial insight, system performance simulations have been performed across a wide variety of system parameters.

Alongside primary system parameter  $b$ , several other important system parameters have been considered, namely  $M$ ,  $K$ ,  $T$ ,  $\tau$  and preprocessing  $SNR = p_u/p_n$  (defined with large-scale fading normalized to the level of thermal noise). In order to reduce the dimensionality of the analysis, two auxiliary system parameters have been introduced, namely *spatial loading* ( $K/M$ ) and *temporal loading* ( $K/T$ ).

In addition to all the assumptions on system setup listed before, it was assumed that perfect power control was performed in the uplink (so all  $\beta_k = 1$ ). In all the analyses, reference bit resolution  $b_{ref}$  was set to 2.

For the first set of results,  $\alpha$  and  $SNR$  were swept together with  $b$ . Additionally,  $M = 100$ ,  $\tau = K$ ,  $K/T = 0.01$  [users/coherence time],  $K/M = 0.1$  [users/antenna]. Results are shown in Fig. 4. Optimal energy efficiency points are denoted by the circular marker.

Results indicate that, as power consumption of ADCs becomes comparable to power consumption of all the other blocks, from energy efficiency point of view it is beneficial to use lower bit resolution. However, in practical system designs it is reasonable to expect that ADC power consumption is only a small fraction of the total power consumption when ADC resolution is low.

Just to provide an illustrative example, BS power model presented in [16] was used with the parameters listed above (additionally, system bandwidth was assumed to be 20 MHz) and yielded  $P_{rest} = 43.3W$ . On the other hand, at  $b_{ref} = 2$ ,

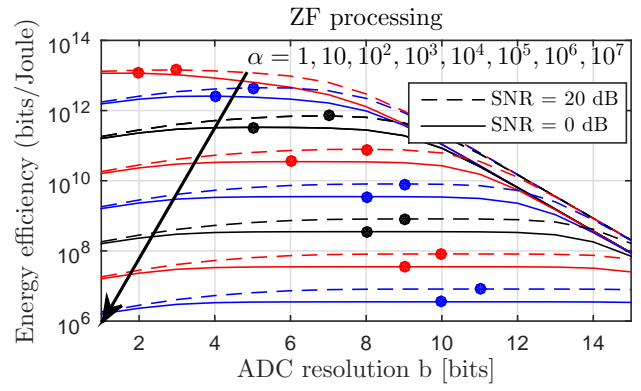
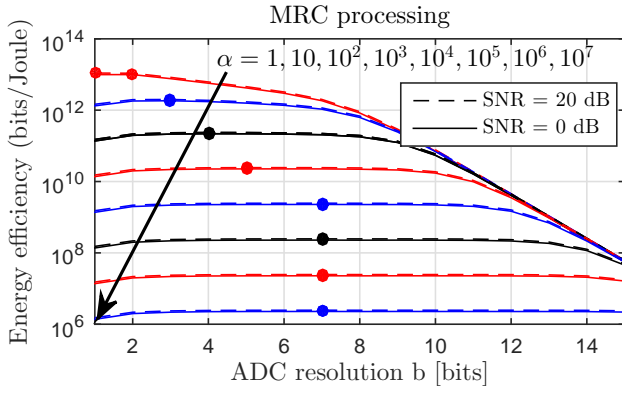


Fig. 4: Energy efficiency as a function of architecture parameter  $\alpha$  and SNR. Left: MRC, right: ZF

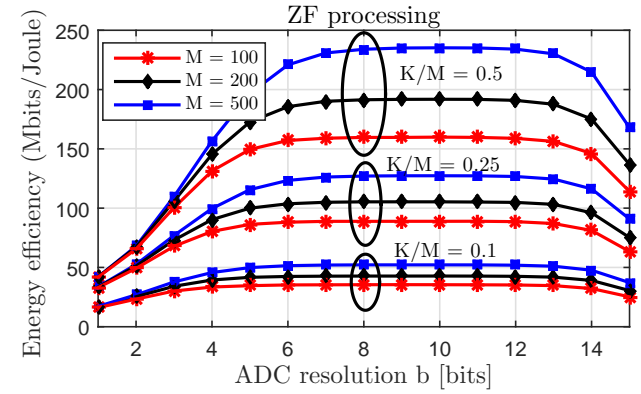
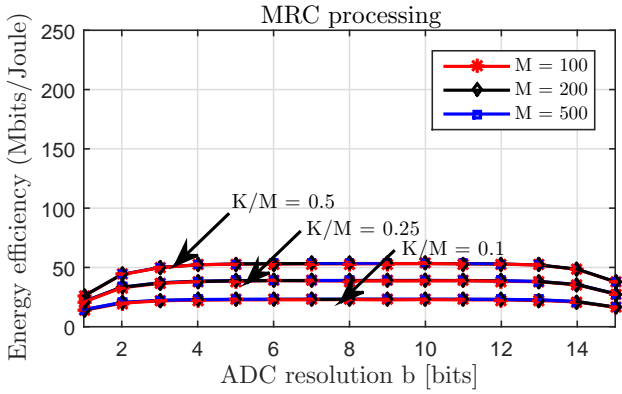


