Energy-efficient Power Control for MIMO Time-varying Channels

Zhijiat Chong and Eduard Jorswieck
Communications Theory, Communications Laboratory,
Dresden University of Technology, 01061 Dresden, Germany.
{zhijiat.chong,eduard.jorswieck}@tu-dresden.de

Abstract—Fast rising data traffic in mobile communication entails an increase in its energy consumption, which may result in climate effects and high operation costs, unless measures are taken to enhance its energy efficiency (EE). As part of the endeavour towards higher EE, we present a power control policy that adapts to a time-varying channel for multiple antenna communication. We provide a power budget model that reflects the power expenditure of a base station, that even accounts for the power used for training which enables channel information at the transmitter. We apply the results to Rayleigh fading channels and compare them with a suboptimal power policy where a constant power allocation is applied at all times. The relationship between the EE and the antenna configuration is discussed.

I. INTRODUCTION

Recently, energy efficiency (EE) in mobile communications is receiving growing attention as its increasing energy consumption raises concern over climate effects, energy availability and operation costs of network providers. Currently, about 0.2% of the global CO₂ emissions are contributed by the mobile telecommunication industry [1]. Without considering users’ phones, the power consumption of a typical mobile communications network in the UK is estimated to be about 40 MW [2]. These values are expected to increase as data traffic rises unless measures are taken to improve the EE. Due to restrictions on power supply on mobile units, their EE is highly optimised to provide customer satisfaction. However, the EE of base stations, whose power consumption constitutes the major portion in a cellular network [3], [4], has been neglected until lately [5].

Studies in wireless communications in the past decades have been focused on maximising spectral efficiency until recently, when EE enhancement has also gained importance. Increasing spectral efficiency (SE) has been done by [6] for time-varying single-input-single-output (SISO) channels through power control, data rate and coding scheme adaptation. Since the pioneering work by Telatar [7], much attention has been given to studying wireless communication systems with multiple antennas because of their capability of increasing SE. As EE is vital in sensor networks, the EE of multiple-input-multiple-output (MIMO) systems was investigated and compared with that of SISO in [8] where circuit energy consumption is also taken into account. In [9], EE in MIMO communication is treated from an information-theoretic point of view. Game theory is applied in [10] to maximise the EE of multiple users in an uplink MIMO scenario. In [11], the technique of switching between SIMO and MIMO to achieve EE in the uplink was investigated.

In this paper we show how EE can be maximised through power adaptation in MIMO time-varying channels, provided that channel statistics are known to the transmitter. We first provide an analysis of an arbitrary fading channel and later focus on MIMO Rayleigh fading channels. Our focus is on downlink transmission in a cell. As for the power budget model, which is necessary for measuring the EE, we use the one given in [12] and extend it to include the power needed for training, such that channel information is available on the transmitter’s side. The power for training is assumed to be dependent on the number of antennas employed. For the computation of the ergodic rate, which involves integrations of multiple variables, we use the result from [13] to reduce it to an expression involving integration with only one variable. In the simulations we show that if the number of antennas of the transmitter \( n_T \) and the receiver \( n_R \) are identical, EE increases linearly with \( n = n_T = n_R \). As a comparison, we also perform an EE optimisation where the power allocation is chosen once for all times. We see that as \( n_R - n_T \) and the inverse noise increases, the difference between the EE through power adaptation and that through static power allocation decreases. For \( n = n_T = n_R \) and a given fixed average sum power, we discuss the choice of an optimal \( n \) for maximum EE.

The rest of the paper is organised as follows. In Section II we introduce the system model and elaborate on the power budget and the EE. In Section III we first formulate and analyse the problem. We then provide the optimal power function that maximises the EE for an arbitrary channel state distribution. Secondly, we explain how the maximum EE for Rayleigh fading channels can be obtained using the optimal power function and the static power allocation, respectively. Simulations for these are shown and discussed in Section IV. Section V concludes the paper.

