How Much Will Tiny IoT Nodes Profit from Massive Base Station Arrays?

Ema Becirovic, Emil Björnson, Erik G. Larsson
Dept. of Electrical Engineering, Linköping University, Linköping, Sweden
Email: {ema.becirovic, emil.bjornson, erik.g.larsson}@liu.se

Abstract—In this paper we study the benefits that Internet-of-Things (IoT) devices will have from connecting to a massive multiple-input-multiple-output (MIMO) base station. In particular, we study how many users that could be simultaneously spatially multiplexed and how much the range can be increased by deploying massive base station arrays. We also investigate how the devices can scale down their uplink power as the number of antennas grows with retained rates.

We consider the uplink and utilize upper and lower bounds on known achievable rate expressions to study the effects of the massive arrays. We conduct a case study where we use simulations in the settings of existing IoT systems to draw realistic conclusions.

We find that the gains which ultra narrowband systems get from utilizing massive MIMO are limited by the bandwidth and therefore those systems will not be able to spatially multiplex any significant number of users. We also conclude that the power scaling is highly dependent on the nominal signal-to-noise ratio (SNR) in the single-antenna case.

I. INTRODUCTION

The International Telecommunication Union specified the key features of the next generation of wireless communication systems in [1]. The category massive machine type communications (mMTC) is mainly focused on having a large number of devices that transmit small amounts of data and require high energy efficiency in order for the devices to last for years on batteries. A subset of mMTC is the Internet-of-Things (IoT).

The use cases of IoT are, for example, energy metering and other scenarios where sensors send small amounts of data to the cloud. The communication is mainly done in the uplink. We refer to these types of small sensors as tiny IoT nodes.

Currently, there is a number of proprietary (and non-proprietary) systems for IoT communications. The requirements put on these systems are high energy efficiency and a large number of users associated with each cell. The systems do not provide a particularly high data rate.

Massive MIMO is a well explored research area in the context of mobile broadband applications, but not many studies have considered the application of massive MIMO to serve IoT devices. The coherence interval is the interval in time and frequency where we can consider the channel to be time-invariant and frequency-flat. The existing IoT systems have a small bandwidth, smaller than the coherence bandwidth, which makes the coherence interval small. Having a small coherence interval makes it difficult to fit many orthogonal pilot sequences such that many users could be spatially multiplexed. Our aim is to, for a required data rate per user, serve as many users and increase the range as much as possible, compared to conventional mobile broadband systems where the goal is to, for a fixed number of users and range, maximize the data rate.

The main goal of this paper is to quantitatively answer the question “How Much Will Tiny IoT Nodes Profit from Massive Base Station Arrays?”; specifically:

- How much can the power of the tiny IoT nodes be lowered by using a multi-antenna base station?
- How much can the range be increased by using a multi-antenna base station?

There are many lower, achievable, bounds on the uplink capacity in massive MIMO [2]. However, these bounds are not tight for a small number of antennas at the base station. In [3] and [4], upper bounds on the single-cell uplink capacity are derived, however these do not hold when the users have equal SNR, which is a typical goal of power control.

To answer the questions above, we derive a new upper bound on the single-cell uplink capacity assuming maximum ratio combining and statistical channel inversion power control, analyzing the power scaling capabilities in tiny IoT nodes and performing a case study where we use the system parameters from existing systems, LoRa [5] and Sigfox [6], in our simulations.

Other works on this topic include [7] where it is investigated how IoT devices and Fourth Generation (4G) devices can coexist with the help of massive MIMO. In [8] the coexistence and performance of LoRa and Sigfox is studied, and in [9] the Chirp Spread Spectrum modulation used in LoRa is studied and compared to the BPSK modulation used in Sigfox. Further, in [10] the authors studied massive MIMO in industrial IoT and presented a discussion on future research of the subject.

A. Notation

The Gamma function is denoted by \( \Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt \), \( \Gamma(s, x) = \int_x^\infty t^{x-1} e^{-t} dt \) is the upper incomplete Gamma function, and \( E_1(x) = \Gamma(0, x) = \int_x^\infty \frac{e^{-t}}{t} dt \) is the Exponential integral.
B. Background on Current IoT Solutions

In this section, we introduce two existing IoT systems. Both of them, Sigfox and LoRa, are proprietary and mainly operate in the ISM bands around 868 MHz in Europe and 913 MHz in the USA. In this paper, we do not consider any particular system but simply use the performance goals of these systems as a benchmark for our conclusions.