Fig. 5: Energy efficiency as a function of  $M$ ,  $K$  and  $b$ . Left: MRC, right: ZF

using a correction factor  $\Omega = 100$ , the ADC power consumption model described above gave  $2MP_{ADC} = 3 \text{ mW}$ , resulting in  $\alpha = 1.5 \times 10^4$ . While this is by no means a definite power number, it serves to illustrate what are reasonable orders of magnitude for  $\alpha$ .

Some other interesting insights can be drawn from this result, for example: system using MRC proves to be quite insensitive to changes in SNR and  $b$ , indicating that an overwhelmingly dominant impairment is the interuser interference and that playing with ADC resolutions will not yield a considerable impact on the energy efficiency; if ZF is used, the dynamics are much more pronounced and show that by going from a system design with a large SNR and large  $\alpha$  (“wasteful” system) to a system where SNR and  $\alpha$  are low (a more “economical” system) allows for choosing ADCs with smaller resolutions. Nevertheless, all systems with a “reasonable”  $\alpha$  (say  $10 - 10^5$ ) should use ADCs with resolutions in the range 4 - 10 bits.

In order to focus more on what are the improvements and degradations of energy efficiency when using different ADC resolutions, we turn to a different analysis where spatial load  $K/M$  and  $M$  are swept together with  $b$ , and additionally  $SNR = 0 \text{ dB}$ ,  $K/T = 0.01$  [users/coherence time] with  $\tau = K$  and  $\alpha = 10^4$ , results shown in Fig. 5.

What these results show is that going from optimal ADC resolution to a very low one can incur a substantial degradation of the energy efficiency (for ZF processing and with assumed values of system parameters, up to 5.5 times). This is due to sumrate being degraded while the overwhelming power consumption of other blocks “drowns” the coincident savings in power consumption of the ADCs. Another interesting observation is that, in the ZF case, increasing the number of antennas can help recover the energy efficiency lost by going to lower bit resolutions.

Finally, we take a look at the interplay between the channel estimation length and  $b$  in the context of energy efficiency. We analyze a system with  $K/M = 0.1$  [users/antenna], while varying  $b$ , SNR and training length. Architecture parameter  $\alpha$  is again fixed to  $10^4$ . What is plotted is the normalized training length  $\tau/T$  that maximizes the energy efficiency, results shown in Fig. 6. Main takeouts from here are that ZF is much more sensitive to quantization noise during training; even in the case of high temporal loading (indicating fast fading), when there is little room to spare for channel estimation, it is beneficial to train the system longer than minimum required time in order to compensate for the effects of quantization. The effect becomes more pronounced as the fading becomes slower and channel estimation is not so costly in terms of time. On the other



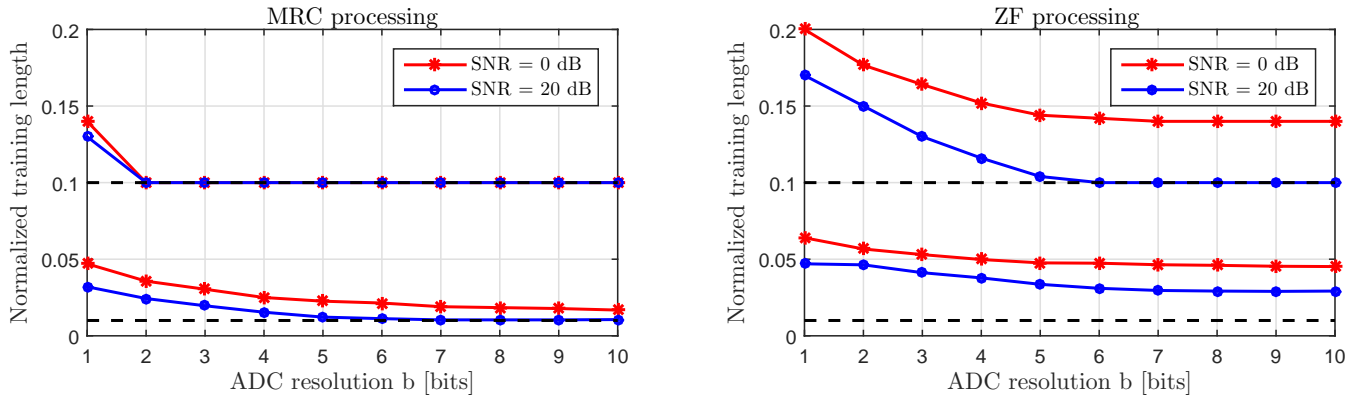


Fig. 6: Normalized training length that maximizes energy efficiency. Left: MRC, right: ZF

hand, we see that the system using MRC is so overwhelmed by interuser interference that additional training does little to improve the energy efficiency.