II. PRELIMINARIES

A. Notation

Lower-case boldface letters like \( \mathbf{x} \) are used for vectors such that \( \mathbf{x} = (x_1, \ldots, x_n) \). They may also represent vector functions such as \( \mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), \ldots, g_n(\mathbf{x})) \). Upper-case boldface letters such as \( \mathbf{X} \) are matrices where \( X_{ij} \) is the entry of the \( i \)th row and \( j \)th column. \( \mathbf{X}^H \) refers to the Hermitian transpose of \( \mathbf{X} \). The determinant of \( \mathbf{X} \) is denoted either by
\( [X] \) or \( \det X \), \( \text{diag}(x_1, \ldots, x_n) = \text{diag}(x) \) symbolises a square matrix \( D \) with entries \( D_{ij} = x_i \) for \( i = j \) and \( D_{ij} = 0 \) for \( i \neq j \). \( \frac{\int_0^\infty g(x) dx}{
abla} \) stands for a multivariable integral \( \int_0^\infty g(x) dx_1 \ldots dx_n \). \( \mathbb{E}_x [g(x)] = \int_0^\infty g(x) f(x) dx \) is the mean value of the function \( g(x) \), where \( f(x) \) is the probability density function (PDF) of the vector \( x \).

B. System model

Consider a single-user MIMO link with \( n_T \) transmit antennas and \( n_R \) receive antennas. Transmission occurs through a flat-fading channel with bandwidth \( W \). The \( n_R \times 1 \) received signal vector \( y \) can be written as

\[
y = Hx + n, \tag{1}
\]

where \( H \) is the \( n_R \times n_T \) channel matrix, \( x \) the \( n_T \times 1 \) transmit vector, and \( n \) the \( n_T \times 1 \) white Gaussian noise vector with variance \( \sigma^2 \) at each component.

Assume a simple block-fading channel, where the channel is constant over a time interval of duration \( T \). After each interval it changes to an independent channel value for a subsequent interval \( T \). We employ a simple feedback model such that during each interval, training symbols are transmitted for channel measurements. A fraction of the interval \( T_{tr} \) is utilised for training where \( T_{tr} < T \). We further assume that perfect information of the channel matrix \( H \) is known to the transmitter through feedback.

For every channel realisation \( H \) the rate is given by

\[
R = \frac{W (T - T_{tr})}{T} \log \det \left( I_{n_R} + \frac{\gamma}{\sigma^2} HQHQ^H \right), \tag{2}
\]

where \( Q = \mathbb{E} [xx^H] / W \) is the \( n_T \times n_T \) covariance matrix, \( \gamma \in (0, 1] \) the SNR gap that accounts for the difference between the channel capacity and the specific modulation and coding scheme implemented. This expression is similar to that shown in [14] in that the rate is reduced by a factor \( \frac{T - T_{tr}}{T} \).

Instead of using the lower bound capacity in [14] for the rate expression, we use the multi-antenna mutual information expression from [7] modified by \( \gamma \). Let \( n_{i} = \min \{n_T, n_R\} \) and \( n_{a} = \max \{n_T, n_R\} \). We can further write (2) as

\[
R = \frac{W (T - T_{tr})}{T} \sum_{i=1}^{n_{a}} \log (1 + \rho \alpha_i p_i), \tag{3}
\]

where \( \rho = \frac{\gamma}{\sigma^2} \), \( \alpha_i \) are the eigenvalues of \( H^H H \) (see [15]). The values of \( \alpha \) are computed using eigenvalue decomposition yielding an eigenvector matrix \( V \) such that \( H^H H = V \text{diag}(\alpha) V^H \). Using these eigenvectors we relate \( p \) and \( Q \) by \( Q = V^H (p) V \). The sum output power, which is dissipated into the air interface, is \( W \text{trace} Q = W \sum_{i=1}^{n_{a}} p_i \). Shadowing, multipath, antenna correlation and pathloss effects are incorporated in \( H \) and thus also in \( \alpha \). Suppose the PDF that accounts for the channel realisations is given by \( f(\alpha) \) and is known to the transmitter. The transmission power allocation is adapted according to each \( \alpha \), represented by the function \( p(\alpha) \). The ergodic rate is then defined as