1) LoRa: LoRa uses a modulation technique based on chirp spread spectrum (CSS) modulation. LoRa uses 14 dBm transmission power, which is the highest allowed transmission power in the operating frequency band [5].

The bit rate achieved by LoRa depends on the bandwidth, B, and the spreading factor, S, of the signal as [11]

\[ R_b = S \frac{B}{2^S}. \] (1)

Each symbol is S bits long, we send \( 2^S \) chips per symbol and the chip rate is 1 chip/Hz. The bandwidth of LoRa ranges from 7.8 to 500 kHz.

2) Sigfox: Sigfox is an ultra narrowband (UNB) system that operates with a bandwidth of only 100 Hz. The bit rate is also very low, only 100 bit/s [6], which is significantly lower than LoRa. The transmission power is 14 dBm, the highest allowed in the frequency band. Sigfox uses differential-binary-phase-shift-keying (DBPSK) modulation.

II. System Model

We consider a single-cell uplink system with M base station antennas and K single-antenna users, which are tiny IoT nodes. We let \( g_{mk} \) denote the channel per coherence interval between base station antenna \( m \) and user \( k \), and we model it by independent Rayleigh fading as

\[ g_{mk} = \sqrt{\beta_k h_{mk}}, \] (2)

where \( \beta_k \) is the large-scale fading coefficient, which is assumed to be known, and \( h_{mk} \sim CN(0, 1) \) is the small-scale fading that needs to be estimated at the base station.

The coherence interval contains \( \tau_c = T_C B_c \) samples, where \( T_c \) is the coherence time and \( B_c \) is the coherence bandwidth. The coherence bandwidth is the bandwidth where the channel is approximately frequency flat, but it will be limited by the available bandwidth of the system. Therefore, \( B_c \) is the minimum of the coherence bandwidth of the physical channel and the available bandwidth of the system. The coherence time depends on the mobility of the user and the propagation environment. If a user is stationary, the coherence time will be limited by the changes in the environment.

In order to learn the channel, the users send \( \tau_p \)-length pilot signals in each coherence interval. The pilot signals are orthogonal and each user gets a unique pilot signal; therefore we need at least as many pilots as users. Also, we cannot have more pilot sequences than there are samples in the coherence interval: \( K \leq \tau_p \leq \tau_c \).

The minimum mean-squared error (MMSE) estimator of the channel is [2, p. 47]

\[ \hat{g}_{mk} = \frac{\sqrt{T_p \rho_p \eta_k \beta_k}}{1 + \sqrt{T_p \rho_p \eta_k \beta_k} \rho_p \eta_k g_{mk} + n_{mk}}, \] (3)

where \( n_{mk} \sim CN(0, 1) \) is denoting the noise, \( \rho_p \) is the nominal uplink SNR and \( \eta_k \in [0, 1] \) is the power control coefficient of user \( k \) for both pilots and data. The MMSE estimator is slightly modified from [2] where the authors assume full power on the pilots, and use power control only on the data. We do this modification to simplify the argumentation of the power scaling later in the paper.

We let \( \gamma_k \) denote the mean square of the channel estimate of user \( k \), which is

\[ \gamma_k = \frac{T_p \rho_p \eta_k \beta_k^2}{1 + T_p \rho_p \eta_k \beta_k}. \]

We write the MMSE estimate of the channel to user \( k \) as

\[ \hat{g}_k = \sqrt{\gamma_k} z_k, \] (4)

where \( z_k \) is a vector of length \( M \) with independent \( CN(0, 1) \) elements.