## VII. CONCLUSION

A parameterized analysis of energy efficiency in the uplink of a MaMi system with varying ADC bit resolutions at the base station has been performed. System setup and models have been chosen with the aim of being close to practical system implementations. Initial results (obtained by simulations) indicate that using ADCs with very low bit resolutions is not an optimal approach from energy efficiency point of view, except for highly specific system architectures. Instead, for a wide variety of systems, ADCs with intermediate bit resolutions (4 - 10 bits) are shown to maximize system energy efficiency. Additionally, it was also shown that systems using MRC uplink processing are quite insensitive to the changes in ADC bit resolution, due to interuser interference being the prime source of impairments in such systems. On the other hand, systems using ZF processing (in addition to showing overall superiority in terms of energy efficiency compared to MRC - based systems) are shown to be rather sensitive to changes in bit resolution.

## REFERENCES

- [1] T. L. Marzetta, "Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas," in *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590-3600, November 2010.
- [2] E. Björnson, M. Matthaiou and M. Debbah, "Massive MIMO with Non-Ideal Arbitrary Arrays: Hardware Scaling Laws and Circuit-Aware Design," in *IEEE Transactions on Wireless Communications*, vol. 14, no. 8, pp. 4353-4368, Aug. 2015.
- [3] C. Risi, D. Persson, and E. G. Larsson. "Massive MIMO With 1-Bit ADC," Apr. 2014 [Online]. Available: <http://arxiv.org/abs/1404.7736>
- [4] S. Jacobsen, G. Durisi, M. Coldrey, U. Gustavsson and C. Studer, "One-bit massive MIMO: Channel estimation and high-order modulations," 2015 IEEE International Conference on Communication Workshop (ICCW), London, 2015, pp. 1304-1309.
- [5] D. Verenzuela, E. Björnson and M. Matthaiou, "Hardware design and optimal ADC resolution for uplink massive MIMO systems," 2016 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM), Rio de Janeiro, 2016, pp. 1-5.
- [6] J. Mo, A. Alkhateeb, S. Abu-Surra, and R. W. Heath Jr, "Hybrid architectures with few-bit ADC receivers: Achievable rates and energy-rate tradeoffs." [Online]. Available: <https://arxiv.org/abs/1605.00668>
- [7] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Boston, MA, USA: Kluwer, 1992.
- [8] A. Sripad and D. Snyder, "A necessary and sufficient condition for quantization errors to be uniform and white," in *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 25, no. 5, pp. 442-448, Oct 1977.
- [9] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," in *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 884-893, March 2006.
- [10] C. Wang, E. K. S. Au, R. D. Murch, W. H. Mow, R. S. Cheng and V. Lau, "On the Performance of the MIMO Zero-Forcing Receiver in the Presence of Channel Estimation Error," in *IEEE Transactions on Wireless Communications*, vol. 6, no. 3, pp. 805-810, March 2007.
- [11] B. Le, T. W. Rondeau, J. H. Reed and C. W. Bostian, "Analog-to-digital converters," in *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 69-77, Nov. 2005.
- [12] C. Svensson, "Towards Power Centric Analog Design," in *IEEE Circuits and Systems Magazine*, vol. 15, no. 3, pp. 44-51, third quarter 2015.
- [13] B. Murmann, "Trends in low-power, digitally assisted A/D conversion," *IEICE Trans. Electron.*, vol. E93-C, no. 6, pp. 718727, 2010.
- [14] T. Sundstrom, B. Murmann and C. Svensson, "Power Dissipation Bounds for High-Speed Nyquist Analog-to-Digital Converters," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 56, no. 3, pp. 509-518, March 2009.
- [15] B. Murmann. ADC performance survey 19972016. [Online]. Available: <http://www.stanford.edu/~murmann/adcsurvey.html>
- [16] C. Desset, B. Debaillie and F. Louagie, "Modeling the hardware power consumption of large scale antenna systems," 2014 IEEE Online Conference on Green Communications (OnlineGreenComm), Tucson, AZ, 2014, pp. 1-6.