\[
\mathbb{E}_\alpha \left[ \frac{W (T - T_{tr})}{T} \sum_{i=1}^{n_{a}} \log (1 + \rho \alpha_i p_i(\alpha)) \right], \tag{4}
\]

which describes the average rate achieved over all channel realisations. The average transmission power allocation is given by

\[
\mathbb{E}_\alpha \left[ \frac{W (T - T_{tr})}{T} \sum_{i=1}^{n_{a}} p_i(\alpha) \right]. \tag{5}
\]

C. Energy efficiency

In a general sense, efficiency can be seen as the ratio of goods produced to the resources consumed. Here, we focus on the EE in the physical and medium access control layers of downlink transmission at a cell of a mobile cellular network. The results can also be applied to the uplink transmission. The goods are effective data transmitted from the base station to the mobile receivers (measured in bits or nats) and the resources are the total energy (Joule) consumed for transmitting the data. For a period of \( t \), the amount of effective transmitted data is given by \( t \cdot R_t \), where \( R_t \) represents the average data rate in this period. The variation of the rate within the period depends on the instantaneous channel realisation, which varies with time. Similarly, the amount of consumed energy is given by \( t \cdot P_t \), where \( P_t \) is the average total power used in this period. Here, the variation of the power can also depend on the channel realisation, if a policy is chosen such that it adapts to the channel state. EE is then the ratio of the average data rate to the average consumed power \( \overline{R_t} / \overline{P_t} \). We assume \( t \) is large such that \( \overline{R_t} \) and \( \overline{P_t} \), which are averaged over time, can be represented by the ergodic rate \( \overline{R_\alpha} \) and power \( \overline{P_\alpha} \), which are averaged over channel realisations. Next, we present the power budget and then describe the explicit expression for EE.

As an extension of the power model for base stations given in [12], we include the power used for training and write the average power budget as \( \overline{P_\alpha} = P_T + P_R \), where

\[
P_T = p_{C,sta}^T + n_T W p_{C, cir}^T + \frac{T_{tr}}{T} n_T R_T n_R \overline{P_{C, tr}},
\]

\[
+ \frac{W (T - T_{tr})}{T} \mathbb{E}_\alpha \left[ \sum_{i=1}^{n_{a}} p_i(\alpha) \right], \tag{6}
\]

\[
P_R = p_{C,sta}^R + n_R W p_{C, cir}^R + \frac{T_{tr}}{T} n_T R_T n_R \overline{P_{C, tr}}.
\]

\( P_T \) and \( P_R \) represents the average power consumption at the transmitter and receiver, respectively. The circuit power coefficients \( p_{C, cir}^T \) and \( p_{C, cir}^R \) are related to the circuitry at the transmitter and receivers, respectively, e.g. for filters, mixers, baseband signal processing and the analogue RF amplifier (see [4], [8], [16] for details). We assume that the actual power consumption in the circuitry grows linearly with the number of antennas at each side respectively and with the bandwidth used, since more antennas and larger bandwidth require more signal processing. Furthermore, we assume that a constant average power is expended for training during period \( T_{tr} \).
which scales with the number of antennas on both sides due to the increasing amount of feedback information but not with \( W \) since we have a flat-fading channel. The scaling magnitude for training is determined by the factors \( p_{C,\text{tr}}^R \) and \( p_{C,\text{tr}}^T \) for the transmitter and receiver, respectively. We assume that both \( T\frac{\epsilon}{R} n T n R p_{C,\text{tr}}^R \) and \( T\frac{\epsilon}{R} n T n R p_{C,\text{tr}}^T \) do not exceed the power constraints, if any are imposed.

The base station also constantly consumes power that does not depend on the transmit power, the number of antennas nor the bandwidth due to equipments such as power supply losses and climate control. We quantify this as \( P_{C,\text{sta}}^R \). Though it is often negligible at the mobile equipment, we account for a similar kind of power consumption at the receiver as \( P_{C,\text{sta}}^R \) for completeness’ sake. The last term in \( P^T \) represents the power used for data transmission. The coefficient \( \eta \) is expressed in (8).