Assuming maximum ratio combining at the base station, from [2, p. 54] the uplink capacity of user \( k \) is lower bounded by \( R_k(K) \) in (5) (at the top of the next page). For convenience we introduce

\[ a = 1 + \rho_p \sum_{K'} \beta_{k'} - \gamma_{k'} \eta_{k'}, \] (6)

\[ b_k = \rho_p \gamma_k \eta_k, \] (7)

and

\[ c_{k,k'} = \frac{\| z_k \|^2}{\| z_{k'} \|^2}. \] (8)

Bounding (5) with Jensen’s inequality in two different ways (as in [2, Appendix C.1]) when \( K = 1 \) gives us

\[ \log_2 \left( 1 + \frac{(M - 1) b_1}{a} \right) \leq R_k(1) \leq \log_2 \left( 1 + \frac{M b_1}{a} \right). \] (9, 10)

When \( K > 1 \) and \( \rho_p \eta_k \gamma_k \) are different for each \( k \), we get the bounds [3] (12) \( \leq R_k(K) \) \leq (13), where (12) and (13) are given at the top of the next page.

Using statistical channel inversion power control [12],

\[ \eta_k = \frac{\min_{k'} \beta_{k'}}{\beta_k} = \frac{\beta_{\min}}{\beta_k}, \] (14)

where we have defined \( \beta_{\min} = \min_{k'} \beta_{k'} \), ensures that all users have the same rate. It also ensures that all users, except the user with the smallest large-scale fading coefficient, lowers its transmit power compared to using full power. With statistical channel inversion power control all \( b_k = \rho_p \eta_k \gamma_k \) are equal. In that case, we cannot use the bounds derived in [3]. However, we can use the following results.
Lemma 1: When using statistical channel inversion power control (14) the uplink rate $R_k(K)$ can be bounded as

$$
\log_2 \left( 1 + \frac{(M - 1) b_k}{1 + \frac{1}{K} \rho \alpha \beta_{\min} - b_k} \right) \leq R_k(K) \leq \log_2 \left( 1 + M \left( \frac{a}{b_k} \right)^{K-2} e^{\frac{a}{b_k} \Gamma \left( 2 - K, \frac{a}{b_k} \right)} \right),
$$

where

$$
a = 1 - \rho \alpha \sum_{k'=1}^{K} \gamma_{k'} \frac{\beta_{\min}}{\beta_{k'}} - K \rho \alpha \beta_{\min}
$$

and

$$
b_k = \rho \alpha \gamma_k \frac{\beta_{\min}}{\beta_k} = \frac{\rho \alpha \gamma_k \beta_{\min}}{1 + \rho \alpha \gamma_k \beta_{\min}}
$$

are obtained by inserting (14) into (6) and (7). The lower bound is proved in [2, Appendix D]. The proof of the upper bound can be found in Appendix A.

The results in Lemma 1 are, as far as we know, new.

III. POWER SCALING

In this section, we will study how much transmit power that could be saved for a user by deploying a massive base station array. We want to find by which factor

$$
\eta = \frac{1}{M^\alpha}
$$

we can scale the transmitted power at the user by adding antennas at the base station while maintaining the same rate at the user as with $M = 1$. This information will give insights into how the use of massive base station arrays can increase longevity of batteries at the tiny IoT nodes.

When $K = 1$ we have $R_1(1)$ from (5). Because we only have one user, we do not need the user index $k$. Therefore, we can rewrite $R_1(1)$ as

$$
E \left\{ \log_2 \left( 1 + \frac{\rho \alpha \gamma \|z_k\|^2}{1 + \rho \alpha \gamma (\beta - \gamma)} \right) \right\} = R_1(K)
$$

and evaluate it in closed form:

$$
E \left\{ \log_2 \left( 1 + \frac{\rho \alpha \gamma \|z_k\|^2}{1 + \rho \alpha \gamma (\beta - \gamma)} \right) \right\} = \frac{1}{\ln(2)} \sum_{l=1}^{M-1} \sum_{i=0}^{l-1} \frac{\Theta^{l-i} e^{\frac{a}{\Theta}} E_1 \left( \frac{1}{\Theta} \right)}{(M - 1 - l)! (M - l - 1)!} + \frac{1}{\ln(2)} \sum_{l=0}^{M-1} \frac{\Theta^{M-l+1} e^{\frac{a}{\Theta}} E_1 \left( \frac{1}{\Theta} \right)}{(M - 1 - l)! (M - l)!}
$$

where

$$
\Theta = \frac{\rho \alpha \gamma}{1 + \rho \alpha \gamma (\beta - \gamma)}.
$$

Note that for $M = 1$, (21) can be written as

$$
\frac{1}{\ln(2)} e^{\frac{a}{\Theta}} E_1 \left( \frac{1}{\Theta} \right).
$$

$R_1(1)$ can be bounded by the Jensen inequality as stated before with $9 \leq R_1(1) \leq 11$.