With this power model and the average rate in (3), we quantify the EE as the ratio of the average rate to the average power w.r.t. the vector function \( \mathbf{p} (\alpha) \):

\[
EE [\mathbf{p} (\alpha)] = \frac{\epsilon E_{\alpha} [\sum_{i=1}^{n} \log (1 + \rho\alpha_i p_i (\alpha))] }{p_C + E_{\alpha} [\sum_{i=1}^{n} p_i (\alpha)]},
\]

where the constant

\[
p_C = \frac{T \epsilon (p_{C,\text{sta}} + n T W p_{C,\text{cir}} + \frac{T\epsilon}{T} n T n R p_{C,\text{tr}})}{W (T - T_{tr})},
\]

with \( p_{C,\text{sta}} = p_{C,\text{sta}}^T + p_{C,\text{sta}}^R \), \( p_{C,\text{cir}} = p_{C,\text{cir}}^T + p_{C,\text{cir}}^R \) and \( p_{C,\text{tr}} = p_{C,\text{tr}}^T + p_{C,\text{tr}}^R \). For simplicity, we henceforth call the mean in the numerator the average SE and the mean in the denominator is called the average transmitted power per unit bandwidth, although to be correct, the factor \( \frac{T - T_{tr}}{T} \) should be multiplied to them. Observe that EE increases with the power amplifier efficiency and decreases with training time, \( p_{C,\text{sta}}^R \) and \( p_{C,\text{cir}}^R \). Also if \( p_{C,\text{cir}}^R \) is small, EE increases with the bandwidth.

III. PROBLEM FORMULATION AND ANALYSIS

A. Optimal power control with link adaptation

Suppose an average sum transmission power constraint per unit bandwidth \( P_{\text{max}} \) is imposed. Given an arbitrary channel state PDF \( f (\alpha) \), we are interested in maximising the EE in (7). The problem is a nonlinear fractional program (see [17] for details):

\[
\max_{\mathbf{p} (\alpha)} \frac{\epsilon E_{\alpha} [\sum_{i=1}^{n} \log (1 + \rho\alpha_i p_i (\alpha))] }{p_C + E_{\alpha} [\sum_{i=1}^{n} p_i (\alpha)]}
\]

subject to \( 0 \leq p_i (\alpha) \) \( \forall i \), \( E_{\alpha} [\sum_{i=1}^{n} p_i (\alpha)] \leq P_{\text{max}} \).

Since the numerator in the objective function is strictly concave in \( \mathbf{p} (\alpha) \) and the denominator is linear (i.e. also convex) in \( \mathbf{p} (\alpha) \), the objective function is strictly pseudoconcave (see [18]–[20] for details). This implies that the optimum can be found at the unique function \( \mathbf{p}^* (\alpha) \). And because the constraint set is convex, it also implies that the KKT conditions are sufficient for optimality. The Lagrangian function is given by

\[
L = \epsilon E_{\alpha} [\sum_{i=1}^{n} \log (1 + \rho\alpha_i p_i (\alpha))] + \frac{\epsilon}{p_C + E_{\alpha} [\sum_{i=1}^{n} p_i (\alpha)]} \int_0^\infty \sum_i \lambda_t (\alpha) p_i (\alpha) d\alpha + \nu (P_{\text{max}} - E_{\alpha} [\sum_{i=1}^{n} p_i^* (\alpha)]),
\]

where \( \lambda_t (\alpha) \) for all \( is \) and all \( \alpha \in \mathbb{R}_+^{n_i} \) and \( \nu \) are Lagrange multipliers. The stationarity condition is derived by setting the functional derivative of (10) to zero, i.e. \( \frac{\partial L}{\partial p_{\alpha}} = 0 \), can be rearranged as

\[
p_k^* (\alpha) = \frac{\epsilon}{(\nu^*- \frac{\lambda^*_t (\alpha)}{f (\alpha)}) P_s^* + E E^* - \frac{1}{\rho\alpha_k}},
\]

for each \( k \), where \( EE^* = \epsilon E_{\alpha} [\sum_{i=1}^{n} \log (1 + \rho\alpha_i p_i^* (\alpha))] \) and \( P_s^* = p_C + E_{\alpha} [\sum_{i=1}^{n} p_i^* (\alpha)] \).

The solution to the KKT conditions can be summarised as the following function:

\[
p_k^* (EE^*, \alpha) = \left( \frac{\epsilon}{\nu^* p_s^* + EE^* - \frac{1}{\rho\alpha_k}} \right)^+,
\]

where \((x)^+ = \max \{0, x\}\) for all \( ks \). Notice that this function is only explicitly dependent on the component \( \alpha_k \) and not on the vector \( \alpha \). Notice also that the optimal vector function as the water-filling solution.

We first check whether \( \nu^* = 0 \), which would eliminate \( P_s^* \) in (12). If there is no sum power constraint, \( \nu^* = 0 \) automatically holds. Otherwise, this is true only if the sum power constraint

\[
E_{\alpha} \left[ \sum_{i=1}^{n} \left( \frac{\epsilon}{EE^* - \frac{1}{\rho\alpha_i}} \right)^+ \right] \leq P_{\text{max}}
\]

is fulfilled while the following holds:

\[
EE^* = \epsilon E_{\alpha} \left[ \sum_{i=1}^{n} \log (1 + \rho\alpha_i (\frac{\epsilon}{EE^* - \frac{1}{\rho\alpha_i}})^+) \right] \left( p_C + E_{\alpha} \left[ \sum_{i=1}^{n} \left( \frac{\epsilon}{EE^* - \frac{1}{\rho\alpha_i}} \right)^+ \right] \right).
\]

If true, we can compute the optimal \( EE^* \) in (14) by finding the root of the function

\[
F (EE^*) = \epsilon E_{\alpha} \left[ \sum_{i=1}^{n} \log (1 + \rho\alpha_i p_i^* (EE^*, \alpha_i)) \right] - EE^* \left( p_C + E_{\alpha} \left[ \sum_{i=1}^{n} p_i^* (EE^*, \alpha_i) \right] \right) = 0,
\]

treating \( EE^* \) as a variable. According to [21], \( F \) is shown to be a strictly monotonic decreasing function of \( EE^* \). Furthermore, it can be solved efficiently using the Newton algorithm.

\[1\]Note that a positively weighted integral (i.e. also the mean) of a convex function is convex.
This optimisation procedure is similar to Dinkelbach’s algorithm and is shown to have a superlinear convergence rate [22].

However, if (13) is not fulfilled with (14), we can conclude from the complementarity condition that $\nu^* > 0$, implying that the sum power constraint is active. Using the variable $\xi^*$ to represent $\nu^* P^*_i + EE^*$, we then compute $\xi^*$ that satisfies the constraint

$$
E_{\xi^*} \left[ \sum_{i=1}^{n_i} \left( \frac{\xi^*}{\rho_{ai}} - 1 \right)^{+} \right] - P_{\text{max}} = 0, \quad (16)
$$
as in (15). The corresponding $EE^*$ is given by

$$
EE^* = \frac{\mathbb{E}_{\xi^*} \left[ \sum_{i=1}^{n_i} \log \left( 1 + \rho_{ai} \left( \frac{\xi^*}{\rho_{ai}} - 1 \right)^{+} \right) \right]}{pc + E_{\xi^*} \left[ \sum_{i=1}^{n_i} \left( \frac{\xi^*}{\rho_{ai}} - 1 \right)^{+} \right]},
$$

Note that $\frac{\xi^*}{\rho_{ai}}$ corresponds to the cut-off value. This means that the data stream $i$ with $\alpha_i \leq \frac{\xi^*}{\rho_{ai}}$ will not be supported in the transmission.