To find the scaling factor $\alpha$ we set up the equation

$$
E \left\{ \log_2 \left( 1 + \frac{\rho \alpha \gamma |z_1|^2}{1 + \rho \alpha \gamma (\beta - \gamma)} \right) \right\} = E \left\{ \log_2 \left( 1 + \frac{\rho \alpha \gamma \|z_M\|^2}{1 + \rho \alpha \gamma (\beta - \gamma)} \right) \right\},
$$

where $z_1 \in \mathbb{C}$ denotes the channel with one antenna at the base station and $z_M \in \mathbb{C}^M$ denotes the channel with $M$ antennas at the base station. In (24), the power control coefficient is set to 1 meaning that the terminal is using full power. In (25), the power control coefficient is set to $\eta = \frac{1}{M^\alpha}$.

The solution to (24)-(25) is obtained numerically by Monte Carlo simulations. $\tau_p$ was set to 1 and the pre-log factor of the rate expression was neglected since the number of users is fixed. From the results in Fig. 1, we can see that for high nominal SNR, we can scale the transmitted power as $\frac{1}{M^\alpha}$ since $\alpha \approx 1$. However, for low nominal SNR, the scaling is limited.
to \( \frac{1}{\sqrt{M}} \) since \( \alpha \approx \frac{1}{2} \). The effect of increasing \( \tau_p \) is that the curves for lower nominal SNR will get a larger slope.

If \( K > 1 \) is studied, the nominal SNR for all the users affects the scaling factors, and a power control scheme would have to be used for fair comparison.

IV. SIMULATIONS

We consider a single-cell system in a multi-cell world, meaning that we consider our hexagonal target cell to be in the center of two rings of hexagonal cells. The users are placed uniformly in the world and assigned to the base station to which the user has the smallest path loss including shadow fading. Because of the shadow fading, a user is not necessarily assigned to the physically closest base station. We continue placing users until we have \( K \) users assigned to our base station of interest.

The large-scale fading, \( \beta_k \), is modeled by [13]

\[
\beta_k = -120.5 - 36.7 \log_{10}(d_k) - \sigma_k \quad [\text{dB}],
\]

where \( d_k \) is the distance in kilometers between user \( k \) and the base station and the shadow fading \( \sigma_k \) is normally distributed with zero mean and standard deviation of 8 dB.

The nominal uplink SNR, \( \rho_{ul} \), is calculated as

\[
\rho_{ul} = \frac{p_t}{k_B T B},
\]

where \( T \) is the nominal noise temperature in Kelvin, \( k_B = 1.38 \times 10^{-23} \) [J/K] is Boltzmann constant, and \( p_t \) is terminal radiated power in Watt. We have assumed that the noise figure and the antenna gains cancel out.

Two different case studies were simulated, one in a UNB Sigfox-like scenario and one in a CSS LoRa-like scenario. The

### Table I

<table>
<thead>
<tr>
<th>Quantity</th>
<th>UNB</th>
<th>CSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bandwidth, ( B )</td>
<td>100 Hz</td>
<td>125 kHz</td>
</tr>
<tr>
<td>Carrier frequency</td>
<td>868 MHz</td>
<td>868 MHz</td>
</tr>
<tr>
<td>Terminal radiated power</td>
<td>14 dBM</td>
<td>14 dBM</td>
</tr>
<tr>
<td>Required rate</td>
<td>100 bps</td>
<td>366 bps</td>
</tr>
<tr>
<td>Coherence time</td>
<td>50 ms</td>
<td>50 ms</td>
</tr>
<tr>
<td>Coherence bandwidth</td>
<td>100 Hz</td>
<td>125 kHz</td>
</tr>
<tr>
<td>Coherence interval</td>
<td>5 samples</td>
<td>6250 samples</td>
</tr>
</tbody>
</table>

The main difference between the two cases is the available bandwidth of the system. The parameters used for the simulations are presented in Table I.