**Remark 1.** With an active sum power constraint, this solution yields the same SE and power allocation as that obtained by maximising the SE for a given $\text{tr} \, Q = P_{\text{max}}$. In this case, maximising the EE is identical to maximising the SE, as discussed in [20].

**Remark 2.** Since $EE^*$ only depends on $n_i$, interchanging $n_T$ and $n_R$ will not affect the results, provided that the channel distribution is not altered by this interchange.

**B. Optimal power control with Rayleigh fading**

In the previous section we derived the solution to the power control problem with an arbitrary channel distribution $f(\alpha)$. Here we investigate the EE particularly for Rayleigh fading without spatial antenna correlation. An extension to Rayleigh fading with antenna correlation is easily possible (see [13] for details). In this setting each component of $\mathbf{H}$ is a complex number with independent real and imaginary parts that are i.i.d. zero-mean Gaussian random variables with variance 1/2 ($H_{ij} \sim \mathcal{CN}(0,1)$). For our problem we need the distribution of $\alpha$ which are eigenvalues of the Wishart matrix $H^*H$ [23]. For sorted eigenvalues ($\alpha_1 \geq \ldots \geq \alpha_{n_T}$), this distribution is given by

$$
f(\alpha) = K \left| \mathbf{V}(\alpha) \right|^2 \prod_{i=1}^{n_T} e^{-\alpha_i} \alpha_i^{n_a-n_i}, \quad (17)
$$

where $K = 1/\prod_{i=1}^{n_T} (n_i-i)! (n_a-i)!$ and $\mathbf{V}(\alpha)$ is the Vandermonde matrix with its components given by $V_{ij} = \alpha_i^{j-1}$ [7], [23]. For a distribution with unsorted eigenvalues, $f(\alpha)$ is simply multiplied by a factor of $1/n_T!$.

**Remark 3.** Note that (17) only depends on $n_i$ and $n_a$. This implies that it yields the same distribution even if $n_T$ and $n_R$ are interchanged.

Using Theorem 3 in [13], we can write the average SE at the optimal EE as

$$
\mathbb{E}_{\alpha} \left[ \sum_{i=1}^{n_i} \log \left( 1 + \rho_{ai} \left( \frac{\epsilon}{\xi^*} - 1 \right)^{+} \right) \right] = \int_0^\infty \frac{K}{n_i} \left| \mathbf{V}(\alpha) \right|^2 \prod_{i=1}^{n_i} e^{-\alpha_i} \alpha_i^{n_a-n_i} \sum_{i=1}^{n_i} \phi(\xi^*, \alpha_i) \, d\alpha,
$$

$$
= K \sum_{i=1}^{n_i} \det \left( \left\{ \int_0^\infty e^{-\alpha_i U_{i,l}(\phi(\xi^*, \alpha))} \, d\alpha \right\}_{k,l=1,\ldots,n_i} \right), \quad (18)
$$

where $\phi(\xi^*, \alpha) = \log \left( 1 + \rho_{ai} \left( \frac{\epsilon}{\xi^*} - 1 \right)^{+} \right)$ and $U_{i,l}(x)$ is defined as

$$
U_{i,l}(x) = \begin{cases} x, & \text{if } i = l \\ 1, & \text{if } i \neq l. \end{cases}
$$

The expression (18) is advantageous for numerical computation since it only involves integration of one variable. This significantly reduces computation time in evaluating (15) in comparison with Monte-Carlo integration. Similarly, the average total transmission power per unit bandwidth $\mathbb{E}_{\alpha} \left[ \sum_{i=1}^{n_i} \left( \frac{\xi^*}{\rho_{ai}} - 1 \right)^{+} \right]$ can be evaluated by replacing $\phi(\xi^*, \alpha)$ with $\tilde{\phi}(\xi^*, \alpha) = \left( \frac{\xi^*}{\rho_{ai}} - 1 \right)^{+}$ in (18).

**C. Static power allocation with Rayleigh fading**

Suppose that the base station is configured to allocate constant transmission power for all times, i.e. with no channel state adaptation. If provided with channel statistics, we can calculate a constant power allocation that maximises the EE. The programming problem is given by

$$
\max_{0 \leq \text{tr} \, Q \leq P_{\text{max}}} \frac{\mathbb{E}_{\mathbf{H}} \left[ \log \det (\mathbf{I}_{n_R} + \rho \mathbf{H}^* \mathbf{H}) \right]}{pc + \text{tr} \, Q}. \quad (19)
$$

We approach this problem by using an optimisation technique given in [20]. The problem above can be reformulated as

$$
\max_{0 \leq P \leq P_{\text{max}}} \frac{\mathbb{E}_{\mathbf{H}} \left[ \log \det (\mathbf{I} + \rho \mathbf{H}^* \mathbf{H}) \right]}{pc + P},
$$

As shown in [7], the solution to the inner maximisation is given by $Q = \frac{P}{n_T} \mathbf{I}_{n_T}$. We thus have

$$
\max_{0 \leq P \leq P_{\text{max}}} \frac{\mathbb{E}_{\alpha} \left[ \sum_{i=1}^{n_i} \log \left( 1 + \rho_{ai} \frac{P}{n_T} \right) \right]}{pc + P}. \quad (20)
$$

The mean in the numerator represents the SE. It can also be shown that the objective function is strictly pseudoconcave (see [20] for details). This implies that the optimum is found at a unique point $P^*$. We can find this point using algorithms such as bisection or interior-point methods. For Rayleigh fading, the SE can be efficiently computed during the optimisation procedure using Theorem 3 in [13] as in Section III-B by utilising the expression (18), replacing $\phi(\xi^*, \alpha)$ with $\phi = \log \left( 1 + \rho_{ai} \frac{P}{n_T} \right)$. 

IV. Simulations and Discussion

In our simulations we compute the solutions to the optimal link adaptation (9) (Optfun) and the static power allocation (19) (StaOpt) with Rayleigh fading for the following settings: (a) $n_T = n_R = 1, \ldots, 8$; (b) $n_T = 8, n_R = 1, \ldots, 8$; (c) $n_R = 8, n_T = 1, \ldots, 8$; (d) $n_R = 2, n_T = 1, \ldots, 8$. No sum power constraint is imposed so that we can observe the maximum achievable EE. We assume that $\gamma = 1$, such that $\rho = 1/\sigma^2$, and $\epsilon = 1$. We repeat the simulations for $\rho = 20$ dB and 10 dB. Furthermore, we assume that $p_{C,sta} + p_{C,sta}^R \gg n_T p_{C,cir} + T_0 n_T n_R p_{C,ir} + n_T p_{C,cir} + T_0 n_T n_R p_{C,ir}$ such that the influence of $n_T$ and $n_R$ on $p_C$ is negligible. We then set $p_C = 1$ mW/kHz. The optimal EE achieved by Optfun are shown in the blue lines in Fig. 1 and Fig. 2 for $\rho = 20$ dB and 10 dB, respectively. The abscissa shows various values of $n_T = n_R$ for (a), $n_T$ for (b), and $n_T$ for (c) and (d). The EE achieved by StaOpt are denoted by red symbols without lines in the same figures. The corresponding average SE and average total transmission power per unit bandwidth $E_a[\sum_{i=1}^n \gamma_i(\alpha)]$ for $\rho = 20$ dB are shown in Fig. 3 and Fig. 4, respectively. The results for $\rho = 10$ dB for these behave similarly and are therefore omitted.