The results for the UNB and CSS scenarios are presented in Fig. 2 and Fig. 3, respectively. The results presented are depicting the number of users that can be simultaneously spatially multiplexed with a specific cell radius while still achieving the specified rate in 99% of the location realizations, using both the lower bound (15) and the upper bound (16) on the achievable rate. The lengths of the pilot sequences, \( \tau_p \), are optimally chosen to achieve the specified rate. The optimal pilot lengths for the UNB case is, in the vast majority of cases, the number of users. On the contrary, the optimal pilot lengths for the CSS case are ranging from the number of users up to 2650 samples.

In Fig. 2 we can see that the coherence interval is a limiting factor in the UNB system, hence we can only spatially multiplex a few users. However, it is clear that by adding many antennas at the base station the range can be increased significantly. For the CSS case in Fig. 3 the number of spatially multiplexed users is greater and therefore more beneficial. The range benefits can be clearly seen in this case as well.

Furthermore, we can conclude that the lower and upper bounds, (15) and (16), on the rate \( R_k(K) \) in (5) are fairly tight for large number of antennas, \( M \).

V. CONCLUSION

In this paper, we asked the question “How Much Will Tiny IoT Nodes Profit from Massive Base Station Arrays?”. To answer this question, we showed that tiny IoT nodes can profit from massive MIMO by spatially multiplexing many users at the same time, depending on the coherence bandwidth. This can be done while either lowering the transmitted power at the users, by at least a factor of \( \frac{1}{\sqrt{M}} \), and thereby increasing battery life, or increasing the range, to more than double with 100 antennas at the base station compared to a single-antenna base station. To reach these conclusions, we derived a
new upper bound on the single-cell uplink capacity assuming maximum ratio combining and statistical channel inversion power control.

APPENDIX

PROOF OF UPPER BOUND IN LEMMA 1

With the power control $\eta_k = \frac{\min_k \beta_k}{\beta_{\text{max}}} = \frac{\beta_k}{\beta_{\text{max}}}$, we have $\alpha$ and $b_k$ as defined in (17) and (18). An upper bound on (5) is achieved by utilizing Jensen's inequality (as in [2, Appendix C.1]) and to move the expectation into the logarithm:

$$
\mathbb{E}\left\{ \log_2 \left( 1 + \frac{b_k \| z_k \|^2}{\alpha + b_k x} \right) \right\} \leq \log_2 \left( 1 + \frac{b_k \| z_k \|^2}{\alpha + b_k x} \right),
$$

where we use the notation

$$
x = \sum_{k' = 1 \atop k' \neq k}^{K} \frac{z_{k'}^H z_k \chi_{k'}}{\| z_k \|^2} = \sum_{k' = 1 \atop k' \neq k}^{K} c_{k, k'}.
$$

Next, we realize that $\| z_k \|^2$ and $\frac{z_k^H z_{k'}}{\| z_k \|^2}$ are independent, which stems from the fact that conditioned on $z_k$, $\frac{z_k^H z_k}{\| z_k \|^2} \sim \mathcal{CN}(0, 1)$ will not depend on the realization of $z_k$, so that we can split the expectation in two parts:

$$
\mathbb{E}\left\{ 1 + \frac{b_k \| z_k \|^2}{\alpha + b_k x} \right\} = 1 + \mathbb{E}\left\{ \| z_k \|^2 \right\} \mathbb{E}\left\{ \frac{1}{\alpha + b_k x} \right\}
$$

Finally, we complete the proof by combining (33) and (36), and inserting the result in (29):

$$
\mathbb{E}\left\{ \log_2 \left( 1 + \frac{b_k \| z_k \|^2}{\alpha + b_k x} \right) \right\} \leq \log_2 \left( 1 + M \frac{a}{\beta_k} \right)^{K-2} e^{\frac{a}{\beta_k} \Gamma \left( 2 - K, \frac{a}{\beta_k} \right)}.
$$

REFERENCES


[13] “Universal mobile telecommunications system (UMTS); radio frequency (RF) system scenarios (3GPP TR 25.942 version 14.0.0 release 14) radio frequency (RF) system scenarios,” European Telecommunications Standards Institute, April 2017.