The EE achieved through Optfun is always superior, if not equal, to that of StaOpt, as expected. We can also observe that it is always better to have more antennas for better EE. Note that this conclusion is made assuming that the number of antennas has no influence on $p_C$. This may not be true for small $p_{C,sta} + p_{C,sta}^R$, e.g. in the case of wireless sensor networks [8] where climate control and power supply are absent. Notice that the result is the same for Optfun even if $n_T$ and $n_R$ are interchanged (compare (b) and (c)), as stated in Remark 2, whereas in StaOpt it is better to have more antennas at the receiver. When $n_R = 8$, we see that for $\rho = 20$ dB the EE of StaOpt comes close the that of Optfun. No difference can be seen for $n_T < 6$. Therefore, it may even be advantageous to apply static power allocation for $n_R \gg n_T$ and high $\rho$, where we can save energy reducing feedback from the receiver to the transmitter. For $n_R = 2$, the EE gain of Optfun in comparison with StaOpt increases significantly as $n_T$ increases. We can conclude that for $n_R > n_T$ ($n_R < n_T$, resp.), as $n_R - n_T$ increases, the difference between the EE of Optfun and StaOpt decreases (increases). This difference is further decreased (increased) as $\rho$ is increased.

It is well-known that when $n_T = n_R = n$, the capacity increases linearly with $n$. We see a similar trait here for both the EE and the average rate in both optimisation problems. We can approximate this behaviour of the EE by the following simple relation: $EE = n\bar{R}(P)/\rho p_C$, where $n\bar{R}(P)$ is the function that yields the maximum ergodic rate (either through Optfun or StaOpt) when applying a total average transmission power $P$, and $\bar{R}(P)$ is the rate function for $n = 1$. For any $n$, maximising this $EE$ would yield the same $P^*$ since $n$ is only a factor in the objective function. We observe in Fig. 4 that the transmission power allocation is almost constant over $n = n_T = n_R$, which supports this deduction.

Using the simple approximation for $n_T = n_R$, we would like to investigate the behaviour of EE when the influence of the number of antennas on $p_C$ is not negligible. Following the
model in Section II-C, we define \( p_c = a + bn + cn^2 \), where \( a = \frac{W T \rho}{P_{\text{max}}}, b = \frac{W T \rho}{P_{\text{cir}}}, \) and \( c = \frac{W T \rho}{P_{\text{tr}}} \). Assume the optimal \( P^* = P_{\text{max}} \) is not affected by any change in \( p_c \). This occurs e.g. when a relatively low average sum power constraint \( P_{\text{max}} \) is imposed, making it likely for the constraint to be active. We would like to find the optimum \( n \) that maximises the EE. We assume that \( n \) is a real positive number. The objective function \( EE = \frac{n R(P_{\text{max}})}{a + bn + cn^2 + P_{\text{max}}} \) which we want to maximise w.r.t. \( n \) is a strictly pseudoconcave function in \( n \), since its numerator is linear (i.e. also concave) and its denominator is strictly convex in \( n \). Thus, its maximum is given through the stationary point where \( \frac{dEE}{dn} \bigg|_{n=n^*} = 0 \). This yields \( n^* = \sqrt{\frac{a+P_{\text{max}}}{c}} \). Because \( n \) has to be an integer, we choose the \( n \) closest to \( n^* \). This result implies the following. On the one hand, the larger the \( p_{\text{C,STA}} \), the more antennas should be employed. On the other hand, the larger the power used for training, the smaller \( n^* \) should be. \( n^* \) should also increase with \( P_{\text{max}} \) but decrease with power amplifier efficiency \( c \). It is interesting to note that \( n^* \) is independent of the circuit power and the bandwidth, which are both represented by \( b \). The analysis of a joint optimisation over both \( n \) and \( P \), i.e. when no power constraint is applied, is left for future work. Additionally, the case where \( n_T \neq n_R \) still needs investigation.

V. CONCLUSION

We present the optimal power control function that maximises the EE in base station utilising multiple antennas for an arbitrary time-varying channel. We applied the result to Rayleigh fading channels and compared it with a suboptimal technique, i.e. the static power allocation. We also discussed the relationship between antenna configuration and EE. Further work includes a joint optimisation over the number of antennas and power allocation.

ACKNOWLEDGEMENT

The authors would like to thank our project partners and Christian Isheden for helpful discussions. This work is supported by the German Federal State of Saxony in the excellence cluster Cool Silicon within the frame of the project Cool Cellular under grant number 14056/2367.

REFERENCES